Impersonal Constructions
Studies in Language Companion Series (SLCS)

This series has been established as a companion series to the periodical Studies in Language.

For an overview of all books published in this series, please see http://benjamins.com/catalog/slcs

Editors

Werner Abraham
University of Vienna

Elly van Gelderen
Arizona State University

Editorial Board

Bernard Comrie
Max Planck Institute, Leipzig
and University of California, Santa Barbara

William Croft
University of New Mexico

Östen Dahl
University of Stockholm

Gerrit J. Dimmendaal
University of Cologne

Ekkehard König
Free University of Berlin

Christian Lehmann
University of Erfurt

Brian MacWhinney
Carnegie-Mellon University

Marianne Mithun
University of California, Santa Barbara

Heiko Narrog
Tohoku University

Johanna L. Wood
University of Aarhus

Volume 124

Impersonal Constructions. A cross-linguistic perspective
Edited by Andrej Malchukov and Anna Siewierska
Impersonal Constructions

A cross-linguistic perspective

Edited by

Andrej Malchukov
Max Planck Institute for Evolutionary Anthropology, Leipzig
Institute for Linguistic Studies, St. Petersburg
Johannes Gutenberg University, Mainz

Anna Siewierska
University of Lancaster

John Benjamins Publishing Company
Amsterdam / Philadelphia
Library of Congress Cataloging-in-Publication Data
p. cm. (Studies in Language Companion Series, ISSN 0165-7763 ; v. 124)
Includes bibliographical references and index.
P293.5.I47 2011
415.6--dc22 2011000415
ISBN 978 90 272 0591 9  (Hb ; alk. paper)
ISBN 978 90 272 8716 8  (Eb)

© 2011 – John Benjamins B.V.
No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands
John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118·0519 · USA
# Table of contents

List of contributors .......................... vii  
Introduction .................................. 1  
   *Andrej Malchukov & Anna Siewierska*  

## Part I. Impersonal constructions: Typological and theoretical aspects  
Towards a typology of impersonal constructions:  
A semantic map approach .................... 19  
   *Andrej Malchukov & Akio Ogawa*  
Overlap and complementarity in reference impersonals:  
Man-constructions vs. third person plural-impersonals  
in the languages of Europe .................. 57  
   *Anna Siewierska*  
Verbs of motion: Impersonal passivization between  
unaccusativity and unergativity ............ 91  
   *Werner Abraham*  
On the distribution of subject properties in formulaic presentationals  
of Germanic and Romance: A diachronic-typological approach ...................... 127  
   *Volker Gast & Florian Haas*  

## Part II. Impersonal constructions: Diachronic studies  
Impersonal constructions and accusative subjects in Late Latin ............ 169  
   *Michela Cennamo*  
From passive to impersonal: A case study from Italian and its implications  
189  
   *Anna Giacalone Ramat & Andrea Sansò*  
Passive to anticausative through impersonalization:  
The case of Vedic and Indo-European ........ 229  
   *Leonid Kulikov*
PART III. Cross-linguistic variation in Impersonal constructions: Case studies

The Maa (Eastern Nilotic) Impersonal construction 257
   Doris L. Payne

Impersonal constructions in Jóola-Banjal 285
   Alain-Christian Bassène & Denis Creissels

Impersonal configurations and theticity: The case of meteorological predications in Afroasiatic 307
   Amina Mettouchi & Mauro Tosco

Revisiting impersonal constructions in Modern Hebrew: Discourse-based perspectives 323
   Ruth A. Berman

The elephant in the room: The impersonal -ne/-te construction in Polish 357
   Anna Kibort

Meteorological verbs in Uralic languages – are there any impersonal structures to be found 395
   Merja Salo

Impersonal constructions in Ket 439
   Edward Vajda, Andrey Nefedov & Andrej Malchukov

Impersonal verbs in Central Alaskan Yupik (Eskimoan) 459
   Osahito Miyaoka

Impersonals in Innu 489
   Lynn Drapeau

A diachronic study of the impersonal passive in Ainu 517
   Anna Bugaeva

Referential impersonal constructions in Mandarin 547
   Yi Yan & Anna Siewierska

Impersonal constructions in some Oceanic languages 581
   Claire Moyse-Faurie

Impersonal constructions in Umpithamu and the Lamalamic languages 607
   Jean-Christophe Verstraete

Author index 627
Language index 633
Subject index 637
List of contributors

Werner Abraham
Universität Wien 1090, Berggasse
11/3 Allg. Sprachwissenschaft
Studies in Language/LA/SLCS/SDG
home: Lindwurmstrasse 120c,
D-80337 München
werner_abraham@t-online.de

Alain Christian Bassène
Département de linguistique
Université Cheikh Anta Diop
B.P. 5005 Dakar-Fann
Dakar
Sénégal
Alain-Christian.Bassene@univ-lyon2.fr

Ruth A. Berman
Department of Linguistics
Tel Aviv University
Ramat Aviv, Israel 69978
rberman@post.tau.ac.il

Anna Bugaeva
Waseda Institute for Advanced Study,
Waseda University
1-6-1 Nishi-Waseda
Shinjuku-ku, Tokyo
169-8050, Japan.
bugaeva@japan.email.ne.jp

Michaela Cennamo
Department of Modern Philology
University of Naples Federico II
Via Porta di Massa 1
I-80133 Naples
Italy
micennam@unina.it

Denis Creissels
DDL ISH
14 avenue Berthelot
69363 Lyon Cedex 07
France
Denis.Creissels@univ-lyon2.fr

Lynn Drapeau
Département de linguistique
Bureau DS-4531 Uqam
C.P. 8888, Succ. CV
Montréal (QC)
Canada H3C 3P8
drapeau.lynn@uqam.ca

Volker Gast
Friedrich-Schiller-Universität Jena
Department of English and American Studies Ernst-Abbe-Platz 8
07743 Jena
volker.gast@uni-jena.de
Anna Giacalone Ramat
Dipartimento di Linguistica
Teorica e Applicata
Università di Pavia
Corso Strada Nuova 65
I-27100
Pavia, Italy
annaram@unipv.it

Florian Haas
Friedrich-Schiller-Universität Jena
Department of English and American Studies
Ernst-Abbe-Platz 8
07743 Jena
florian.haas@uni-jena.de

Anna Kibort
University of Cambridge
Department of Linguistics
Faculty of Modern and Medieval Languages
Sidgwick Avenue
Cambridge CB3 9DA
UK
A.Kibort@surrey.ac.uk

Leonid Kulikov
Faculty of Humanities,
Institute of Linguistics (LUCL)
PO Box 9515
2300 RA Leiden, The Netherlands
L.Kulikov@hum.leidenuniv.nl

Andrej Malchukov
Max Planck Institute of Evolutionary Anthropology
Department of Linguistics
Deutscher Platz
04103 Leipzig
Germany
andrej_malchukov@eva.mpg.de

Amina Mettouchi
Professional
Ecole Pratique des Hautes Etudes
45-47 rue des Ecoles, 75005 Paris, France
home (use this one for correspondence)
14 rue Brossolette, 44400 REZE, France
Amina.Mettouchi@univ-nantes.fr

Osahito Miyaoka
1-7-809, Semiyama,
Fukiai-cho, Chuo-ku,
Kobe 651-0058,
Japan
miyaoka@cup.ocn.ne.jp

Claire Moyse-Faurie
8 rue Auguste Blanqui
94250 Gentilly,
France
moyse@vjf.cnrs.fr

Andrey Nefedov
Department of Siberian Languages,
TPGU, pr. Komsomolsky, 75,
Tomsk, 634041,
Russia
nefedov@eva.mpg.de

Akio Ogawa
Kwansei Gakuin University
School of Humanities
Uegahara 1-1-155
Nishinomiya
Japan 662-8501
akiogawa@kwansei.ac.jp
Doris L. Payne
Department of Linguistics
University of Oregon
Eugene, OR 97403
USA
dlpayne@uoregon.edu

Merja Salo
Department of Finnish, Finno-Ugrian and Scandinavian Studies
University of Helsinki
PO Box 24, FL-00014 Helsinki
Finland
mesalo@mappi.helsinki.fi

Andrea Sansò
Dipartimento di Linguistica Teorica e Applicata
Università di Pavia
Corso Strada Nuova 65
I-27100
Pavia, Italy
asanso@gmail.com

Anna Siewierska
Department of Linguistics and English Language
Lancaster University
Lancaster, LA1 4YT
United Kingdom
A.Siewierska@lancaster.ac.uk

Mauro Tosco
Department of Oriental Studies
University of Turin
via Giulia di Barolo, 3A
I-10124 Torino
Italy
mauro.tosco@unito.it

Edward Vajda
Department of Modern and Classical Languages
Western Washington University
Bellingham, WA 98225
USA
edward.vajda@wwu.edu
eddievajda@yahoo.com

Jean-Christophe Verstraete
Department of Linguistics
University of Leuven
Blijde-Inkomststraat 21,
PO Box 3308
3000 Leuven
Belgium
Jean-Christophe.Verstraete@arts.kuleuven.be

Yi Yan
National Center for Foreign Language Education
Beijing Foreign Studies University
Beijing, Xi San Huan Bei Lu No. 2.
People’s Republic of China
www.yiyan@yahoo.com.cn
Introduction

Andrej Malchukov & Anna Siewierska
Max Planck Institute for Evolutionary Anthropology, Leipzig & Institute for Linguistic Studies, St. Petersburg & Johannes Gutenberg University, Mainz/University of Lancaster

1. Introductory remarks

While impersonal constructions (such as Latin *Me pudet* lit. ‘me shames’, German *Mich friert* lit. ‘me freezes’, or Russian *Svetaet* ‘It dawns’) have been a regular topic of investigation in Indo-European studies (see, e.g. Seefranz-Montag 1983; Lambert 1998; Bauer 2000; Barðdal 2004; Siewierska 2008), they have not been the subject of detailed cross-linguistic research (the work of Lehmann et al. 2000 is a notable exception). The lack of cross-linguistic studies of impersonals may be attributed to the difficulties involved in identifying impersonal constructions on a cross-linguistic basis. These difficulties stem from different interpretations of the term impersonal, as well as from the heterogeneity, both semantic (weather verbs, experiential predicates, presentational constructions) and structural (basic impersonals vs. impersonal passives) of impersonal constructions.

The volume is intended to remedy the above situation by bringing together scholars interested in various aspects of the structure of impersonal constructions, viewed broadly as constructions lacking a referential subject, both in individual languages and cross-linguistically. The contributions address different issues in the typology of impersonal constructions across a wide range of languages. Care has been taken to ensure that the selected contributions (some stemming from a Societas Linguistica Europea 41 workshop organized by the editors in Forli in September 2008, others specially invited) are representative of the structural and genealogical diversity across languages in the impersonal domain. Thus, the present volume is unprecedented in bringing a typological perspective to the topic of impersonal constructions. It is expected that it will be of central interest to all scholars and advanced students of linguistics, especially to those working in the field of language typology and comparative syntax.

2. Some conceptual considerations

As noted above, the study of impersonal constructions has not been attempted so far in language typology, in spite of the popularity of this topic in Indo-European Studies (starting from classic work in 19th century by Herman Paul, Aleksandr Potebnja,
among others). This has to do with the heterogeneity of the constructions noted above, but also with the lack of consensus concerning the definition of impersonal constructions. As discussed in Siewierska 2008 impersonality has been traditionally viewed from two partially overlapping perspectives, a communicative-functional one which defines impersonalization in terms of agent defocusing/backgrounding, and a more structure based perspective which associates impersonality with the lack of a (referential) subject. The subject-based approach is the narrower of the two in that it conceives of impersonality as involving elements of or operations on argument structure. Under the agent-defocusing approach on the other hand impersonality is conceived of more widely as involving speaker choice with respect to the construal of an event and is regarded as sensitive to the effects of discourse. Under the subject-centred view the constructions which are considered to be impersonal may be grouped into four broad types: (a) those with an argumental subject which is not fully referential, (b) those with a subject which does not display canonical subject properties, (c) those with a subject which is not a verbal argument but merely a place filler manifesting no semantic or referential properties, i.e. an expletive subject, and (d) those with no overt subject at all. From the semantic perspective, the type of constructions involved are those featuring non-referential subjects, those depicting meteorological phenomena, those expressing sensations, emotions, need, potential and other modalities and existential and presentative constructions. The agent-defocusing approach adds to the domain of impersonality constructions in which an argument other than the agent has been selected for subject in preference to the agent such as personal passives (*The manuscript was sold for 100.000 pounds (by an unknown collector)*)) and locative subject clauses (e.g. *The garden is swarming with bees*). Note that personal passives, unlike so-called impersonal passives, do possess a subject and one which typically displays the full set of subject properties in a language and as such cannot be seen as impersonal under the subject-centred view. The agent-defocusing approach also identifies as being impersonal constructions which depict events logically involving agents but construed as lacking them such as anticausatives (e.g. *The vase broke*) or action nominalizations (e.g. *the circling of the camp*). Given that the presence of agent-defocusing may be discernible only in a wider discourse-pragmatic context, as is the case with the use of existentials in European Portuguese (Afonso 2008), this approach is more amenable to language-specific, discourse-based investigations than typological studies. It should therefore come as no surprise that most contributions to this volume take the structure based approach as their starting point. Nonetheless several contributions to the volume, most notably the typological study by Malchukov & Ogawa and the descriptive study by Payne, try to combine both approaches to arrive at a new synthesis.

Apart from this major conceptual divide there are more controversial issues around impersonal constructions (and the notion of subject on which these constructions are predicated) which are tied to specific theoretical approaches or models of
Introduction

Thus, for example, within the mainstream generative tradition of particular relevance to the subject-based notion of impersonality has been Chomsky’s (1981) famous taxonomy of syntactic zeros (pro, PRO, NP-trace; operator-trace) and its subsequent extensions dealing with expletive zero subjects, EPP-violations, non-canonical subjects etc. as reflected most recently in the papers in Svenonius (2002), the work of Mendikoetxea (2008) and especially the analyses presented in Biberaruer et al. (2010). Similar accounts have been developed in other frameworks such as Lexical Functional Grammar (LFG) (Kibort 2008), Head-Driven Phrase Structure Grammar (HPSG) (Blevins 2003), Optimality Theory (OT) (Grimshaw & Samek-Lodovici 1998), and OT-LFG Bresnan (2001). An original approach to impersonal constructions has been developed by Melčuk (e.g. Melčuk 1979) embedded in his Meaning-Text Theory. It postulates several zero elements (zero lexemes) including Ø(elements) and Ø (people) to account for the wide variety of impersonal constructions in Slavic (see Zimmerling 2007 for a cross-theoretical discussion). Highly interesting are also the OT approaches to impersonality as they view the emergence of impersonal constructions as the result of the interaction of different violable constraints. Thus, for example, the emergence of a dummy subject in constructions with meteo-verbs which do not license arguments is attributable to the higher ranking of the constraint penalizing subjectless clauses over the constraint requiring faithful encoding of roles/arguments. Such an account can be translated into a functional-typological framework, if the postulated constraints can be shown to have a functional basis. Within the functional-cognitive paradigm the preference has been for the broader functional definition which takes its origin in approaches like that of Keenan’s functional decomposition of subjects (see Malchukov & Ogawa for discussion).

The controversies surrounding the notion of zero-element which have played such a dominant role in formal grammatical frameworks have been primarily driven by theoretical assumptions rather than language data (see also remarks in Creissels 2007). In a few cases where the generative approach explicitly addressed typological issues the results have been rather inconclusive disappointing (see Newmeyer’s 2005 critical discussion of the pro-drop parameter). Yet there seems to be a renewed interest in explicating the empirical consequences of certain assumptions and in applying them to different languages reflected especially in the work of Holmberg (2005, 2010 and Biberauer et al. (2010). One area which has been intensively explored from this perspective is impersonal passivization, the different instantiations of which present challenges to some basic assumptions of generative grammar, such as Burzio’s (1986:178) generalization (see Abraham 2006, this volume et passim for a critical review of the literature on impersonal passives). One reaction, to the empirical challenges has been a proposal to reclassify some of the impersonal passive constructions (those which derive from unergatives and unaccusatives indiscriminately), as impersonals proper (or impersonal active constructions); see (Blevins 2003). While for some languages this reanalysis is
justified, it needs to be recognized that the distinction between the two types is gradient rather than clear cut (cf. also the underspecification analysis of Kibort, this volume).

The existing functional-typological approaches to impersonal constructions are few in number: they either take a very broad function-based definition of impersonalization, or else are focused on some specific varieties of impersonal constructions, as these constitute a more concrete point of departure for function-based typological studies. The impersonal constructions which have received the greatest amount of attention are arguably reference based impersonals (see e.g. Afonso 2007; Langacker 2006; Myhill 1997; Słon 2003) and so-called dative subject or psych-verb constructions (e.g. Barbådal 2004; Divjak & Janda 2008). Given the heterogeneity of the impersonal domain, there is no question that individual subdomains such as the above or meteor-verbs or presentational constructions need to be studied in their own right. However, as shown in Malchukov & Ogawa (this volume), it is also the case that these subvarieties of impersonals are not orthogonal to each other but rather are united by family resemblances spanning several domains, which the authors label A-impersonals, T-impersonals, and R-impersonals. (A-impersonals include varieties of experiential predicates, where a notional subject lacks agentivity; T-impersonals embrace presentational constructions with a notional subject deficient in topicality, and R-impersonals cover constructions with a notional subject deficient in referential properties, with a further subdivision into constructions with a non-referential subject of meteor-verbs as opposed to man-impersonals.) Further while the family resemblances among the different domains of impersonals have eluded a characterization in terms of a single general meaning approach or a single proto-type approach, Malchukov & Ogawa show that they can be insightfully captured by a semantic map (see also Payne, this volume, and Bassene & Creissels, this volume, for a similar approach).

Another problematic issue in typology concerns the applicability of the notion of impersonality to languages of different types, in particular to ergative languages (see Lazard 1998; Creissels 2007; Siewierska 2008). Here the general consensus seems to be that even for ergative languages lack of an A-argument rather than of a P-argument should be definitive of impersonality (for the latter the term anti-impersonal has been coined by Lazard). Indeed the opposite assumption would result in impersonal constructions in accusative and ergative languages being functionally incompatible since the semantic domain of “anti-impersonal” constructions is quite different from impersonal ones. Although this question has not been systematically followed up in contributions to the present volume, this discussion is echoed in a number of contributions involving ergative languages, in particular by Moyse-Faurie’s discussion of differential case marking in Oceanic languages, and Verstraete’s discussion of the passive-like behavior of certain ergative constructions in the Australian languages. The issue, is, however, broader and not confined to ergative languages, as it concerns the distribution of subject properties in some constructions in accusative languages.
as well, in particular those which display non-canonical subject-marking, but also in other cases labeled “covert impersonal constructions” by Creissels. In the present volume the question of the distribution of subject properties is explicitly addressed by Gast & Haas who discuss the distribution of subject properties between an expletive and postverbal NP in presentational structures and also in the contribution by Salo who considers the issue in relation to a variety of meteo-constructions in Uralic.

One recent trend in the study of impersonal constructions is the renewed interest in the diachronic paths leading to or from impersonal constructions. This has been a popular topic in Indo-European studies (Seefrantz-Montag 1983; Bauer 2000; Koch 2000), but has recently been also extended to other languages. In particular a number of the contributions to Donohue & Wichmann 2008 show that impersonal (or rather ‘transimpersonal’ constructions with a transitive verb taking an indefinite A argument) have played an important role in the rise of split intransitive patterns (see Malchukov 2008; cf. Mithun 2008; Holton 2008). Earlier a path from transimpersonal active to impersonal passive has been equally well documented (by Greenberg 1959; Shibatani 1985 and others), even though not all of the examples cited are upon closer inspection entirely unproblematic (see Siewierska 2010 for a typological approach). Yet another grammaticalization path, from locative inversion constructions via locative subject constructions to existential impersonals has been recently proposed by Croft (forthcoming). In the present volume, the contributions by Payne on Maa, Bugaeva on Ainu and to some extent also by Siewierska deal with the impersonal to passive path, the contributions by Vajda et al. on Ket and Miyaoka on Yupik deal with the impersonal to intransitive path, and the contribution by Gast & Haas touches upon the locative to existential developments.

The above diachronic analyses point to an interesting perspective on impersonals as transitional structures or an intermediate stage of a more basic diachronic change be it from transitive to intransitive, or from active to passive or participant to event-centered construction etc. To the best of our knowledge, such a view of the whole domain of impersonality, as opposed to specific types of impersonal constructions, has not been explicitly advanced before. Yet it provides one potential answer to the heterogeneity of the constructions within the domain of impersonality, on the one hand, and the functional, structural and semantic connections that can be forged among them, on the other. We contend that this transient view of the impersonal domain deserves further consideration.

3. Contributions to this volume

Given that the domain of impersonal constructions is highly complex and heterogeneous, the contributions to this volume address a wide variety of topics pertinent to
impersonal constructions. They also differ in empirical focus, perspective, theoretical orientation and range of languages considered. This diversity poses obvious problems for the structuring of the volume, as contributions can be classified and cross-classified along different dimensions.

We opted for a volume structure which opens with general theoretical and typological studies (chapters by Malchukov & Ogawa, Siewierska, Abraham, Gast & Haas), which is then followed by diachronic studies (chapters by Cennamo, Giacalone Ramat & Sansò, and Kulikov), and continues with a sample of case studies of impersonal constructions in individual languages (or language groups). The latter chapters are roughly ordered on geographical grounds, starting with African and Afro-Asiatic languages (Payne on Maa, Bassene & Creissels on Jóola-Banjal; Mettouchi & Tosco on Afro-Asiatic in general and Berman on Modern Hebrew), European or Eurasian (Kibort on Polish, Salo on Uralic, Vajda et al. on Ket), continued by American Indian languages (Miyayoka on Yupik, Drapeau on Innu), and then followed by the chapters from East-Asia (Bugaeva on Ainu, Yi & Siewierska on Mandarin) and finally the Pacific and Australia (Moyse-Faurie on Oceanic languages, and Verstraete on Australian Pama-Nyungan languages).

But even this division is somewhat arbitrary, given that many chapters offer an areal outlook (e.g. Moyse-Faurie on Oceanic languages, or Salo on Uralic), and many others make important general (theoretical) points. Further while those in the diachronic part of the volume mostly address Indo-European languages with a better documented history, diachronic issues also feature in other papers, most notably the previously mentioned papers in part three dealing with the reanalysis of transimpersonal constructions as well as the papers by Gast & Haas and Siewierska in part one.

Berman’s chapter on Modern Hebrew has a strong corpus orientation. Although many of the diachronic studies (by Giacalone Ramat & Sansò on Italian, Cennamo on Late Latin, and Kulikov on Vedic) of Indo-European languages with documented histories also make ample use of corpus data, Berman’s contribution is unique within the context of this volume in that it combines the corpus perspective with data from language acquisition (see also Berman 2005 for further discussion).

More importantly, individual contributions focus on different functional varieties of impersonal constructions. If we adopt the general distinction between R-impersonals, T-impersonals and A-impersonals (as characterized earlier), the contributions to the volume may be grouped as follows:

- Contributions which address the whole broad domain of impersonals: Malchukov & Ogawa, Moyse-Faurie, Bassene & Creissels, and Drapeau;
- Contributions on R impersonals of the meteo-type: Mettouchi & Tosco, Salo;
- Contributions on indefinite subject constructions: Siewierska, Yi & Siewierska, Giacalone Ramat & Sansò, partially also Berman;
– Contributions focusing on T-impersonals, namely subvarieties of presentational structures: Gast & Haas, Bassene & Creissels;
– Two types of contributions addressing A-impersonals, those mainly involving experiential predicates: Verstraete and partially Cennamo; and those centered on transimpersonal constructions at later stages of reanalysis: Miyaoka, Bugaeva and Vajda et al.

Another divide is between basic vs. derived impersonals including impersonal passives. The latter constitute the primary focus of the contributions by Abraham and Kibort, but they are also considered in the papers by Payne and Bugaeva, in which transimpersonal constructions are identified as the likely sources of impersonal passive structures. This also suggests that the basic vs. derived division is gradient rather than clear-cut (as also manifested in the controversial status – impersonal passive or active impersonal – of some constructions in Slavic, Baltic and Uralic languages; cf. Blevins 2003; and also Kibort, this volume).

More information on how the individual contributions relate to each other and fit into the overall structure of the volume can be gathered from the brief presentations of the contents of each chapter provided below.

The first part of the volume Impersonal constructions: typological and theoretical aspects, opens with a paper by Andrej Malchukov & Akio Ogawa “Towards a typology of impersonal constructions: a semantic map approach”. The authors argue that in order to capture selective similarities and affinities between different varieties of impersonals it is necessary to combine a function-based and a structure-based approach to impersonality. Building on the work of Keenan (1976), they propose to distinguish R-impersonals (with a notional subject lacking in referential properties), A-impersonals (with a notional subject lacking agentivity), and T-impersonals (with a notional subject lacking in topicality). Their investigation of the most common strategies of encoding these functional varieties culminates in a proposal of a semantic map of the impersonal domain which is designed to restrict the existing polyfunctionality within the domain of impersonality.

Many of the papers in subsequent chapters zoom in on particular subvarieties of impersonal constructions. Anna Siewierska in her contribution “Overlap and complementarity in reference impersonals: man-constructions vs. third person plural-impersonals in the languages of Europe” maps out the distribution of the two types of R-impersonals and considers to what extent the patterns found can be related to an important typological parameter within the European context, namely the formal realization of pronominal subjects. She shows that while the association between non-pro-drop and the presence of MAN-IMPS (man-impersonals) in a language posited by Holmberg (2005, 2010) in the main holds, the presence of 3PL-IMPS (third person plural impersonals) in a language is not dependent on the pro-drop parameter.
Nonetheless, the relative range of uses of both man-imps and 3pl-imps is argued to correlate with the realization of pronominal subjects; man-imps are shown to exhibit a wider range of referential uses (not only quasi-generic and existential but even specific) in non-pro-drop languages than in pro-drop ones, and 3pl imps in pro-drop languages than in non-pro-drop ones. However, the actual frequency of use of the two types of impersonals within languages is seen to be heavily dependent on the range of alternative R-impersonalizing strategies available. Only in very few languages does either emerge as the favored forms of R-impersonalizing.

The next chapter by Werner Abraham “Impersonal passivization between unaccusativity and unergativity” addresses the challenge posed by motion verbs to the definition of unaccusativity. The problem arises under impersonal passivization, as motion verbs may, on the one hand, undergo impersonal passivization, but, on the other hand, may emerge as ergative verbs/unaccusatives in directional use. In-depth semantic analysis of the aspectual properties of these predicates reveals that motion verbs are split unaccusatives: unergatives in the present, ergatives in the preterit participle. The chapter also contains an interesting discussion relating the availability of impersonal passives to a distinction between be-perfects and have-perfects, which can be equally traced to and explained by the aspectual properties of these constructions.

The final chapter in part one by Volker Gast and Florian Haas describes the distribution of subject properties in formulaic presentationals of Germanic and Romance. After introducing the formal and functional types of presentational structures in different varieties of Germanic and Romance, the paper argues for a connection between the availability and syntactic properties of dummies and the word order patterns found in thetic sentences in these languages. The authors relate differences in subject properties to word order constraints observed in thetic sentences in two group of languages. They show that only languages which allow verb-initial order in thetic sentences (‘thetic-V1 languages’) tend to use expletives in their existential formulas, and that the expletives tend to attract subject properties. This paper constitutes a bridge to the following chapters as it adopts a diachronic-typological approach.

The second part of the volume includes three chapters dealing with diachronic studies of impersonal constructions in Indo-European. It opens with Michela Cennamo’s paper on “Impersonal constructions and accusative subjects in Late Latin”, which examines the role played by impersonal constructions in the rise of the accusative-subject construction in Late Latin. The author argues that impersonal constructions with accusative arguments are likely to have played an important role in the use of the accusative in intransitive S function and the changes in morphological alignment thus manifested. She thus relates the establishment of the ‘extended accusative’ (Plank 1985) pattern in Latin to the ‘transimpersonal’ scenario of the rise of split-intransitive systems, as discussed in the typological literature.
The next chapter by Anna Giacalone Ramat & Andrea Sansò “From passive to impersonal. An Italian case study and its implications” continues the discussion of the historical developments of impersonals in Romance. The authors show that the emergence of the impersonal si-construction in Italian is due to reanalysis of si as a marker of generic human agency. On the basis of a large corpus of literary and non-literary documents, they document the initial stage of this process, namely the extension of the si-construction to intransitive verbs, and the emergence of the non-agreeing pattern with transitive verbs.

The paper by Leonid Kulikov “Passive to anticausative through impersonalization: The case of Vedic and Indo-European verbs of perception and speech”, also deals with the mechanisms of semantic change involving passives and their impersonal uses. It is argued that the rise of anticausative usages with these groups of verbs involves a stage of ‘impersonalization’. The focus of this chapter as well as the previous one is on the mechanisms of linguistic change, but this paper is different insofar as it adopts a function-based rather than structure-based definition of impersonality. The chapter ends with a brief discussion of the relationships between ‘agentless’, ‘impersonalized’ and ‘impersonal’ passives.

Part three of the volume is entitled Cross-linguistic variation in impersonal constructions: case studies. It begins with a paper by Doris L. Payne on impersonal construction in Maa. This construction has figured in the literature (since Greenberg’s 1959 seminal study) as a prime example of impersonal to passive reanalysis. Yet the author concludes that the construction is still rather impersonal and reanalysis to personal passive if underway is not complete. This paper makes also one important methodological point arguing that the study of impersonals should take into account both structure and function. The author uses a “hybrid” definition of impersonal constructions, which first identifies an unspecified agent, but then explores the range of functions that this constructional form has. This approach yields a “family” of constructions which have some conceptual or historical core, and lines up with semantic map approach as advocated by Malchukov and Ogawa.

Alain Christian Bassene & Denis Creissels provide a comprehensive account of different varieties of impersonal constructions in Jóola-Banjal (an Atlantic language spoken in Senegal). They document several constructions, not mentioned in previous works on Jóola languages, which include a non-canonical subject. This paper puts forward a formal classification of these constructions and a description of their functions, emphasizing the contribution of this West African language to the general typology of impersonality. The article can serve also as an introduction to different dimensions of or connections between impersonal constructions, related to semantic classes (meteor-verbs), discourse functions (presentational focus structures), word order restrictions (postverbal word order), and also more complex phenomena (related to raising and extraposed structures).
The next paper by Amina Mettouchi and Mauro Tosco “Impersonal configurations and theticity: the case of meteorological predications in Afroasiatic” addresses a wider range of languages, but focuses on meteo-predicates. The authors provide a survey of impersonal predication involving meteorological predicates, in various Afroasiatic idioms, and further argue that the subject or agent may not be centrally associated to the notion of impersonal. Rather, they suggest that impersonal patterning of meteo-constructions is related to the thetic nature of such sentences. They sketch a typology of meteo-constructions where defocusing or backgrounding can involve either the entity or the event, resulting not only in subjectless structures and non-canonical subjects, but also in verbless structures and non-canonical predicates.

Ruth A. Berman's paper "Revisiting impersonal constructions in Modern Hebrew: discourse-based perspectives” stands out methodologically in being a corpus based study which further explores impersonality from psycholinguistic and sociolinguistic perspectives. In particular the study serves to demonstrate the effects of such usage-based factors as genre, age-schooling development, as well as target language typology for the domain of impersonal constructions. The study concludes by arguing for a confluence of structural devices that combine to form a cline of impersonalization in the expression of a more or less depersonalized discourse stance.

The chapter by Anna Kibort “The elephant in the room: the impersonal -ne/-te construction in Polish” offers an overview of the little studied Polish impersonal construction (which has been overshadowed by the widely discussed -no/-to impersonal constructions). The paper discusses its relation to the impersonal passive of the intransitive as well as to the impersonal predicative adverbiaial construction, concluding that the form of the impersonal -ne/-te construction is underspecified with regard to which of the two syntactic constructions (passive verbal or non-passive adverbiaial) it instantiates. It seems that an underspecification analysis might be an optimal synchronic analysis of gradient phenomena resulting from diachronic developments, which frequently give rise to controversies (cf. the abovementioned controversy concerning the status of the morphologically related -no/-to constructions in Polish and Ukrainian, which has been alternatively analyzed as impersonal passive or impersonal active construction; Blevins 2003).

Merja Salo's chapter “Meteorological verbs in Uralic languages – are there any impersonal structures to be found?” resembles the chapter by Mettouchi & Tosco in addressing meteo-verbs and having an areal orientation, but the languages addressed are Uralic languages. The paper provides a detailed documentation of the diversity of constructions (and their distribution) formed by meteo-verbs in Uralic languages, with a focus on lesser studied minority languages. A recurring theme throughout the paper is the subject vs. object status of the explicit arguments occurring with these type of predicates and the lexical differences that are displayed in this regard both
within and across the Uralic languages. In conclusion an areal distribution of individual patterns is provided.

The chapter by Edward Vajda, Andrej Malchukov & Andrey Nefedov “Impersonal constructions in Ket” discusses a subclass of verbs with which the intransitive subject appears in slots normally reserved for object marking, while the slot normally reserved for transitive subject marking contains the invariant prefix da-. This construction has long puzzled students of Ket, and has been analyzed as instantiating a patient-subject or an ergative pattern. Yet, diachronically, it is shown to result from the reanalysis of a transimpersonal construction, expanding the range of languages for which such a development has been documented. Different formal varieties of this construction are shown to represent different stages of reanalysis.

Osahito Miyaoka deals with “Impersonal Verbs in Central Alaskan Yupik (Eskimoan)”, which is also a highly polysynthetic language, but which otherwise appears to have little in common with Ket. Yet, it seems that a similar development from transimpersonal to intransitive construction has occurred in this language as well. After presenting impersonal constructions with verbs of natural force and derived impersonals of necessity, the author proceeds to argue that the polysemy of the impersonal marker may be regarded as a result of a (historically) secondary change of the impersonal suffix into a modality marker. This might be regarded as a version of the transimpersonal-to-intransitive reanalysis also documented in the literature (Malchukov 2008).

Lynn Drapeau’s chapter provides an overview of “Impersonal constructions in Innu”, an indigenous Algonquian language spoken in Northern Quebec. This study offers fascinating data showing how the structural characteristics of a language may affect the way in which the distinction between personal and impersonal constructions is spelled out. In particular, the Inuu data shows the importance of animacy as a contributing factor insofar as suppressing a human agent requires special derivation on the verb, while suppressing a lower inanimate need not require special morphology, verbs with inanimate subjects being open to both a personal and an impersonal reading (lexical impersonals). This discussion brings to mind the explanation for the asymmetry between meteo-verbs which lack a referential argument and intransitive verbs featuring a subject argument: only the latter may need to employ special verbal morphology to block a referential interpretation (see the studies by Holmberg, Siewierska and others).

Anna Bugaeva’s chapter on “Impersonal constructions in Ainu” takes up the diachronic line of investigation of impersonal constructions in Ainu. It argues that impersonal (or maybe, impersonal passive) constructions in Ainu can be regarded as intermediate stages in the cross-linguistically frequent grammaticalization pathway, viz. IMPERSONAL > IMPERSONAL PASSIVE > PASSIVE. Yet, the author shows
that contrary to some earlier proposals none of the constructions can be regarded as personal passive because reanalysis of O to S is incomplete, thus a construction qualifies as an impersonal rather than personal passive. This echoes Doris Payne's conclusion on Maa.

The chapter by Yan Yi & Anna Siewierska “Referential/nominal impersonal constructions in Mandarin” considers the range of R-impersonals available in Mandarin. The authors demonstrate that like many European languages Mandarin utilizes generalized nouns, regular personal pronouns and a zero pronoun in subject position to encode impersonality. However, unlike in many other languages, the 3PL form ta-men is used impersonally at best only in very restricted circumstances. Significantly, it is not used to render any of the five types of 3PL impersonals identified by Cabredo Hofherr (2006): the universal, corporate, vague, inferential and specific existential. The most preferred strategy for conveying these senses in Mandarin is the you-ren construction which is a combination of the existential verb you, a generalized noun ren and a predicate whose agent is replaced by you-ren.

The chapter by Claire Moyse-Faurie “Impersonal constructions in Oceanic languages” provides an overview of different varieties of impersonal verbs in Kanak and Polynesian languages, focussing on constructions which exceptionally feature no person marking on the verb. It takes into account constructions with no (overt) arguments, with non-referential ‘dummy’ pronouns, valency alternations involving labile verbs, as well as cases of differential agent marking. The latter two cases are controversial as they are open to a different analysis. In particular, instances of the intransitive use of labile verbs are amenable to an impersonal analysis. In fact, this interpretation appears to be more motivated for ergative languages, than for accusative language, where the corresponding constructions of the “anticausative” type can be considered as impersonal only under a function-centred rather than a structure-centred approach. Similarly, the author argues that at least in some languages constructions interpreted in terms of differential agent-marking or even differential subject marking might be better interpreted as impersonal (or transimpersonal) constructions (cf. Malchukov & Ogawa, this volume, for some suggestions how these different interpretations can be reconciled in a diachronic perspective).

The final chapter by Jean-Christophe Verstraete on “Impersonal constructions in Umpithamu”, analyzes a set of impersonal constructions in four Pama-Nyungan (Paman) languages of the Cape York Peninsula. It is shown that there is a core of impersonals centered around experiencer object constructions, describing involuntary physical processes. Formally, these constructions are characterized by the systematic absence of nominative pronominal cross-reference, and usual absence of ergative agent nominals. Interestingly, in some languages, the same construction is also regularly used with inanimate subjects (agents), and in some other languages such
as Umpithamu it is regularly used as a kind of a functional equivalent of a passive construction. It would be interesting to speculate about the diachronic connections between these varieties. Under one scenario, the connection might be interpreted as involving reanalysis of a transimpersonal construction into a passive construction (reanalysis is incomplete though, as observed by Verstraete).

4. Conclusions

The present volume is intended as a cross-linguistic contribution to the study of impersonal constructions. Its main value consists in careful documentation of (varieties of) impersonal constructions, across a wide range of languages from all over the world (most of them non-Indo-European). In addition to this foundational descriptive goal, the volume attempts to arrive at cross-linguistic generalizations by synthesizing different traditions in the research of the impersonal domain:

- It unifies the domain of impersonal constructions, which apart from obvious similarities displays a certain structure where particular domains (A-impersonals, T-impersonals, R-impersonals) are related by family resemblances;
- It brings together and synthesizes both function-centered and structure-centered approaches; obviously both are needed for cross-linguistic studies;
- It integrates typological and theoretical studies (as is also clear from mutual references to theoretical work – syntactic like Holmberg’s or semantic like Cabredo-Hofherr’s in some of typological contributions, and references to typological work in theoretical studies, such as Abraham’s);
- It combines synchronic and diachronic approaches, as is obvious not only from the studies which take a historical perspective, but also in many contributions which attempt to provide diachronic scenarios for explaining intricacies of the language data (in particular the chapters on Maa, Ket, Ainu, Yupik).
- It suggests a potentially promising new view of impersonality, according to which impersonal constructions are by definition transient structures, intermediate stages of the shift between more basic structures such as transitive and intransitive clauses or active and passive constructions or participant-centered and event centered constructions etc.

Yet, we are aware that this volume is only a first step towards a comprehensive typology of impersonal constructions, which should be followed up by subsequent in-depth studies of certain subdomains as well as systematic studies of the domain in general.
References


Croft, W. Forthcoming. Locative subjects, impersonals and related constructions. Ms, University of New Mexico.


PART I

Impersonal constructions
Typological and theoretical aspects
Towards a typology of impersonal constructions
A semantic map approach*

Andrzej Malchukov & Akio Ogawa
Max Planck Institute for Evolutionary Anthropology, Leipzig & Institute for Linguistic Studies, St. Petersburg & Johannes Gutenberg University, Mainz/Kwansei Gakuin University, Nishinomiya

The present article proposes a semantic map for the domain of impersonal constructions. Following Keenan, we assume that different varieties of impersonal constructions register loss of the functional subject properties (such as definiteness, topicality and agentivity). It is argued that different functional varieties of impersonal constructions (where the subject lacks some of the functional properties of a prototypical subject) are coded by different types of impersonal constructions, although certain constructions have a broader range of applications. On the basis of this partial overlap between different coding strategies for different functional varieties of impersonal constructions, we construct a semantic map for the impersonal domain, which can be used to represent and constrain variation of coding of impersonal constructions across languages.

Keywords: A-impersonals; R-impersonals; T-impersonals; transimpersonals; semantic maps

1. Introductory

The fate of the concept of impersonal construction remains somewhat controversial in modern linguistics. On the one hand, there are numerous studies of impersonal

* We are very grateful to Anna Siewierska, Denis Creissels and Werner Abraham for valuable comments on the prefinal version of this paper. The usual disclaimers apply. The research on this paper was initiated in 2004 when Andrzej Malchukov was a JSPS (Japanese Society for the Promotion of Sciences) fellow at the University of Kobe, and the paper was finalized during Andrzej Malchukov’s stay as a visiting professor at the National Institute for Japanese Language and Linguistics in Tokyo in 2010. The financial support of both institutions is gratefully acknowledged.
constructions in Indo-European languages (see Seefranz-Montag 1983; Lambert 1998; Bauer 2000), on the other hand, there are very few studies which approach this topic from a typological perspective (cf. also Creissels 2007; Siewierska 2008). This seems to be largely due to the fact that the term “impersonal” is notoriously ambiguous pertaining to different structures often not directly related: constructions with meteo-verbs with a generic (“ambient”) subject or no subject at all, presentational structures with downgraded subjects, but also other types of oblique subject constructions as often found with emotion verbs. While we are aware of the heterogeneity of the domain, and concur that many of the constructions implicated should be studied in their own right (atmospheric verbs, presentational structures, psych-verbs, etc.), we will try to show that the domain of impersonal constructions, broadly conceived, has a certain structure. Yet, this structure is complex: as the relations between individual subtypes may be more or less direct, the whole network is more reminiscent of ‘family resemblances’. For such domains the semantic map approach (as developed by Anderson, Croft, Haspelmath, van der Auwera and others) has proven to be particularly fruitful, and for this reason this approach will be also pursued here.

In this paper we adopt the definition of the impersonal constructions as constructions lacking a referential subject (cf. Siewierska 2008). This formal-syntactic definition conveniently covers different subtypes of constructions both lacking a subject (as in some constructions with atmospheric verbs) and those with a non-referential (dummy) subject, as well as constructions where the notional subject lacks typical subject properties, as is often attested in constructions with psych-verbs. We also follow Siewierska (2008) in her view that formal-syntactic (“subject-centered”) definitions of impersonal constructions and functional definitions appealing to “agent/instigator defocusing”, frequently confused in the literature, should be kept apart. Under the subject-based view of impersonal constructions, only those constructions qualify as impersonal which lack a grammatical (referential) subject. Under the functional view, impersonalization is viewed as under-elaboration or demotion of the agent/instigator. This latter approach is more inclusive as it also covers constructions which feature a subject but where an agent is suppressed or demoted (as e.g. in agentless passive constructions). We agree that a strict separation of the two approaches is needed. Moreover, we go a step further in actually separating the two approaches of impersonal constructions (note that the term demotion which enters into the functional definition of impersonal construction still has a syntactic flavor). Following Keenan (1976), we will also assume a richer set of functional subject properties, than under some versions of the agent-centered approach. Generally, a strict separation between the two approaches will be crucial for our study, as it sets out to investigate how these two conceptions of impersonality (the formal and the functional) relate to each other. (See also Payne, this volume, for an application of a similar approach, appealing both to form and function of impersonal constructions, to a single language).
Towards a typology of impersonal constructions

Note that the adoption of the formal definition of impersonal constructions (which relies on the notion of a grammatical subject) raises the issue of its cross-linguistic applicability. Nonetheless, even though a grammatical subject might have to be identified on the basis of language-particular diagnostics, this approach opens a way to study co-variation of form and meaning across languages in a non-circular way. In this respect we follow important functionalist work done in the 1970s–1980s on multifactorial analyses of grammatical concepts, in particular, Keenan’s (1976) seminal analysis of the notion of subject. In line with functionalist approaches, the main research question to be addressed in this article is: What is the relation between functional and formal varieties of impersonal constructions? The hypothesis that we pursue – following Keenan – is that different formal subtypes of impersonal constructions register different deviations from the subject prototype. Further, we will study formal similarities in encoding of different functional varieties of impersonal constructions in order to construct a semantic map for the impersonal domain. The established semantic map can further be used to capture and constrain possible types of impersonal constructions cross-linguistically.

The paper is structured as follows. In §2 we discuss the functional definition of subject, taking Keenan’s work as a point of departure, and distinguish between different functional varieties of impersonal constructions that fail to conform to the subject prototype on different counts. In §3 we provide a brief cross-linguistic survey of the formal coding of these different functional varieties of impersonal constructions. The partial similarities in encoding of particular constructions form the basis for the semantic map proposed in §4. §5 addresses a diachronic connection between the functional subvarieties of impersonals which is not captured by the semantic map approach. Finally, §6 summarizes the main results of the paper.

2. Subject prototype

Given that the notion of subject enters into the “subject-based” definition of impersonal constructions adopted here, it is appropriate to start with a definition of the subject. As noted above, we take as a starting point Keenan’s (1976) multifactorial definition of the concept of ‘subject’. In this seminal study, Keenan proposed 30-odd properties which are cross-linguistically characteristic of “basic” (i.e. non-derived) subjects, in the sense that they may affect subject selection and/or subject encoding in individual languages. Keenan’s paper is well-known so it need not be presented in detail here. Suffice it to say that the proposed subject properties span different domains, including both functional properties (such as autonomous reference/specificity, topicality, referentiality and definiteness), semantic properties (e.g. agency), coding properties (e.g. unmarked case, control of verb agreement), structural
properties (being dominated by an S-node in a clausal structure), as well as a host of behavioral properties (accessibility to relativization, raising, being the target of advancement, control of raising and equi, etc.). Keenan has convincingly shown that these properties are critical for subject choice at least in some languages (for example, he cites the Philippine and Bantu languages, where the subject must be definite). He also demonstrated that if one of the subject properties is not met, this may result in a diathetic shift to a passive construction and in addition discussed the availability of subject properties on the part of derived subjects in these constructions. Similar decompositional approaches have been developed by other scholars working in the functional-typological paradigm, even though they did not always reach the level of detail of Keenan's analysis (cf., e.g. Comrie's (1989: 107) definition of the prototypical subject in terms of the intersection of agent and topic properties).

Keenan's work has been very influential in the functional-typological tradition, and has been used to account for different types of voice alternations, but also to explain categorizing languages into different subtypes in terms of conflation vs. differentiation of subject properties (as especially relevant for ergative and "Philippine-type languages"). Even though Keenan did not focus on impersonal constructions per se, his work is obviously relevant for the study of impersonal constructions. Yet, this work has also certain problematic aspects, which need to be addressed before this framework can be used for the purposes of our study. Thus, it is clear that Keenan's list of subject properties conflates different properties which need to be differentiated (cf. Givón 2001: 195–196):¹

- functional properties (topicality, definiteness, agentivity, etc.)
- behavioral (syntactic) properties (relativization, target of raising, etc.)
- coding properties (case/flagging, agreement/indexing, word order)²

In this paper we will have little to say about behavioral properties, which will be disregarded for the time being (as shown by Givón 1997, these properties can be further

---

¹ For example, the by far largest subgroup of “autonomy properties” on Keenan's list includes both functional properties ("independent existence"), coding properties (control of verbal agreement), as well as a host of syntactic properties (control of reflexivization, etc). In fairness, it should be added that in subsequent discussion (in particular, concerning the acquisition of subject properties), Keenan does distinguish between coding, behavioral and semantic properties.

² We will (occasionally) refer to case marking in a broader sense (i.e. dependent marking of argument relations by case or adposition) as 'flagging' and to agreement in a broader sense (including cross-referencing in head-marking languages) as 'indexing' following conventions introduced in (Haspelmath 2005).
grouped in terms of parameters such as topicality, agentivity and the like). Instead we will focus on the relation between functional and formal (coding) properties in the impersonal domain. In particular, we will study the correlation between functional properties (a decrease in functional properties contributing to subjecthood: agentivity, definiteness, etc.), on the one hand, and coding properties (in particular, deviation from the canonical encoding of subjects), on the other hand. This approach is thus similar in spirit to the multifactorial approaches to transitivity of Hopper & Thompson (1980) and Tsunoda (1981) which aim at uncovering the functional factors contributing to the more transitive or less transitive encoding of events across languages. This approach is also implicit in Keenan's work, although it has not been applied specifically to the study of impersonal constructions (recall that deviations from canonical subjecthood need not necessarily yield an impersonal structure, but may rather lead to a diathetic change). The adopted approach departs from previous approaches in that it considers how different deviations from the functional prototype may yield somewhat different structural outcomes. This is so since the coding properties themselves (case, agreement, and word order) need not always line up in a particular construction, but may show mismatches. So the next step in our enterprise is to uncover the similarities between the formal encoding of different functional types of impersonal structures.

As noted above, a semantic map approach will be used to this end, as it allows capturing partial similarities between different types of impersonal structures and locating them in semantic space. (Compare Malchukov 2006 for a similar approach applied to transitivity gradients which captures covariation between transitivity parameters and transitivity alternations in a form of a scale; see also Beavers & Zubair (2010) for an insightful follow-up discussion).

In what follows we will confine ourselves to a few functional properties which seem to have wide cross-linguistic applicability and work with ‘crude’ functional dimensions, which sometimes are used as a cover term for several features on Keenan's list. Following Keenan, we assume that a prototypical subject is:

a. a referential argument;
b. a definite NP;
c. topical;
d. animate;
e. agentive.

Note that the last two properties on this list (properties d–e) are a little special as they pertain to transitive subjects (As) but not necessarily to intransitive ones (Ss), an issue to be taken up later. Of course, many of the features on the list can be further decomposed, as is well known from the vast literature on the topic (see, e.g. Dowty 1991 on agent properties, Silverstein 1976 on animacy (hierarchy), Chafe 1976 on topicality
and related notions, Siewierska 2004 and Cabredo Hofherr 2006 on definiteness and referentiality), but for the purposes of this general study aiming to provide a general outline of the impersonal domain this crude classification will do.

3. **Functional varieties of impersonal constructions and their encoding**

In what follows we address different functional varieties of impersonal constructions and provide a cross-linguistic overview of some common strategies of their encoding. The focus will naturally be on subvarieties of constructions which count as impersonal on formal grounds, but also alternative constructions will be mentioned for particular functional subtypes in the impersonal domain. As in Keenan’s (1976) approach different departures from the basic construction with a full-fledged subject are seen here as a reflection of the fact that the notional subject fails to display some functional subject properties. These features of (viz. departures from) the subject prototype will define what we refer to as the ‘impersonal domain’. Our typology is thus purely functional (for a different, syntax-based, typology of impersonal constructions see Creissels 2007).

Next we discuss the most important functional varieties of impersonal constructions, including constructions lacking a referential subject (weather verbs), constructions with indefinite (human) subjects, constructions with non-topical subjects (in particular, presentational constructions), as well as constructions with inanimate and non-volitional subjects. The aim of this brief cross-linguistic survey is not to provide a detailed discussion of particular subtypes but rather to establish – in reliance to the literature on the topic (including contributions to this volume) – the preferred strategies of coding individual functional subtypes of impersonal constructions. The point of the presentation is that these encoding strategies are, as one might expect, partially different, but, importantly, also exhibit a certain degree of overlap. The recognition of these selective formal similarities which in turn suggest that certain functional types of impersonal constructions can be grouped into ‘natural classes’ is crucial for constructing a semantic map of the impersonal domain.

3.1 **Impersonal constructions with non-referential subjects**

As noted by Keenan (1987 [1976]: 102) and others, grammatical subjects are normally referential (have “absolute reference”, in terms of Keenan). Yet, like other features on his list this feature is not universally characteristic of subjects as it does not pertain to (dummy) subjects of weather verbs. Also, Bauer (2000:100) in her discussion of impersonal constructions in Indo-European languages notes that weather verbs have a
special status among impersonal verbs as they do not feature argument structure at all. Indeed, these constructions frequently either lack a subject altogether as in the Russian example (1), or have a dummy subject as in the English example (2).

(1) Svetaet.
dawn.pres.3sg
'It dawns.'

(2) It dawns.

Constructions with zero subjects of impersonal verbs are very common cross-linguistically, as illustrated by examples from Guarani and Kiowa. Note that if a language features person agreement, the default 3rd person (singular) forms are usually used (as in (3)); less frequently some other form is used (cf. 3rd person inanimate plural forms in Kiowa in (4)):

(3) Guarani
O-kí.
3sg-rains
'It rains.'

(4) Kiowa
Khí-dél gyà-sál.
yesterday 3pl-hot
'Yesterday it was hot.'

Constructions with “dummy subjects” as familiar from Germanic languages (see the English example above) seem to be less common cross-linguistically and are areally restricted (cf. Dryer 2007). It has been suggested that they are confined to so called ‘configurational’ languages that require a subject slot to be filled. Most researchers seem to agree that dummy subjects developed in various Indo-European languages with the rise of ‘configurational’ syntax with an obligatory subject position (Faarlund 1990: 187; cf. Seefranz-Montag 1983; Bauer 2000). Somewhat similarly, in Optimality Theory, the rise of dummy pronouns is explained by a constraint requiring a subject slot to be filled (Extended Projection Principle) outranking a semantic “faithfulness” constraint (Full Interpretation) which requires that only arguments licensed by a verb should be encoded in the syntax (Grimshaw & Samek-Lodovici 1998). Yet, as we will see shortly, a tension between a requirement for the filled subject position, characteristic of some languages, and the semantics of weather expressions, which do not license arguments, can be also resolved in other ways.

As is also well-known from the literature, languages may have other means of encoding atmospheric events (see, e.g. Ogawa 2005; 2006; Dryer 2007; Erikson et al. 2009; see also Mettouchi & Tosco this volume, on Afroasiatic, and Salo this volume, on
Uralic languages). Apart from zero subject and dummy subject constructions, there is the “dummy (auxiliary) verb” construction (cf. Ogawa’s (2006) “nouny constructions”), as illustrated below on the basis of Russian and Japanese:

(5) Russian
   *Dožd’ idět.*
   rain goes
   ‘It rains.’

(6) Japanese
   *Ame-ga furu.*
   rain-nom new falls
   ‘It rains.’

Another strategy, which is also occasionally found in Indo-European (cf. Russian: *Grom gremit* lit. ‘Thunder thunders’), but seems to be more widespread in other languages is the use of a “cognate construction”. This construction, where the subject and the predicate are lexically cognate, is characteristic for weather expressions in, for example, Even (Tungusic).

(7) Even
   (Malchukov 2002 and f.n.)
   a. *Udan udna-n.*
      rain rain-aor.3sg
      ‘It rains.’
   b. *N’ölten n’ölten-ni.*
      sun sun/shine-aor.3sg
      ‘The sun shines.’

And of course at least some atmospheric events will be rendered by a full-fledged clause displaying a regular subject-predicate structure (cf. *The sun shines*, where unlike in (7b) the subject is not etymologically related to the verb). Some languages have been reported to lack impersonal constructions altogether, even in the domain of weather expressions (see, e.g. Lichtenberk 1983: 224 on Manam). There may be further certain lexical preferences for particular types of atmospheric events for particular encoding, as tentatively proposed by Ogawa (2005, 2006), but these will not be addressed here (but see Ogawa & Malchukov in preparation).

These different strategies should be further viewed as idealized types, with some intermediate types in between. Thus, Ogawa (2005, 2006) argues for a cline between ‘verby’ and ‘nouny’ constructions with the ‘cognate’ type in between. One construction which has recently received some attention in the literature (Erikson et al. 2009), which might be seen as intermediate between a construction between dummy and a lexical subject is where a nominal meaning ‘world’, ‘weather’ or ‘sky’ appears as a formal subject with at least some atmospheric events.
Towards a typology of impersonal constructions

Another species of weather constructions with a semi-formal subject lit. is a construction with a mythological subject, as is familiar from Latin (cf. *Jupiter tonat* lit. 'Jupiter thunders'), but which otherwise seems to be marginal in Indo-European.³

It may well be that other constructions qualified as impersonal in some languages, may on closer inspection also display such semi-formal subjects or mythological subjects. Nonetheless, there is no doubt that that some constructions with atmospheric expressions do not allow for either a lexical subject or semilexical subject, as explicitly stated by Donohue (1999) for Tukang Besi:

As our study focuses on impersonal constructions those constructions which lack a subject altogether or have a formal (dummy) subject would be of most immediate relevance. These are also the types which we encounter in some other functional varieties of impersonal constructions to be considered next, while some other types (e.g. cognate constructions) are restricted to constructions with atmospheric events and will be of less relevance here.

### 3.2 Impersonal constructions with indefinite subjects

The constructions with indefinite (mostly human) subjects also commonly involve the use of zero pronouns or formal subjects, as exemplified by familiar examples from
Russian (zero subject in the context of 3rd person verb form) and German (the formal subject is expressed by the indefinite pronoun man):

(11) Russian

\[
\text{Govor\textacute{a}t} \quad (\text{on uexal}).
\]

say.pres.3pl he leave.pst.3sg

‘They say he left.’

(12) German

\[
\text{Man sagt (er ist gegangen).}
\]

‘They say (he left).’

This type of impersonal constructions have been extensively discussed in the literature, in particular, in the work by Siewierska (2004; this volume, \textit{et passim}), so it need not be rehearsed here. Siewierska (this volume) identifies third-person plural impersonals (as in the Russian example (11)), and \textit{man}-constructions (as in the German example (12)), as representative of two main strategies for encoding of indefinite constructions in European languages. As also shown by Siewierska (this volume; cf. Ramat & Sansò 2007) both strategies are also common elsewhere. On balance, \textit{man}-impersonals (as opposed to regular 3rd person plural pronouns) seem to be more restricted cross-linguistically, but are common, for example, in African languages, as illustrated here for Babungo:

(13) Babungo

\[
\text{Vi jia n\text{\textae}.}
\]

they.imps catch.pf him

‘They (indefinite) caught him.’

The last example can be also translated as ‘He was caught’, a point to which we return in our discussion of transimpersonal construction in §5. What is of relevance here is that the indefinite 3rd person plural pronoun is a form distinct from a ‘personal’ 3rd pl pronoun (which is \textit{v\text{\textae}} ‘they’; Schaub 1985: 194).

As shown by Holmberg (2006) and Siewierska (this volume) the choice of a particular strategy for the encoding of indefinite agents partially correlates with the “pro-drop parameter” (i.e. possibility of pronominal ellipsis). Thus, constructions with indefinite impersonal pronouns (Man-Imps in Siewierska, this volume) have been shown to be common in non-pro-drop languages (Holmberg 2010; Siewierska, this volume). On the other hand, as also argued by Holmberg (2010) and Siewierska (this volume), languages which do allow pro-drop often feature special verbal impersonal forms. An explanation proposed by Holmberg (2010) appeals to the blocking of a non-referential interpretation by a referential one (Siewierska 2004:213 makes the same point to explain a rare occurrence of 3rd person pronouns with impersonal reference). In some cases both interpretations do coexist (as in the Russian (11) which can also be understood as involving a definite subject), but indeed most of the languages in question make
use of special constructions (in particular special verbal form) to exclude a referential interpretation. (Holmberg 2006, 2010) cites Italian and Spanish as examples; below the much discussed -no/-to construction in Polish is used to illustrate this point:

(14) Polish (Siewierska 1984: 113)

\[
\begin{array}{ll}
\text{Podano} & \text{kawe.} \\
\text{served.IMPS.PART} & \text{coffee.ACC}
\end{array}
\]

‘Coffee was served.’

This participial construction in Polish as well as some other special (dedicated) impersonal forms found in other languages (e.g. Finno-Ugrian), have been also analyzed as impersonal passives, but this analysis is controversial (see Blevins 2003). On the other hand, Siewierska (1988) notes that a distinction between dedicated indefinite subject forms and impersonal passives is gradient. Generally, the functional similarity between impersonal passives usually implying an indefinite human agent and man-impersonal constructions can’t be overlooked. Indeed, on some accounts (Frajzyngier 1982), impersonal passives are viewed as a variety of indefinite subject constructions. We will disregard impersonal passives for the time being, but will return to this topic in §3.6 below.

### 3.3 Impersonal constructions with non-topical subjects

Constructions with non-topical notional subjects also come in several guises (Sasse 1987; Lambrecht 2000; Gast & Haas this volume). One case in point (considered above) is when the subject is necessarily non-topical since it is indefinite/generic; accordingly it is frequently omitted. Another case, to be considered next, is when the subject is non-topical by virtue of being new (subject-focus), or the whole clause being in focus (thetic or sentence-focus utterances). The most important subtype of sentence-focus constructions, which introduces the subject referent, is generally known as a presentational construction (see Gast & Haas this volume, for a detailed discussion). One very common strategy of encoding presentational constructions (in SV-type languages) is through word order inversion (cf. Hetzron 1975 for an early account).

(15) Russian

a. \[\text{Mal’čik vošēl.} \]

\[
\begin{array}{ll}
\text{boy} & \text{come.PST.3SG}
\end{array}
\]

‘The boy came in.’

b. \[\text{Vošēl mal’čik.} \]

\[
\begin{array}{ll}
\text{come.PST.3SG} & \text{boy}
\end{array}
\]

‘A boy came in.’ or ‘There entered a boy.’

In the Russian example, the change in word order does not result in any ostensible change in structure. English is different from Russian insofar as it uses an expletive
in clause-initial position, but as in Russian the postverbal subject retains its case and control of agreement (see the translation of the Russian example above).4 Yet, as noted by Givón 1997, a postverbal subject in English loses some of the syntactic properties (such as control of the infinitive, etc.) In some other languages, however, word order inversion does involve loss or redistribution of subject coding properties. A well-known case is French, where the postverbal actant (dubbed ‘actant H’ by Lazard 1998) loses some subject properties including the control of agreement:

(16) French

a. *Une femme viendra.
   INDEF.SGF WOMAN.SG COME.FUT.A3SG
   ‘A woman will come.’

b. Il viendra une femme.
   3SGM COME.FUT.3SG INDEF.SGF WOMAN.SG
   ‘A woman will come.’ (lit. ‘It will come a woman.’)

In some other languages, the loss of subject properties may depend on further conditions. Thus, in Russian loss of agreement with non-topical/indefinite/new subjects is restricted to constructions with numeral phrases (where the numeral has certain head properties and assigns genitive case to the noun). While with definite subjects in preverbal position plural agreement is obligatory, with a postverbal indefinite subject agreement may fail resulting in an impersonal structure:

(17) Russian

a. Tri čeloveka prišli. (*prišlo)
   three men COME.PST.3PL (COME.PST.3SG)
   ‘(The) three men came.’

b. Prišli / prišlo tri čeloveka.
   COME.PST.3PL/ COME.PST.3SG three men.gen
   ‘The three men came.’

Note that in both French and Russian the structural changes involve word order and agreement, but not case (the accusative form of the numeral phrase would have been trëx čelovek in the Russian example above). This suggests, as also argued by Givón 1997, that change in agreement predates change in case in this kind of impersonal structures triggered by loss of topicality. Although Givón’s point related to the order of acquisition of subject properties (as in the earlier studies by Keenan 1976 and Cole et al. 1980), the same order is observed with respect to loss of subject properties (a point to which we return in §3.6 below). Thus, if case is affected, other coding properties

4. Note, however, substandard forms of English cited by Bolinger (1977): There’s him, and there’s you and me. See Gast & Haas (this volume) for discussion of such examples.
including word order and agreement would normally be affected as well. This is found, for example, in Finnish, although the construction itself is not presentational. In this construction the focal pronoun in the postverbal position appears in the accusative and does not control verbal agreement:

(18) Finnish (Itkonen 79: 83)

Niin kauan kuin minulla on sinut tunnen itseni onnelliseksi.

as long as I.ADESS is.3SG you.ACC feel.I myself happy

'As long as I have you I feel happy.'

Looking beyond presentational constructions, we can also observe that loss of agreement is a common reaction to non-topical subject. This is most obvious for languages like Oromo where agreement is restricted to topical subjects. Note that a non-topical subject (lacking a topic marker in (19b)) doesn’t control agreement in Oromo:

(19) Oromo (Clamons et al. 1999:60)

a. Intal-t-ií-n hoolaa bit-t-e.
girl-FM-SBJ-TOP sheep buy-FM-PAST

'The girl bought a sheep.'

b. Intala takka-á hoolaa bit-e.
girl one.FM-SBJ sheep buy-PAST

'A girl bought a sheep.'

Also for Somali, a similar connection has been observed: as noted by Lazard (1998: 171), the verb agrees with the subject if it is topical, there is no agreement if the subject is in focus, and agreement is optional in verb-focus constructions. Further examples of the loss of agreement with focal NPs are discussed in Siewierska (2004: 159–161).

Bantu languages also provide evidence for topical NPs as privileged agreement controllers. In some of these languages loss of topicality on the part of the subject in a locative construction might lead to “locative inversion”, as exemplified for Swahili below (cf. Bresnan & Kanerva 1989 on Chichewa):

(20) Swahili (Givón 2001:260)

child(MS) MS-PAST-be house-LOC

'The child is in the house.'

house-LOC LOC-PAST-be with child(MS) one

'There is a child in the house.'

In the latter example, the verb shows class agreement with the topicalized locative NP, while the subject is demoted to a (prepositional) oblique. Thus, the construction counts as impersonal at least with respect to ‘flagging’. Another type of impersonal construction is found in Sesotho (discussed in Baker 1992), where the class prefix
on the verb is the default agreement marker rather than agreement marker normally
found with a topicalized locative phrase:

\[(21)\]  
Sesotho  
(Baker 1992: 35)  
\[\text{Montse-\text{ng} \ ho-\text{fihl-ile} \ baeti.}\]  
village-LOC IMPs.CL-arrive-PF visitors  
'To the village arrived visitors.'

Note again that loss of agreement predates a change in case marking in Bantu lan-
guages: in Sesotho only agreement is affected, and if case is affected, as in Swahili,
agreement is affected as well.

Diachronically, a connection between agreement and topicality is of course
expected given that agreement frequently arises through topic resumption (Givón
2001 \textit{et passim}), and at least for some Bantu languages it is claimed that also syn-
chronically we are dealing with topic agreement rather than subject agreement (cf.
Morimoto 2008 on Kinyarwanda).

3.4 Impersonal constructions with inanimate subjects (agents)

Apart from topicality, control of agreement may be determined by animacy. This is so,
for example, in Tlapanec (Wichmann 2008), in which only animate/human arguments
can trigger verbal agreement/cross-referencing.

\[(22)\]  
Tlapanec  
(Wichmann 2008)  
a. \[\text{Ni-\text{ngahtá} d3amà.}\]  
PFV-fall banana  
'The banana fell.'

b. \[\text{f\text{abù} ra ni-\text{ngaht-á}.}\]  
man top PFV-fall.to.side-3SG.G.ABS  
'As for the man, he fell to the side.'

This case seems to be similar to the previous one, as in both we are dealing with
absence of agreement, even though in the former situation it is triggered by lack of
topicality and in the latter by lack of animacy (in the example above, the subject takes
a topic marker but the rules of agreement are not contingent on that). The two features,
topicality and animacy, can be brought together under the heading of ‘prominence’
and can easily be offered a functional explanation: animate beings are higher on the
empathy hierarchy; hence they are more topical in discourse. Siewierska (2004: 154–6)
discusses further examples of languages where agreement is controlled by animate
arguments, but in most cases the argument is a P rather than an A/S.

On the other hand, constructions with inanimate subjects/agents can also license
Differential Case Marking (cf. Malchukov 2008 for discussion of animacy effects
in differential case marking). Consider the case of Samoan, where an inanimate argument, unlike an animate one, has the choice of appearing in the Ergative or in an Oblique case:

(23) Samoan (Mosel & Hovdhaugen 1992:423)

\[Na\ tapuni\ e\ /i\ le\ matagi\ le\ faitotoa.\]
\[PAST\ close\ \text{ERG} /\text{LOC} \ \text{ART} \ \text{wind} \ \text{ART} \ \text{door}\]

’The wind closed the door.’

A seemingly similar construction is attested in Russian where the animate “force” argument can appear either in the nominative case or the instrumental case:

(24) Russian

a. \(Sneg\ zanes\ dorog-u.\)
\[\text{snow} \ \text{cover-PAST.3SG.M} \ \text{road-ACC}\]

’The snow covered the road.’

b. \(Dorog-u\ zanes-lo\ sneg-om.\)
\[\text{road-ACC} \ \text{cover-PAST.3SG.N} \ \text{snow-INST}R\]

’The road was covered by snow.’

Note that the Russian construction is more obviously impersonal: the A appears in the instrumental case and fails to control verbal agreement (note that agreement is neutral in (24b) rather than masculine as in (24a)). The situation is less clear in Samoan, which is an ergative language and does not show subject agreement.

Generally, ergative languages present a challenge for a typology of impersonal constructions (see also Creissels 2007; Siewierska 2008). Thus, it is well known from the literature on ergativity (see Kibrik 1997 for a recent overview) that the concept of (grammatical/syntactic) subject is somewhat problematic for ergative languages, as different diagnostics might identify either A or P as a syntactic subject (pivot). Dixon (1994) in his treatment of ergativity uses the notion of subject but in a semantic (A/S-argument) rather than syntactic sense. Yet, the data from ergative languages are instructive as they show that animacy effects can be registered in the choice of case rather than, say, agreement or word order, an option which was rare in the functional varieties of impersonal constructions considered above. The data from ergative languages are also important given that ‘differential case marking’ of the type illustrated for Samoan seems to be more frequent in ergative languages than in accusative (Malchukov 2006). In the latter languages the decrease in animacy is more likely to lead to a diathetic shift (e.g. to the passive construction). We will have more to say about this phenomenon, when we consider constructions with non-volitional As in the next section.
3.5 Impersonal constructions with non-volitional subjects (agents)

Constructions with non-volitional agents can be marked in a variety of ways across languages (see Kittilä 2005 for an overview). One type of construction, familiar from the previous discussion, displays a change in case marking, as illustrated for Lezgian below:

(25) Lezgian (Haspelmath 1993:292)

   child-ERG pot(ABS) break-AOR
   ‘The child broke the pot.’

   Zamira-AdEl pot(ABS) break-AOR
   ‘Zamira broke the pot (accidentally/involuntarily).’

In example (25a) in which the A is in the ergative case the action may be either volitional or not, in example (25b) where the A argument is in the oblique (ad-elative) case, it is explicitly signalled as non-volitional. It seems that differential case marking is still more common as a strategy of encoding non-volitional agents as compared to non-human agents. For example, in Lezgian a change from the ergative to the ad-elative case is used in involitive construction, while constructions with inanimate agents are coded as ergatives (somewhat unexpectedly, given that inanimates can’t be volitional). An explanation for that should be sought in the factor of economy: as noted by Kittilä (2005) and Malchukov (2008) it is redundant to mark inanimate arguments for the feature of (in)volitionality.5

The pattern attested above for dependent marking, can also be attested for head-marking languages. Malchukov (2006) cites Abkhaz (Hewitt 1979:195), where ‘ergative agreement’ switches to ‘oblique agreement’ in situations when the verb takes a non-volitional form. Another example is Koasati (Kimball 1991), where a change from agentive to dative agreement also signals non-volitionality:

(26) Koasati (Kimball 1991:253)

a. Ca-yawoplíc-iska-hónk.
   1SG.PAT-scarele-2SG.AG-ADV
   ‘You scared me (deliberately).’

b. Cin-ca-yawópka-hónk.
   2SG.DAT-1SG.PAT-scarele-ADV
   ‘You scared me (accidentally).’

5. De Hoop & Malchukov (2007) discuss a number of cases where inanimate As show a regular transitive (ergative) pattern, while non-volitional As show a deviant pattern and explain them (in Optimality-Theoretic terms) as an economy effect (which overrides role-based faithfulness constraints).
Towards a typology of impersonal constructions

We will refer to this kind of phenomenon involving a choice between several different agreement series (associated with A, P, sometime also Oblique arguments) as ‘Differential Indexing’ as opposed to Differential Case Marking (or differential ‘flagging’, in terms of Haspelmath 2005). Note that this pattern is different from the loss of agreement (indexing) as attested for Tlapanec and some other languages, insofar as here the change is not confined to loss of ergative agreement but rather to change to a non-agent agreement set.

Again, accusative languages turn up clearer instances of impersonal constructions, as impersonality may be evidenced by both flagging and indexing. Consider Bengali, where a non-volitional agent may be coded by the genitive case as in (27b) and in contrast to the nominative counterpart in (27a) fail to control verbal agreement:

(27) Bengali (Onishi 2001: 120)

a. Tumi (bondhu-ke) Teliphon kor-ech-o?
you (friend-obj) telephone do-PF-PRES.3
‘Have you made a telephone call to your friend?’

b. Toma-R (bondhu-ke) Teliphon kør-ά ho-ech-e?
you-gen (friend-obj) telephone do-NZR become-PF-PRES.3
‘Have you finished a telephone call to your friend?’

Yet, as observed earlier, differential case marking of the type illustrated above is dispreferred in many accusative languages. A more common outcome seems to be a diathetic shift, to a passive or, rather, an “experiencer-anticausative construction” (Malchukov 2006) or “dative-anticausative construction” (Creissels 2007), as found in Russian (28b) and many other languages:

(28) Russian

a. Ja slomal zu b.
I.nom broke.3sg tooth.acc
‘I broke (my) tooth.’

b. U menja slomal-sja zu b.
to me broke.3sg-refl tooth.nom
‘I have a tooth broken.’

Note that while the regular transitive construction (in (28a)) can be interpreted either volitionally or non-volitionally, the “experiencer-anticausative construction” in (28b) has only the non-volitional reading.

Malchukov (2006) (following Shibatani 1977 and Tsunoda 1981) observed that ergative languages are more prone to exhibit ‘differential subject marking’ without further measures being taken, while accusative languages take recourse to a diathetic shift. Malchukov (2006) explained this asymmetry by a constraint against a change of case on the part of the (primary) argument in the unmarked case (Primary Actant Immunity Principle; cf. Unmarked Case Constraint in Tsunoda 1981). This constraint,
penalizing impersonal constructions, is also assumed in different frameworks (cf. the ‘final-1 Law’ of Relational Grammar, Extended Projection Principle in GB, and the ‘subject condition’ in LFG).

3.6 A note on passives and derived strategies

Above, we have noted that certain coding strategies are restricted to particular functional types of impersonal constructions. Some constructions have a broader range of applications. Especially passives, which come in different functional and structural subtypes, are notorious in that respect (see, in particular, Xrakovskij 1974; Comrie 1977; Siewierska 1984; Keenan 1985; Shibatani 1985; Keenan & Dryer 2007 for overviews). Typological properties of passives are a topic too vast to be addressed in any detail in this paper. What is of relevance here is that many ingredients of the passive prototype as described by Shibatani (1985) and others are also characteristic of impersonal constructions. Shibatani (1985), in part following Comrie (1977), ascribes to passives a general function of Agent-defocusing. He also shows that particular modes of defocusing may involve different parameters including (in)definiteness/genericity, but also topicality and agentivity. Thus, the functional domain of Shibatani’s study of passives reveals significant overlaps with the functional domain of impersonals as conceived here (which is maybe unsurprising given that Shibatani included also impersonal passives in his account). Also Sansò (2006) concludes that it is difficult to distinguish between two domains in functional terms.

Thus, passives in toto can be viewed as a maximally general “across-the-board” strategy registering different deviations from the subject prototype. In this respect the passive is less instructive for establishing particular connections (clustering) between individual subject properties. Nonetheless, certain formal varieties of passives might be more useful in that they may display a narrower range of applications than the passive per se and thus be more directly comparable to individual subtypes of impersonals. The agentless passive is a case in point. As is well known, constructions with indefinite or generic agents are commonly rendered by agentless passives. It is equally well-known (Xrakovskij 1974; Siewierska 1984: 35) that ‘short’ (agentless) passives are more common cross-linguistically, than ‘long’ passives having an overt agent phrase. A connection between indefinite agent constructions and agentless passives is further strengthened if one takes impersonal passives into account (see Siewierska 1984: 93–126 for discussion of impersonal passives). As observed in the literature (Frajzyngier 1982; cf. Blevins 2003), impersonal passives are usually agentless and imply an indefinite human subject.6

6. There are also aspectual constraints on formation of impersonal passives which won’t be addressed here; but see Abraham, this volume, Abraham 2006 for an extensive discussion (see also Ogawa 2005b).
For example, the “new” Polish impersonal passive, which is confined to intransitive verbs, normally does not permit an expression of agent (see, however, Kibort, this volume, for further discussion and qualifications):

(29) Polish

(T. Siewierska p.c.)

Tu było już sprzątane. (*przez pokojówki)
’Here it has already been cleaned. (*by the maids)’

On these grounds, Frajzyngier (1982) proposed to regard impersonal passives as a variety of indefinite subject construction. Yet, according to Frajzyngier (1982), this characterization only pertains to agentless impersonal passives. Impersonal passives with an overt agent (as in the following Dutch example) are considered by Frajzyngier as a variety of presentational sentences.

(30) Dutch

(Frajzyngier 1982: 281)

Er wordt hier door de jongelui veel gedanst.
Lit. ‘It is danced here a lot by the young people.’

In sum, impersonal passives cover the domain of constructions with indefinite or non-topical (new) subjects. Some other varieties of passives to be considered next may show sensitivity to other subject properties.

As has already been observed by Croft (2001), some passives are similar to inverse constructions in that they are used when the P ranks higher than A on the Animacy Hierarchy, and do not involve a clear reduction of transitivity (see Croft 2001 on a passive-inverse continuum). Shibatani (1998) also observes that accessibility of an NP to the Agent-of-passive function is determined by the “reversed empathy scale” where (non-volitional) Force and Instrument arguments rank higher than (volitional) Agents. Indeed, some languages like (Arizona) Tewa disallow (volitional) Agents in a passive construction, but allow for inanimate A arguments pertaining to natural forces and instruments. A somewhat similar case, but involving volitionality rather than animacy, is to be found in some New Indo-Aryan languages which feature ‘involitive passive’ forms. The involutive passive construction in Dhivehi is exemplified in (31b) and contrasted with the transitive construction in (31a).

(31) Dhivehi

(a) Aharen doru leppin.
I door close.pst non-3
‘I closed the door.’

(b) Ahannaš doru leppunu.
i.dat door close.inv.pst
‘I closed the door (involuntarily).’

7. Xrakovskij (1974), by contrast, regarded man-constructions to be a variety of passives.
Note that the verb in (31b) takes a special involitive form, while the A of the transitive construction (31a) is demoted to the dative (note also that subject agreement on the verb is hereby lost). And in Taba (Keenan & Dryer 2007 citing Bowden 1997) the passive marker used for agentless passives with transitives is used with intransitives to indicate diminished agency (non-intentional action).

In short, while passive formation in general seems to be “promiscuous” in functional terms, with agent-defocusing targeting different features from the subject prototype (Shibatani 1985), some individual passive subtypes might have a narrower range of applications. For example, agentless passives are used with indefinite/generic subjects, but the same forms may be used in presentational functions when the subject is expressed. And the passive form in Tewa is used when the agent is either inanimate or non-volitional. These specific varieties of passive are therefore more informative for uncovering connections between individual subject properties and individual functional varieties of impersonal constructions.

3.7 Conclusions

As shown above, a decrease in different subject properties may lead to different structural outcomes, in particular, to different formal varieties of impersonal constructions. In what follows, the common strategies are summarized in tabular form. In the table, we make a further distinction between impersonal constructions (here in a broad sense, at least showing some deviation from canonical subjects in their encoding), and alternative strategies which need not be impersonal. Our focus will be expectedly on the former, but the latter may also be interesting insofar as they reveal particular connections between subtypes of impersonals, as summarized later on a semantic map.

<table>
<thead>
<tr>
<th>Subject properties</th>
<th>Common outcome</th>
<th>Alternative strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) No referential argument</td>
<td>subject omitted/dummy subject</td>
<td>dummy verbs, cognate constructions (agentless) passive</td>
</tr>
<tr>
<td>(2) Subject is indefinite</td>
<td>indefinite pronouns, dedicated impersonal (verbal) forms; subject omission;</td>
<td>passive (with an agentive phrase) passive/inverse</td>
</tr>
<tr>
<td>(3) Subject is non-topical</td>
<td>word order inversion; agreement loss</td>
<td></td>
</tr>
<tr>
<td>(4) Subject is inanimate</td>
<td>differential case marking, agreement loss</td>
<td></td>
</tr>
<tr>
<td>(5) Subject is non-agentive</td>
<td>differential case marking; ‘differential indexing’</td>
<td>passive/anticausative</td>
</tr>
</tbody>
</table>
As is also clear from the table above, individual functional varieties of impersonal constructions prefer certain coding strategies. Thus, constructions with non-agentic or inanimate subjects (agents) may display differential case marking, with agreement and word order being unaffected. Constructions with non-topical (new) subjects, on the other hand, most frequently show word order inversion and sometimes also agreement loss, but hardly differences in case marking. Finally constructions with indefinite/generic subjects feature omission rather than demotions of notional subjects. Some other coding strategies (e.g. passive) have a broader domain of application, but subdivisions into further structural types (e.g. agentless passives) will pertain to a narrower range of functions.

Thus, with different functional types of impersonal constructions different coding properties seem to be primarily affected. The correlation between functional subtypes and coding strategies is not completely random. In fact, preferred strategies (favorite constructions) reflect most frequent functions of the respective coding mechanisms: case marking most typically reflects role properties, agreement typically indexes a prominent NP, while word order usually reflects information structure (see Lehmann 1988 and Malchukov & Spencer 2009 for a comparison of different functions of coding strategies). Functional considerations can also explain counterexamples to such hierarchies, as when the case is sensitive to information structure (see Ogawa 2009 for discussion of the famous ga/wa alternation on subjects in Japanese, and Malchukov & Spencer 2009 for discussion of other languages with “pragmatic cases”).

Switching to a diachronic perspective, these data suggest that different varieties of impersonals will lose their subject properties in a different order. As shown for the varieties of impersonals sensitive to topicality distinctions (e.g. presentational constructions), the loss of (coding) properties of subjects proceeds along the following hierarchy: position > agreement > case. This is in accordance with the hierarchy of acquisition of subject properties posited by Givón (1997: 29) who also discusses the data from presentational constructions. As is clear from the above, either the changes affect word order alone (as in Russian or English), or both position and agreement (as in French or in the Russian construction with numeral phrases), and if case is involved other features should be involved as well (as in Finnish in (18)). The same can be observed in instances of locative inversion, where postposed (non-topical) subjects lose agreement before they lose case (on the assumption that Sesotho (in 21), where

---

8. Berman (1980) discusses a similar case of loss/redistribution of subject properties in a possessive construction in Hebrew, which is reminiscent of the Finnish pattern in (18). Also in Hebrew the initial subject (the possessed item) loses its subject position and agreement control, but may also optionally take an object-marker (preposition).
agreement but not case is affected manifests an earlier stage of reanalysis than Swahili (in (20b)).

On the other hand, impersonal constructions featuring non-agentive subjects rather shed the subject case marking at early stages. This is most clear for ergative languages (like Lezgian, as in (25) and Samoan, as in (23)), where ergative marking is not coupled with agreement, so agreement is left unaffected. So the hierarchy of loss of coding properties looks rather like: case > agreement > position. Finally, with respect to impersonal constructions with indefinite/generic subjects, the hierarchy is irrelevant, at least in constructions in which the notional subject is not expressed. So it seems that both acquisition and loss of subject properties (which are, of course, two sides of the same process) would crucially depend on the functional characteristics of the construction in question.

4. Towards a semantic map for the impersonal domain

Above we have provided a short summary of some preferred strategies of encoding of particular types of impersonal constructions displaying different deviations from the subject prototype. It has also been noted that certain constructions (like passives) have a broader application range, while some other strategies are rather dedicated to a particular type of impersonal construction (e.g. cognate constructions are found with atmospheric events). For such domains where individual categories display both differences and selective similarities (shared constructions) a semantic map approach has proved to be particularly fruitful.

The semantic map approach aims at capturing functional similarities between individual categories through a study of cross-linguistically recurrent polysemy patterns (see Anderson 1982; Croft 1990; 2001; van der Auwera & Plungian 1998; Haspelmath 2003; Malchukov 2010). On iconicity assumptions, cross-linguistically recurrent identity in form is assumed to reveal semantic similarities. In this way even a limited sample of (genetically and areally unrelated) languages may reveal a certain pattern, which not only can capture cross-linguistic variation but also make certain predictions. The latter is possible inasmuch as there is a ‘contiguity requirement’ on semantic maps which disallows gaps on the map: if two non-adjacent functions share the same form, the intermediate functions will do so as well. In this way the semantic

9. This does not necessarily mean that impersonal constructions develop at intermediate stages of realignment of a structure undergoing locative inversion. For example, Creissels (2009) suggests that the impersonal construction in Tswana (a language closely related to Sesotho) represents rather fossilization of the erstwhile construction with locative inversion.
map methodology can be used to constrain typological variation in a certain domain. In this paper we will use this method for establishing the semantic map of impersonal constructions.10

In what follows we will sketch the semantic map for the impersonal domain, based on the previous discussion (in particular on the data summarized in Table 1 above). The map is arranged so as to represent the formal similarities (shared constructions) between the individual functional subtypes of impersonal constructions. Of special interest here are instances where two functional varieties share a similar coding strategy which strengthens the assumption that the two functional varieties are treated as one natural class cross-linguistically. In other words, while differences in encoding provide evidence for recognizing them as particular subtypes, similarities, insofar as they are cross-linguistically recurrent, are assumed to reveal functional similarities.

As is clear from the previous discussion certain strategies are restricted to certain subtypes of constructions: thus differential case marking is found with non-agentive, but also with inanimate subjects/agents (recall the case alternation in the ergative languages Lezgian (25) and Samoan (23)). Non-prominent subjects (non-agentive and/or non-topical) are often not cross-referenced through agreement (as exemplified above for Tlapanec (22b) and Oromo (19b)). Non-topical and new subjects are frequently postposed (in a presentational construction), as discussed in §3.3 above. Constructions with non-definite (human) subjects may take the form of an impersonal passive construction, but impersonal passives can be also used in the presentational function

10. Although, to the best of our knowledge, the semantic map approach has not been attempted for the impersonal domain, in the literature we do find ‘scales of impersonality’ aimed to model a gradience from personal to impersonal structures. An early attempt is a typological study by Moreno (1990), where the author proposes a scale including a variety of impersonal constructions ranging from less agentive (weather verbs) to more agentive (specific agents in a passive construction). Yet, Moreno’s scale, as also more sophisticated scale recently proposed by Siewierska (2008), relying on notions of instigator elaboration and instigator demotion, is designed to model impersonality gradients but cannot (and is not intended to) capture formal similarities between individual subtypes of impersonal constructions. The recent study by Sansò (2006), on the other hand, is more similar to the approach advocated here but more distinct in the domain of application. Sansò (2006) pursued a semantic map approach with respect to agent-defocusing constructions, but his study is more concerned with passive constructions in general rather than impersonals per se. The two types of agent-defocusing identified in this study are associated with promotional/personal passives (‘patient-oriented events’), on the one hand, and with demotional passives/impersonals (‘agentless generic events’), on the other hand. Barðdal (2004) adopts a semantic map perspective in her discussion of impersonal constructions in a number of Germanic languages, but her discussion centers on psych-verbs. Finally, Ogawa (2005, 2006) suggested a scale of impersonal constructions (which may be viewed as a one-dimensional semantic map), but only for the domain of the meteo-expressions.
when the subject is overt (see, e.g. Payne this volume, on Maa). And in Mandarin, the presentational marker *you-ren* has extended its use to an indefinite agent marker (Yi & Siewierska this volume). Finally, non-referential and indefinite subjects are frequently omitted, as is familiar from European (pro-drop) languages. All these shared strategies support the configuration of the semantic map in Figure 1 below.

- Differential case marking:  
- Agreement loss:  
- Word order inversion:  
- Impersonal passivization:  
- Subject omission:  

![Figure 1. Semantic map of the impersonal domain](image)

In the semantic map above, those functions which share the same strategies are placed adjacently on the map, and joined by connecting lines. (A particular connection displayed by an arrow from non-referential to non-volitional As has not be discussed so far, and will be taken up in the next section.) The oval shapes show the extension of the particular constructions for individual functional types of impersonal constructions. Note that this two-dimensional semantic map is ‘well-formed’ insofar as it contains no gaps: as is clear from the above strategies extend contiguously to adjacent categories leaving no gaps.

While the configuration on the semantic map is arrived at empirically through the study of cross-linguistic polysemies of particular forms/constructions, semantic maps are designed to capture semantic relatedness of individual functions. Recall that on iconicity assumptions, which underlie the semantic map methodology, recurrent identity of form reflects similarity of meaning. In our case the resulting configuration clearly makes sense in functional terms, as adjacent categories share certain semantic and/or pragmatic features. Thus, animacy correlates with volitionality insofar as only animate entities can be volitional. A connection between animacy and topicality is also well-established in the functionalist literature even though its nature is pragmatic.
rather than semantic. It builds on the notion of empathy which ensures that animate entities are more topical in discourse than inanimate ones (cf. e.g. Givón 2001). In the literature, both dimensions are sometimes conflated under the notion of ‘prominence’. As also clearly reflected by the map, non-topical arguments may be non-topical either by virtue of being indefinite/nonspecific, or by virtue of being new (as in presentational constructions). The connection between topicality and definiteness/referentiality is also familiar from the functionalist literature: in fact, Givón’s (2001) concept of ‘referential importance’ or Payne’s (1997) concept of ‘discourse referentiality’ covers both phenomena. Finally, (in)definiteness and (non)specificity are semantically related insofar as definiteness presupposes specificity.

The semantic map in Figure 1, as semantic maps in general, can be used not only for the description of individual patterns but also for constraining possible patterns. This is due to the abovementioned contiguity requirement which implies that if two nonadjacent categories on the map share the same strategy (construction), the same strategy should be found with intermediate categories (functions) on the map. This can be also demonstrated with respect to some unusual varieties of impersonal constructions which might be potentially problematic for the proposed map. Thus, Kambera (Klamer 1998), has been described as a split intransitive language where a choice of the subject vs. object agreement is sensitive to features of control/volitionality:

(32) Kambera (Klamer 2009: 294)
   a. Hi na=hi=ma=a=ka  i Mada una…
      and 3SG.NOM=cry=EMPH=MOD=PFV ART Mada EMPH.3SG
      ‘… and Mada was just crying…’
   b. Hi hi=ma=a=ya=ka  i Mada una…
      and cry=EMPH=MOD=3SG.ACC=PFV ART Mada EMPH.3SG
      ‘… and Mada just cried and cried.’ (i.e. could do nothing else)

This pattern of ‘differential indexing’ driven by volitionality is cross-linguistically not unusual, as illustrated by Koasati above. What is less usual though is that a switch to object agreement can also signal indefiniteness on the part of the subject:

(33) Kambera (Klamer 1998: 166)
     jåka nda yumu meti-ya-ka lâti.
     CNJ NEG you die-3SGACC-PFV in_fact
     ‘Without you, we (lit one) would have died.’

As noted by Klamer, inclusion of a lexical subject (for example, a proper name) is impossible in this construction. This kind of polyfunctionality is both unusual and problematic for the semantic map insofar as the functions of indefiniteness and
volitionality are non-adjacent on the map. Yet, it seems that the problem is more apparent than real since apart from the dimensions of definiteness and volitionality, ‘differential indexing’ can also signal intermediate functions related to prominence/topicality. As reported by Klamer (1998) object-agreement is also used in cases when the subject is backgrounded, as for example in verb-focus constructions with reduplicated predicates.

Such “global effects” on semantic maps can be related to a general notion of subject-prominence, which invokes a concept of a canonical subject which is agentive, topical and definite. As noted above, one of these global constructions are passives which have been shown to have a general function of Agent-defocusing (Keenan 1976; Shibatani 1985). Such global strategies are not in itself problematic for a semantic map, rather they don’t provide supporting evidence for particular connections. Yet, as observed above, once we distinguish between particular subtypes of passive constructions (agentless vs. agentive passives, impersonal passives, ‘inverse-like’ passives, which do not show any obvious decrease in transitivity, ‘anticausative-like’ passives, etc.) these structural subtypes have a narrow domain of application on the semantic map, and can support particular links between individual functional subvarieties of impersonal constructions.

Of course, many of the notions on the semantic map above can be further decomposed and reveal on closer inspection a richer structure, with further links which need to be investigated. Yet, already in this form the map provides a viable hypothesis concerning the limits of cross-linguistic variation in the domain of impersonal constructions.

5. Transimpersonal constructions: An additional link on the map?

Above we have shown that different functional varieties of impersonals show preferences for different coding strategies. In most general terms, one could distinguish between the domain of A-impersonals (sensitive to agentivity/animacy properties), T-impersonals (sensitive to topicality), and R-impersonals (sensitive to referentiality/definiteness properties). We further suggested that T-impersonals provide a bridging domain (intermediate link) between the domains of A-impersonals and R-impersonals (as is also clear that the notion of ‘prominence’ pertains to the A-T

11. Thus, Mettouchi & Tosco (this volume), point out a connection between one type of weather-expressions which feature an overt subject and an auxiliary verb (“dummy verb”) to thetic utterances. This connection is not represented on our map as it focuses on impersonal varieties of weather expressions.
Towards a typology of impersonal constructions

Yet, there appears to be a further link between the domains of R-impersonals and A-impersonals, which has not been discussed before, yet which is assumed or proposed in specific analyses. For example, Mel’čuk (1979; cf. Mel’čuk 1995) analyzed the Russian construction such as (24b) as involving a zero ‘Force’ argument. In other words, (24b) is assigned the structure in (ii) below rather than the one in (i):

(i)  \[ \text{Dorogu}_p \text{ zaneslo } \text{snegom}_A \]
(ii)  \[ \text{Dorogu}_p \text{ zaneslo } \emptyset_A \text{snegom}_{\text{Instr}} \]

While some arguments for this analysis may be theory internal (and follow from Mel’čuk’s conception of zero-signs; cf. Zimmerling 2007 for cross-theoretical discussion), similar analyses have been proposed in other frameworks. For example, working in Lexical Functional Grammar, Kibort (2008) proposes a similar account for Polish, which also implicates a null pronominal element (\textit{pro}) with a generic meaning (‘something’). And, while we analyzed the Samoan pattern involving an alternation of an ergative and oblique marking in terms of differential case marking, Moyse-Faurie (this volume) proposes to analyze a similar construction in East Futunan as a transitive construction with a covert (ergative) agent and an overt adjunct (referring to a natural force).

This structural ambiguity might suggest an additional link to be drawn between non-volitional and non-referential subjects. Yet, this link we would argue is of a different, diachronic, nature. By now, there is ample evidence that constructions with an indefinite or generic subject tend to be reanalyzed into intransitive or passive constructions. It is more convenient to start with discussion of a related path of reanalysis of indefinite subject constructions into passives, as it is better documented in the literature (see Greenberg 1959; Keenan 1976; Givón 1979; Siewierska 1984; Keenan 1985; Shibatani 1985; Heine & Claudi 1986; Haspelmath 1990; Palmer 1994; Creissels 2007; Keenan & Dryer 2007; Malchukov 2008; but see Siewierska 2010, and Payne this volume, for cautionary remarks). As amply demonstrated in this body of work, indefinite subject constructions tend to be reanalyzed into passive constructions through a stage

\[ \text{12. } \text{Compare a somewhat similar distinction proposed in the generative framework by Platzack (2001) and Svenonius (2002), who distinguish between V-domain, I-domain and C-domain subject properties. V-domain properties refer to thematic/aspectual subject properties and correspond roughly to our A-impersonals, while C-domain properties pertain to discourse-informational subject properties and correspond to our T-domain impersonals. I-domain properties, however, relate to coding properties (by case and agreement) and are treated differently in our approach.} \]
of impersonal passives. The source construction might involve an indefinite subject pronoun (as in Babungo in (13)), or an agreement marker signaling the presence of an indefinite subject. At later stages, the construction is reanalyzed into an impersonal passive. Both stages can be illustrated by the data from Ainu, which features a special impersonal (indefinite subject) prefix a(n)- (Shibatani 1985; Tamura 2000; Bugaeva this volume)

(34) Ainu

   INDF-2SG.P-scold MOD
   'You will be scolded/one will scold you.'

b. Unuhu oro wa an-kóyki.
   mother place from INDF-scold
   'He was scolded by (his) mother.' (lit.?? 'We/one scolds him by his mother')

The construction (34a) is a transitive impersonal construction with an indefinite A argument (signaled by the a(n)- prefix), and with the object experiencer indexed in the regular way (through an object agreement prefix). The example (34b) is similar as far as the cross-referencing is concerned, but is different inasmuch as the agent is introduced here by the oblique (prepositional) phrase. This means that the construction can no longer be analyzed as an active construction with an indefinite A, but is on the way to being reanalyzed as a passive construction. Still at this stage, the construction is an impersonal rather than a personal passive, as there is no conclusive evidence that P is promoted (see Bugaeva, this volume, for the arguments, that it is not). Moreover, Siewierska (2010) argues that, in general, cases where the reanalysis into personal passive is complete are relatively rare.\(^\text{13}\) An important point here is that demotion of an A predates promotion of a P, and also that constructions with a demoted A defy the analysis of the construction as an active transitive clause. In the same way, we would argue that an indefinite subject construction with the structure in (ii) can be further reanalyzed into the oblique-A construction in (i).

A related path of reanalysis, which has attracted attention more recently, is the reanalysis of ‘trans impersonal’ constructions (i.e. constructions with an experiencer object and an indefinite subject, frequently referring to an inanimate force) into patient-subject constructions in split intransitive languages (Malchukov 2008; see also Mithun 2008; Holton 2008; Donohue 2008). At early stages such constructions can still

\(^{13}\) Dryer & Keenan (2007), however, mention that while for several Algonquian languages (like Ojibwa) it is not clear whether one is (still) dealing with an “unspecified subject construction” or with a (promotional) passive, for other languages (like Kutenai) the passive analysis is uncontroversial.
be regarded as elliptical versions of a transitive structure with an impersonal agent. This source construction is found, for example, in Turkana:

(35) Turkana (Dimmendaal 1983:73)

\[ K.\text{à}-\text{bur-un-it’} \ a-yøf’ \ (i-\text{børe}). \]

obj.1sg-tire-VEN-ASP me (thing)

‘I am tired’ (lit. It/(some)thing tires me.’)

The construction above still qualifies as a transitive structure with the inanimate generic subject ‘thing’ and an object experiencer (indexed by the object agreement marker \( k.\text{à} \)). Yet, already at this early stage the expression of the generic A is optional (as indicated by the brackets), and the construction itself is prone to further reanalysis. Indeed, for many split intransitive languages, it can be shown that the patient subject construction developed from reanalysis of the transitive construction with inanimate and indefinite agent (Malchukov 2008; see also Mithun 2008 on Native American languages). Both Mithun (2008) and Malchukov (2008) noted the presence of object agreement on the verbs and zero expression of (inanimate) agent as factors facilitating this reanalysis. Thus, (36) from Koasati (Kimball 1991), could still be analyzed as a transitive clause (‘It/something burned me’), yet the construction is currently intransitive as is clear from the form of the subject pronoun (which if overt is nominative):

(36) Koasati (Kimball 1991:251)

\[ \text{Ca-libátli-t.} \]

1sg.p-burn-PAST

‘I got burned.’

Similar examples of reanalysis of transimpersonal constructions into patient-subject intransitives have been reported elsewhere (see, e.g. Miyaoka, this volume, on Central Alaskan Yupik, and Vajda et al. this volume, on Ket).

Thus, reanalysis of transimpersonal constructions provides evidence for yet another connection between transimpersonal constructions with indefinite A arguments, and constructions with non-canonical subjects lacking agentivity (indeed they are experiencers or patients rather than agents). In fact, this connection seems even more robust here, insofar as constructions with patientive subjects can be construed as constructions with an unnamed external cause. Yet, it is still clearer in this case that the relation between the two parameters is diachronic rather than synchronic, as the parameters pertain to different NPs (to an indefinite/generic agent in the source construction, and to the non-volitional subject-patient in the target construction).

In a recent paper, Creissels (2007) takes this analysis one step further. He notes that impersonal constructions with ‘oblique subjects’ as familiar from European languages
can be also analyzed as a case of alignment split, similar to cases of split intransitivity considered above. In particular, Creissels discusses three different usages of the Russian verbs *tr'astî* 'shake' as illustrated below:

(37) Russian
   a. *Men’â tr’asët lîxoradka.*
      1SG.ACC shake.PRS.A3SG fever.GEN
      'Fever shakes me.'
   b. *Men’â tr’asët ot lîxoradki.*
      1SG.ACC shake.PRS.A3SG from fever.GEN
      'I am shaking with fever,' lit. 'It shakes me from fever'
   c. *Men’â tr’asët.*
      1SG.ACC shake.PRS.A3SG
      'I am shaking/trembling/shivering,' lit. 'It shakes me'

While (37a) is a regular transitive construction with an inanimate “Force” argument, the construction in (37b) is impersonal: the verb appears in the default agreement form and the cause of the event is in the instrumental case. Finally, (37c) is a transimpersonal construction with an unexpressed A argument (it need not pertain to a disease but may imply any unnamed external force). Creissels (2007) further points out that the construction in (37b) with a demoted A argument cannot be regarded any longer as either a version of the transitive construction (cf. (37a)) or of the transimpersonal construction (37c), but rather represents a subsequent stage of reanalysis into a subject-patient construction.

While this latter conclusion might be questioned on syntactic grounds (the accusative experiencer does not show any syntactic subject properties typical of canonical subjects in Russian, like converbial control; cf. Donohue 2008 for a general discussion), the general diachronic interpretation is valid: the expression of the A through an oblique is a first step in the reanalysis. One could further observe that the conventionalization of the intransitive subject ellipsis leading from an elliptical version of the transitive construction to transimpersonal construction needs to predate the construction with

---

14. Creissels 2007: “It seems reasonable to assume that the impersonal construction developed from the reanalysis of a null-subject construction with an arbitrary reading: ‘[An unspecified cause] shakes me’. But the fact that the cause is encoded as an oblique introduced by the ablative preposition *ot* proves that, in the present state of Russian, this construction is no longer an elliptical variant of the transitive construction.”
a demoted A. Admittedly, a shift from a transitive construction with subject ellipsis to a transimpersonal construction can’t be easily diagnosed in most languages, but some languages do provide evidence of reanalysis. Thus, in Finnish we can tell the difference since transimpersonal constructions license nominative marking of objects (see the “impersonal passive” in (37)), while transitive constructions (including “zero constructions” with indefinite subject ellipsis) take a (definite) object in the accusative (Helasvuo & Vilkuna 2008):

(38) Finnish

\[ Kone\ ostet-t-i-in. \]
\[ \text{machine(NOM) buy-PASS-PST-PERS} \]
‘The machine was bought.’

A final development on this path of reanalysis would be promotion of the notional patient/experiencer to subject status. This stage arguably has not been reached yet in Russian, but is attested in some other languages such as Icelandic (also discussed in Creissels 2007), as well as in the split-intransitive languages mentioned above.

Thus, there is a clear connection between transimpersonal constructions with an indefinite/generic subject and constructions with non-agentive subjects, be it primarily a diachronic one. Generalizing across the two scenarios (from indefinite subject to passive, and from transimpersonal to the intransitive pattern), reanalysis seems to involve the following stages: (a) conventionalization of ellipsis of the generic/indefinite subject; (b) the possibility of reintroducing the notional A as an oblique; (c) promotion of the notional P to subject position (a gradual process in itself which starts from behavioral properties, and is then followed by coding properties; Cole et al. 1980; Givón 1997). At intermediate stages, the construction might give rise to alternative analyses.

---

15. Siewierska (2010) in the discussion of conventionalization of impersonal forms based on 3rd person plurals noted that on earliest stages the meaning of the zero pronouns is most closely related to a range of meanings of a 3rd person plural pronoun, which we take as indication of the fact that the source construction involved subject ellipsis.

16. Kulikov (this volume) posits a similar development from agent ellipsis (in a passive construction) to zero agent to account for a shift from passives to anticausatives.

17. It is a moot point whether an oblique argument can still be analyzed as A at this stage, as implied by the analysis of such patterns in terms of differential case marking. Thus, ergative-oblique variation in Lezgian (illustrated in (25) above) has been analyzed by Melčuk (1988) as differential coding of agents, but Haspelmath (1993) gives arguments for analyzing the pattern with the oblique Force as an intransitive structure. The latter analysis is clearly preferable for Akhvakh (D. Creissels p.c.), where transitivity distinctions are signaled by the presence of the causative marker on the verb in the ergative pattern. On the other hand, Ganenkov et al. (2008) in a discussion of Agul (still another Daghestanian language) observe that the change of case-marking (from ergative to ad-ative) does not ostensibly affect (syntactic) subject...
however already at the second stage when the A can be (re)introduced as an oblique
it seems unnatural to analyze the construction as an active transitive clause with an
indefinite/generic A. In both cases reanalysis is driven by functional pressure to syn-
tactically upgrade a prominent (animate/topical) object-experiencer, and downgrade
the non-prominent (inanimate/generic/non-topical) actor (Malchukov 2008; see also
Haspelmath 2001 on the role of topicality in creation of non-canonical subjects).

A question which is relevant for our study is whether this additional diachronic
link between indefinite/generic subjects and inanimate/non-volitional subjects
should be represented on our semantic map. At first sight the answer appears to be
yes, given the current consensus that semantic maps can also be viewed as synchronic
representations of grammaticalization paths (see van der Auwera & Plungian 1998;
Narrog & van der Auwer 2010). Yet, this conclusion does not seem to apply here. Note
that while there is a clear diachronic connection between transimpersonal construc-
tions with object experiences and subject-experiencer constructions which would
necessarily qualify as non-volitional, no single construction displays both meanings.
Rather, the two constructions prior to reanalysis or after that, are mutually disjunctive.
Thus although a possibility of viewing semantic maps diachronically (“dynamiciza-
tion”) is beyond dispute, it is not always the case that a particular diachronic develop-
ment is restricted to adjacent categories on the map. This is true for standard cases of
grammaticalization (where the old and a new reading more often than not coexist),
but less so for scenarios involving reanalysis where the two uses are mutually disjunc-
tive. This has been also observed by Narrog (2010) for a particular case of reanalysis
involving case markers (from a passive agent marker to an ergative case) which does
not leave any traces on a synchronic semantic map, but it seems to be true for (most)
other cases of reanalysis as well. Therefore this connection is of different nature, and is
represented on our semantic map in a different way as compared to other connections
(as indicated by the arrow).

6. Conclusions

The present article proposes a semantic map for the domain of impersonal construc-
tions. It started from the observation (originally due to Keenan) that different variet-
ies of impersonal constructions reflect the loss of functional subject properties (such
as definiteness, topicality and agentivity). We further showed that different functional
varieties of impersonal constructions (where the subject lacks some of the functional
properties. This seems to be more in line with Mel’čuk’s analysis (except for the fact that
Mel’čuk treated both patterns as intransitive rather than transitive).
properties of a prototypical subject) are coded by different varieties of impersonal constructions, although certain constructions have a broader range of applications. On the basis of this partial overlap between different coding strategies for different functional varieties of impersonal constructions, we proposed a semantic map for the impersonal domain, which can be used to represent and constrain variation of coding of impersonal constructions across languages. The main functional varieties in this conceptual space of impersonal constructions are A-impersonals (sensitive to agentivity properties), T-impersonals (sensitive to topicality), and R-impersonals (sensitive to referentiality/definiteness of the subject referent). We further argued that T-impersonals constitute a bridging domain between A-impersonals, on the one hand, and R-impersonals on the other hand. A further connection from R-impersonals to A-impersonals has been shown to be of a diachronic nature, and does not result in a regular polysemy pattern.

Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABS</td>
<td>absolutive</td>
</tr>
<tr>
<td>ACC</td>
<td>accusative</td>
</tr>
<tr>
<td>ADEL</td>
<td>adelative (case)</td>
</tr>
<tr>
<td>ADESS</td>
<td>adessive (case)</td>
</tr>
<tr>
<td>AOR</td>
<td>aorist</td>
</tr>
<tr>
<td>ART</td>
<td>article</td>
</tr>
<tr>
<td>ASP</td>
<td>aspect (marker)</td>
</tr>
<tr>
<td>CL</td>
<td>class (marker)</td>
</tr>
<tr>
<td>CNJ</td>
<td>conjunction</td>
</tr>
<tr>
<td>DAT</td>
<td>dative</td>
</tr>
<tr>
<td>IMPS</td>
<td>impersonal</td>
</tr>
<tr>
<td>IND(e)F</td>
<td>indefinite</td>
</tr>
<tr>
<td>ERG</td>
<td>ergative</td>
</tr>
<tr>
<td>EMPH</td>
<td>emphatic</td>
</tr>
<tr>
<td>FM</td>
<td>feminine</td>
</tr>
<tr>
<td>FUT</td>
<td>future</td>
</tr>
<tr>
<td>GEN</td>
<td>genitive</td>
</tr>
<tr>
<td>LOC</td>
<td>locative</td>
</tr>
<tr>
<td>MS</td>
<td>masculine</td>
</tr>
<tr>
<td>MOD</td>
<td>modal</td>
</tr>
<tr>
<td>NOM</td>
<td>nominative</td>
</tr>
<tr>
<td>NZR</td>
<td>nominalizer</td>
</tr>
<tr>
<td>OBJ</td>
<td>object</td>
</tr>
<tr>
<td>P, PAT</td>
<td>patient</td>
</tr>
<tr>
<td>PART</td>
<td>participle</td>
</tr>
<tr>
<td>PASS</td>
<td>passive</td>
</tr>
<tr>
<td>PL</td>
<td>plural</td>
</tr>
<tr>
<td>PRES</td>
<td>present</td>
</tr>
<tr>
<td>PFV</td>
<td>perfective</td>
</tr>
<tr>
<td>P(A)ST</td>
<td>past</td>
</tr>
<tr>
<td>REFL</td>
<td>reflexive</td>
</tr>
<tr>
<td>SBJ</td>
<td>subject</td>
</tr>
<tr>
<td>SG</td>
<td>singular</td>
</tr>
<tr>
<td>TOP</td>
<td>topic</td>
</tr>
</tbody>
</table>

References


Towards a typology of impersonal constructions


Malchukov, A. 2008. Split intransitives, experiencer objects, and ‘transimpersonal’ constructions: (Re-) establishing the connection. In M. Donohue & S. Wichmann (eds), 76–100.


Ogawa, A. & Malchukov, A. In preparation. Lexical typology and atmospheric events.


Overlap and complementarity in reference impersonals

Man-constructions vs. third person plural-impersonals in the languages of Europe

Anna Siewierska
Lancaster University

This paper is concerned with two kinds of so-called R-impersonals, i.e. impersonals triggered by a reduction in referentiality, namely third-person plural impersonals (3pl-imp) and man-constructions (man-imp). In the languages of Europe, both of these impersonal constructions display a wide range of uses including referential uses and therefore enter into competition with topicality and definiteness based-impersonals. Building on the work of Giacalone Ramat & Sansò (2007), the current paper maps out the distribution of 3pl-imp relative to man-imp and relates these patterns to an important typological factor in the European context, namely the formal realization of pronominal subjects (see e.g. Holmberg 2005; Cabredo Hofherr 2006; Siewierska 2008). This relationship will, however, be argued to be somewhat more nuanced that has been previously assumed.

Keywords: generic; episodic; existential; passive; agentivity; pro-drop; voice

1. Introduction

As discussed by Malchukov & Ogawa (this volume), impersonalization, if viewed as a departure from canonical subjecthood, may be triggered by a variety of properties often working in tandem such as agentivity, animacy, topicality, definiteness and referentiality. This paper will be concerned only with impersonals triggered by a reduction in referentiality, R-impersonals, for short.¹ R-impersonals have the appearance of regular, personal constructions but feature a subject which is human

¹ There is an obvious functional overlap between R-impersonals in the sense of the term discussed here and agentless passives, both those with and without a thematic subject. Under the approach to impersonalization espoused here, however, passives are primarily impersonal by
and non-referential. The non-referential human subject may be expressed lexically, pronominally or by the whole construction. The subject of lexical R-impersonals is typically the word for ‘person’ or ‘people’, for instance, \textit{a-way} in Abkhaz (1) or \textit{ludzie} in Polish.

(1) Abkhaz \quad \text{(Hewitt 1979: 157)}
\begin{align*}
a-way & \quad \text{aray} \quad a \quad \z^0q^0 \quad d\-\text{-a}-px\-a\-r \quad ak\-\text{or} \quad \text{Ø-ey-la} \\
\text{art-person} & \quad \text{this} \quad \text{art book} \quad \text{he-it-read something} \quad \text{it-preverb} \\
y & \quad k\-t \quad aa\-\text{-}we-yt' \\
\text{he-learnt-} & \quad \text{suff-} \quad \text{dyn-finite} \\
\end{align*}

‘If one reads this book, one will learn something.’

Pronominal R-impersonals come in a variety of guises. The subject may be rendered by a pronominalized form of the numeral ‘one’ such as the Basque \textit{bat} (2) or Spanish \textit{uno}, a pronominalized form of the noun for ‘person’ or ‘human’ as in the case of the Danish \textit{man} (3), a regular personal pronoun, such as the English \textit{we}, \textit{you} or \textit{they}, be it used non-referentially as in (4), or a special nonspecific pronominal form such as \textit{kas nors} in Lithuanian (5).

(2) Basque \quad \text{(Saltarelli 1988: 208)}
\begin{align*}
Bat\-ek & \quad ba \quad al\-\text{daaki} \quad bada \quad zer \quad eg\-n? \\
one\-\text{erg} & \quad \text{EMP} \quad Q \quad \text{knows anyhow} \quad \text{what} \quad \text{do-prf} \\
\end{align*}

‘Does one really know what to do anyhow?’

(3) Danish \quad \text{(Jensen 2009: 5)}
\begin{align*}
\text{Man kunne ikke komme ind i ungdomsklubben} \quad \text{om aftenen.} \\
\text{MAN could not come in} \quad \text{in youth club} \quad \text{in the-evening} \\
\end{align*}

‘You couldn’t get to the youth club in the evening.’

(4) a. \textit{We find this species in South America.}

b. \textit{You would not get away with such a sexist remark in Britain.}

c. \textit{They’re going to raise taxes.}

(5) Lithuanian
\begin{align*}
Mayt, & \quad \text{kas nors} \quad juo \quad \text{bus} \quad \text{pasiskundęs.} \\
\text{apparently} & \quad \text{who-nonspec} \quad \text{her was complaining} \\
\end{align*}

‘Apparently, someone (non-specific) complained about her.’

Arguably falling also under pronominally-based R-impersonals are instances of subject ellipsis which contextually receive necessarily an unspecified human interpretation, as in Lezgian (6), for example.

\begin{align*}
\text{Mayt, kas nors juo bus pasiskundęς.} \\
\text{apparently who-nonspec her was complaining} \\
\end{align*}

Apparent cases of subject ellipsis which contextually receive an unspecified human representation, as in Lezgian (6), for example.
Overlap and complementarity in reference impersonals

(6) Lezgian (Haspelmath 1993: 288)

Hi μού hul ø sirnaw t-aw-r-ta-ni
which see ERG swimming NEG-do-AORIST.P.-CND-even
ø batmiš ž e-da-c?
ABS sink ANTIC-FUT-NEG

‘In which sea does one not sink even if one does not swim?’

The third type of R-impersonals, where the human non-referential subject is not tied directly to a specific morpheme are constructions such as the impersonal reflexives of Romance or Slavic, as in (7) or the participal impersonals of Balto-Finnic, as in (8).

(7) Italian (Cinque 1988: 522)

Si lavora sempre troppo.
si work:SG always much
‘One always works too much.’

(8) Estonian (Torn-Leesik 2007: 3)

tollal leö-ti peamiselt ilukirjandust.
then read-IMP:PAST mainly fiction:PAR
‘At that time one mainly read fiction.’

While all R-impersonals, by definition, may be used in generic contexts to express non-referential human subjects, some are also used in episodic contexts to denote partially specified sets of human subjects or even individuals. For instance in (9) man denotes the teachers in Hogwarts, and in (10) the referent of the 3pl is clearly someone known to the speaker.

(9) German

Ich war selbst in Hogwarts, doch ich – ähm -man
I was myself at Hogwarts but I ehm MAN
hat mich rausgeworfen, um dir die Wahrheit zu sagen.
had me expelled to you the truth to say
‘Oh, well – I was at Hogwarts meself but I – er -got expelled, ter tell yeh the truth.’

(10) Russian (Perlmutter 2001: 10)

Pozvonili čtoby skazat’ čto vystupenie budet v sredu.
called:3PL in-order say:INF that performance be:FUT on Wednesday
‘Someone called in order to say that the performance will be on Wednesday.’

In the case of the latter, the non-referential use may even wane in favour of the episodic. This in turn may lead the relevant R-impersonal to enter the territory of
one of the other type-of impersonals, especially of the topicality and definiteness based types.

In the languages of Europe, chief among the R-impersonals which evince not only non-referential but also referential uses and enter into competition with topicality and definiteness based-impersonals are man-constructions (man-imps) and third-person plural impersonals (3pl-imps). The term man-imp is used here to refer to an impersonal construction which denotes an unidentified human subject expressed by a word etymologically related to ‘human’ or ‘man’, such as on in French, hom in Catalan and man in German, illustrated above in (3) and (9). Significantly, not included under man-imps are the previously mentioned impersonal constructions with an unidentified human subject expressed by the word for ‘one’ or a word etymologically related to ‘one’ as illustrated earlier in (2) on the basis of Basque. The latter typically have a narrower range of uses than man-imps. By 3pl-imp, will be meant in this study a construction with an unspecified human subject expressed by a third person plural pronoun which, in contrast to what is the norm for personal pronouns, lacks an overt antecedent in the discourse, even a non-specific one such as everyone, all or anybody. The third person plural subject may be a morphologically free though phonetically unstressed (and often reduced) form, as in the case of the English they exemplified in (4c) above or a bound form, as in the Russian (10).

Man-imps and 3pl-imps are common in the languages of Europe. The vast majority of European languages manifest at least one of these constructions and some have both. Being both R-impersonals, they often function as translational equivalents of each other and also of other impersonals (see e.g. Jisa & Viguié 2005: 130; Tolchinsky & Rosado 2005: 15). For example, the translation of the English get-passive in (11a), taken from the first book of Harry Potter, is the man-imp in French (11b) and a 3pl-imp in Polish (11c).

(11) a. “Red card!” said Dean furiously. “In soccer you get shown the red card and you’re out of the game!”

b. On n’est pas au football, l’interrompit Ron. On ne peut
man neg:is neg in football him-interrupted Ron man not can
pas renvoyer les joueurs.
not send away the players

‘We are not at a football game, interrupted Ron. We/one cannot send away the players (here)’.

2. One-impersonals in Europe are found mainly in Germanic and Romance. They also occur in Greek where according to Maria Papastathi (pc), they are rather common in the spoken language. In Romance they are typically derived from the Latin unus.
Overlap and complementarity in reference impersonals

The defining characteristic of man-imps and 3pl-imps is the non-referential nature of their subjects. It should therefore come as no surprise that referential range is the

2. The typology of man-imps and 3pl-imps

The defining characteristic of man-imps and 3pl-imps is the non-referential nature of their subjects. It should therefore come as no surprise that referential range is the
property which constitutes the basis for a typology of these constructions. In terms of the referential range of their subjects, MAN-IMPS and 3PL-IMPS may be grouped into three types, (quasi)-generic, episodic and specific (cf. e.g. Alonso-Ovalle 2000; Malamud 2004, 2005; Cabredo Hofherr 2004, 2006).

2.1 Generic uses

Generic MAN-IMPS and 3PL-IMPS reflect the readings of these constructions in non-assertive, irrealis contexts where the referent of the subject corresponds to anybody or some unspecified or loosely specified group of individuals. The range of generic contexts in which MAN-IMPS are used are somewhat wider than those in which 3PL-IMPS are found. MAN-IMPS occur in all types of non-assertive contexts, habitual, potential, conditional, future, negative and deontic. Some relevant examples are given below.

(12) Norwegian (Johansson 2007: 188)

Overalt i koralberget ser man skelettstrukturene av de samme korallater som ennå i dag. ‘One can actually see in the ancient formations the structure of coral species still extant today.’

(13) Catalan (Wheeler et al. 1999: 518)

Hom no parla angles aquí. ‘English is not spoken here.’

(14) Icelandic (Egerland 2003: 78)

Ef madur er ekki giftur. ‘If one is not married.’

(15) Swedish (Egerland 2003: 76)

Man måste arbeta till 65. ‘People have to work till the age of 65.’

3PL-IMPS appear to be found mainly in habitual contexts, as in (11) and also the English original of (16) further below. Another point of difference between the generic uses of the two constructions relates to the speaker and addressee. The subject of a 3PL-IMP excludes the speaker and typically also the addressee, reflecting the third person origins of this construction. By contrast, the subject of a MAN-IMP in its generic usage is claimed to always at least potentially include the speaker and addressee. Nonetheless,
there are instances where the speaker and hearer are clearly not included. This is so with respect to *on* which in (16) is used rather than the impersonal *they* of the English original in the French translation of Harry Potter.

(16) French

\[ \text{Là où tu vas, on met la tête des nouveaux dans} \]

there where you go \text{MAN put:3SG the head of newcomers in} \text{le trou des toilettes, dit-il à Harry.} \]

the hole of toilets said-he to Harry

“They stuff people’s heads down the toilet the first day at Stonewall,”

he told Harry. \hfill (HP p. 40)

Interestingly the generic readings of both \text{MAN-IMPS} and \text{3PL-IMPS} are closely tied to the presence in the construction of some type of adverbial restrictor typically a locative such as \text{here} in (13) or \text{at Stonewall} in (16) or a temporal one such as \text{in the evening} in (3). In the absence of such a restrictor only a definite reading of the \text{3PL} is possible (see below). The same holds for \text{MAN} unless the utterance contains other indicators of genericity such as modals, conditionals etc.

### 2.2 Episodic uses

In their episodic uses, also referred to as arbitrary or existential, \text{MAN-IMPS} and \text{3PL-IMPS} describe actual events which are anchored in time, typically in the perfective past as in (17) and (18), less frequently in the ongoing present as in (19).

(17) Dutch \hfill (Weerman 2006:7)

\[ \text{Men heeft mij gevraagd iets te zeggen.} \]

\text{MAN have me asked something to say}

‘They’ve asked me to say something.’

(18) Portuguese \hfill (Cavadas Afonso 2003:49)

\[ \text{Nem sequer puseram placards pela cidade desabafa} \]

\text{not even put: past:3PL posters all.over city comments}

\text{este comerciante.} \text{this merchant}

‘They didn’t even put posters around the city comments the merchant.’

(19) Spanish \hfill (Alonso-Ovalle 2000:3)

\[ \text{Preguntan por ti.} \]

\text{ask:3PL for you}

‘Somebody is asking for you.’
The anchoring in time narrows down the range of potential referents of their subjects from anyone or anyone within a given location to some set of individuals situated in space and time.

Depending on whether further information is provided with respect to the identity of the individuals in question, the episodic uses of MAN-IMPS and 3PL-IMPS can be divided into vague, as in (17) and (18) and corporate, as in (20) and (21).

\[(20)\] French
\[\text{On va encore augmenter les impôts.}\]
\[\text{man go:3sg again raise:inf the taxes}\]
\[\text{‘Taxes will be raised again.’}\]

\[(21)\] Hungarian
\[\text{Megerősítették a hírt a rádió-ban.}\]
\[\text{confirm-past-3pl the news the radio in}\]
\[\text{‘They confirmed the news on the radio.’}\]

The vague 3PL-IMPS carry no additional information about the identity of the referents of the subject, though they may be deducible from the wider discourse setting. The so-called corporate, on the other hand, identify the group to which the referents of the subject belong but not the identity of the individuals in question. The identification of the group is achieved via the lexical effects stemming from the predicate argument combination in the construction. Thus, for example, raising taxes, as in (20) immediately brings to mind the government. By contrast raising the blinds would not, under normal circumstances, invoke any group of referents. The identification of the group to which the referents of the subject belong in corporate MAN-IMPS and 3PL-IMPS unites them with the generic in which the group is identified via a locative expression. Due to this partial identification of the referent, these two types of 3PL-IMPS are sometimes referred to as semi-impersonal as opposed to the full impersonal vague and also specific (see below) types. Another property of corporate MAN-IMPS and 3PL-IMPS, however, unites them with the vague, namely, the referent of the subject may be a group or an individual acting on behalf of the group. Thus in principle the clause is open to a paraphrase with someone. This does not hold for generic MAN-IMPS and 3PL-IMPS.

2.3 Specific

In the last type of MAN-IMPS and 3PL-IMPS, the specific, as one would expect, the potential referents of the subject are somewhat narrower. Actually the subject of specific MAN-IMPS is not only specific but definite as it is generally the first person, either first person plural, as in (22) or first person singular as in (23).
Overlap and complementarity in reference impersonals

(22) French

_French_ (Creissels 2009: 6)

_Avec Jean on _ira _au théâtre ce soir._

With Jean _MAN_ will go to theatre this night

‘Jean and I will go to the theatre tonight.

(23) Polish

(_Polish_)

(…) _i _człowiek _sobie _mysli _a _wejdę, _zobaczę._

_and _MAN_ self _think:3SG and _enter:FUT:1SG see:FUT:1SG_ 

‘so one thinks to oneself, I’ll go in and see.’

It may even be the second person as in (24).

(24) Danish

(Danish) (Jensen 2009: 8)

_hvad _har _jeg _være, _ja, _der _i _midten _af _tresserne _har_ 

what _have I been yes there in the- middle of the sixties have

_man _være _en _fem _seks _år _ikke._

_MAN_ been _a five six years not

‘How old was I yeah in the middle of the sixties, you were about five or six years, right?’

And also the third.

(25) German

(German) (Cabredo Hofherr 2004: 9)

_An der Rezeption hat _man _mir _gesagt, _ich _könne _den _Aufzug _benutzen._

_at the reception had _MAN_ me _said _I _could _the lift _use

‘At the reception someone told me I could use the lift.’

**3PL-IMPS** do not exhibit any variation with respect to the person feature of the subject. They do, however, exhibit variation in regard to number. What are termed specific **3PL-IMPS** denote a single individual who is specific indefinite, i.e. known to the speaker. Specific **3PL-IMPS** are typically used in situations where the identity of the subject derives not from the verbal but from the situational context, as in (10) and (19) cited earlier. Specific **3PL-IMPS** thus appear to be used in contexts involving some form of contact of the speaker with the referent of the subject, for example, visual (at the door, on TV), as in (10) and (19), audio (on the phone, radio), written (via an e-mail or letter) etc. It is because of this contact that the referent of the subject is known to the speaker.

3. **The salient properties of MAN-IMPS and 3PL-IMPS**

In addition to the considerable overlap in the referential range of their subjects, discussed above, **MAN-IMPS** and **3PL-IMPS** exhibit several other properties which need to be briefly mentioned. We will begin with the relevant morpho-syntactic properties, then turn to the pragmatic and finally the stylistic.
3.1 Morpho-syntactic properties

The group or individual denoted by 3PL-IMPS always corresponds to that of the clausal subject. The man of MAN-IMPS, however, may in some languages occupy the position of a non-subject. For instance, we see in (26) that mađur in Icelandic may be used as a direct object in addition to its use as a subject.

\[(26)\] Icelandic \quad \text{(Egerland 2003: 91)}

\[
\text{Svona tőlur segja manni ad eithvad sé i ólagi.}
\]

such figures tell MAN that something is wrong

'Such figures tell one that something is wrong.'

Whereas morphologically 3PL-IMPS are always plural, man is typically singular. In some languages, however, such as Swedish and French, it may also trigger plural agreement in certain contexts, as is the case in (27) in which the predicative adjective may be singular or plural.

\[(27)\] Swedish \quad \text{(Egerland 2003: 78)}

\[
\text{om man inte är gift/ gifta måste man ha}
\]

if MAN is not married;SG/ married:PL must man have

\[
\text{skilda rum på detta hotell.}
\]

separate rooms in this hotel

'If one is not married, one must have separate rooms in this hotel.'

An important morpho-syntactic reflex of the generic vs. episodic distinction, which both MAN-IMPS and 3PL-IMPS share, noted originally by Cinque (1988), concerns the range of predicates that they may occur with. The generic readings are open to all types of predicates, transitive, unergative, unaccusative, copulative, passive etc. However, the episodic readings are restricted only to transitive and unergative predicates, i.e. those featuring non-derived subjects. Note the contrast between the generic interpretation of the second man-clause in (28), where man is the subject of a passive, and the impossibility of an arbitrary, as opposed to a definite reading of the passive clause in (29).

\[(28)\] German \quad \text{(Malamud 2004: 10)}

\[
\text{Wenn man mit der Mafia verhandelt, wird man}
\]

if MAN with the mafia deals will man

\[
\text{normalerweise ermordet.}
\]

usually killed

'If one deals with the Mafia, one will usually get killed.'

\[(29)\] Spanish \quad \text{(Alonso-Ovalle 2001: 3)}

\[
\text{Estan siendo golpeados.}
\]

be:3PL being beaten

'They (*someone is) are being beaten.'
3.2 Discourse-pragmatic properties

Neither man-imps nor 3pl-imps have the ability to introduce new discourse referents (see e.g. Prince 2002; Malamud 2004, 2005). Recall that in episodic contexts the subject of man-imps and 3pl-imps may receive a reading corresponding to someone. As is well known an indefinite pronoun may introduce a new discourse referent which may then be subsequently picked up by a regular singular anaphoric pronoun. However, this is claimed to be not possible with either man-imps or 3pl-imps. Compare (30) with (31) and (32).3

(30)  I noticed someone, on the stairs. He, was going down.

(31)  German

Gestern, hat man ein Haus abgebrannt. Er wurde verhaftet.
Yesterday has man a house burned. He was arrested.
‘Yesterday someone, set a house on fire. He, was arrested.’

(32)  Spanish

Tocan a la puerta Quiere entrar.
knock:3pl at the door want:3sg come in:inf
‘Someone’s, knocking at the door. He, wants to come in.

It is less clear whether non-generic man-imps and 3pl-imps, unlike indefinite pronouns, can participate in regular cross-sentential anaphora. As shown in (33), indefinite pronouns cannot; the two instances of someone can only be interpreted as non-co-referential.

(33)  Someone, came into the room. Someone, said that it was stuffy.

Cross-sentential co-reference with non-generic man seems to be also impossible. Whether the same applies to 3pl-imps may depend on the type of 3pl-imp and the language in question. Carbedo Hofherr (2006) argues that Spanish allows cross-sentential co-reference of non-generic 3pl-imps but only under the corporate reading not the specific. Alonso-Ovalle (2000), however, maintains that the co-reference in (34) where the 3pl-imp receives a specific interpretation in both the first and second clause is just as good as in the case of the corporate (35).

(34)  Spanish

Tocan a la puerta Quieren entrar.
knock:3pl at the door want:3pl come in:inf
‘They’re, knocking at the door. They, want to come in.’

3. According to Malamud (2005), in Russian the referent of a 3pl-imp can be taken up by an overt plural pronoun, namely oni. However, her corpus investigations reveal that this happens extremely rarely.
(35) Spanish

Ayer arreglaron mi computadora. Llegaron a las 10 de la manana y se fueron a las 6 de la tarde.

'yesterday repair:3pl my computer. came:3pl at the 10 in the morning and refl left:3pl at the 6 in the afternoon'

'yesterday they repaired my computer. They came at 10 in the morning and (they,) left at 6 in the afternoon.'

In Russian (Perlmutter 2001:11; Malamud 2005), on the other hand, co-reference seems to be possible for 3pl-imps in all types of contexts, generic, corporate and specific. For instance in (36) the individual who called could also be the same individual who had changed the schedule.

(36) Russian

Pozvonili čtoby skazat’ čto izmenili raspisanie
called:3pl in-order say:inf that change:3pl shedule

'someone called in order to say that the schedule had been changed [lit that they had changed the schedule.]

The restricted referential potential of man-imps and 3pl-imps makes these constructions ideal vehicles for expressing event-centred as opposed to participant-centred situations. While topicality based impersonals such as the passive express patient-oriented as opposed to agent-oriented situations, man-imps and 3pl-ims tend to be used for bare happenings (Sansò 2006). Thus not only the referent of the subject but also of the object, if there is one, tends to be neither topical nor focal, i.e. likely to be taken up in subsequent discourse. Nonetheless, as eluded to already in the introduction, both constructions, particularly 3pl-ims may be put to use in ways similar to passives in some languages (see §5).

4. The genetic and areal distribution of the two constructions

Before taking a close look at the distribution of man-ims and 3pl-ims among the languages of Europe, let us first have a global impression.

4.1 The global perspective

Neither man-ims nor 3pl-ims have been the subject of a detailed global cross-linguistic investigation. This is hardly surprising as most grammars contain no or very little information on these types of R-impersonals. In their excellent overview of man-ims in Europe, Giacalone Ramat & Sansò (2007) take as an
approximation of the global distribution of man-imps the distribution of generic-noun based indefinite pronouns in Haspelmath’s (1997) sample of indefinite pronouns. The sample identifies three areas in which man-based indefinite pronouns are attested with some frequency, Europe, Africa and South-East Asia. My own perusal of the world’s languages suggest that in addition to the above three areas man-based R-impersonals are also likely to be found among the language of Oceania (e.g. racu in Araki, geba in Buru, armaj in Mokilese, tino in Tuvaluan), though perhaps not as frequently as in Europe, Africa or South-East Asia. However, to the best of my knowledge, none of the potential man-based impersonals outside of Europe in either Haspelmath’s sample nor mine has been investigated in any detail. Consequently, they do not yet constitute a potential source of comparison for the European data.

3pl-imps are much more widely attested than man-imps. Outside Europe, they are found in Eurasia (e.g. Kashmiri, Persian, Tamil, Singhala) though only rarely in South-East Asia. They are common in African languages: Afro-Asiatic (e.g. Arabic, Coptic, Hebrew, Godie, Konso, Mupun), Niger-Congo (e.g. Babungo, Joola-Banjal, Nkore- Kiga, Fonge, Koromfe, Zande), Nilo-Saharan (e.g. Arusa, Baka, Kunama, Mundani, Ngiti, Tirmanga). They occur in Oceania (e.g. Adzera, Ambryn, Lewo, Nélémwa, Tawala, Paamese, Uma), among the languages of New Guinea (e.g. Amele, Kobon, Kuot, Olo), Australia (e.g. Marunguku, Maung) and the Americas (e.g. Copala Trique, Creek, Hopi, Lakhota, Mountain Maidu, Pipil, Southern Piaute, Tetelcingo Nahuatl). As in the case of potential man-imps, none of these constructions has been yet investigated with respect to its actual usage.

The areas where both man-imps and 3pl-imps are found are Europe, Africa and to a lesser extent Oceania. However, to the best of my knowledge, only in Europe do we have languages which evince both constructions.

4.2 Europe

In considering the distribution of man-imps and 3pl-imps one can focus on the mere presence in a language of the constructions, the range of referential uses of the constructions that each language allows (as characterized in §2) or the actual frequency of use of the construction in comparison to the other constructions within the repertoire of R-impersonals available in a language. We will first concentrate on

4. The languages in Africa and South-East Asia that may have man-based impersonals that I have come across are: in Africa: Kisi, Lele, Ndonga, Ngizim and Madi; in South-East Asia: Kashmiri, Marathi, Punjabi (all Indo-European) Mandarin (see Yi & Siewierska this volume) and other varieties of Chinese, Japanese, Korean and Thai-Nuea.
the presence and range of referential possibilities that the constructions display and leave the matter of the status of the construction in relation to other R-impersonals to §6.

4.2.1 MAN IMPS

Beginning with man-imps, Giacalone Ramat & Sansò (2007) posit man-imps as a hitherto unnoticed feature of Standard Average European (SAE). They argue that the current distribution of man-imps in Europe falls into three areas: a core, a periphery and a non-Man area. In the core area man-imps are used generically and episodically and typically also specifically. In the periphery they are used only generically. And in the non-Man area they do not occur, though the word for ‘people’ or ‘human’ may be used in a species specific sense. Giacalone Ramat & Sansò document that the core-Man area embraces the so-called Charlemagne area (cf. van der Auwera 1998:823ff, Haspelmath 2001:1493) plus Mainland Scandinavia and the South West Romance area. This core area of man-imps thus includes: French, German, Dutch, Frisian, Danish, Swedish and Norwegian and potentially Catalan. Their investigation of historical material reveals that this core area was once more extensive embracing also Spanish, Occitan, Italian and the North Italian dialects as well as English, all of which (but for some Italian dialects) nowadays do not have any man-imps at all. Boarding on the current and former core areas is the periphery where man-imps are used only generically. The periphery is comprised of three areas. These embrace in the north-west Faroese and Icelandic, in the west the Celtic languages, Irish, Welsh and Breton, in the east the West and South Slavic languages as well as Albanian and in the south Maltese. No man-imps are attested in the Indo-European languages further to the east, south or south-west i.e. in East Slavic, Lithuanian, Rumanian, Greek, Sardinian5, Galician, Portuguese nor in Basque. My investigations suggest that man-imps are also not a feature of Uralic, Turkic or the languages of the Caucuses, though they do occur in Hungarian and also sporadically in the languages of the Caucuses as evidenced by (1) from Abkhaz.

The above distribution is shown in the map in Figure 1 which is an enriched version of the map in Giacalone Ramat & Sansò (2007).

5. According to D’Alessandro & Alexiadou (2006) Sardinian does have two man-forms deriving from the Latin homo, namely omine which is used as a generic NP and omo which can also be used in episodic contexts. The latter, however, is claimed to be disappearing.
4.2.2 3PL-IMPS

Turning to 3PL-IMPS, it is not always easy to determine which combination of the four types of 3PL-IMPS distinguished in §2 a language has. On the basis of the data that I do have, I have allocated the languages in my sample into four groups:

- **Group I** languages manifesting generic, corporate, vague and specific uses of 3PL-IMPS
- **Group II** languages manifesting generic, corporate and vague uses
- **Group III** languages manifesting less than the above three uses
- **Group IV** languages with no 3PL-IMPS or only lexically restricted uses

The distribution of the above four groups of languages is shown in the map in Figure 2. The squares denote group I languages, the circles the languages belonging to group 2, the triangles languages of group 3 and the hexagonals languages of group 4.

We see that the languages of group I which manifest all four uses of 3PL-IMPS are within a north-east to south-west band. They are: the East Slavic languages (Perlmutter 2001; Malamud 2004), Polish, Serbo-Croatian, Hungarian (Tóth 2000), Greek (Papastathi pc), Italian (Cinque 1988), Sardinian (Jones 1993), Portuguese (Cavadas Afonso 2003), Galician (Gonzalez-Diaz pc), Catalan (Wheeler et al. 1999) and Spanish (Alonso-Ovalle 2000; Cabredo Hofherr 2006). Within group II we definitely have English, French, Dutch and Turkish. Subsequent research may prove that this group is...
in fact wider and also includes some of the languages of group III. The latter is a big and very heterogeneous group embracing: languages for which I have no information apart from the fact that they do utilize the 3pl impersonally (Albanian, Breton, Irish, Faroese, Maltese, Veps and Livonian); languages for which at least the vague use is possible such as Latvian, the Permic languages (Komi, Udmurt) and the Volgaic languages (Mari, Erzyya Mordvin), the North-West Caucasian languages (Abkhaz, Abaza, Kabardian), Kartvelian (Georgian) and Basque; and languages which manifest the vague and corporate uses (Danish, Swedish) or vague and generic (Icelandic German, Norwegian). The last group which (virtually) lacks 3pl-imps, group IV embraces Lithuanian and the Baltic Finnic languages, i.e. Estonian and Finnish. In Standard Finnish and Estonian, according to Holvoet (2001: 381), 3pl-imps are occasionally found but only with one or two speech act verbs, particularly in reporting rumours, as in (37).

(37) Standard Finnish (Holvoet 2001: 381)

Siellä kuuluvat tienaan hyvin.

‘It is said that one earns well here.’

---

6. The exceptional nature of 3pl imps in Standard Finnish is confirmed by the fact that not even a single instance of the impersonal use of the 3pl was found in the Finnish translation of Harry Potter. The main impersonalizing strategy in the language in addition to the impersonal passive is the so-called zero construction (Helasvuo & Vilkuna 2008), an instance of which is provided further below in (48), which involves the use of 3sg person marking on the verb and no overt subject.
In Finnish dialects, 3PL-IMPS are used somewhat more widely, as documented in Forsberg et al. (2009). In addition to occurrences with speech act verbs, we find also corporate uses, as in (38).

(38) Finnish, Eastern Savo, Kontiolahti (Forsberg et al. 2009)

siinä oli niinikkiään sillon ‘ensi’ alusta sita
there be:PAST:3SG also then first beginning-ELATIVE it-PART
‘It [the school] was there [in the house] at first, before they built any school buildings.’

However, since all the other languages have been allocated to 3PL-IMP groups on the basis of the use of the 3PL in their standard varieties, I have assigned Finnish to group IV.

4.2.3 Comparing the distribution of the two

If we compare the distribution of 3PL-IMPS in Figure 2 with the distribution of MAN-IMPS in Figure 1, we see that none of the languages in 3PL-IMP group I overlaps with those of the core-Man area. On the other hand, none of the languages of the core-Man area is to be found in 3PL-IMP group IV i.e. among the languages which lack 3PL-IMPS. In short, all the MAN-IMP languages also display 3PL-IMPS be it not the full range.

Of the languages within the core-Man area the widest range of uses of 3PL-IMPS appears to occur in French and Dutch. According to Cabredo Hofherr (2006) in French the 3PL ils can receive generic (39a) corporate (39b) and also vague readings (39c).

(39) French (Creissels 2009: 15, 21; Cabredo Hofherr 2006)

a. En Italie ils savent préparer les pâtes.
   ‘In Italy they know how to prepare pasta.’

b. Ils vont encore augmenter les impôts.
   ‘They will raise taxes again.’

c. Ils ont trouvé une moto dans la cour.
   ‘They have found a motorbike in the courtyard.’

The corporate and vague uses are also mentioned in Judge & Healey (1985:70) and the corporate and generic in Creissels (2009:15, 21). It is worth pointing out though that on seems to be very much preferred over ils in generic uses. Thus the French translation of the generic (16) in Harry Potter features on.
The 3pl in Dutch when used impersonally occurs in its reduced form, namely ze (the strong form being zij). The example in (40) is from the Harry Potter corpus.7

(40) Bestonden er echt winkels waar ze boeken met toverspreuken
exist there really shops where 3pl books with spells
en bezemstelen verkochten?
and broomsticks sold

‘Were there really shops that sold spell books and broomsticks?’

The impersonal use of ze in Dutch is considered to be an informal substitute for men and an alternative to je, the Dutch equivalent of generic you. Interestingly, in the study by van Hell et al. (2005:253) ze, while considerably less frequent than je, was found to be relatively common in the written expository texts of younger children. In the Harry Potter corpus, there are instances of generic, corporate and vague uses of ze. The impersonal uses of ze in this corpus (N = 20) are considerably more common than the impersonal uses of ils in French (N = 9). Significantly, as (40) documents, not all of the occurrences of ze in the Dutch translation correspond to the use of impersonal they in the English original.

The Mainland Scandinavian languages appear to allow a limited vague use of the 3pl with verbs such as ‘say’ and ‘call’ and also the corporate use, as shown in (41) from Danish8 and (42) from Swedish.9

(41) Danish (Kim Ebensgaard Jensen: p/c)

a. De kalder ham Olav.
   3PL call he:ACC Olaf
   ‘They call him Olaf.’

b. Nu har de igen havet billetpriserne.
   now have 3PL again raised ticket-prices.the
   ‘They’ve raised the ticket prices.’

7. Occasionally rather than ze the object form hun is used, as in (i).

   (i) [‘Oh, well – I was at Hogwarts meself but I – er – got expelled, ter
tell yeh the truth. In me third year]
   hun hebben me toverstaf in tweeën gebroken.
   they have my wand in two broken
   ‘They snapped me wand in half.’

8. Jensen (2009: Footnote 2) observes that in the variety of Danish spoken on the isle of Funen, de rather than man is the primary form used for generic reference.

9. Askedal (1994:234) mentions that in Nynorsk, as opposed to Bokmål, the 3pl dei is more widely used for impersonal reference.
Overlap and complementarity in reference impersonals

(42) **Swedish**

(Holmes & Hinchliffe 1993:431)

a. *De säger han är rik.*
   *3pl say he is rich*
   ‘They say he’s rich.’

b. *De kommer att höja skatterna, det är jag säker på.*
   *3pl going to increase taxes:the, that am I sure of*
   ‘They are going to increase taxes, that am I sure about.’

According to Askedal (1994:234), in Norwegian the use of the 3pl-imps is more common in Nynorsk than in Bokmal. And indeed of the three instances cited in Faarlund et al. (1997:329) two are from Nynorsk. Significantly of the three illustrations given, one involves the verb ‘call’ and another ‘say’ just as in the above examples from Swedish and Danish. The third example presented in (43) is essentially generic.

(43) **Norwegian, Nynorsk**

(Faarlund 1997:329)

*I gamle dagar trudde dei at jorda var flat.*

in old days believed 3pl that earth:the was flat

‘In the old days, they believed that the world was flat.’

In contrast to French, Dutch and the Scandinavian languages, in German the impersonal use of the 3pl is not mentioned in grammars of the standard language. However, my investigations suggest that a phonetically reduced version of *Sie*, namely */za/ may be used impersonally in colloquial speech. Such usage appears to be more likely when *Sie* occurs in the so-called middle field rather than in sentence initial position, as in (44), taken from the Harry Potter corpus. 10

(44) *“In Stonewall stecken sie den Neuen am ersten Tag den Kopf ins Klo” eröffnete er Harry.*

*p. 43*

‘They stuff people’s heads down the toilet the first day at Stonewall,”

he told Harry.

In fact all of the examples of impersonal *Sie* in the Harry Patter corpus are either in the middle field or following a conjunction, as in (45).

---

10. I am grateful to Volker Gast for the observation that impersonal uses of *Sie* are much better in the middle field than in the fore-field, which is not very tolerant of clitics in German.
Gleis neun, Gleis zehn. Dein Gleis solte irgendwo
platform nine platform ten your platform should somewhere
dazwischen liegen aber sie haben es wohl noch nicht gebaut, oder?
in-between lie but 3pl have it possibly still not built or

“Platform nine – platform ten. Your platform should be somewhere in the middle, but they don’t seem to have built it yet, do they?”

My German informants agree on the acceptability of only the generic uses with an accompanying locative, as in (44) and corporate uses of sie to the exclusion of the other uses.

As for the languages in the peripheral man areas, some display very restricted use of 3pl-imps, others evince all four uses. Icelandic belongs to the first type. Of the various examples that I offered to my informants, only the generic and corporate (46a) were accepted as possible, though in the case of the corporate the expletive subject construction in (46b) was strongly preferred.

(46) Icelandic
  a. þeir eru búñir að loka veginum einu sinni enn.  
     3pl.m are:3pl finished to close road-the.dat one time again
     ‘They’ve closed the road once again.’
  b. það er búð að loka veginum einu sinni enn.
     it. is:3:sg finished to close road-the.dat one time again
     ‘They’ve closed the road once again.’

Polish and Serbo-Croatian belong to the second type. Examples of the generic, corporate, vague and specific uses of 3pl-imps from Polish are given in (47).

(47) Polish
  a. W Francji jedzą ślimaki.
     in France eat:3pl snails.
     ‘In France, they eat snails.’
  b. (…) myślę, że przez środek miasta nie bedą ciągnęli
     think:1sg that through centre town not be:3pl pull:3pl
     obwodnicy.
     circular road
     ‘I don’t think they’ll build a circular road through the centre of town.’
  c. Muszę iść, czekają na mnie.
     must:1sg go wait:3pl on me
     ‘I have to go. They’re waiting for me.’
Whereas all of the languages in the core man area exhibit some range of uses of 3pl-imp, there are 3pl-imp languages in group I which have no man-constructions. These are Russian, Italian, Portuguese, Spanish, Galician and Greek, all of which are geographically further removed from the northern-European Man-core. These languages also differ typologically from those in the core Man area, namely with respect to the presence and nature of pronominal subjects. Let us take a closer look at this property.

5. The realization of pronominal subjects: Pro-drop

In the generative literature attempts have been made to tie the formal realization of non-referential subjects to the obligatoryness vs. non-obligatoriness of pronominal subjects, i.e. to the so-called pro-drop properties of languages. The most influential typology has been that of Holmberg (2005, 2010). Taking as a point of departure the assumption that referential subjects should be clearly distinguished from non-referential subjects, Holmberg argues that non-pro-drop languages have special forms for non-referential subjects such as man (and also one) which are necessarily overt. Partially pro-drop languages, i.e. those which manifest pro-drop in the first and second person but not the third, in turn, evince covert non-referential subjects. And full pro-drop languages express non-referential subjects by verbal morphology or some other non-pronominal morphology.

In Europe the distribution of pro-drop strongly overlaps with that of man-imp. The pro-drop properties of European languages are depicted in the map in Figure 3. The squares indicate the full-pro-drop languages, the circles the partially pro-drop ones and the triangles the non-pro-drop ones.

We see that all the non-pro-drop languages in Figure 3 belong to the core-Man area with the exception of Icelandic and Faroese, which nonetheless also have man-imp be it more restricted ones. And none of the pro-drop languages belong to the Man-core. Both facts are captured by Holmberg’s three-way typology. The typology also correctly reflects the absence of any overt third person pronoun in Standard Finnish (48), Hebrew (49) and Russian (50) all of which qualify as partially pro-drop languages.11

11. The person marking on the verb which occurs in Hebrew in the past and future but not in the present and in Russian in the past as opposed to the present is treated under this analysis.
Figure 3. Distribution of pro-drop among the languages of Europe

(48) Standard Finnish\(^{12}\) (Holmberg 2005: 544 (ex.30))

Tässä istuu mukavasti.
here sit:PRES:3SG comfortably
‘Here one sits comfortably.’

(49) Hebrew (Berman 1980: 762)
šot-im hamon mic ba arec.
drink-PL lots juice in country
‘They drink lots of juice in Israel.’

(50) Russian
Da tut vsakyju erundu peredajut v novostjach,
yes here give:PRES:3 in news
promjamlil mister Durslej.
munnled mr Dursley

“They funny stuff on the news,” Mr. Dursley mumbled.’ (HP p. 23)

not as person marking but as default marking. The same analysis is extended to the Standard Finnish (48) in which the 3SG agreement is considered to be default agreement.

12. Whereas Standard Finnish qualifies as partially pro-drop in Holmberg’s terminology, Finnish dialects are essentially non-pro-drop. They do not conform to his typology as many of them use overt 3PL forms in 3PL-IMP constructions, as illustrated in (ii) from Forsberg et al. (2009).

(ii) ne sannoot et se nousoo puu
3PL say-3PL that it rises up tree
‘They say that it [snake] rises up a tree.’
Further, it is consistent with the fact that virtually all of the languages outside of the core-man area are not only fully pro-drop but also utilize special verbal morphology to express non-referential subjects, for example, reflexive marking or impersonal participle marking, as previously illustrated in (7) and (8) in the introduction to the paper. Reflexive impersonals are widely used in West and South Slavic languages and Romance as well as in Lithuanian. In addition Polish, Ukrainian and the Baltic Finnic languages have special impersonal participle forms used in constructions with non-referential human subjects.

The above notwithstanding, the typology does not lead us to expect the existence of languages such as those in the peripheral Man-area in Figure 1, all of which, do utilize special non-referential forms but with the exception of Icelandic and Faroese, are also pro-drop. More importantly though, it has nothing to say about the distribution of 3PL-IMPS. In fact it is incompatible with the existence of 3PL-IMPS. Recall that 3PL-IMPS may be rendered by weak free forms such as the English they or Dutch ze or by bound affixes. The former occur in non-pro-drop languages, the latter in pro-drop languages. If non-pro-drop languages are argued to require special forms for non-referential subjects, they should not also employ their 3PL pronouns for this purpose. Yet they do so, as is the case in English, Dutch, to some extent in German, the Mainland Scandinavian languages and French. Nor should pro-drop languages exhibit 3PL-IMPS in addition to the special non-referential subject expressing constructions that they are predicted to require. Yet again they do. At least the West and South Slavic languages, the Romance languages and the Permic and Volgaic languages display 3PL-IMPS in addition to their reflexive impersonals.

Although contrary to what Holmberg and others have suggested, there is no one-to-one relation between the pro-drop properties of a language and the type of non-referential subjects that they have, there is a correlation between pro-drop properties and the range of uses the different types of non-referential subjects display. This follows from the overlap between absence of pro-drop and man-IMPS documented in Figures 1 and 3. In Europe, non-pro-drop languages make more use of man-IMPS than of 3PL-IMPS, and pro-drop languages make greater use of 3PL-IMPS than of man IMPS. More concretely, in non-pro-drop languages man-IMPS are used in generic and episodic contexts as well as specifically, while in pro-drop languages they are typically confined to generic contexts. 3PL-IMPS, on the other hand, may evince generic, corporate, vague and often also specific uses in pro-drop languages but display only a subset of these in non-pro-drop languages. These differences may be seen to follow from the degree of grammaticalization of the respective impersonal constructions in pro-drop and non-pro-drop languages.

The referential range of man-IMPS discussed in §2 is viewed by Egerland (2003), D’Alessandro & Alexiadou (2006) and Giacalone Ramat & Sansò (2007) as defining the grammaticalization cline shown in (51).
According to the cline in (51), a word originally used to denote the whole species of humans, as in (a), becomes to be used for any loosely characterized set of humans, as in (b), and then acquires the possibility of denoting a single individual, as in (c) and/or being used even definitely as in (d). This semantic bleaching or schematization is accompanied by changes on the morpho-syntactic and phonological planes. The form for ‘man’ develops pronominal properties and is restricted to the function of subject. It also tends to become phonologically reduced, as in the case of Latin homo giving rise to French on. In addition, the range of co-reference possibilities that man displays decreases as also does the range of thematic properties that it may evince.

A grammaticalization cline comparable to that for man-imsps in (51) can also be posited for 3pl-imsps. Of the readings found with the different types of 3pl-imsps, the semi-impersonal ones characteristic of generic and corporate 3pl-imsps are arguably the closest to that of the normal definite plural reading of a third person plural anaphoric pronoun. Recall that in such 3pl-imsps the group is identified, only the individuals within it are not. And significantly the impersonal reading is induced by elements within the construction itself. By contrast, the existential reading of 3pl-imsps, where the referent of the subject is an individual, especially a specific one is semantically quite removed from the necessarily plural interpretation associated with a third person plural person form. And in between the two we have the non-specific plural readings found with vague 3pl-imsps. These distinctions may be captured in the cline in (52).

There are several facts which can be cited in support of the cline in (52). The first is that the existence of the singular specific reading is predicated on the existence of the other three readings. Recall from §3.2 that all the languages which have the specific reading, i.e. those in group I also have the generic, corporate and vague readings. Whether the possibility of the vague reading also presupposes the corporate and/or generic is less clear. This does appear to be so in the European languages of group II and III for which data are currently available, French, Dutch, German, Turkish and the Scandinavian languages. The second piece of evidence in favor of (52) is the narrowing down of the thematic properties that may be borne by the subjects of the 3pl-imsps as we progress down the cline. Recall from §2.3 that the episodic readings, as opposed to the generic and corporate, are available only for non-derived subjects. Such narrowing

13. Giacalone Ramat & Sansò (2007) consider step (d) in (51) to be independent of step (c). However, whether this is indeed so is not clear.
down of the range of subjects which may occur in a construction is characteristic of an increase in degree of grammaticalization (see e.g. Traugott & Hopper 1991; Lehmann 1995). The following two properties which correlate with the cline in (52) are particularly important in the context of this discussion of pro-drop as they relate to the formal realization of the third plural morpheme. In all of the examples of 3PL-IMPS from fully pro-drop or partially pro-drop languages given so far the form of the 3PL is a bound one. In some languages, for instance Italian and Spanish, this is necessarily the case. If in addition to the bound marking on the verb a free form of the 3PL is used, the only interpretation possible is a definite one. However, in other languages, for example Russian, according to Franks (1995) and Livitz (2006:18), the bound 3PL-marking on the verb can co-occur with an accompanying free 3PL form but only in generic and corporate 3PL-IMPS. In the case of vague and specific 3PL-IMPS only the bound form alone can occur. The higher grammaticalization on the referential dimension of 3PL-IMPS is thus reflected in the reduced possibilities of morpho-phonological realization of the 3PL. Yet another morpho-phonological reflection of the grammaticalization cline is that all the languages that exhibit the right-most stage on the cline in (52) are pro-drop. And significantly none of the non-pro-drop languages in which the 3PL is realized by a weak form rather than an affixal one display the singular specific reading. This dichotomy with respect to the possibility of a specific reading is again fully in line with what the grammaticalization approach would lead us to expect. A weak pronominal is more closely tied to the person/number features of the anaphoric pronoun from which it originates than the affixal form. As such, it is less likely to be used for evidently singular as opposed to plural or potentially plural referents. And given that the weak forms are found in non-pro-drop languages and the affixal ones in pro-drop languages, non-pro-drop languages may be expected to manifest a more restricted range of 3PL-IMPS than pro-drop ones. And this is indeed so.

The two grammaticalization clines in (51) and (52) together with the difference in the formal realization of 3PL subjects in pro-drop and non-pro-drop languages go a long way in explaining the distribution of different types of MAN-IMPS and 3PL-IMPS among the languages of Europe. MAN-IMPS are preferred over 3PL-IMPS in non-pro-drop languages since the former provide an unambiguous means of identifying non-referential subjects. Give that MAN-IMPS are highly grammaticalized, there is less scope for the development of 3PL-IMPS. If these do arise, their grammaticalization is restricted to semi-impersonal contexts in which they are least likely to be confused with referential 3PL forms. 3PL-IMPS, on the other hand, are favoured over MAN-IMPS in pro-drop languages by virtue of the overall preference for bound pronominal subjects that these languages display. Note that although a subject realized by MAN has the potential of becoming bound to the verb, if frequent enough, it is not very likely to do so in a language that already displays third person marking on the verb. The affixal expression of the third plural in turn facilitates its use with a wide range of referents
even specific individual ones leading to the high grammaticalization of the construction. In fact the grammaticalization of the 3pl can proceed even further.

Although we have no clear instances in Europe, there are languages in which the 3pl has developed into a passive marker. Such is the case, for example, in the Mayan language Kaqchikel, as a comparison of (53a) and (53b) suggests.

(53) Kaqchikel (Broadwell & Duncan 2002: 4)

a. \textit{Ritzi’} x-ki-chap r-oma’ ri achim
   the:dog COM-PASS-grab 3SG-by the man
   ‘The dog was grabbed by the man.’

b. X- in- ki- k’utuj
   COM- 1SA- 3SPL- ask
   ‘They asked me.’

What are the grounds for considering passives such as those in (53) to be a further potential development of 3pl-imps?\textsuperscript{14} Siewierska (2010) argues that a 3pl-to-passive reanalysis is predicated on the availability of specific readings of a 3pl-imp. If this is indeed so, we might expect the factors which facilitate a 3pl-to-passive reanalysis to be already in evidence in the earlier stages of grammaticalization of 3pl-imps, especially in languages which display specific 3pl-imps. One of the factors that constitute a precondition for the 3pl-to-passive reanalysis documented by Siewierska (2010) is the presence of bound 3pl marking on the verb. We have already discussed this property in relation to our 3pl-imps. Recall that in all the languages which display specific 3pl-imps the 3pl is a verbal affixal form. Another facilitating factor noted by Siewierska, which in fact often goes hand in hand with bound person marking, is word order flexibility, in particular the possibility of locating a nominal object in typical subject-topic position. She argues that such placement of the object makes the 3pl-construction much more like a participant-centred rather than an event-centred one and over time renders the object more prone to reinterpretation as a subject. How thus do our 3pl-imp languages fare in relation to word order flexibility?

All of the group I languages have flexible or highly flexible word order. And none of the group III languages for which we have sufficient data do. Significantly all the group I languages allow preverbal objects particularly when the subject is realized just by bound morphology. Thus all display some of the prerequisites for a 3pl-to-passive reanalysis. Moreover, in two of these languages 3pl-imps appear to be actually used for participant-centred events. The languages in question are Russian and Greek.

\textsuperscript{14} As I argue in Siewierska (2010), the 3pl-to-passive reanalysis results in impersonal passives which only rarely develop into personal passives. The Mayan example in (53) is one of the rarer instances of the latter development.
Particularly in Russian the object of 3PL-IMPS is often preverbal and clearly topical as in (54) taken from the Russian translation of Harry Potter.

(54) Russian

[His blue eyes were light, bright, and sparkling behind half-moon spectacles]

balanced on long nose crooked so much that seemed

as if this nose broke:3PL in at least in two places

'his nose was very long and crooked, as though it had been broken at least twice.' [lit. as if this nose they broke…]

Interestingly in none of the other languages of the parallel corpus is a 3PL-IMP used in this context. And it is not irrelevant that the use of 3PL-IMPS in the Russian translation of Harry Potter is considerably higher than in the translations into any of the other languages. In Greek too 3PL-IMPS are sometimes used for patient-centred events. In Sansò’s (2006) investigation of the Greek translation of Umberto Eco’s *The Name of the Rose*, 3% of the patient-oriented processes in the analysed section of the text were expressed by means of 3PL-IMPS. While such usage is clearly the minority type in Greek, as compared to the generic uses and bare-happenings, the fact that it exists is significant.

In the light of the above, the grammaticalization cline in (52) should be substituted by that in (55).

(55) (a) plural definite > (b) plural-semi-definite > (c) plural non-specific >
(d) singular specific > (e) impersonal passive > (f) passive

It is important to note, that MAN-IMPS, unlike 3PL-IMPS, do not appear to develop into passives. Given that the languages which favour MAN-IMPS lack the range of factors which facilitate the development of 3PL-IMPS to passive, this is not surprising.

In the preceding discussion we have concentrated on the correlation between MAN-IMPS and non-pro-drop on the one hand, and 3PL-IMPS and pro-drop on the other, skimming over the differences in the actual use of both MAN-IMPS and 3PL-IMPS in the two groups of languages. To get a fuller picture of the distribution of the two constructions, we will briefly consider this issue below.

6. The actual use of MAN-IMPS and 3PL-IMPS

In the case of MAN-IMPS, the range of referential uses that they can display in a language is a good indicator of their frequency of occurrence and stylistic associations. In French, German and the Mainland Scandinavian languages MAN-IMPS are clearly the dominant construction for human generic reference, at least in writing. In French
impersonal *on* is widely used in writing. However, according to Jisa & Viguié (2005), its episodic use in writing decreases with age in favour of other constructions with a similar functional load especially the passive. In speech, on the other hand, the generic uses of *on* are being taken over by the second person singular *tu*. Fonesca-Greber & Waugh (2003) document that in their spoken corpus 68% of impersonal references were realised by *tu* as opposed to only 31% by *on*. They also show that of the impersonal uses of *on*, the vast majority (75%) correspond to what we have been calling generic as opposed to episodic. Given the dominance of non-personal *on* in writing, it seems clear that the construction is acquiring a formal flavour. In German *man-imps* are widely used in different modes and registers and do not carry strong stylistic or formal associations. According to Zifonun (2001), their dominant use is also unequivocally generic; the episodic account only for 18% of the instances of *man* in her corpus. In Swedish *man* is highly frequent as well. Ragnarsdóttir & Strömqvist (2005:148) state that *man* is the 12th most frequent single word form in spoken Swedish and the 33rd most frequent form in written Swedish. It is especially common in non-fiction and its use in spoken expository texts increases with age (Ragnarsdóttir & Strömqvist 2005:154). In Danish too *man* “is undoubtedly the most frequent and the most generally usable pronoun for generic reference” in the language (Jensen 2009:4). In Sansò’s (2006) investigation of the Danish translation of sections of the text of Umberto Eco’s *The Name of the Rose*, just over 50% of the agentless generic situations considered are expressed by means of *man*-constructions.

By contrast, the status of *men* in Dutch is quite different. Although it can be used generically and episodically its use is evidently in decline. The construction is only rarely found in speech and if so essentially only in academic discussions, which is why, Weerman (2006) argues, there are no instances of *men*-constructions in the Dutch CHILDES corpus, as compared to its German equivalent. The different status of *men* in Dutch as opposed to *man* in German and *on* in French is also reflected in my parallel translation corpus of Harry Potter. There are no instances of *men* in the first six chapters of the Dutch translation of *Harry Potter and the Philosopher’s Stone*, while there are 37 instances of *man* in the German translation and 101 instances of *on* in the French, (half of the instances of *on* correspond to *we*). In Catalan, on the other hand, in which *homo* is also considered to be formal and bookish (Wheeler et al. 1999:518), attempts are being made to revive the construction.

---

15. The first person plural uses of *on* and also its other definite uses (it can in fact be used with reference to any of the three persons) are even more common. In informal spoken French *on* has virtually replaced *nous* (Coveney 2000).

16. Jensen provides examples from Danish, of the specific 1pl reading of *man*. It thus appears that Egerland (2003:84) is incorrect in asserting that Danish differs from Swedish in this respect.
Of the man-imps in the languages outside of the core area, only the Icelandic 
mađur has received some detailed attention (Egerland 2003:84; Ragnarsdóttir & 
Strömqvist 2005:148). Unlike the Mainland Scandinavian man, mađur appears to be 
decidedly colloquial. Ragnarsdóttir & Strömqvist state that it is frowned upon and 
associated with "low style" but this negative assessment is not shared by all speakers 
(Jóhanna Barđdal p.c). In more formal registers and written texts, the passive is clearly 
preferred ((Barđdal & Molnar 2003). The status of Faroese man is also seen to be low. 
It is viewed as a borrowing from German via Danish (Barnes & Weyhe 1994:203) and 
its use is thus discouraged.

Unlike man-imps, 3pl-imps do not tend to be the main vehicles of either 
generic or episodic non-referential uses in the languages in which they occur. This 
holds even for most of the languages in group I. In Italian, Sardinian, Spanish, Cata-
lan, Galician, Polish, Serbo-Croatian and Slovene the primary means of expressing 
generic impersonals is via the reflexive impersonal active construction. Impersonal 
episodic reference is preferentially rendered via a host of passive constructions or 
in the case of Polish also by the active impersonal no/to construction (see Kibort, 
this volume). The use of 3pl-imps is considered to be colloquial and/or dialectal. 
The same applies to the Germanic and Scandinavian languages, as well as to French 
(Creissels 2007:21).

Quite exceptional in this context are Russian, Greek, Hungarian and a number of 
Italian dialects. In Russian, Greek and Hungarian 3pl-imps are used not only in col-
loquial speech or renditions of such in fiction but also more widely both in fiction and 
expository texts. For example, in Sansò’s (2006) previously mentioned investigation of 
The Name of the Rose of the five languages considered, Italian, Spanish, Danish, Polish 
and Greek, Greek is the only language which used 3pl-imps often enough for them to 
be taken into account as an impersonalizing strategy. And if frequency of use in the 
Harry Potter corpus is anything to go by, then 3pl-imps in Russian are even more com-
mon than in Greek. I have no frequency data for Hungarian. However, as the passive is 
highly restricted in this language, 3pl-imps are seen to be very widely used. In the case 
of varieties of Italian such as Sicilian (Giorgio Iemmolo p.c) Napalese and Cosenza 
(Ledgeway 2000:237–239) 3pl-imps are claimed to be unequivocally preferred to the 
passive which is considered a borrowing from Italian.

If we allocate to each language the R-impersonalizing strategy used most fre-
quently and widely in that language, considerably fewer languages emerge as man-imp 
languages and especially as 3pl-imp languages than in Figures 1 and 2. The relevant 
data are presented in the map in Figure 4 where the squares denote languages with 
dominant man-imps, the circles the languages with dominant 3pl-imps, the triangles 
the languages for which 3pl may be dominant (but we do not know for sure due to lack 
of data) and the hexagonal shapes the languages where some other R-impersonalizing 
strategy is dominant.
7. Concluding remarks

The paper sought to establish the patterns of distribution of man-imps and 3pl-imps among the languages of Europe. We have found that the distribution of the two types of impersonals is effected by genetic, areal and typological factors. man-imps are clearly currently most strongly associated with the Germanic languages and French. Whether there is a genetic skewing to 3pl-imps is less obvious. They are most probably best viewed as a pan-European or Eurasian phenomenon which has receded from the Germanic languages due to the expansion of man-imps and which failed to fully develop in Baltic Finnic and also Celtic due to the dominance of verbal impersonals. There is also a strong typological factor affecting the distribution of the two: man imps strongly favour non-pro-drop languages and 3pl-imps undergo higher degrees of grammaticalization in pro-drop languages, especially those which freely allow OV order. As for the actual frequency of use of the two impersonal constructions in languages, this appears to be highly dependent on the range of alternative R-impersonalizing strategies available. The major competitors to man-imps appear to be passives. 3pl-imps, by contrast, compete mainly with active impersonals, especially reflexives, as in Romance and Slavic. Only in the languages which lack such impersonals, Russian, Greek and Hungarian, do 3pl-imps emerge as the favoured forms of R-impersonalizing.
References


Verbs of motion

Impersonal passivization between unaccusativity and unergativity

Werner Abraham
Universities of Vienna and Munich

Motion verbs (VoM) are discussed as examples of exceptions to the definition of unaccusativity. The problem arises under impersonal passivization (IPass): VoM may, on the one hand, be subjected to impersonal passivization, but, on the other hand, they turn out as ergative verbs/unaccusatives in directional use. As laid out in §2, the clash consists in the fact that IPass is always imperfective, whereas unaccusatives are perfectives. The solution is sought in the assumption that VoM are split unaccusatives: unergatives in the present, ergatives in the preterit participle. The main body of this article is devoted to illustrating and discussing in detail this assumption. §3 is on split auxiliary selection and the Unaccusative Hypothesis, the search for a uniquely motivated auxiliary selection as well as transitivity and mutativity parameters. Russian verbs of motion (VoM) turn out to specify the essential distinctions of split unaccusativity (in 3.3). Event semantics and event syntax as well as event decomposition and the underspecificity of VoMs (3.3 and 3.4) further pave the pathway to an understanding of the specifics of VoM as ergatives/unaccusatives and unergatives simultaneously. §5 presents Scandinavian as an alternative encoding to impersonal passivization. §7, finally, summarizes the findings to give an answer to the question: How do agentivity and unaccusativity align? A brief discussion of Aspect-based account vs. argument-based account for passivization ($8$) leads to the unified IPass criterion in §9. §10 furnishes the results of our discussion.

Keywords: unaccusative; unergative; passive; aspect; agentivity

1. Outlook

There is an interesting relational question pertaining to impersonal passivization, namely what do ergative verbs have to do with impersonal passives? More precisely, this question subsumes several more specific ones. First, given a certain language, why is it
that not all verbs, and in particular not all e(rgative)V(verbs),\(^1\) can passivize impersonally; and, second, why is it that a number of languages disallow impersonal passivization altogether? Furthermore and third, if we think of narrative functions of impersonal active and passive constructions (Sansò 2006, among others), one may ask what the place of impersonal passives (IPass) is in such a functional scenario and how do they relate form and function with explanatory force. In seeking to answer the above, we will concentrate on data from German(ic) and Slavic. The main tenet of our explanation will be that only languages that have sufficient means of distinguishing verb-diathetic functions will allow impersonal passives in the first place. However, the “sufficient means” cannot be simply equated with the presence of a distinction in terms of \textit{be} and \textit{have} as a comparison of German and English might suggest. Although German has a tripartite Aux division between \textit{be}, \textit{have}, and \textit{become} and has impersonal passives while English lacks the Aux distinction and impersonal passives, the Scandinavian languages have IPass in the absence of the Auxiliary tripartition. Moreover, while, almost expectedly according to our tenet, Russian shies away from the IPass radically, Polish does in fact have IPass (Frajzyngier 1982), albeit quite infrequently in usage.

Another question that arises in connection with impersonal passivization is what do verbs of motion have in common with eV and why can they in fact passivize impersonally – at least under a certain derivation.

We begin with a brief discussion of what IPass should be. And we continue with verbs of motion (VoM) as a means to touch upon the core questions of IPass.

2. \textbf{Motion verbs (VoM) as exceptions to unaccusativity: The problem under impersonal passivization}

Impersonal passives are passives without any subject. In other words, either the original external argument of the active version of a monovalent verb appears only in demoted status (its theta role being suppressed, but not deleted; see (5c) below for the monovalent verb \textit{laufen} “run”), or any object in the original active construction of a bivalent verb has not been theta promoted to subject status (transitive IPass; rarely represented across languages), i.e. remains in situ (see (5b) below). See also (4a,b) for the Icelandic IPass (at least (4b) is an IPass). An IPass, thus, either subcategorizes

\(^1\) The following abbreviations will be used:

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>iV</td>
<td>intransitive verb</td>
</tr>
<tr>
<td>tV</td>
<td>transitive verb</td>
</tr>
<tr>
<td>eV</td>
<td>ergative (=unaccusative) verb</td>
</tr>
<tr>
<td>ueV</td>
<td>unergative verb</td>
</tr>
<tr>
<td>IPass</td>
<td>impersonal passive</td>
</tr>
<tr>
<td>ongPass</td>
<td>ongoing event passive</td>
</tr>
<tr>
<td>DO</td>
<td>direct/structural object</td>
</tr>
</tbody>
</table>

I use “ergative verb/eV” and “unaccusative verb” indiscriminately in this paper.
for no arguments at all or may do so just for the original structural object staying unpromoted in situ. It suppresses, but does not delete, its demoted external/subject argument – which remains intact for semantic interpretation. From this follows that unaccusatives (eV) do not passivize – this is for the simple reason that the subject of eV has been promoted from its original internal argument position already and, under the normal freezing conditions ("no two derivative operations of the same kind"), is no longer accessible for another promotion.

Notice that under such a grammatical scenario, Burzio’s Generalization (Burzio 1986) does not hold since the passive participle should have lost its force of governing an accusative object.

1. “A verb which lacks an external argument fails to assign accusative case.”
   (Burzio 1986: 178f.)
2. “A verb which fails to assign accusative case fails to theta-mark an external argument.”
   (Burzio 1986: 184)
3. “All and only the verbs that can assign a theta role to the subject can assign (accusative) Case to an object”
   (Burzio 1986: 184)

We claim here, as we have claimed earlier (Abraham 1993) and as was discussed later in detail (Alexiadou et al. 2004), that Burzio’s Generalization covers only part of the case related passive conversion. Any passives with unpromoted objects fall outside the range of Burzio. Needless to say, Burzio’s Generalization holds as long as there is a promoted nominative (more precisely, as long as the passive participle has lost its case government of the object accusative) irrespective of its linear place in the clause.2 Thus, the Icelandic indefinite passive in (4a) does not demonstrate an unpromoted theta bearer in the nominative (counter to Barðdal & Molnár 2003: 232). See further the ‘New passive’ in (4b) (Árnadóttir & Sigurðsson 2008).

4. a. það var laminn strákur á leikvellinum
   there.EXPL was hit a boy.NOM on the playground
   "A boy was being hit on the playground"

   b. það var barið hana
   it.EXPL was beaten her.ACC
   "There was beating her" = “She was being beaten"

We shall return to (4b) later. With respect to (4a), the bare fact that the nominative does not topicalize has nothing to do with functional promotion as a consequence of

---

2. Alexiadou, Anagnostopoulou & Everaert (2004) provide a succinct overview of the developments in syntactic theorizing over the last two decades (conducted mostly in the transformational/Minimalist tradition) which have resulted from Burzio’s Generalization (Burzio 1986). Their conclusion is that ‘the syntax of unaccusativity drawn in Burzio can no longer be maintained’ (Alexiadou, Anagnostopoulou & Everaert 2004: 13).
theta promotion and the subsequent case conversion. See German (5a) which is not an impersonal passive, nor an abnormal passive in any sense despite the fact that the nominative subject, *(die) Zähne* “(the) teeth”, is not in topical sentence position.

(5) a. *Heute werden endlich einmal anständig (die)*

   tonight become.pl finally once optimally (the)

   Zähne geputzt          ... ongPass/from bivalence

   teeth.nom.pl cleaned.part

   “Tonight, the teeth will be brushed really well”

b. *Heute wird endlich einmal anständig (?*die)*

   tonight become.sg finally once optimally (the)

   Zähne geputzt          ... IPass/unpromoted DO

   teeth.acc.pl cleaned.part

   “Tonight, there will be perfect teeth brushing”

c. *Heute wird wieder einmal gelaufen ... IPass/from monovalence*

   today becomes.sg once again run.part

What is at stake with the linear position of any argument may have a bearing on whether the construction is *thetic* (as (4)–(6) are) or not and, consequently have an inceptive narrative function. Such discursive criteria have no bearing on the grammatical status.

Impersonal passives derive from one place predicates under passive conversion, thus prototypically from VoM. See (6a) below for such a typical IPass. The incipient question is why (6a) is good, while (6c) is not and the acceptability of (6b) is between (6a) and (6c).³ Note the case difference between (6a) and (6b) in the PP.

(6) a. *Es wird im Saal getanzt*

   it becomes in the hall danced

   “There is dancing in the hall”

b. *Es wird in den Saal hineingetanzt*

   it becomes into the hall danced

   “There is dancing into the barn”

c. *Im 1. Weltkrieg schon wurde viel gestorben*

   in WWI already was a lot died

   “There was much dying in WWI”

³. I have enjoyed invaluable discussions with Wladimir and Gerda Klimonov as well as Leonid Kulikov on Russian and with Jadwiga and Kinga Piskorz on Polish. Joan Maling had a critical look at the Icelandic examples. Elisabeth Leiss, Jóhanna Barðdal, and Elly van Gelderen were helpful with bibliographical data. Furthermore, two anonymous reviewers as well as the volume editors have contributed numerous and extremely valuable remarks that helped reshape this paper.
German is one of the not so frequent languages that allows for impersonal passives (Abraham & Leiss 2006). While the absence of a stative/adjectival impersonal passive for imperfective *t*anzen can be attributed to the fact that the stative/adjectival impersonal passive for imperfective tanzen (in German/Dutch: *Es ist getanzt/*Er is gedansd) fails – see Abraham 2006 as well as Abraham & Leiss (2006) – verbs of motion present problems in their own right, namely the following ones.

Certain verbs which qualify as eV by all control properties such as auxiliary selection, participial attributivity to the external argument (subject) etc. (see Abraham 2002 for German) show an exceptional behavior since they are definitely agentive. Consider *laufen/springen/ tanzen* “run/jump/dance”. While clearly being intransitive verbs, they fail the attribute test, as shown in (7a). On the other hand, the derived directional verbs *hineinlaufen/in den Graben springen/heraustanzen* “in-run/into the moat jump/dance out” meet the attribute test as (7b) demonstrates.

(7)  

a. *das gelaufene/gesprungene/getanzte Mädchen*  
the run/jumped/danced girl  
although *sein/be*-verbs

b. *das hinein/in den Saal/gelaufene/gesprungene/ getanzte Mädchen*  
the in/into the stable run/jumped/danced girl  
although *AGENT*-subject

The crux is that not only is the iV agentive (has an Agent subject theta), but the supposed eV is also – which is counter to the non-Agent requirement for eVs. Agentives cannot be internal arguments. Consequently, while *laufen* “run” with the external argument, eA, expectedly disqualifies for the attribute test in (2a) above, *hineinlaufen* “in-run” is an eV under the force of the attribute test, but it is not for the agency control property. Notice what this means: *hineinlaufen* as a directional motion verb takes an agent – a semantic role which disqualifies for iA status – but the attribute property points to the opposite, i.e. that it has clausal object status. What is at the bottom of this apparent incompatibility?

One way out of the apparent dilemma would be to assume that the past participle attribute property only yields unaccusativity for the result phase – not, however, for the incremental phase of perfective verbs. This would mean that directional motion phrases and particle verbs like *hineinlaufen/-springen* “in-run/-jump” are split ergatives: iV in the present tense since agentive, but unaccusative only in the past participle, where agentivity no longer holds for the mere reason that results are adjectival (consider the copula auxiliary *sein* “be” for adjectives) and, consequently, non-agentive.

Notice that this solution of split unaccusativity does not hold for simple, underven eV such as *sterben* “die” and *ankommen* “arrive” for the simple reason that no exception has to be taken for the present tense realizations of these verbs: they are undeniably non-agent. No split of the unaccusativity property has to be foreseen.
The fact that English cannot make use of this control property (cannot apply it to classify eVs) is due to the principled inapplicability of past participles as nominal attributes (which, in turn, may be due to the fact that English has no passive auxiliary options between be and werden “become” thereby preempting the distinction between statal/adjectival passive participles and ongoing passive participles).

It is not without interest and general appeal that the list of illustrations, (4)–(9) as well as others, and the ensuing discussion are similar to the line of argument in the Siewierska & Malchukov (in this volume).

3. Split auxiliary selection and the Unaccusative Hypothesis

3.1 The search for a uniquely motivated auxiliary selection:

This section deals with auxiliary selection in the present perfect of monadic verbs which appears to at least partly reflect a pattern grounded in lexico-aspectual properties of individual verbs though mediated through their syntax. It may seem that the fact that a language distinguishes stative BE and ongoing HAVE qualifies it for the IPAss construal. But this is too meager a Null Hypothesis as the ensuing discussion will show.

See the distribution of German constituents of VoM which is inseparably linked to the Aux sein/be, a choice from sein, haben, werden, and bekommen next to haben. The Aux split of German and Dutch – one of the determining control properties of transitivity (always haben, with but a few exceptions) vs. intransitivity (alternates between sein and haben), eventive passive (only werden) vs. adjectival passive (only sein), and the so-called dative, or recipient, passive (bekommen) as well as aspect choices – is shared by Italian as well as other languages. To what extent is this unified across languages? It will be seen that it is not – although there are important generalizations.

<table>
<thead>
<tr>
<th>Italian eV: essere</th>
<th>German eV: sein</th>
<th>Italian iV: avere</th>
<th>German iV: haben</th>
</tr>
</thead>
<tbody>
<tr>
<td>è caduto “fell”</td>
<td>ist gefallen</td>
<td>ha sorriso “smiled”</td>
<td>hat gelacht</td>
</tr>
<tr>
<td>è partito “left”</td>
<td>ist weggegangen</td>
<td>ha mentito “lied”</td>
<td>hat gelogen</td>
</tr>
<tr>
<td>è sparito “disappeared”</td>
<td>ist verschwunden</td>
<td>ha tremato “trembled”</td>
<td>hat gezittert</td>
</tr>
<tr>
<td>è rimasto “remained”</td>
<td>ist geblieben</td>
<td>ha sanguinato “bled”</td>
<td>hat geblutet</td>
</tr>
<tr>
<td>è scoppito “exploded”</td>
<td>ist explodiert</td>
<td>ha schiumato “foamed”</td>
<td>hat geschäumt</td>
</tr>
<tr>
<td>è morto “died”</td>
<td>ist gestorben</td>
<td>ha nuotato “swam”</td>
<td>(hat)/ist geschwommen(^4)</td>
</tr>
<tr>
<td>è arrossito “blushed”</td>
<td>ist errötet</td>
<td>ha peccato “sinned”</td>
<td>hat gesündigt</td>
</tr>
</tbody>
</table>

Figure 1. Intransitives and Aux selection in Italian and German (partly following Aranovich 2006: 2) – underscored verbs are VoM (eV=ergative V, iV=intransitive/unergative V)

\(4\) Dutch uses Auxes, zijn/wezen “be” and hebben “have”, identical to those in German except for verbs of motion, where hebben accounts for the durative use as in heeft de hele dag gezwommen “has swum all day” as opposed to is tot het andere oever gezwommen ‘is swum
Since Perlmutter (1978) first pointed out the importance of split auxiliary selection phenomena for modern theoretical linguistics, careful research has shown variation in auxiliary selection to be widespread.

Syntactic analyses of split auxiliary selection are based on evidence that the subject of \textit{BE}-verbs shares properties with direct objects of transitives, while the subject of \textit{HAVE}-verbs shares them with the subject of transitives. Recall the discussion around (4)–(7) above. The formal expression of this observation is the Unaccusative Hypothesis (UH; see Perlmutter 1978; Hoekstra 1984; Burzio 1986). Unaccusative clauses have an underlying object, but no underlying subject. The UH, originally formulated in Relational Grammar, assigns an unaccusative clause as in (8) in modern formal syntax.

The fact that verbs with external theta-roles select \textit{HAVE} is explained in Hoekstra (1984) as a result of the fact that \textit{HAVE} differs from \textit{BE} in its ability to assign accusative case. When \textit{HAVE} as a full verb combines with a transitive participle, the direct object of \textit{HAVE} receives its case not from the participle but from the auxiliary verb. \textit{BE} lacks this case assigning property. From this follows that in conjunction with a transitive participle it can only yield a passive construction. If the participle is unaccusative, \textit{BE} must be selected. With the auxiliary \textit{HAVE}, on the other hand, the internal argument would receive accusative case, and it could not move to the subject position.

After such an analytic explanation as that by Hoekstra (1984), the question arises why the distribution of Aux is different across languages after all. Recall that we suggested sufficient auxiliary distinctions to both legitimize IPass and the formal distinction of unaccusativity vs. unergativity.

\begin{center}
\begin{tikzpicture}[level distance=1.5cm, sibling distance=1.5cm]
  \node (ip) {IP}
    child {node (np) {NP}
      child {node (i) {I'}}
      child {node (case) {case\textbackslash e}}
      child {node (v) {V}}
      child {node (vp) {VP}}
    }
    child {node (v') {V'}}
  child {node (vot) {\textit{caduta}}}
  child {node (var) {\textit{\theta-role}}} 
\end{tikzpicture}
\end{center}

until the other bank’ with a telic adverb. There is a like tendency in the north of the German speaking area (Hamburg).
The expectation of sufficient auxiliary distinction based on a unified semantic list of criteria turns out to be disappointing. The solution to this aporia has been sought in assemblies of criteria for Aux assignment different from those in Figure 2 above. I will briefly discuss the criteria proposed by Shannon (1992) and by Sorace (2000). The first of the two studies focuses on sensitivity to the degree of transitivity of a clause as well as three transitivity parameters: the number of participants, telicity, and punctuality. The methodological assumption behind Shannon’s study is his conviction that the syntactic phenomenon of auxiliary selection is gradient and not binary. Since, according to Shannon, a binary classification scheme cannot predict which verbs in German will take haben and which will take sein, Shannon (1992, 1995) took the route that Hopper & Thompson (1980) had taken on developing transitivity parameters. Hopper and Thompson’s parameters are number of participants, kinesis, aspect, punctuality, volitivity, affirmation, mode, agency, and affectedness of the object and separateness of the object. The underlying principle in Hopper & Thompson (1980) and Shannon (1992, 1995) is that it is not a verb that is transitive or intransitive, but a clause. Shannon claims that prototypically transitive and prototypically mutative events can be represented by expanding Hopper & Thompson’s parameters. Mutativity is a new term and it represents clauses that have one participant and are punctual. Shannon’s (1992: 103) classification scheme is given in Figure 8. See also Arnett (2006: 27).

![Figure 2](image)

**Figure 2.** Cross-linguistic variation in auxiliary selection (abbreviations: dir. = directional, ch. = change, loc. = location, cont. = continuation)

![Figure 3](image)

**Figure 3.** Transitivity and mutativity parameters
Sorace (2000) highlights her finding that, within the Constructional Model of unaccusativity (see e.g. Borer 1994; McClure 1995; van Hout 1998), the syntactic status of verbs is unstable. Thus, if agentivity gets added to change of location verbs like fall ‘fall’, the verb’s syntactic structure is affected, not its auxiliary choice – the auxiliary is always be. It is the presence or absence of an aspectual functional head that determines unergativity versus unaccusativity. If this aspectual head (which corresponds to the agent-role-assigning v-head) is present, the verb has unergative status. If it is not present, the verb has unaccusative status.

Needless to say that the criteria in Figure 3 share the unaffirmative quality that we observed for those used in Figure 2. They yield no uniform class distinctions. Now see Figure 4 below. Both Sorace and Keller & Sorace are concerned with identifying the cross-linguistic parameters that determine auxiliary selection for monadic non-reflexive intransitive verbs. They propose the hierarchy shown in Figure 4 (Sorace 2000:863), a structural hierarchy of verb types, defined by aspectual and thematic parameters.

<table>
<thead>
<tr>
<th>Verb Class</th>
<th>Auxiliary Selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHANGE OF LOCATION</td>
<td>selects BE (least variation)</td>
</tr>
<tr>
<td>CHANGE OF STATE</td>
<td>↓</td>
</tr>
<tr>
<td>CONTINUATION OF A PRE-EXISTING STATE</td>
<td>↓</td>
</tr>
<tr>
<td>EXISTENCE OF STATE</td>
<td></td>
</tr>
<tr>
<td>UNCONTROLLED PROCESS</td>
<td></td>
</tr>
<tr>
<td>CONTROLLED PROCESS (MOTIONAL)</td>
<td>selects HAVE (least variation)</td>
</tr>
<tr>
<td>CONTROLLED PROCESS (NONMOTIONAL)</td>
<td></td>
</tr>
</tbody>
</table>

Figure 4. Auxiliary Selection Hierarchy (ASH):

Cross-linguistically, verb classes which fall at the ends of ASH are predicted to show the least variation with respect to auxiliary selection. At the extreme higher end, auxiliary selection (in languages which have a choice of auxiliary) is invariably BE, and at the extreme lower end, it is invariably HAVE. The further away from the extremes, the greater the variation in auxiliary selection a verb displays. “Parameter sensitivity of verb classes (which are not at either end of the hierarchy) differs from language to language, and even within a verb class of one language, from verb to verb” (quote from Lee-Schoenfeld 2006:138). This, then, is at the bottom of the variation as sketched in Figure 4 above – not much of a relief to the aporia we found ourselves at given Figure 4.5 What we appear to miss are fundamentally different criteria for

5. Lee-Schoenfeld (2006) questions Sorace’s (2000) results in the Auxiliary Selection Hierarchy. Her main claim is that agentivity and auxiliary choice do not correlate and that
distinguishing ongoingness, on the one hand, and stativity or verbal adjectivalhood, on the other hand.\(^6\)

In conclusion one can say that there are distributional correlates of unaccusativity and perfectivity, relating to Auxiliary selection, across languages, but none is reliable in an absolute way as Figures 2–4 demonstrate. Without doubt, there is interaction between any of the three different lists of categorial correlations, but no such interaction appears to be universal. Moreover, not a single one of these lists and their correlates offers an explanation to our VoM-problem: Agent subjecthood and nevertheless unaccusativity status, or, by the same problematic token, *sein/be* auxiliary selection paired with Agent subjecthood. We appear to be in need of a quite different corollary and explanation. Notice that none of the listed illustrations and distributional criteria contains the correspondent Aux for German *werden* “become”. We believe that this is a crucial factor accounting, in ways still to be explained, for the occurrence of IPass.

The fact that German VoM and their narrow constituents show irregular behavior has a direct replica in Russian. Moreover, the emergent documented history of Russian clearly instantiates the transition from lexical spatial and temporal boundedness to perfective aspect.

### 3.2 Russian verbs of motion (VoM)

According to Russian grammar (Tauscher & Kirschbaum 1983: 271ff.), there is a class 14 of verbs (*Doppelzeitwörter* “double verbs”) that have the following properties: They are imperfective, unprefixed, and express some motion. They occur in pairs of far-related forms whose specifics are contingent upon the feature \([\pm \text{definite}]\) or \([\pm \text{determined}]\), respectively, which defines Russian Aktionsarten (as opposed to Aspect). See (11) below. An event is definite if its duration is specific or definite/...
Verbs of motion

101

determined (indicated, for example, by adverbial expressions such as “a little”, “for a certain period of time” as opposed to specifying a goal to be reached). They are usually prefixed with po- as in (9). Needless to say, the verbs in (9) do not represent VoM, while (10) indeed do.

(9) posidet’ “sit for a while”
pospat’ “sleep a little”
pochitat’ “read a little”

(10) poplyt’ “begin to swim”, “start swimming” … Ingressive with [+det]
as opposed to
poplavat’ “swim a little/a while” … Delimitative with [–det]

The definite Aktionsart comprises also ingressesives and resultatives, but not iteratives which are indefinite. Most relevantly, the following non-prefixed verb pairs go by the criterion of definiteness (Tauscher & Kirschbaum 1983:270f.; I restrict myself to a selection):

(11) ±definiteness/determination
    of direction or goal

+    -
idi  xodit’ “go”
lejet’ letat’ “fly”
bežat’ begat’ “run”
bresti brodit’ “stroll”
plyt’ plavat’ “swim”
nesti nosit’ “bring, carry”

The definiteness criterion as in (11) occurs also under the terminology of telic/atelic (Russian predel’nyj/nepredel’nyj, i.e. literally, [goal delimited] vs. [goal unde-limited]). Etymologically, number of (but not all) atelic verbs have a reflex of the

Markedness of the primary imperfective verbs occurred by stem extension through /-a-/ as in pomoch’ – pomogat’. See also the forms in (11) in the undetermined branch. See Regnéll (1944:29ff. and 55). Other philologists such as Delbrück (1897:75), Kuiper (1937:203), and Poulteney (1937) point at nasal-infixing present encoding terminativity. This is reminiscent of the class of -nan-verbs encoding perfective intransitivity (= unaccusativity/ergativity) in Gothic (Leiss 1992).

7. The distinction is also captured by the terminologies of ‘primary perfective’ (or ‘determined’) as opposed to ‘perfectiv(iz)e(d)’ and ‘primary imperfective’ (‘undetermined’) as opposed to ‘imperfectiv(iz)e(d)’. The distinction is not without good philological reason as there are nowadays two types of verbs: the ones where the unmarked form of the pair is imperfective (as with delat’ as opposed to the perfective sdelat’), and the other ones where the unmarked form of the pair is perfective (as with past’ as opposed to the marked padat’). Markedness of the primary imperfective verbs occurred by stem extension through /-a-/ as in pomoch’ – pomogat’. See also the forms in (11) in the undetermined branch. See Regnéll (1944:29ff. and 55). Other philologists such as Delbrück (1897:75), Kuiper (1937:203), and Poulteney (1937) point at nasal-infixing present encoding terminativity. This is reminiscent of the class of -nan-verbs encoding perfective intransitivity (= unaccusativity/ergativity) in Gothic (Leiss 1992).
Proto-Indoeuropean suffix *-eio/- (identical with the causative suffix) as in nesti – nosi(t’) (< *nos-ei- …) (Kulikov 2008).

Definite motion verbs in Russian designate a unique (singleton) uninterrupted event with a specific direction extended over a limited period of time and with a specific goal. By contrast, indefinite VoM are undetermined with respect to direction, the motion event occurs more than once, they are not goal-directed; they may also denote a habitual event which may be directionally and temporally determined; they may characterize an ability of someone to execute an action; and they may designate a singleton event of motion to-and fro. (12)–(13) illustrate this with one of the most frequent double verbs of motion, idti “go”.

(12) Mal’chik idets v gorod … “The boy is going to town”
(13) Mal’chik xodits po dvoru … “The boy is walking around in the yard”

What does this paradigm of Russian double verbs of motion tell us in view of the problem indicated for German VoM? The interesting fact is that the definiteness distinction was the diachronically primary one. In other words, the modern aspectual distinction is a reanalysis of the original one by the definiteness criterion.

(14) a. polzti “creep” [+det] as opposed to polzati “creep around” [–det]
⇒ by way of reanalysis and grammaticalization
b. padti [+perf] as opposed to padati [–perf], both “fall”

In the case that prefixes were involved, reinterpretation took place in terms of grammaticalization of the prefixes. See (15) below where “V-” stands for the lexical preposition v “into”, whereas “v-” is the prefix category of the same morpheme – much like German hinein in hineinkriechen ‘in-creep’ as opposed to the unprefixed kriechen “creep”.

8. One reviewer points out that nesti and nosit’ are both atelic (given that they both combine with duration adverbs like dva chasa “for two hours”). Given the clash of positions on this issue with the literature cited, I will not attempt to solve this contradiction.

9. Examples in (14)–(15) and discussion relate to Old Russian. The simplex pasti (< *padti) and the prefigated derivate upasti, both equally for “fall”, compete with one another. As distinct from Old and Modern Greek, grammaticalization (i.e. the perfectivizing effect) of the prefixes decided this competition in favor of prefigated upasti. The latter form made it into Modern Russian, while the old perfective past’ (< pasti < *padti) was retained only with a slightly archaic touch.
a. the original opposition pair

\[ V_{polzti} [+det] \rightarrow V_{polzati} [-det] \]

for “creep in” vs. “creep around”

⇒ under grammaticalization of the prefixes

b. the secondary opposition pair

\[ vpolzti [+perf] \rightarrow vpolzati [-perf] \]

“creep in(to)” vs. “creep around”

3.3 Event semantics and event syntax

These facts in the previous §3.2., trigger the following idea. Take preterit presents (preterit-in-form, but present-by-meaning) such as the primary forms and meanings of the Germanic modal verbs and German wissen “know” and the English cognate wit (as in to wit); comparable Latin forms are the class of inchoatives such as (cog) nōscō “learn ≡ get to know” – (cog)nōvi “gotten to learn ≡ know”. The crucial difference between the present and the preterit tenses is their distinct meanings that are based on the corresponding event properties: Presents of inchoatives (telics) denote the incremental event phase of what is necessarily to be reached eventually as the result of that incremental phase: the resultant phase. With this type of event, one implies the other and vice versa: no result without the preceding approach event, and no approach event without the resultant state. The two tense-aspects in German and English are essentially congruent sharing an incremental first phase and a consequent result due to the perfective verlassen/leave embedded under the Aux have. See (16a) below for the biphasic representation of perfectives as opposed to that of imperfectives in (16b), the latter being monophasic (Abraham 1990). [eΘ=external theta role; iΘ=internal theta role; |>>>| = incremental event phase; |----| = resultative/statal/adjectival phase].

(16)

a. eΘ leaving iΘ iΘ being left

\[ \text{for perfective tV, leave:} \]

\[ \exists X \forall t \forall r(1-n) [X(t) \leftrightarrow Y(r)] \]; i.e.

biphasic event: the two event phases imply each other; t and r do not mark identical event properties.

b. eΘ contemplating iΘ

\[ \text{for imperfective tV, contemplate:} \]

\[ \forall Y [ev \subseteq t \rightarrow \neg \forall f \subseteq ev. \exists i \subseteq t(Y(ev, f) \leftrightarrow Y(t, i))] \]; monophasic event or state: t and r mark identical event properties throughout.

Modal verbs are of essentially the same make as (16a), i.e. biphasic, although futural instead of incremental. Thus, it does not come as a surprise that they designate future events much in line (though not identical) with (16a).

To sketch the two aspects, or Aktionsarten, as in (16a,b) is in line with a syntax that mirrors eventive predication in the sense of Ramchand’s First Phase Syntax
(2008). In this syntax the argument relations are tied to event decomposition that includes (maximally) a causing subevent, a core process or transition, and a result state. See (17) below.

While I do not share Ramchand’s 2008 precondition that these relations are the only ones possible (see Borer 1998, among others, for alternatives while sharing the basic event predicative notion) the syntactic projection of arguments based on event structure is in line with the notion pursued from scratch in Abraham 1993. There, the semantics of resultatives (as opposed to that of imperfectives; see (16b)) was conceived of in terms of secondary predications (i.e. an incremental event phase followed by a result phase). Their syntax mirrors these semantic relations as closely as possible – with the fundamental idea behind it that, if we draw but a close line between aspect and Aktionsart (Abraham 1993, 2006 a,b), the syntax should mirror the lexical (as opposed to the functional) properties of their morphologies. See (17) for a structure that resembles, but is not identical with, Ramchand’s (2008) proposal featuring three sub-event projections for the representation of all the possible components of the event structure building processes of dynamic predicates.

This is what the present structural division amounts to: the topmost for lexically determined causatives, an intermediate one for eventively undetermined processes, and the deepest level for morphologically determined results, each in a separate projection and with the result phase as the lexically deepest (or most lexical) of the eventive projections (see a similar reasoning in Abraham 1990).

(17) $vP \ (= AspCP, \text{causing projection})$

\[ \begin{array}{c}
\text{NP3} \\
\text{subj of ‘cause’}
\end{array}
\begin{array}{c}
v' \\
\text{V'}
\end{array}
\begin{array}{c}
\text{VP} \ (= AsppP, \text{imperfective process projection})
\end{array}

\begin{array}{c}
\text{NP2} \\
\text{subj of ‘process’}
\end{array}
\begin{array}{c}
V
\end{array}
\begin{array}{c}
\text{RP} \ (= AsprP, \text{result state projection})
\end{array}

\begin{array}{c}
\text{NP1} \\
\text{subj of ‘result’}
\end{array}
\begin{array}{c}
R'
\end{array}
\begin{array}{c}
\text{R} \\
\text{XP}
\end{array}
\cdots

In (17), the verb phrase contains three different projections with each projection as an instantiation of a (possible) subpart of the whole event, corresponding to the semantic decomposition of the imperfective and the perfective event properties described in (11a,b) above.
Verbs of motion

- **vP** introduces the causation event and licenses different types of external argument (‘subject’ of cause): e.g. for the transitive verb *move* (a huge weight) or German *einschlafen lassen/einschlafen* “make go to sleep/put to sleep”.

- **VP** specifies the nature of the change or process and licenses the entity undergoing change or process (‘subject’ of process): e.g. for the intransitive verb *move* (away from) or German *schlafen* “sleep”.

- **RP** represents the ‘goal state’ or ‘result state’ of the dynamic event and licenses the entity that comes to hold the result state (‘subject’ of result): e.g. for the unaccusative predicate *move into* or German *einschlafen* “fall asleep” (with the two eventive subcomponents ‘fall more and more asleep’ and ‘(have fallen)/be asleep’).

Totally in line with Ramchand 2008, we assume that **VP** is at the core of dynamic predicates, since it represents change through time (see \( t_1/t_{n+1}, t_n/t_{n+m} \) in (11a,b)). It is present in every dynamic verb. It may have any type of external theta argument, especially Patients and Experiences – thus, non-Agents. By contrast, **vP** exists for verbs with an ‘external’ argument identifiable as the causer or initiator, i.e. the ‘theme’ or subject of the causative subevent. The **RP** only exists when there is a result state explicitly expressed by the lexical predicate. Ramchand highlights her conclusion that it does not correlate with semantic/aspectual boundedness in a general sense. But this is not the present conclusion which is focused on morphologically rich languages such as German.  

3.4 Event decomposition and the underspecificity of VoMs

Notice that this allows for a decision on the alternative typological motion distinctions discussed by Talmy (1985): It appears that the German type of decomposition sketches the optimal way to describe analytically VoM. Based on the illustrations in (18) below, the French intransitive/transitive verb \( (i/t)V \), henceforth) *monter* needs to be decomposed into three covert components: one for the type of movement, the second for the direction or orientation in space, and the third one for the ending location of the predicated argument. Notice that the argument predicated of the result state cannot be the original agent theta, but needs to be a patient. For originally intransitives, this means that the only and overt external argument – let us say of *hin-ein-laufen* ‘thither-in-run’ – needs to be split up into an Agent e-theta for the first, incremental phase, while the

---

10. If telicity arises because of the entailments based on DP structure, resultativity arises out of the wholeness relation of the entire VP despite the fact that the verb by itself may not be lexically specified as expressing a result state. Conversely, the expression of result can further be triggered by auxiliaries, PPs etc. outside the first phase syntax to create predications that are atelic. This alone does not warrant the removal of **RP** in the syntactic representation as Ramchand notes correctly.
coreferential i-theta for the result phase is in fact a patient (since only internal arguments are allowed for states).

\[
\text{(18) } \text{monter} \equiv \text{[type of motion aller [direction en [location haut]] “go up”, where the interior bracket with en haut “up” modifies the outer one, aller “go”.}
\]

<table>
<thead>
<tr>
<th>French type</th>
<th>German type</th>
<th>Direction</th>
<th>location</th>
<th>decomposition</th>
<th>V-diathesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>il monte l’escalier “he walks up the stairs”</td>
<td>er steigt die Treppe</td>
<td>hin-</td>
<td>-auf</td>
<td>il va en haut descaler</td>
<td>eV</td>
</tr>
<tr>
<td>il monte la valise “he carries the suitcase”</td>
<td>er bringt den Koffer</td>
<td>hin-</td>
<td>-auf</td>
<td>il porte la valise en haut</td>
<td>causV</td>
</tr>
<tr>
<td>les prix montent “prices rise”</td>
<td>die Preise steigen/gehen</td>
<td>hin-</td>
<td>-auf</td>
<td>les prix vont en haut</td>
<td>eV</td>
</tr>
<tr>
<td>ils montent les prix “he raises the price”</td>
<td>sie -höhen die Preise</td>
<td>er-</td>
<td></td>
<td>ils font (les prix (sont) en haut)</td>
<td>causV</td>
</tr>
</tbody>
</table>

Figure 5. Decomposition of VoM: French monter vs. German prefix equivalents

Basically, all causatives are resultatives and, consequently, perfectives (Abraham 1995/2005; Schäfer 2008). Lexically derived decausatives may, but need not be perfective. Figure 5 shows to what extent French monter is underspecified with respect to the codes of direction and location that German encodes explicitly.

Let us return to our point of departure. If, on the one hand, VoM are agentive, as they are, and likewise meet one main distributional property of unaccusative/ergative verbs, eV, then this sketches a clear incompatibility (eV are non-agents) – unless VoM represent bipartite verb pairs much in the sense of the Russian “double verbs” illustrated above in (4)–(10). The question, then, is whether these verbs derive historically in any way from something like the event typology of inchoatives/telics in (11). There is no doubt that they do so synchronically. But their morphemic makeup does not show this, and historically we trace them back no further than signaled by (9)–(10). In fact, no such temporal, or aspectual, distinction is apparent that would relate the Russian double verbs with preterit presents of Germanic.

The solution may lie in the terminological division between the primary verbs as (in)definite/(un)determined (with respect to length of time and/or goal limitation and the modern ones as aspectually distinct (im)perfectives. Keep in mind the emergence of perfective aspect in Old Russian. Determination, or definiteness, is a crucial criterion since it applies to VoM as the prototype of determinable verbality. Definiteness takes exception for VoM from the class of intransitive perfectives to begin with, and yet it makes them similar to perfectives in exactly the same way as holds for the
German VoM: There is agentivity of the external argument, but simultaneously there holds agreement with the distributional properties of unaccusatives. These properties are to be aligned with what is goal determination in the sense of the Russian VoM and German constituents of VoM, on the one hand, and with the eventive telicity/perfectivity design in (11) and the atomic elements of decomposition in Figure 5 above. But before we go into this a view at Scandinavian impersonal passivization need be taken, which will turn out to be revealing.

5. Scandinavian: Alternative ways to encode impersonal passivization

The following (19)–(24) is taken from Abraham & Leiss (2006:501).

(19) a. Es wird gelaufen … German  
    it becomes run = “There is running” from NP[+human] läuft ‘NP runs’

b. Wird (*it) heute gelaufen?  
    becomes (*it) today run  
    “Is there running today?”

(20) Currit-ur  
    run.pres-pass.3.sg … Latin

(21) Er/*Het wordt gelopen  
    there/*it becomes run … Dutch

(22) *It is run … English

(23) *Il est couru  
    it is run from On[+human] cours  
    “one runs”

(24) a. Það er dansað í skólanum  
    it is danced in school.the  
    “There is dancing going on at school”

b. Í skólanum er (*pað) dansað.  
    in school.the is (*it) danced  
    … Icelandic

c. Er (*pað) dansað í skólanum?  
    is (*it) danced in school.the  
    … Icelandic

11. It remains to be seen to what extent there is resemblance with the VoM-participles with unaccusative properties and the nan-verbs of Gothic. One thinks of the nasal presents with perfective meaning in Indo-European. See Poultney 1934; Regnèll 1943, and Kulikov 2008.

12. The following insights are due to Jóhanna Barðdal. There is one major difference between the stative adjectival passive and the ordinary passive in Icelandic, namely that the
s-diathesis

\begin{align*}
(25) & \quad a. \quad \text{Det dansas i skolan.} \\
& \quad \text{it dance.REFL-PASS in school.the} \quad \ldots \text{Swedish} \\
& \quad b. \quad \text{I skolan dansas det.} \\
& \quad \text{in school.the dance.REFL-PASS it} \quad \ldots \text{Swedish} \\
& \quad c. \quad \text{Dansas det i skolan?} \\
& \quad \text{dance.REFL-PASS it in school.the} \quad \ldots \text{Swedish}
\end{align*}

s-diathesis\textsuperscript{13}

\begin{align*}
(26) & \quad a. \quad \text{Det danses på skolen.} \\
& \quad \text{it dance.REFL-PASS on school.the} \quad \ldots \text{Norwegian} \\
& \quad b. \quad \text{På skolen danses det.} \\
& \quad \text{on school.the dance.REFL-PASS it} \quad \ldots \text{Norwegian} \\
& \quad c. \quad \text{Danses det på skolen?} \\
& \quad \text{dance.REFL-PASS it on school.the} \quad \ldots \text{Norwegian}
\end{align*}

blir-Aux (*være as Aux is impossible)

\begin{align*}
(27) & \quad a. \quad \text{Det blir danset på skolen.} \\
& \quad \text{it is danced on school.the} \quad \ldots \text{Norwegian} \\
& \quad b. \quad \text{På skolen blir det danset.} \\
& \quad \text{on school.the is it danced} \quad \ldots \text{Norwegian} \\
& \quad c. \quad \text{Blir det danset på skolen?} \\
& \quad \text{is it danced on school.the} \quad \ldots \text{Norwegian} \\
& \quad d. \quad *\text{Under krigen ble det forsvunnet ofte} \\
& \quad \quad \text{during the war was it.EXPL disappeared often} \\
& \quad \quad \text{uten spor} \quad \ldots \text{Norwegian} \\
& \quad \quad \text{without a trace} \\
& \quad \quad \quad \text{“During the war people went underground often without leaving a trace”}
\end{align*}

stative adjectival passive can only occur with nominative subjects, while the ordinary passive differentiates between accusative, dative and genitive object verbs. This means that dative and genitive subjects of the ordinary passive do not maintain their oblique case marking with adjectival passives; see Barðdal & Molnár 2003. There is of course also a corresponding difference in agreement. This pattern is a common Germanic pattern, also found in Gothic. It also means that there is ambiguity with accusative object verbs, which show up with a nominative subject in both constructions, while no ambiguity is found with dative and genitive object verbs.

\textsuperscript{13} Examples are due to Terje Lohndal (p.c.); (27d) is from Maling (2006:220).
Note that the unacceptability of (27d) is in line with our claim that IPass must be imperfective – which is counterfactual to forsvunnet “gone underground”. The glosses in (23)–(24) use be, but the verb bli really is the German werden “become”, both as an Aux and a full (inchoative) verb. be is German sein, in Norwegian være. One cannot use være ‘be’ in these cases of impersonal passives. As for the meaning of Norwegian IPass, these cases have a habitual meaning as opposed to a momentary reading. The variation between the middle form (see Abraham 1994; Schäfer 2008), or -s-form, and the bli-IPass is telling in that the Norwegian -s-form, or -s-passive, is generic or habitual just like the analytic bli-IPass. This means that either is non-perfective – in line with the German and Dutch IPass (Abraham & Leiss 2006).

It is to be noticed that what is called -s-passive in Norwegian has emerged from reflexive diathesis accountable also as a middle construction (cf. Abraham 1994). Such middles cross-linguistically encode habituals, generics, or even properties of individual arguments represented in such constructions.

IPass is cross-linguistically discursively thetic or presentational (i.e. non-categorical, or non-argument predicative) in nature as well as imperfective or non-goal oriented. From this follows as well as from the two variant constructions in Norwegian, (23) and (24), we may draw the generalizing conclusion that only languages with the bipartite intransitive Aux division be-become are able to functionally distinguish IPass. Furthermore, the aspectual characteristics of both (23) and (24) – atelic ongoingness for the blir-passive as well as genericity, or property assignment in the case of the -s-middle construction – are totally in line with the generalizations about IPass in Abraham & Leiss 2006. Since the adjectival/stative passive is possible only under verbal or syntactic (constituent) perfectivity conditions, predicational sein-være are out as Aux choices. However, unless a language possesses the sein-alternative with inchoative werden/bli, IPass remains unencodable vis-à-vis any of the sein-using diatheses. English and the Romance languages are cases in point, and so is Russian.

All that these languages have at their disposition as Aux is a be-form. The ongoing event passive with be in these languages is due to an implication: The state expressed by the code be + past participle implies the prior incremental phase to reach this stative expression. There is no non-homonymic code for the ongoing passive, and, thus, no direct and unambiguous way to distinguish it from the adjectival passive.14

14. Except for English be beaten vs. be being beaten.
This yields the following diathetic division and distribution of criteria.

<table>
<thead>
<tr>
<th>Perfectivity</th>
<th>Discourse theticsity</th>
<th>Ongoingness (vs. stativity)</th>
<th>Aux choice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal passive</td>
<td>±</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>Adjectival passive</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>IPass</td>
<td>–</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Middle construction</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Dative passive</td>
<td>+(±)</td>
<td>–</td>
<td>+/-</td>
</tr>
<tr>
<td>Goal oriented VoM</td>
<td>+</td>
<td>–</td>
<td>+/-</td>
</tr>
</tbody>
</table>

Figure 6. Basic verb-diathetic division in German (and Dutch)

The fact that each verbal diathesis has a different distribution of functional criteria legitimizes the 5 different forms in German and Norwegian.15

Icelandic (24a–c) appears to be a different case. But notice that, while the language has no alternative auxiliary beside be, it has another, synthetic passive – the s-passive, which characterizes the ongoing one as opposed to the periphrastic stative one. We will argue that it is either the lexical alternation or the constructional one that enables a language to exhibit IPass.

6. Addendum: Decomposition of auxiliaries

The link between Aspect and verbal diathesis/voice sketched in Figure 6 above is mirrored by the connections between the three auxiliary verbs sein “be”, werden “become”, and haben “have”. The remarkable fact about the composite passive is its growth out of diachronic aspectual conditions (Abraham 1987, 1992, 1993; Andersen

---

15. To arrive at any wider cross-linguistic generalization one may have to consider Uralo-Altaic languages where IPassives are rather pervasive, but go without a concomitant Aux split. Thus, as one reviewer points out, the impersonal/personal passive split is better conceived of as a one-way correlation: A language with an Aux split is likely to have a passive split, but not the other way around. This concurs with the observation that the use of HAVE as an Aux is cross-linguistically infrequent.
1994; Beedham 1981, 1998; Toyota & Mustafovic 2006). This decomposition of auxiliary verbs is the following one (from Abraham 2006: 482).

\[\text{werden}_1(e) \iff \lambda e \lambda e_1 \lambda e_2: e(e_1, e_2) \land \text{Inchoative}(e_1) \land [\text{Res State}(e_2) \land \text{Culm}(e_1) \land e(e_1 < e_2)]\]

\[\text{werden}_2(e) \iff \lambda e \lambda e_1 \lambda e_2: e(e_1, e_2) \land \text{Ongoing}(e_1) \land \text{Ongoing}(e_1 < e_2)]\]

\[\text{haben/eignen}(e_2) \iff \lambda e_2 \lambda x \lambda y: \text{State}(e_2) \land e_2(x, y) \land \text{Poss}(x) \land \text{Th}(y)\]

\[\text{sein}(e_2) \iff \lambda e_2 \lambda y: \text{State}(e_2) \land e_2(y) \land \text{Th}(y)\]

\[\text{bekommen}(e) \iff \lambda e \lambda e_1 \lambda e_2: \text{haben werden}(e) \iff [e(e_1, e_2) \land \text{Inchoative}(e_1) \land \text{Culm}(e_1) \land e(e_1 < e_2) \land \text{haben}(e)]\]

The two semantics of \text{werden}, (25) and (26), bear on the distinction between the full verb with adjectival, infinitival and nominal selection, on the one hand, and the auxiliary with passive participle selection, on the other hand. Only the second one in (26) can be used for the ongoing passivization. Notice, furthermore, that \text{eignen/haben} "own" have an argument structure which is converse to that of \text{sein} "have" as in \text{i}ch. \text{n}om, \text{habe etwas.acc} \equiv \text{mir.dat} \text{e}ignet/ist \text{etwas.nom}. There are thus the following lexical links:

\[\text{between haben} "have" and \text{sein} "be"; illustrated by Late Latin, where \text{habēre} (x-Nom-inative, y-Accusative) equivocates \text{esse}(y-Dative, x-Nominative) as in Russian "to-me-is";\]

\[\text{between bekommen} "get" (crucial for the Dative passive) and \text{haben werden} "will have";\]

\[\text{summarily, between haben} "have", \text{werden} "become", and \text{sein} "be" to the extent that the distribution of the sub-events is specific of the three predicate types. Cf. (23)–(26) above.\]

7. VoM: How do agentivity and unaccusativity align?

Recall the problem: There is an unbridgeable incompatibility between the fact that telic VoM-constituents are agentive and yet are in line with the distributional property of unaccusatives – i.e. with a verbal diathesis that does not allow an agent as external argument, since this external argument is an internal one by prederivational status.

Now, the directional PP in telic VoM-constituents, much in line with the early telic verbal constituents of motion in Old Russian before the emergence of telic aspect), are at the base of the distributional attribute property of VoM constituents. They are perfectives. Perfectives have the following event structure (see Abraham 1990, 1994). See (36a).
biphasic event structure for \([_{pp} \text{ins Ziel/(ein-)}]_{[_{iV} \text{lauf}]}\) “run across the finishing line” (consisting of an unergative verb, \(iV\), + a terminative PP): the event structure consists of two lexically inherent components, \(E_1\) and \(Zu_2\), carrying together the Aktionsart reading and, consequently, the telic lexical meaning it has. \([t_1, t_m, t_n] = \text{temporal points on the event-constituting axis; } E_1 = \text{‘approach, or incremental, event component, } ResZu_2 = \text{‘resulting state’}\) = state phase resulting from \(E_1\). Read: \text{ins Ziel laufen ‘across-the finishing line-run’ is characterized by the telic, incremental, event phase (>>>), whereas eingelaufen sein ‘across the finishing line/in-run’ is the result (–––) of the prior incremental event component.}

\[
\begin{align*}
\text{ins Ziel/ein-} & \quad \text{ins Ziel/ein-} \\
\text{lauf} & \quad \text{gelaufen (sein)} \\
\mid & \quad \mid \\
\text{t}_1 & \quad \text{E}_1 & \quad \text{t}_m & \quad \text{ResZu}_2 & \quad \text{t}_n
\end{align*}
\]

b. By contrast, the untelic simple VoM has a monophasic event structure for \(\text{lauf} “\text{run}” \) (intransitive verb and intransitive): the event structure is restricted to one single homogeneous, durative component, inherently not complex and further analyzable (>>>). The graph below represents the temporal points of reference, \(t_1\) and \(t_m\) as well as the relative point of the speech act, \(t_s\) (relation of anteriority), thus a temporal relation, that for \(\text{lauf “run}”\) and that for \(\text{gelaufen sein (/haben) ‘be (/have) run’}\). Note that one does not have to distinguish two inherent event components as in (36a), i.e. \(E_1\) and \(Zu_2\).

\[
\begin{align*}
\text{lauf} & \quad \text{gelaufen (sein/haben)} \\
\mid & \quad \mid \\
\text{t}_1 & \quad \text{t}_m & \quad \text{t}_s
\end{align*}
\]

Now heed that the thetas distribution over the two phases in (36a) is the following, (36c), as opposed to (36d) below. [IncrE=incremental event (phase), \(iV=\text{intransitive V, } ueV=\text{unergative V, } einglaufen=’in-run’, eingelaufen sein=’be an in-runner’\)].

c. AG \quad PAT

\[
\begin{align*}
\text{for } i/ueV \text{ einglaufen “run in”} : & \quad \text{einglaufen} & \quad \text{eingelaufen (sein)} \\
\mid & \quad \mid \\
\text{t}_1 & \quad \text{IncrE}_1 & \quad \text{t}_m & \quad \text{ResZu}_2 & \quad \text{t}_n
\end{align*}
\]

d. AG+PAT \quad PAT

\[
\begin{align*}
\text{for tV umbringen “kill”} : & \quad \text{umbringen} & \quad \text{umgebracht (sein)} \\
\mid & \quad \mid \\
\text{t}_1 & \quad \text{E}_1 & \quad \text{t}_m & \quad \text{ResZu}_2 & \quad \text{t}_n
\end{align*}
\]

The past/passive participle of tV in (36d) is the valence of the active \(\text{umbringen “kill”}\) reduced by the original external argument, \(eA\), which is an Agent. The fact that the
event characteristic of ResZu is no longer verbal, but adjectival, forbids an agent theta status. It has to be a Patient. In other words, the diathesis of the resulting participle changes the original valence characteristic. The very same diathetic shift holds for i/ueV in (36c). Recall that the verbal particle, ein-, is a pronominal equivalent for the telic PP ins Ziel “across the finishing line”. The fact that object predications (in German) are contingent not only on the immediate linear neighborhood of the governing full verb participle to the direct object, DO, as in (37b)) as opposed to (37a), but also on aspectual telicity of the lexical (resultative) participle corroborates our analytic account.

(37)  a. daß die Mutter Juden
    that mother jews
    versteckt hatte = [IP Mutter [VP DO Juden] versteckt hatte]]
    hidden had
    (i) ‘that mother had hidden jews’ event reading
    (ii) *‘that mother had jews hidden’ no object predicate reading

b. daß die Mutter Juden
    that mother jews
    versteckt hatte = [IP Mutter [VP hatte [DO Juden (als) versteckte]]]
    hidden had
    (i) ‘that mother had hidden jews’ event reading
    (ii) ‘that mother had aliens hidden’ also object predicate reading

c. [CP daß [TP die Mutter [VP *[DO Überirdische gesehen] hatte]]
    that mother aliens seen had
    “that mother had seen aliens“ no object predicate reading possible,
    event reading only

It is important to note that sehen “see” is not telic and thus does not qualify for object predication despite the favourable linear neighbourhood between object and adjectival participle.

Notice that our distinction of event telicity in (36a–d) covers up a complication in the following sense: We may want to distinguish the temporal boundedness that can arise even from the presence of adjuncts (external aspect) compositionally added after merge, on the one hand, and the temporal boundedness that comes from the identification of an RP in the event structure syntax (Ramchand 2008). German and Dutch are useful languages because their variable Aux selection is contingent upon the event structure of the predicate. By hypothesis, anything that affects auxiliary selection must be causing differences at the level of merge (the ‘first phase’ in Ramchand’s (2008) terms). Ramchand is probably right in claiming that telicity constructed at later stages of the derivation does not change the prior auxiliary selection. We shall argue (as Ramchand does) that this diagnostic offers useful solutions to the
questions we asked at the outset of this paper – questions that are beyond the squishy, ununifiable accounts provided on the basis of Figures 1, 3, and 4 above.

8. Aspect-based account vs. argument-based account

It is assumed as part of a Diathetic Null Hypothesis that IPass is atelic, and durative, cross-linguistically. This is what one may call the Diathetic Aspect-based Null Hypothesis. Since the argument-based account of Burzio’s Generalization – see (1)–(3) above – has nothing to say about aspectual criteria as regards this type of impersonal passivization and, much less, about the particular behaviour of VoM and their telic constituents, we shall, in what follows, weigh arguments pro and contra.

There is a class of verbal prefixes, and verbal particles, in the continental West Germanic languages which change the non-perfective meaning of a verbal simplex to a perfective one. This class is open and actively productive to the extent that it is enriched by adjectivals and those adverbials that designate states to be reached through some prior (approach) event. Among such adverbials are adverbial constituents (PPs) designating direction (without exception P+accusative in German, as in Latin and Ancient Greek). See the following groups of examples, where the first one is the one-place durative, whereas the second one represents the perfective (with the typical, obligatory diagnostic emphasis on the verbal prefix).

(38) schlafen “sleep” – ein+schlafen (verbal prefix) “fall asleep”, welken “wilt”
weg+ welken (adverbial) “away-wilt”, laufen “run” – sich müdelaufen (adjectival prefix or ‘verbal affixoid’) “to run oneself tired”, (im Garten) laufen “in-the-garden-run” (stative location) as opposed to in den garten laufen “into-the-garden-run”.

For such perfectivization by means of lexical prefixation, see the productive prefixal paradigms in Modern Russian, among which po- or na- (Gawronska 1993 as well as Smith 1997):

(39) Ona pro-stoja-l-a na uglu celyi čas (Smith 1997)
    she pro-stand-PAST-AGR on corner-P entire-ACC hour-ACC
    “She stood-perf on the corner for an entire hour”

Any of these derived predicative perfectives are covered without exception by the participial and auxiliary diagnostics of ergatives (see Burzio 1986 for Italian; Haider 1985; Abraham 1985 for German; Haider & Rindler-Schjerve 1987 for the contrast between Italian and German). Note that this is an empirical finding, which is in need of a structural account just as much as Burzio’s generalisation. But the claim, according to these exceptionless observations, is that ergatives in the languages under inspection (languages with an actively productive derivative system producing perfective verbs) are perfectives. Note, further, that this is in line, in a fashion yet to be explained, with the
equally general observation that in split ergative vernaculars, the ergative is bound to the perfect(ive) forms. See the following example from Hindi (Mahajan 1994: 6).

(40) **raam-ne** vah kitaabĕ par^p^iī hĕ (Mahajan 1994: 6)
Ram-**erg**(M) those books(F.PL) read-**perf**-F.PL be-**pres**-PL
“Ram has read those books”

These perfectivized constructions can most adequately (i.e. rendering the semantic interpretation of an accomplishment or achievement semantics in the most direct fashion) be represented as predicatives (i.e. copular predicates + prefix/adjectival/adverbal) for the object in transitive two-place constructions or the subject in one-place constructions. See the following representations in (b)–(c). Note that such perfectivizing small clause constructions derived from transitives can be paraphrased systematically in accordance with the following pattern: subject-verbal event-such THAT-[sc direct object’ sc-subject copula+ tense+ agreement XP], where XP’[adjective, verbal prefix, verbal affixoid, NP, perfect participle (‘adjective)]. Note, at the same time, that one-place ergatives such as sterben/die, representing the approach as well as the resultative phases in (33a) below, do not take a stative predicate in the small clause (object predication).

It has to remain open for the time being whether there are other one-place perfectives/ergatives that do in fact take such a small clause predicate. [t/riV = transitive or reflexive intransitive verb as in (35b) below; eV = ergative verb as in (35c)].

(41) a. Θ1 tV [Θ2 AgrP ADJ] [Maria[θ1] [iV tanzte [sc Peter[θ2] [ADJ müde]]]]
M. danced herself worn out

b. Θ1 tV [Θ2 AgrP PP_DIR]
[Mary[θ1] [iV walked [sc her daughter[θ2] [PP_DIR home]]]]
M. danced herself worn out

c. Θ1-1 riV [Θ1+ AgrP ADJ]
[Maria[i-θ1] [iV tanzte [sc sich (selbst)[i-θ2] [ADJ müde]]]]
M. danced herself worn out

d. Θ2 eV [t2 AgrP] [Maria[θ2] [iV fiel ([sc t[θ2] [ADJ zu Tode]])]]
M. fell to death

where Θ2 necessarily = TH, and AgrP is restricted to states; Θ2 with iV = reflexive pronouns with subject coreference.

e. for eV: NPj [VP [SC t j AgrP] V ]
f. for t/riV: NPj [VP [SC NPj AgrP X(P)] V ]

where NPj necessarily = TH; AgrP is restricted to states; NPj with iV = reflexive pronouns with subject coreference.

g. note the bracketing below, which is fully in line with the predicative representation in (41a–e) above (from Mahajan 1994: 6; my bracketing, WA):

[CP/IP raam-ne [SC vah kitaabĕ par^p^iī he]]
Ram-**erg**(M) those books(F.PL) read-**perf**-F.PL be-**pres**-PL
These are the corresponding features between (41a–e) above: Copular reading in the small clause (whether hidden or not) – see hĕ in Hindi; agreement not with the matrix subject, but with the object (subject in the small clause) as well as the participial adjective and the be-copula. Cf. the bracketing for French with identical structural identifications:

(42) \[ \text{[CP/IP Jean [SC/AGRO les, [tp a, [VP t, t repeintes,]]]]} \]

Where no proclitic is present as in (43) below, no such agreement features are exhibited:

(43) Jean a repeint-0 les bouteilles

See the beginning of passivization in Latin and in one of the diachronically earliest documents in the diachrony of German (see in more detail Abraham 1996). The morphology indicative of object predication is object case and number agreement on the adjectival participle (underlined) in line with the nominal object (see, in this context, Cennamo (in this volume), in particular, her concept of “extended accusative”).

(44) a. (Habet filia) arbores transplantatos Late Latin

has-3sg daughter trees-acc-m.pl transplanted-acc-m.pl

“The daughter planted a tree”

b. (Tho quad her thesa ratissas:) Old High German
dicebat autem hanc similitudinem:

then spoke he this simile

phigboum habeta sum giflanzotan (Tatian Γ 102, 2, CII 23; Sievers 1960: 146)

arborem fici habebat quidam plantatum

dfigtree-4sgM had a certain.1sgM planted-4sgM

in sinemo uuingarten

in vinea sua

have-3pl tree-acc planted-acc in vineyard

“I planted a fig tree in his vineyard”

Almost needless to say, constructions as (41a–c) resemble lexical causativity as opposed to decausatives in that Mary as external argument is the Agent in the true sense of a transitive construction, while her daughter is the Patient/TH, or Undergoer, of the walking. The ‘walking’ describes the ‘incremental’ causing subevent in the sense of (36a,b,d) above, while the process of ‘walking’ is also undergone by her daughter, but in the sense of an iV. This analysis makes the whole object predication in (41b) synonymous with Mary accompanies her daughter walking until home. And her daughter being home is the result of the object predication, with her daughter is the Resultee in the sense that she comes to hold the result state of being ‘home’. Clearly, causatives are argument raisers with any Agent in the embedded ‘decausative’ construction retaining its theta status. As soon as the ResultP is present in the merge
phase, the Resultee specifier position becomes available and the extra argument is possible but only with Resultee semantics (Ramchand 2008: 59). Result phrases always licenses an extra argument, all in line with (36a,b,d) and (41a,b,c) above as well as the nature of what sort of real world relation it will bear to the walking event in (41b).\textsuperscript{16} See (45) for this.

\begin{enumerate}
\item[(45)] \textbf{Cause-Process-Result} for \textit{x walk y to-DP} in (41b) (in line with Ramchand 2008: 60); also pertaining to (41a,c,d)
\end{enumerate}

\begin{center}
\begin{tikzpicture}
\node (VP) at (0,0) {\textit{v} \textit{walk}};
\node (vP) at (-1.5,1.5) {\textit{vP}};
\node (INITIATOR) at (-2.5,3) {\textit{INITIATOR}};
\node (UNDERGOER) at (-1.5,3) {\textit{UNDERGOER}};
\node (RESULTEE) at (-0.5,3) {\textit{RESULTEE}};
\node (x) at (-2,2) {\textit{x}};
\node (y) at (0,2) {\textit{y}};
\node (RP) at (-2,1) {\textit{RP}};
\node (V) at (0,1) {\textit{V}};
\node (R) at (0.5,2) {\textit{R}};
\node (PP) at (1,1) {\textit{PP}};
\end{tikzpicture}
\end{center}

We repeat: Simple goal-incorporating VoMs merge in the early merge phase, while directional VoM-constituents receive their telic status only until the goal-PP is inserted – that is, at a later stage above vP. There also seem to be differences between languages as to the nature of prepositional and R(esult) licensing heads available.

\section{The unified IPass criterion}

We have already suggested that IPass is made possible by lexical auxiliary distinctions such as between German/Dutch \textit{werden/worden} and \textit{sein/zijn} or Scandinavian \textit{bli} and \textit{vaera}. On the other hand, the fact that Icelandic has a distinct IPass despite the fact that the Mainland Scandinavian lexical distinction is absent in Icelandic, has

\textsuperscript{16} This appears to be in line with Legendre’s finding (Legendre 2007) that, while the fundamental syntactic distinction made by the Unaccusative Hypothesis is confirmed by the formal analysis, the English present perfect (with its anterior participle) imposes restrictions of its own on the argumental status of any verb entering the construction, on a par with constructions such as passives and resultatives. See the detailed discussion of the past/anterior and passive uses of the participle in Abraham (2006a, b).
prompted the assumption that alternative passive construals may achieve the same. The gist is thus to have a means to overtly distinguish ongoingsness and stativity. The Scandinavian languages have chosen the synthetic s-passive to encode ongoingsness (or even, as in the case of Danish and, somewhat less reliably in Norwegian, genericity) as opposed to the analytic form with be. This is the specific code that Icelandic has opted for. Other Scandinavian languages that possess the auxiliary bli as opposed to væra have a lexical means and do not have to resort to the Icelandic solution.

Now see Polish (Frajzyngier 1982; see also Comrie 1977), which surprisingly has IPass despite its lack of auxiliary distinctions.17

(46) Zapukano do drzwi
knocked.\text{pret} at door
“There was knocking at the door” = “Es ist geklopft worden”

(47) Zaszczekano do drzwi
barked.\text{pret,perf} at door
“There was barking at the door” = “Es ist gebellt worden”

(48) Szczekano do drzwi
barked.\text{pret,imperf} at door
“There was barking going on at the door” = “Es wurde gebellt”

Such Auxless IPass construals did in fact have the \text{BE}-Aux być originally with the participle in the function of a predicate noun. Today’s IPass is a relic of the old passive participle ending in -t, -n and -to, -n and being stuck to the old nominal agreement declension. By and by, the Aux być was deleted as already in the 15th century on a regular basis. Polish IPass on -no, -to sounds formal and is used more or less exclusively in the written and in the press media.

However, there is also another way to express IPass in Polish: the reflexive passive (like (46)–(48) taken as an active construction) with sie ‘oneself’ formed by

---

17. Two reviewers point out that synchronically (46)–(48) are not impersonal passives, but rather active impersonal constructions. They are widely used in written registers of all types, and they are also used in speech. Diachronically, though, they are clearly construed with passive participles of the -no/-to type, which is reflected in Germanic by the strong -en- and the weak -(e)t type of anterior and passive participle. Note that, although Polish distinguishes być ‘be’, bywać ‘be (iterative)’, zostać ‘become (perfective)’, zostawać ‘become (iterative)’ – all of which are in use even though they differ in frequency (Górski 2008:48–49) – their systematic use for the analytic preterit is questionable. To the extent that there appears to be evidence that a new Polish Perfect, with the auxiliary mieć ‘have’, is emerging from the resultative, I believe there is reason to doubt the fact that mieć ‘have’, has reached the auxiliary state of German haben “have” in its employment in the analytic preterit.
imperfective verbs in the majority of cases. This is good colloquial Polish and is used frequently.

(49) *Obiad* go tuje sie
   lunch cooks itself
   “There is lunch-cooking going on” = “Es wird mittagessengekocht”
   *(obiad as much as lunch and mittagessen are V-incorporated objects)*

Icelandic, on the other hand, has IPass for different reasons. Recall (24a–c) above. Icelandic has the middle option next to the analytic passive (from Thráinsson 2007:284).

(50) a. *Dyrnar voru opnðar* (*af dyraverði*)
    doors-the.n.pl.f were opened-n.pl.f (by doorman)
    “The doors were opened by the doorman”

b. *Dyrnrar opnuðust* (*af dyraverði*)
    doors-the.n.pl.f opened-pl (*by doorman)
    “The doors opened”

(50a) is a syntactic derivation as opposed to the lexical one in (50b), a conclusion which is supported by the deleted (not only suppressible) agent, *af dyraverði*, in (50b). The verbal st-middles in Icelandic – all of which have passive readings next to their reflexive decausative interpretations – closely resemble German middle constructions (Abraham 1994) as modalized property predications (‘the door being openable/can be opened (easily)’).

In fact, case assignment in passives often makes very clear the true nature of Icelandic passives as IPass. The following examples are from Thráinsson (2007:154).

(51) a. *Þessu rúmi* hefur *(aldrei)* verið sofð í
    this.dat bed has (never) been slept in

b. *Báð* hefur *(aldrei)* verið sofð í *Þessu rúmi*
    there has (never) been slept in this bed.dat

c. *Í* *Þessu rúmi* hefur *(aldrei)* verið sofð
    in this bed.dat has never been slept

d. *Aldrei hefur* *(aldrei)* verið sofð í *Þessu rúmi*
    never has been slept in this bed.dat

The word order characteristic is like that of German: clearly V2 with each of the nominal clause members topicalizable irrespective of whether of subject or other clausal status. Since dative *Þessu rúmi* does not have subject properties, any of (51 a–d) has the configuration of IPass.

This throws light on the so-called dative-subject constructions in recent developments of Modern Icelandic. As claimed by Árnadóttir & Sigurðsson (2008:9f.), there
are similarities between the development of dat-nom verbs in Icelandic and passives of ditransitive verbs:

(52) a. Mér var gefinn bíll
    me.DAT was given.M.SG car.M.nom.sg
    ‘mir war gegeben ein Wagen’ = “I was given a car”
    no V-obj. gender agreement

b. Mér var gefðí bíll
    me. DAT was given.N.SG car.M.nom.sg.nom.-obj.

Finite verb and the participle gefinn agree with the nominative object. Then the Nominative is structural Case (cf. the line of argument in Jónsson 1996). But then there are examples similar to the ones in (52a), but without agreement, as (52b), which is equivalent in meaning with (52a).

What (52b) looks like is a re-interpretation of dat-nom verbs (such as Mér likar bilarnir ‘me. DAT like.3 SG cars-the.nom.pl’).

(53) a. Mér lika bílar
    ‘mir gefallen Autos’ = “I like cars”
    agreement nom.-object

b. Mér likar bílar
    no agreement nom.-object

c. Mér lika bílarna
    cars.accusative

In sum, Árnadóttir & Sigurðsson’s (2008) proposal is that the nominative case of an object in dat-nom verbs in Icelandic is not structural. They argue that it is an inherent nominative case, but has underlying accusative Case. The indication of this is the lack of agreement with the nominative object and the change of the nominative into accusative for some speakers of Modern Icelandic. Furthermore, the authors’ discussion goes about how changes in the passive of ditransitive verbs resemble the changes in dat-nom verbs. The passive construction, like the dat-nom construction, is void of agreement, and it shows a change from nominative object to accusative object. While the authors do not claim that the new passive in Icelandic emerged via – or in close analogy with – ditransitive verbs (with the lack of agreement as an important step), the fact that this sheds some light on the dat-nom construction may help to understand the nature of the new Icelandic passive. Not only is this solution different from Barðdal & Molnár’s (2003) and would seem to relieve the latter authors from assuming that passives in Icelandic cannot be atomized into smaller structural pieces and

---

18. There is a decades-long literature on the similarities (and differences) in agreement between dat-nom verbs in Icelandic and passives of ditransitives; see Árnadóttir & Sigurðsson (2008) for a more thorough list of work that should be referred to in this context.

19. See Maling 2006 for a somewhat different picture.
be derived accordingly. The insight beyond that, in the present context, is that IPass is a ubiquitous construction even in languages that have no such distinction between different auxiliaries as in German and Dutch (notably werden/worden and sein/zijn, which serve to keep apart ongoing and resultative passivization).

10. Results of our discussion

This is what we pointed out to be the problem in the inceptive part of the present article: Motion verbs (VoM) are exceptions to the definition of unaccusativity. The problem arises under impersonal passivization (IPass): VoM may, on the one hand, be subjected to impersonal passivization, but, on the other hand, they turn out as ergative verbs/unaccusatives in directional use. The clash consists in the fact that IPass is always imperfective, whereas unaccusatives are perfectives. The solution is sought in the assumption that VoM are split unaccusatives: unergatives in the present, ergatives in the preterit participle.

These are our solutions to the problems. Impersonal passives are imperfectives. But they are not quite. For sure, this is why unaccusative verbs cannot undergo impersonal passivization. And, given its zero valence status, IPass is thetic discourse-functionally. Their valence status is zero since they are derived from intransitives or from transitives where nominal objects have remained derivationally in situ and in possession of their original Case government (accusative DO). Given the typical intransitive status of verbs of motion and their verbal complexes with locative and directional prepositional phrases IPass of VoM appeared to be a natural goal for discussion. Yet, directional constituents of VoM put forward a special aporia in that they have retained their original agentive status, while simultaneously obeying distributional properties of unaccusatives. Unaccusatives cannot be passivized by definition, let alone undergo IPass, simply because their internal argument cannot be agent and their derivational diathetic status is that of a passive in the first place. We have found the solution in the fact that VoM directional phrases only are ergative/unaccusative in the resultative status (i.e. not in any of the present tense forms). One can speak of split ergativity with constituent VoM much in line with typologically split ergativity in languages like Hindi, Urdu, Balochi and a number of other languages (Abraham 2000). This is why directional VoM constituents allow for I-passivization: They are hybrids between imperfectives (in accordance with their simple VoM component), and they are telics and unaccusatives given their motion direction expressed in the PP component (with accusative P-government). We have considered scales of verb diatheses and their subdividing criteria in Figures 1, 3, and 4 and found that none of them explains in an empirically satisfactory way which auxiliary can be used for which type of verbal event or eventual characteristic. What, much rather, counts is
the fact that there is only one option to involve an Aux: i.e. BE subcategorizing for the resultative (since categorially adjectival) participle.\textsuperscript{20} This, then, is a categorically decidable fact outside any squishy or scalable list of properties. However, in view of Burzio’s Case assignment and clause member Generalization for the assessment of IPassivizability and unaccusativity we came to the conclusion that aspect (specifically, telicity as opposed to imperfectivity) plays the ground-shaping role. Finally, we have argued, somewhat circuitously, that recent developments in oblique subject constructions in Modern Icelandic support the present analysis of IPass to the extent that, given their distributional properties, such “oblique” passives may be considered as normal (im)personal passivizations.

References


\textsuperscript{20} Heed the observation made in Note 13 above, however.


Verbs of motion


On the distribution of subject properties in formulaic presentational constructions of Germanic and Romance

A diachronic-typological approach

Volker Gast & Florian Haas
Friedrich-Schiller-Universität Jena

The present study deals with the distribution of subject properties in 'formulaic' presentational constructions such as Engl. there is NP, Fr. il y'a NP and Span. hay NP, approaching the problem from a diachronic-typological perspective. Nine major types of presentational constructions are distinguished by cross-classifying the three-valued parameters 'type of existential predicate' and 'type of expletive'. Moreover, a language-level parameter is introduced which distinguishes languages allowing verb-initial order in thetic sentences ('thetic-V1 languages') from languages disallowing such an order ('thetic-XV languages'). It is shown that thetic-XV languages tend to use expletives in their existential formulas, which attract subject properties and qualify as impersonal. By contrast, thetic-V1 languages often do not use expletives at all, and if they do, these do not attract subject properties. The corresponding constructions are consequently not impersonal. Accordingly, a correlation can be established between the parameters 'thetic-XV' vs. 'thetic-V1', on the one hand, and 'impersonal presentational' vs. 'personal presentational', on the other.

Keywords: presentative; presentational; expletive; existential; thetic

1. Introduction

This study deals with the formal means used for the introduction of new discourse referents in selected Germanic and Romance languages. Speech events in which a

---

1. This study reports on results obtained in a research project entitled 'English-German contrasts – A comprehensive survey of major differences between English and German', granted to E. König and V. Gast by the German Science Foundation. The financial support from this institution is gratefully acknowledged. We are indebted to Ruth Berman, Andrej Malchukov.
speaker “call[s] the attention of an addressee to the hitherto unnoticed presence of some person or thing in the speech setting” (Lambrecht 1994: 39, 177) will be called ‘presen-
tative utterances’, or simply ‘presentatives’. Structural configurations conventionally used to encode presentative utterances will be called ‘presentational constructions’, or simply ‘presentationals’. The concept of a ‘presentative (utterance)’ will function as a tertium comparationis of the study, and the ‘presentational constructions’ of selected Germanic and Romance languages will constitute the objects of comparison.

It is well known that presentationals are often characterized by a non-canonical distribution of subject properties over the (pro)nominal constituents of a sentence (e.g. Seefranz-Montag 1983; Lambrecht 1986, 2000; Lazard 1994; Koch 2003; Börjars & Vincent 2005, among many others). The ‘novel NP’ – the NP introducing a new discourse referent (Milsark’s 1974 ‘pivot’) – exhibits properties of both subjects and objects in many languages. Moreover, many presentational constructions contain an ‘expletive subject’, i.e. a (mostly pronominal) element which does not play an obvious role in the argument structure or interpretation of a sentence while still exhibiting some (or all) subject properties (see for instance Askedal 1986; Lazard 1994; Vikner 1995; Lødrup 1999). Such presentationals belong to the class of ‘impersonal constructions,’ understood as constructions lacking a referential subject (cf. Siewierska & Mal-
chukov this volume). In terms of Malchukov & Ogawa (this volume), they qualify as ‘T-impersonals’. Given that there are also presentationals with a referential subject, a distinction can be made between ‘personal’ and ‘impersonal’ presentationals. As will be seen, this distinction provides a useful parameter of variation in the typology of presentationals, as it correlates with an important structural property of the languages under investigation, i.e. the (im)possibility to use verb-initial order in thetic sentences.

The main objective of our investigation is to formulate generalizations concerning the distribution of subject properties over the two candidates for subject status in presentational constructions, i.e. the ‘novel (postverbal) NP’ and the expletive. We address this question from a diachronic-typological point of view, i.e. by devising a typology of presentationals which is historically motivated. Two major developments can be observed: (a) a shift of subject properties from the novel NP to the expletive, and (b) the acquisition of subject properties by a (formerly non-subject-like) postverbal NP. The former type of development, which leads to the formation of impersonal

and Anna Siewierska for valuable comments on an earlier draft of this paper. Moreover, we wish to thank Heide Wegener and Denis Creissels as well as all the participants of the workshop on ‘Impersonal constructions’ (Forlí, Sept. 2008) for helpful comments and discussion. Any remaining inaccuracies are our own.


3. The term ‘construction’ is used according to Croft (2001).
presentational, is found in those languages that do not allow verb-initial order in thetic sentences (the so-called ‘thetic-XV languages’, i.e. the Germanic languages under consideration and French). The second type of development, which results in personal presentational, can be observed in those languages which do allow verb-initial order in thetic sentences (Romance languages except French, subsumed under the term ‘thetic-V1 languages’). At a general level, our conclusion is that the presentational of thetic-XV languages show a tendency to becoming ‘more impersonal’, while the presentational of thetic-V1 languages tend to change towards ‘more personal’ constructions. This broad generalization emerges from a more fine-grained typology of presentational constructions which makes use of the parameters ‘type of existential predicate’ and ‘type of expletive’.

The discussion starts in §2 by delimiting the object of study, i.e. presentational. Presentational constructions of German, Spanish and German are illustrated and two major types of presentational are distinguished (‘formulaic’ and ‘non-formulaic’ ones). In §3, the morphosyntactic properties of the novel NP in presentational are surveyed, and an explanation of their ‘hybrid’ nature provided by Lambrech (2000) is discussed, which serves as a blueprint for our own analysis. In §4, a diachronic typology of formulaic presentational is presented. Presentational are classified along two parameters, i.e. (i) the type of predicate and (ii) the type of expletive used. Moreover, the aforementioned classification of Germanic and Romance languages into thetic-V1 languages and thetic-XV languages is introduced. §5 discusses the presentational of selected Germanic and Romance languages against the background of the typology presented in §4 with a focus on matters of argument structure. The results of this study are summarized in §6.

2. What is a presentational?

2.1 Presentational in English, Spanish and German

At least two constructions can be regarded as being more or less conventionally used to encode presentative utterances in English: (i) sentences introduced by the ‘existential formula’ there is/are or there’s, and (ii) the syntactic rearrangement rule of ‘main verb inversion’, often alternatively called ‘(locative) inversion’ (e.g. Levin 1993; Birner & Ward 1998; Ward & Birner 2004). The two constructions are illustrated in (1). The ‘novel NP’ – the NP introducing the new participant – is indicated by a subscript ‘NOV’.

\begin{align*}
\text{(1) a. } & \quad \text{There is } [\text{a man in your kingdom who has the spirit of the holy gods in him…}]_{\text{NOV}} \quad \text{(there-existential)} \\
\text{b. } & \quad \text{Until the end of the war so very few folk had beards, and then only short ones nicely trimmed, but into the room came } [\text{a young man with a black fuzz of over eight inches}]_{\text{NOV}} \quad \text{(BNC) (locative inversion)}
\end{align*}
While both structures illustrated in (1) occur in presentative utterances, they have strikingly different distributions. The ‘existential formula’ there is/are is often used as a ‘generic’ existential predicate, e.g. in There are more than 1,200 species of bamboo in the world, understood as an answer to the question How many species of bamboo are there in the world? In this case, the sentence is not used to “call the attention of an addressee to the hitherto unnoticed presence of some person or thing in the speech setting” (Lambrecht 1994: 39), but to answer a question about some previously established topic.

Locative inversion as illustrated in (1b) above covers a more restricted range of contexts than there-existentials (cf. Levin 1993; Bresnan 1994; Birner & Ward 1998; Ward & Birner 2004). In this construction, a directional complement (typically a PP like into the room) occupies the preverbal position, while the novel NP – the prime candidate for subject status – occurs in a postverbal position. One of the most prominent semantic or pragmatic conditions on locative inversion is that “the referent of the inverted subject is introduced or reintroduced on the (part of the) scene referred to by the preposed locative” (Bresnan 1994: 85). Even though this condition may not apply in all cases (cf. e.g. Birner 1994), it is certainly a prominent factor in the textual usage of this construction.4

Presentationals similar to the English structures illustrated in (1) can be found in all major Germanic and Romance languages, but the constructions in question differ considerably in their structural and distributional properties. For example, Spanish has an ‘existential formula’ hay which does not require an expletive corresponding to English there (but note that hay [aj] contains a phonological remnant of a former expletive, i.e. the final glide [j] < Lat ibi ‘there’). The use of hay in presentational utterances is illustrated in (2) with the past tense form había.

(2)  Standard Iberian Spanish

<table>
<thead>
<tr>
<th>Standard Iberian Spanish</th>
</tr>
</thead>
<tbody>
<tr>
<td>Había [muchos estudiantes que trabajaban en bares]_{NOV}.</td>
</tr>
<tr>
<td>‘There were many students who worked in bars.’</td>
</tr>
</tbody>
</table>

Spanish also has a construction that is comparable to locative inversion in English; it can also use VS-order with other types of predicates. Unlike in English, where the preverbal position is invariably filled, either with a locative expletive (there) or with another preposed constituent (e.g. a prepositional phrase, cf. (1b)), no preverbal element is required in Spanish (cf. (3)).

4. B. Birner and G. Ward have repeatedly pointed out that the main condition on inversion is that the preverbal NP must be more ‘discourse familiar’ than the postverbal one.
Prima facie English locative inversion and Spanish VS-order are rather similar in both functional and structural terms. However, they differ considerably with respect to their distribution. While locative inversion in English is (more or less) restricted to the function of a presentative utterance, VS-order in Spanish is also commonly used in other, ‘non-presentative’ utterances with ‘sentence focus’ (Lambrecht 1986, 1994, 2000), i.e. in thetic sentences (cf. Sasse 1987). For example, it is used in contexts like (4), where locative inversion would be impossible in English (cf. (5)).

Yet another type of presentational is found in verb-second languages like German. On the face of it, (6) seems to instantiate a structure completely parallel to locative inversion in English.

However, the German sentence in (6) simply represents the canonical verb-second order that this language commonly uses in main clauses. From a purely structural point of view, there is nothing special about (6). German also has an existential formula similar to the English and Spanish ones pointed out above, i.e. es gibt ‘it gives’ or, alternatively (and more commonly used in Southern varieties), es hat (‘it has’). It is illustrated in (7).5

What the examples given in (1)–(7) show is that the function of a presentative utterance may be encoded with different degrees of specificity. The ‘existential formulas’ there is, hay and es gibt are often, but not necessarily, used with the function of a pre-

5. There is also a formulaic presentational es war einmal ‘there once was’, which is, however, only used in fairy tales, like Engl. once upon a time (cf. Berman & Slobin 1994:74–75).
sentative utterance. The three syntactic configurations illustrated in (1b), (3) and (6) also differ in their distribution. The German example (6) simply uses the ‘generic’ verb-second order that is used in all types of main clauses, and that is not associated with any specific type of information structure. Verb-initial (V1) order in Spanish as illustrated in (3) has a more specific function, covering as it does a class of contexts which can broadly be subsumed under the notion of ‘thetic judgement’. Finally, English locative inversion (cf. (1b)) is distributionally heavily restricted, with presentative utterances providing the most important context of use.

As has been seen, we regard the category of ‘presentative utterance’ as a special case of the more inclusive class of thetic judgements. It is important to bear in mind that ‘thetic’ and ‘presentative/presentational’ do not mean the same thing, even though Lambrecht (2000: 623) remarks that “the overriding function of the SF [sentence-focus/thetic] category is presentational”. To what extent this is true obviously depends on the exact definition of these terms. We will use ‘thetic’ in the traditional sense (cf. Sasse 1987), and ‘presentative/presentational’ only for those cases where a new discourse referent (in the sense of Karttunen 1976) is introduced.

2.2 Formulaic vs. non-formulaic presentational

In order to introduce a new discourse referent, a presentational sentence must necessarily contain a predicate either expressing or entailing existence. As was pointed out in §2.1, existence is often predicated by ‘(explicit) existential formulas’, e.g. Engl. *there is*, Span. *hay* and Germ. *es gibt/hat*. These formulas are highly conventionalized and largely bleached semantically, thus coming close to being natural language correlates of an existential quantifier, commonly represented as ‘∃’ in predicate calculus. Presentational containing an explicit existential formula will be called ‘formulaic presentational’. This type was illustrated in (1a), (2) and (7) above. Two additional examples from French and Italian are given in (8) and (9), respectively (‘ex’ stands for ‘existential formula’).

(8) French

\[
\text{there is a man nude in my garden} \\
\text{‘There is a nude man in my garden.’}
\]

(9) Italian

\[
\text{there is a man nude in the my garden} \\
\text{‘There is a nude man in my garden.’}
\]

Presentational without an explicit existential formula like the ones illustrated in (1b), (3) and (6) will simply be called ‘non-formulaic presentational’, as there is no
‘segmental constant’ in the relevant constructions, which are characterized only by specific word order configurations. In the remainder of this study, we will concentrate on formulaic presentational. This is not to say that non-formulaic presentational are uninteresting, or less relevant to the questions under investigation. However, as a source of insights into diachronic processes, formulaic presentational have the advantage of providing relatively transparent evidence concerning the changes in argument structure accompanying the process of conventionalization. For example, the Spanish formula \textit{hay} \text{NP}_{\text{NOV}} originated as a transitive construction in which the novel NP had the function of an object (< Lat. \textit{habet ibi} \text{NP}_{\text{ACC}}, cf. §5.3.1). In specific varieties of Spanish (e.g. Mexican Spanish, cf. §3.3), this NP now has the status of a subject. Such processes of change can be more easily identified when the predicate of the construction is kept constant. Moreover, existential predicates are arguably the most typical predicate used in presentative utterances and, hence, provide a reasonable point of reference for follow-up studies taking a broader view on the matter.

3. Non-canonical subjects in presentational

Having delimited the object of study, we can now turn to a discussion of the main issue dealt with in this study, i.e. the distribution of subject properties over the nominal constituents of a presentational construction. §3.1 provides a brief overview of the properties generally assumed to be typical of (non-)subjecthood. In §3.2, the ‘hybridity’ of postverbal NPs (with respect to their status as a subject or non-subject) is related to a conflict between form and function which emerges in the presentational of specific Germanic and Romance languages. §3.3 critically reviews an influential explanatory approach to the problem of ‘non-canonical subjects’ in presentational, i.e. the one advocated by Lambrecht (1986, 2000). Lambrecht’s hypotheses – which we summarize under the label ‘global repulsion hypothesis’ – will serve as a background for our own analysis presented in §4 and 5.

3.1 On subjects with object properties

Lambrecht (1986, 2000) has argued that the novel NP in presentational is basically a subject that is stripped of (some or all of) its subject properties. Consequently, it has a status somewhere between subjecthood and objecthood. Lazard (1994) has coined the term ‘actant H’ for such ‘hybrid NPs’, which is an entirely arbitrary label meant to prevent any type of \textit{a priori} assumptions about their status as a subject or object. In a

6. Note that Lambrecht (2000) does not restrict this claim to formulaic presentational but also takes non-formulaic ones into consideration.
nutshell, ‘actant H’ stands for a single argument of a predicate which occupies a post-
nominal position.

Lambrecht (2000: 625) identifies the properties listed in (10) as being typically
associated with (focal) objects. Subjects are characterized by the complementary set
of attributes.

(10) Object properties according to Lambrecht (2000)
   (i) prosodic prominence,
   (ii) specific linear position relative to the verb,
   (iii) cooccurrence with ‘focus particles’,
   (iv) absence of grammatical agreement with the verb,
   (v) non-nominative case marking,
   (vi) single constituent status of the verb-object sequence,
   (vii) constraints on null anaphora.

According to Lambrecht (2000), the sole argument in a ‘sentence focus’ construction
(‘thetic sentence’, in our terminology) tends to exhibit some or even all of the proper-
ties listed in (i)–(vii) above, while functionally being more similar to a subject than
to an object. Lambrecht’s criteria can roughly be grouped into three major categories:
(a) discourse-related properties (criterion [i]), (b) properties concerning constituency
and locality (criteria [ii], [iii], [vi] and [vii]), and (c) morphosyntactic properties (cri-
teria [iv] and [v]). In the following, we will concentrate on morphosyntactic properties,
as they are most relevant to the study of argument structure. We do not regard matters of
word order or constituency as very good indicators of subject or object status. In fact, our
typology will make use of a parameter that groups languages into those that have a post-
verbal subject position (thetic-V1 languages) and those that do not have such a position
(thetic-XV languages). Moreover, we will use one type of locality restriction as a criterion
of subjecthood, i.e. S-X-raising (subject-to-subject or subject-to-object/exceptional case
marking). Our diagnostics indicative of a subject status are listed in Table 1.

Table 1. Criteria for subjecthood and objecthood used in this study

<table>
<thead>
<tr>
<th></th>
<th>subject</th>
<th>non-subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>agreement with verb</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>nominative case marking</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>S-X-raising</td>
<td>yes</td>
<td>no</td>
</tr>
</tbody>
</table>

We will briefly illustrate the three criteria using presentational constructions exhibit-
ing a ‘non-canonical’ distribution of subject properties. There’s NP in spoken English is
often used as an example of a formulaic presentational suspending agreement between
On the distribution of subject properties in formulaic presentational

the verb and the postverbal NP (cf. Lazard 1994; Lambrecht 2000). Examples like (11a) are commonly used in specific registers of English, as an alternative to (standard English) (11b) (cf. Rupp 2005 for details; the register restriction is here indicated by ‘%’). A similar phenomenon can be observed in (specific varieties of) Italian. In (12a), the novel NP dei contadini ‘farmers’ does not agree in number with the copula era (‘was. sg’), which forms part of the existential formula c’era. The corresponding sentence in the standard language, where the existential formula contains a plural form of the copula, is given in (12b).

(11) English
    a. %There’s three students waiting outside.
    b. There are three students waiting outside.

(12) Italian
    a. %C’-era dei contadini.
       there-was INDEF.PL farmers
    b. C’-erano dei contadini.
       there-were INDEF.PL farmers
    ‘There were farmers’
     (Koch 2003: 158)

While in the b-sentences the novel NPs qualify as subjects, controlling as they do verbal agreement, in the a-sentences they have lost this property and are therefore ‘less subject-like’.

The two constructions illustrated in (11) and (12) differ with respect to their behaviour in S-S-raising, e.g. with epistemic predicates like ‘seem’. While those varieties of English that allow (11) (mostly) also allow (13) – where the raising verb does not agree with the postverbal/novel NP – the Italian example (14) is ungrammatical with singular agreement, and a plural form of the raising verb sembrare is required. We regard this as evidence that ci – unlike Engl. there – does not undergo raising, and that Engl. there is thus ‘more subject-like’ than Italian ci (in fact, we will argue in §5.2.2 that Italian ci is not a subject at all).

(13) English
    %There seems to be some students waiting outside.

(14) Italian
    Ci sembrano/*sembra essere molti studenti.
    there seem.pl/seem.sg to.be many students
    ‘There seem to be many students.’
     (Federica da Milano, p.c.)

Non-nominative case marking on the novel NP can be observed in German. The novel NP in (15) (großen Ärger ‘big trouble’) is in the accusative case, as is witnessed...
by the *n*-suffix on the attributive adjective (though, in this particular case, not on the noun itself).

\[(15)\] German

\[
\text{Es gab großen Ärger.}
\]

\[
\text{it gave big.\text{ACC} trouble}
\]

‘There was great trouble.’

Note that postverbal NPs of the type of *großen Ärger* in (15) do not exhibit any subject property at all; they are clearly objects. We will return to this construction in §5.3.2. For the time being, suffice it to say that novel NPs in presentational constructions of Germanic and Romance languages may display subject and non-subject properties to varying degrees.

### 3.2 A conflict between form and function

The heterogeneous distribution of subject properties in presentationals can be regarded as a reflex of a conflict between form and function. Presentative utterances present an interesting problem for the mapping from information structure to syntax, insofar as they impose ‘conflicting demands’ on the structure of a sentence. At least three important ‘functional requirements’ of a presentative utterance can be identified (cf. (16)).

\[(16)\] Functional requirements of a presentative utterance

(i) the ‘novel NP’ must be stressed,

(ii) it should (ideally) be the only argument of the sentence, and

(iii) it should (ideally) come late in the sentence.

Point (i) can be regarded as a universal principle of discourse organization and is related to the fact that the processing of new information requires more attention on the part of the hearer than the processing of given information. Point (ii) is a ‘desideratum’ of discourse organization which is closely related to point (i): The fewer arguments a clause contains, the more attention can be paid to each one of them. Point (iii) can be motivated in terms of several principles of discourse organization. For example, it has been observed that the later a new referent is introduced in the clause, the closer it will be to its next mention in the following sentence, where it is typically picked up in the form of a pronominal subject (cf. Hetzron’s 1975 notion of ‘presentative movement’). Moreover, the NP introducing a novel referent typically comes with additional (postnominal) specifications and is thus rather ‘heavy’. In accordance with the ‘principle of end weight’ (cf. Behaghel 1909; Hawkins 1994) it is thus expected to take up a right-marginal position in the sentence. Point (iii) is also obviously related to point (i), as stress, in itself, implies a certain weight and, hence, a preferred positioning at the right margin of a sentence.
Quite obviously, two of these ‘functional requirements’ – those under (ii) and (iii) – conflict with the syntactic ‘default rules’ of some Germanic and Romance languages, as the ‘ideal’ presentational is expected to exhibit VS-order, and to contain no argument other than V and S (often there are, of course, locative and temporal adjuncts). While most Romance languages allow VS-sentences (cf. §4), Germanic languages as well as French require the presence of some element in the slot preceding V. We will call this slot an ‘obligatory preverbal slot’ (which is actually an abbreviation for ‘obligatorily filled preverbal slot’). In verb-second (V2) languages (West Germanic except English), the obligatory preverbal slot is typically taken by topics (and sometimes foci), while in the SVO-languages English and French it is taken by subjects. As will be seen, the presence or absence of an obligatory preverbal slot has important consequences for the ‘dynamics of subjecthood’, as only expletives in obligatory preverbal slots appear to attract subject properties, whereas in languages without an obligatory preverbal slot the subject properties tend to go to the postverbal NP. Accordingly, it is only in languages with an obligatory preverbal slot (thetic-XV languages) that impersonal presentational emerge.

3.3 Lambrecht’s ‘global repulsion hypothesis’

The ‘hybrid’ nature of novel NPs has been explained by Lambrecht (1986, 2000) in terms of two principles which are based on the assumption of comparison of, or competition between, alternative syntactic and prosodic structures made available by a given grammar. Lambrecht argues that the make-up of subject-focus sentences – which he regards as being largely specialized to the function of a presentative utterance, cf. the quotation from Lambrecht (2000: 623) given above – is crucially determined by the need to be distinct from other, more ‘canonical’ structures, in particular predicate focus structures (i.e. categorical statements). He formulates the two principles in (17) and (18).

(17) The Principle of Detopicalization

SF [sentence focus] marking involves cancellation of those prosodic and/or morphosyntactic subject properties which are associated with the role of subjects as topic expressions in PF [predicate focus] sentence.  

(Lambrecht 2000: 624)

(18) The Principle of Subject-Object Neutralization

In a SF [sentence focus] construction, the subject tends to be grammatically coded with some or all of the prosodic and/or morphosyntactic features associated with the focal object in the corresponding PF [predicate focus] construction.  

(Lambrecht 2000: 626)

The principle in (17) says that sentence-focus marking crucially involves the absence (or loss) of features characterizing predicate-focus marking, especially the features
typical of a subject (e.g. those of being unaccented and pronominal). Accordingly, the subject – which is still required by the syntax of the relevant languages – will shed typical subject properties and become more object-like. The distinction between subject and object is therefore blurred (cf. (18)).

We refer to Lambrecht's hypothesis as the 'global repulsion theory' because it assumes that subject properties are repelled by the postverbal NP (which presupposes that this NP must have had subject properties at some point), and because the principle is formulated in such a way that it is expected to apply universally, i.e. independently of the specific language and construction at hand.

Lambrecht's account is basically a diachronic one. He uses the Saussurean notion of 'motivation' to account for the instantiation of his principles in natural languages. With respect to the 'Principle of Detopicalization' (cf. (17)), he points out:

It goes without saying that the dynamic terms 'detopicalization' and 'cancellation' [...] are not meant to suggest a synchronic derivational relationship between a PF [predicate-focus] and a corresponding SF [sentence-focus] construction. The relevant explanatory concept is the Saussurean notion of motivation [...] What I am trying to capture here are natural paths of grammaticalization, not rules of sentence formation. Accordingly, the principles I am postulating have limited predictive power, since motivations can compete with each other [...].

(Lambrecht 2000: 624–5)

Lambrecht substantiates his argument with data from Germanic and Romance languages, but he also refers to other languages like Kinyarwanda and Russian. One example of a simultaneous loss of subject properties and the acquisition of object properties is provided by there-existentials in English. As pointed out in §3.1, in the spoken language the postverbal constituent – originally a subject – often does not agree in number with the copula. A similar lack of agreement can be observed in parallel examples from French (cf. (19)) and Italian (cf. (12a) above, repeated here as (20)). According to Lambrecht, these examples are instances of subjects becoming more object-like.

(19) French

\[Il \text{ est}/**sont \text{ venu drei femmes.}\]

\[it \text{ is}/**are \text{ come three women}\]

'There came three women.'

(Lambrecht 2000: 643)

(20) Italian (spoken)

\[%C-\text{era dei contadini.}\]

\[there-was \text{ indef.pl farmers}\]

'There were farmers.'

(Koch 2003: 158)

While Lambrecht's 'global repulsion hypothesis' seems to be compatible with a large amount of data, it also faces a serious problem: There are languages in which original
(morphosyntactic) objects have acquired subject properties, rather than the other way around (cf. also Koch 2003). For example, in Mexican Spanish the existential predicate *hay* normally agrees with the postverbal constituent, unlike in (the more conservative construction of) Standard Iberian Spanish (cf. (21) vs. (22)). The same phenomenon can be observed in many other Romance languages, e.g. in (spoken) Iberian Portuguese, in Catalan and in some dialects of Italian (cf. Koch 2003 and below).

(21) Standard Mexican Spanish

\[ \text{Habían muchos problemas.} \]

‘There were many problems.’

(22) Standard Iberian Spanish

\[ \text{Había muchos problemas.} \]

‘There was many problems.’

Moreover, there is a second problem. Our data suggest that (original) subjects only lose subject properties when there is an expletive. This seems to indicate that subject properties are not repelled by postverbal NPs, but attracted by expletives taking a position that is canonically associated with subjecthood. The assumption that subject properties are attracted by specific constituents or positions will be called the ‘attraction hypothesis’ in the following.

4. A typology of formulaic presentational

As has been pointed out, in our view generalizations concerning the distribution of subject properties in presentationals of Germanic and Romance languages need to be (i) based on a structural typology, and (ii) stated in historical terms. §4.1 and 4.2 present a typology of presentationals that is historically motivated and that distinguishes three types of existential predicates and three types of expletives (for a recent synchronic comparative overview of existential constructions, see McNally 2011). Given that the morphosyntactic properties of presentationals depend, to a considerable extent, on more general syntactic properties of the languages concerned, §4.3 introduces a distinction between (i) languages with an obligatory preverbal slot (thetic-XV languages), and (ii) languages without such a slot (thetic-V1 languages).

---

4.1 Types of existential predicates in formulaic presentational of Germanic and Romance languages

Formulaic presentational by definition contain explicit existential formulas. The type of predicate forming part of the existential formula constitutes the first major parameter of variation in our typology of presentational. Diagram 1 illustrates a sub-classification into three major types.

![Diagram 1. Types of existential predicates](image)

We can make a first distinction between one-place predicates and two-place predicates. One-place predicates – simple predicates of existence – are used relatively rarely in the existential formulas of European languages, with Rom. există and perhaps Port. and Span. existe being the most prominent representatives. Comparable existentials are also found in French (il existe) and English (there exists), where they have a more marginal status, however. An example of Rom. există is given in (23). Unlike Span. existe (pl. existen) and Port. existe (pl. existem), există is number invariant.

(23) Romanian

_există bărbați care să înțeleg femeile._
exist men who SBJ understand women
‘There are men who understand women.’

(Andreea Dumitrescu, p.c.)

Among the two-place predicates, two types can be distinguished, (i) those based on a copula, and (ii) those based on a (transitive) predicate, typically one of possession. Copulas form part of the existential formulas of Engl. ([there] is), Scandinavian languages (e.g. Norw. [det] er, Sw. [det] är, Dan. [der] er, Icel. [það] er) and Dutch ([er] is) and are also found in the Romance family (It. [ci] è, Sard. [bi] est). Romanian also uses the copula fi (3rd sg. este, 3rd pl. sunt) as an alternative to există.

The distinction between one-place predicates and copulas is of course not always clear-cut, as some Germanic and Romance languages use the same verb for both
functions (e.g. Germ. sein, Engl. be). As will be seen, the differences between the two types of presentationals are also very minor, and the main division is between transitive presentationals, on the one hand, and the other two types, on the other. With respect to the question of what counts as a copula and what as an existential predicate, we have applied the criterion of frequency. For example, in the overwhelming majority of cases, Engl. be functions as a copula (Hamlet is unhappy), and existential uses of this verb are very rare and sometimes archaic (to be or not to be, God is).

Existential formulas deriving from predicates of possession are widespread among Romance languages, e.g. Span. hay (< ha-y < Lat. ha[bet] i[bi] ’has there’), Port. há, French (il y’a) (’[it there]has’) and Cat. (hi) ha. There are also (more sporadic) instances of such transitive presentationals in Germanic languages, e.g. (Southern) German (es) hat (’[it] has’).

In addition to the major types of predicates surveyed in Diagram 1 there are some minor ones. For example, Germ. (es) gibt (lit. ’it gives’) appears to be derived from a ditransitive verb. However, it seems likely that this use of geben goes back to a (dynamic) mono-transitive use of that verb, roughly meaning ’give rise to’ (cf. §5.3.2). The argument structure of es gibt is thus basically identical to the one of Span. hay etc. Many Scandinavian languages have existential predicates whose valency is reduced by a middle marker. Some of these predicates are derived from a ditransitive verb meaning ’give’, e.g. Dan. (der) give-s, Norw. (det) gi-s, Sw. (det) ge-s (’give-mid/ psv’; 3→2-place). In other cases locative predicates are derived from transitive verbs, e.g. Sw. (det) finn-s, Dan. (der) finde-s, Norw. (det) finne-s (’find-mid’). These predicates can be regarded as copulas with a restriction to locative predications.

4.2 Cross-classifying types of predicates and types of expletives

There are two major types of expletives in Germanic and Romance languages, i.e. (i) weak pronominal expletives, and (ii) locative expletives (cf. also Vikner 1995: Ch. 7 for discussion of the two types of expletive in Germanic languages). Weak pronominal expletives are typically third person neuter pronouns like Germ. es, Dt. er and Norwegian det. In French, which does not have separate neuter pronouns, the masculine (or unmarked) form il is used. Locative expletives are mostly derived from adverbials meaning ’there’ (e.g. Dan. der) or ’here’ (It. ci < hicce, the spoken form of Lat. hic ’here’). There is one language that combines a weak pronominal expletive with a locative one, i.e. French (il y-a ’it there-has’). In those cases where no overt expletive can be identified, we could assume a phonologically empty syntactic element. Given that expletives do not seem to have any semantic function, however, we will simply assume that there is no expletive at all in these cases.

We can now cross-classify the three types of existential predicates and the three types of expletives, thus distinguishing nine types of existential formulas. A sample
of relevant constructions from various Germanic and Romance languages is given in Table 2. Note that French *il y’a* appears twice in the table because it contains two expletives, a locative and a weak pronominal one. Spanish *hay* contains a trace of a former expletive (the final glide [j] < Lat. *ibi* ‘there’), which is no longer recognizable as such, however. Note also that the past tense form *había* does not contain any such trace. We have therefore listed this formula under ‘no expletive’.

**Table 2.** Existential formulas of selected Germanic and Romance languages

<table>
<thead>
<tr>
<th></th>
<th>no expletive</th>
<th>weak pronominal expletive</th>
<th>locative expletive</th>
</tr>
</thead>
<tbody>
<tr>
<td>one-place</td>
<td>Romanian <em>există</em> NP</td>
<td>French <em>il existe</em> NP</td>
<td>Engl. <em>there exists</em> NP</td>
</tr>
<tr>
<td></td>
<td>Portuguese <em>existe</em> NP</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Spanish <em>existe</em> NP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>copular</td>
<td>Latin <em>est</em> NP</td>
<td>Dutch <em>er is</em> NP</td>
<td>English <em>there is</em> NP</td>
</tr>
<tr>
<td></td>
<td>Romanian <em>este</em> NP</td>
<td>Icelandic <em>það er</em> NP</td>
<td>Danish <em>der er</em> NP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Norwegian <em>det er</em> NP</td>
<td>Italian <em>c(i) è</em> NP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Swedish <em>det är</em> NP</td>
<td>Sardinian <em>bi est</em> NP</td>
</tr>
<tr>
<td>transitive</td>
<td>Portuguese <em>há</em> NP</td>
<td>German <em>es hat</em> NP</td>
<td>Catalan <em>hi ha</em> NP</td>
</tr>
<tr>
<td></td>
<td>Spanish <em>hay</em> NP</td>
<td>French <em>il (y’a)~a</em> NP</td>
<td>French <em>il y’a</em> NP</td>
</tr>
</tbody>
</table>

4.3 Thetic-V1 and thetic-XV languages

As has repeatedly been pointed out, generalizations concerning the argument structure of presentational – especially when approached from a diachronic point of view – can only be made if the ‘syntactic architecture’ of the relevant languages is taken into account. In this section we establish a parameter which plays an important role in the diachronic development of existential formulas. It concerns the (im)possibility for the languages in question to form verb-initial clauses for the encoding of thetic judgements. Most Romance languages allow this type of configuration (cf. Kato 1984; Hulk & Pollock 2001 and many others). An example from Spanish with an intransitive predicate was given in (4) above. (24) is an example with a transitive predicate.

(24) Spanish  
*Espero que te devuelva.* *el libro* *Juan.*  
I hope that you return the book Juan  
‘I hope that Juan will return the book to you.’ (Ordoñez 2000: 26)

Italian likewise allows verb-initial order in thetic judgements, as is illustrated in (25).

(25) Italian  
*Capisírá* *tutto* *Maria.*  
will understand everything Maria  
‘Maria will understand everything.’ (Belletti 2009: 184)
By contrast, all the Germanic languages under consideration as well as French do not have corresponding verb-initial structures. This is illustrated in (26) for German, and in (27) for French.8

(26) German

*Lebt ein Mann in dieser Stadt, der alles weiß.
lives a man in this city who everything knows
int.: ‘There lives a man in this city who knows everything.’

(27) French

*Vive un homme en Paris qui sait tout.
lives a man in Paris who knows everything
int.: ‘There lives a man in Paris who knows everything.’

Put differently, languages like German and French have an obligatory preverbal slot in their structural clause templates, while languages of the Spanish and Italian type do not have such a slot. The first type of language will be called ‘thetic-V1 languages’ – as verb-first order is allowed in thetic sentences – and the latter type ‘thetic-XV languages’ – where ‘X’ stands for the element occupying the obligatory preverbal slot.

5. Presentationals in Germanic and Romance languages

Having established a typology of presentationals and one structural parameter classifying the languages under investigation into two types, we now turn to the argument structure of each presentational by using examples from the sample of languages under investigation. The discussion is organized around the three types of existential predicates identified in §4.1, i.e. one-place predicates (§5.1), copulas (§5.2) and (transitive) predicates of possession (§5.3).

5.1 One-place presentationals

As pointed out in §4.1, one-place presentationals are found in some Romance languages. They are invariably based on the Latin verb ex(s)istere < ex-sistere, which originally had a dynamic meaning ‘emerge, come into existence’ (ex ‘from, out, sistere ‘stand’). An example from Romanian was given in (23) above. (28) illustrates the relevant construction of Portuguese.

---

In one-place presentenitals of the type shown in (23) and (28), the novel NP is the only candidate for subject status and accordingly exhibits all of the relevant properties. Most importantly, it controls verbal agreement and also exhibits other ‘behavioural’ subject properties like the ability to undergo raising. Note, however, that even in raising configurations the subject occupies a postverbal position, as in the Spanish example (29).

(29) Spanish

\[
\text{Parecen existir muchos problemas en este país.}
\]

'seem.pl exist many problems in this country'

Muchos problemas in (29) functions as a semantic argument of the existential verb existir, but it is the syntactic subject of the higher (epistemic) predicate parecen. This is the main criterion for a raising analysis, irrespective of its representation in a syntactic model. (30) provides a simplified representation of the mapping from semantic to syntactic argument structure (with \(e\) in the syntactic representation corresponding to the ‘missing’ locative argument of existir, i.e. muchos problemas).

(30) \[
\begin{align*}
\text{SEEM} & \quad \text{EXIST} (\text{many problems}) \\
\downarrow & \quad \downarrow \\
\text{[parecen PL [existir } & e] \text{]} \quad \text{[muchos problemas]}_{\text{UPL}} \\
\text{agreement}
\end{align*}
\]

According to Lambrecht (2000), occupying a postverbal position for a subject amounts to exhibiting non-canonical (since object-like) behaviour. However, we do not agree with Lambrecht (2000) in this respect. As was shown in §4.3, in those languages that we call ‘thetic-V1 languages’, postverbal subjects are regularly found in thetic sentences. Accordingly, the sole argument of the type of presentational illustrated in (29) can be regarded as a ‘genuine’ subject.9

9. Note that the ‘repulsion hypothesis’ runs into problems even if we assume that postverbal positioning is non-subject-like. In Latin, subjects were not tied to any particular syntactic position. The major syntactic change from Latin to Romance languages in this area consisted in fixing the (canonical) subject position to a preverbal slot. It is, accordingly, not the properties of (postverbal subjects in) presentenital constructions that have changed in any way;
The situation is different in the French existential formula *il existe*, which is illustrated in (31).

(31) French

*Il existe beaucoup de langues différents en Afrique.*

It exists many **part** languages different in the Africa

‘There exists many different languages in Africa.’

As pointed out by Lazard (1994: 8), one can “analyser *il* comme un sujet, peut-être un sujet réduit, mais en somme un sujet” (witness the agreement relations in (31)), whereas “[l]’actant H n’a donc rien d’un sujet” (Lazard 1994: 9). Given that in the (Latin) source structure, the postverbal NP was clearly a subject (agreeing with the predicate), there has been a shift of subject properties from the postverbal NP to the (preverbal) expletive. A similar development can be observed in the English construction in (32).

(32) **There exist many different models of syntax.**

Unlike in French, in (standard) English agreement is normally with the postverbal subject in these cases, though singular agreement is also possible in specific varieties (cf. (33a)). Note that a subject status of *there* is also indicated by the subject-to-object raising construction in (33b).

(33)

a. *It is probably not coincidental that there exists these two areas of human interest, namely, the areas of religion on the one hand and that of the healing power of the mind on the other, …*  

b. *We expect there to exist differences between public schools and private schools.*

A comparison of the one-place presentationals of Romanian, Spanish and Portuguese, on the one hand, and those of French and English, on the other, is instructive with respect to the status and role of what we have called ‘obligatory preverbal slots’. In the former group of (thetic-V1) languages, there is no expletive, and the postverbal NP has retained its status as a subject; in the latter (thetic-XV) languages French and English, subject properties have been transferred from the novel NP to the expletive. This suggests that it is only in languages with an obligatory preverbal slot (such as

what has changed is the criteria for subjecthood in Romanian and other Romance languages. While the presentational construction has remained the same, its status within the linguistic system has thus changed, and it has become ‘less canonical’ from a synchronic point of view. Still, it seems counterintuitive to us to say that postverbal subjects have shed subject properties (due to a ‘Principle of Paradigmatic Contrast’), given that they themselves have remained stable. Imagine a pair of twins one of which stops growing at the age of ten. When she is grown up, she is forty centimetres shorter than her sister – and, in a way, forty centimetres shorter than she is supposed to be – but does that mean that she has shrunk?
French and English) that expletives attract subject properties. As will be seen, the other constructions investigated below confirm this impression, and the parameter ‘thetic-V1 vs. thetic-XV’ is one of the most important determinants of the way subject properties are distributed over the arguments in a presentational construction.

The argument structure of presentational based on one-place predicates without an expletive can be represented as shown in (34). Obviously, this structure is restricted to thetic-V1 languages. The single argument of the existential predicate – here classified as a Theme – takes up the only argument position in the sentence and is a genuine subject. The representation of argument structure used in the following is inspired by Functional Grammar (e.g. Dik 1997), but simplifies matters for convenience. In the (lexical-) semantic layer, predicates are written in small caps and arguments are indicated by variables with a subscript indicating their semantic role (‘TH’ stands for Theme). In the syntactic layer, either specific forms from the languages under consideration are indicated or syntactic category labels. Grammatical relations are indicated by subscripts. Note that we follow Croft (2001) in assuming that grammatical relations are defined relative to the constructions in which they occur. Still, specific relations (most notably the ones of predicate and subject) can be generalized across constructions. Strictly speaking, ‘NP\textsubscript{SUBJ}’ in (34) thus means ‘NP with the grammatical relation of a subject within the relevant one-place presentational’. More loosely speaking, it stands for ‘NP with a sufficient number of properties characteristic of subjects across constructions’.

\[(34)\] One-place existential without an expletive

<table>
<thead>
<tr>
<th>Semantic structure</th>
<th>EXIST</th>
<th>(x_{TH})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syntactic realization</td>
<td>exist-</td>
<td>NP\textsubscript{SUBJ}</td>
</tr>
</tbody>
</table>

In thetic-XV languages, subject properties are attracted by the expletive in the obligatory preverbal slot. The English case and the French case are slightly different. In English, the expletive was originally a locative element, and the underlying argument structure can be assumed to be as shown in (35), which is a canonical construction in V2-syntax. The parentheses around the second argument \((y_{LOC})\) in the semantic representation indicate its optionality.

\[(35)\] One-place presentational with a locative expletive

Source construction

\[
\begin{align*}
\text{EXIST } x_{TH} & \quad (y_{LOC}) \\
\text{XP}_{LOC} & \quad \text{exist-} \\
\end{align*}
\]
The locative constituent was often realized twice in the sentence, once as a lexical locative specification (e.g. in the form of a PP), and once as a deictic pronoun functioning as a cataphoric copy of the former (cf. the discussion of copular presentational in §5.2.2). For example, sentences like (36) can be regarded as (originally) containing two syntactic arguments realizing the Location argument, i.e. a PP in Mexico and there as a cataphoric copy of the PP.

(36) There, once existed a mighty Mayan empire [in Mexico],

Accordingly, the (original) argument structure of a construction with two (syntactic) locative specifications can be assumed to have been as shown in (37). The dotted line indicates a redundant syntactic realization (i.e. we assume that the Theme argument is primarily realized as a lexical NP within PP_{LOC}, while there is linked to the semantic argument position more indirectly, i.e. via PP_{LOC}).

(37) One-place presentational with a locative expletive

Source construction with there as cataphoric copy

\[
\begin{array}{c}
\text{EXIST} \quad x_{TH} \quad \left( y_{LOC} \right) \\
\text{there} \quad \exists \quad \text{NP}_{\text{SUBJ}} \quad \text{PP}_{\text{LOC}} \\
\quad \text{cataphoric copy}
\end{array}
\]

Given the strong association between subject function and preverbal positioning in modern English, the cataphoric copy in clause-initial position – the pronoun there or its historical precursors – acquired some (though not all) relevant subject properties (cf. (33) above; see also Seefranz-Montag 1983: 138; Milsark 1974; Bolinger 1977; Lazard 1994). As a consequence, it turned into an expletive and was no longer associated with a locative argument in the semantic layer. The resulting presentational is impersonal, with an argument structure as shown in (38).

(38) One-place presentational with an expletive

Target construction

\[
\begin{array}{c}
\text{EXIST} \quad x_{TH} \quad \left( y_{LOC} \right) \\
\text{there}_{\text{SUBJ}} \quad \exists \quad \text{NP}_{\text{NOV}} \quad \text{PP}_{\text{LOC}}
\end{array}
\]

The syntactic function of the novel NP is not specified in (38). As was pointed out above, in keeping with basic assumptions of Radical Construction Grammar (Croft 2001) we assume that grammatical relations are basically functions of the constructions in which they occur. While specific grammatical relations can be generalized across constructions (e.g. subjects) – as they display specific ‘cross-constructional properties’
(cf. the criteria for subjecthood pointed out in §3.1) – other types of grammatical relations are largely construction-specific.

An argument structure parallel to the one shown in (38) can be assumed for French *il existe* NP. However, in French the historical development is of course different, as *il* cannot be analyzed as a former locative adjunct like Engl. *there* (for a comprehensive account of the historical development of expletives in French, cf. Zimmermann 2009). We will return to the development of such weak pronominal expletives in §5.2.4, where we will argue that they originate as cataphoric copies of the novel NP. Accordingly, the source construction of the French presentational can be represented as shown in (39). There is just one underlying Theme argument which is realized twice in surface syntax, once in the obligatory preverbal slot (in the form of a cataphoric copy), and once in the postverbal position taken by the novel NP (cf. §5.2.4 for more details concerning the relationship between a cataphoric copy and the novel NP).

(39) One-place presentational with a weak pronominal expletive
Source construction

\[
\begin{array}{c}
\text{EXIST} \quad x_{\text{TH}} \\
\downarrow \quad \downarrow \\
\text{PRO} \quad \text{exist-} \quad \text{NP}_{\text{SUBJ}} \\
\end{array}
\]

\text{cataphoric copy}

5.2 Copular presentational

5.2.1 ... without an expletive: Latin, Portuguese, Spanish, Romanian
A copular presentational without an expletive provided the most important presentational strategy of (classical) Latin (cf. (40)). The copula can be regarded as having a locative function in these cases, i.e. as taking as its arguments a Location and a Theme (recall from §4.1 that the copulas of most Indo-European languages also function as existential predicates).

(40) Latin
\[\text{Est [puella]}_{\text{NOV}} \quad [\text{in via}]_{\text{LOC}}\]
is girl in street
'There is a girl in the street.' (Cicone 2009: 186)

In Latin, the novel NP clearly has subject status, and there is no other candidate for it. (41) is an example illustrating subject-to-object raising of *aliquos milites* (‘some soldiers’, the underlying subject of the infinitival clause and the syntactic object of the matrix verb).
On the distribution of subject properties in formulaic presentationals

(41) Dixit mihi alius [esse [in exercitu]LOC [aliquos milites, he.told me some be in army some soldiers
qui consultant, quid agendum sit]NOV].
who debate what to do is.SBJ
'Someone told me that there are soldiers in the army who are debating
what is to be done.' (Livius XLIV 34)

The Latin construction has been preserved in Romanian, where it is used as an alternative to the existential predicate există. The copula agrees in number with the novel NP (infinitive a fi, 3rd sg. este, 3rd pl. sînt/sunt; 10 cf. Cornilescu 2009).

(42) Romanian
Sînt [mulți ținări]NOV [aici]LOC
are many mosquitos here
'There are many mosquitos here.' (Lazard 1994: 22)

The argument structure of copular presentationals without an expletive can simply be represented as in (43). The copula is taken to have a locative function ('be located'), thus requiring two arguments, a Theme (syntactically, the subject) and a Location. The Location argument functions as a syntactic complement of the copula.

(43) Copular presentational without an expletive

\[
\begin{array}{c}
\text{BE.LOCATED } x_{\text{TH}} \downarrow y_{\text{LOC}} \\
\text{COP} \downarrow \text{NP}_{\text{SUBJ}} \downarrow \text{XP}_{\text{COMP}}
\end{array}
\]

5.2.2 … with locative expletives: English and Danish

Copular presentationals with a locative expletive are closely related to the type of presentational considered in the previous section. They are found inter alia in English (there is), Danish (der er), Italian (c’è) and Sardinian (bi est). In spite of superficial similarities, the Germanic cases differ strikingly from the Romance ones. The two cases are therefore dealt with in separate sections. We will start with the Germanic languages. In English, there-existentials can be regarded as remnants of verb-second structure in English, very much like instances of locative inversion (cf. Breivik 1990: 181–249; Pfenninger 2009: 49–74; cf. also §5.1). An example of a presentational from Old English is given in (44).

---

10. In contemporary Romanian the spelling sunt is more common (Andreea Dumitrescu, p.c.).
(44) Old English

þær is mid Estum an mæʒð.
there is with Estonians a girl

‘There is a girl with the Estonians …’

(Ælfred, Orosius I, i, 22; c893)

In this case, the locative pronoun þær can be regarded as a ‘copy’ of the Location mid Estum ‘with the Estonians’. Such ‘intra-clausal copies’ are still commonly used in contemporary German, which has preserved verb-second syntax (cf. (45)).

(45) German

Da₁ sitzt ein Spatz [auf dem Dach]ᵣ.
there sits a sparrow on the roof

‘A sparrow is sitting (there) [on the roof]ᵣ.’

Accordingly, presentationals of the there is-type exhibit the same type of argument structure as copular presentational without an expletive (cf. §5.2.1). The only difference is that there is a cataphoric copy of the Location argument. When English lost its verb-second structure, this element was gradually reanalyzed as a subject. As has been pointed out, the ‘increase’ in subject status of there (in Standard English) is reflected in some behavioural properties, e.g. in its ability to undergo raising, in its participating in subject-auxiliary inversion and in its use in question tags (cf. Milsark 1974: 14–15, Bolinger 1977; Lazard 1994).

(46) English

a. There seems to be a problem here.

b. I believe there to be a problem.

c. Is there a problem?

d. There is a problem, isn’t there?

However, there is certainly not a ‘perfect subject’ in standard English, as it does not display agreement with either the main predicate (in simple clauses) or the raising verb in constructions like (46a). As has been mentioned, this is different in the spoken language, see the examples in (47).

(47) English (spoken)

a. %There’s some students waiting outside.

b. %There seems to be some students waiting outside.

The situation in Danish seems to be largely parallel to the one found in English. As Jespersen (1924:155) points out with respect to (48), “the verb was here put in the singular before a plural verb, even at a time when the distinction between sg. er and pl. ere was generally observed.” In other words, agreement between the predicate and the postverbal NP was suspended.
On the distribution of subject properties in formulaic presentationals

(48) Danish

Der er dem som tror ...
there is those who believe
lit.: ‘There is those who believe …’  (Jespersen 1924: 155)

Given that both Engl. there and Dan. der seem to have acquired subject properties, the copular presentationals of English and Danish provide clear examples of a ‘transfer’ of subject properties from the novel NP to the expletive. In fact, both languages have been used to substantiate the ‘(global) repulsion hypothesis’ by Lambrecht (1986, 2000). However, the loss of subject properties by the novel NP can obviously just as well be explained in terms of what we have called the ‘attraction hypothesis’ (cf. §3.3). Given that this type of change can only be observed when there is an expletive, we will argue that the attraction hypothesis can explain the facts from the Germanic and Romance languages investigated in this study better than the repulsion hypothesis.

It should be noted that the two cases illustrated above – the copular presentationals of English and Danish – have probably resulted from slightly different historical developments. In English, any (nominal) constituent occupying the preverbal position eo ipso turns into a subject (with systematic exceptions like locative inversion). Danish, by contrast, is a verb-second language, so taking a clause-initial position does not, in itself, turn a constituent into a subject. As pointed out by Faarlund (1989: 63ff.), the clause-initial locative element was first reanalyzed as a ‘topic expletive’. The following example from 15th cent. Danish apparently is the first attested occurrence of the locative expletive ther in Scandinavian languages.

(49) Danish (15th cent.)

ther kom ey een tijl lande hiem
there came not one to country home
‘No one came back home to his country.’

(Danish Chronicle, 1495; quoted from Faarlund 1989: 66)

Later, the ‘expletive topic’ was reanalyzed as an expletive subject: “When the time comes for the expletive topic to become an expletive subject, one feature of þar helps to promote the process. In Old Norse the adverbial þar would often occur immediately after the finite verb, probably cliticized to the verb. And as we have seen, this is also the position of the nonfronted subject” (Faarlund 1989: 71). Even though in Danish there is no one-to-one correspondence between structural positions and syntactic functions, the association between ther and subject properties was apparently strong enough for the originally locative element to acquire subject status.

The change in argument structure of English and Danish copular presentationals is summarized in (50) and (51). In the source construction (cf. (50)), the Theme functions as a subject and there is a Locative complement (required by the copula), which is (redundantly) replicated by a locative adverb (cf. also the German example (45)).
In the target construction, the expletive functions as a (syntactic) subject without a semantic role. The construction is thus impersonal. The novel NP is realized as a non-subject constituent. Again, we will not take a stance with respect to its precise syntactic function (e.g. object, complement), as we follow Croft (2001) in assuming that grammatical relations are basically construction-specific.

It should be noted that it is of course conceivable to regard the copulas of English and Danish as having the function of an existential predicate in the relevant constructions. In this case, their argument structure would be as described in §5.1, i.e. the Location argument would not be required by the predicate, but would function as an optional specification.

### 5.2.3 with locative expletives: Italian and Sardinian

As has been pointed out, Italian c'è and Sardinian bi est differ in important respects from the English and Danish existential formulas discussed in the previous section (cf. Remberger 2009 on Sard. bi est). In the present context, the main difference between English and Danish, on the one hand, and Italian and Sardinian, on the other, is that the former languages are thetic-XV while the latter are thetic-V1. Even though the constructions are superficially similar, they have, accordingly, different histories, as is also reflected in their synchronic behaviour.

Note first that existential constructions based on a copula and a locative adverb were widespread in (postclassical) Latin, as in the following example from Apuleius’ *Metamorphoses*.

(52) Latin

Quicquid fieri non potest ibi est.

whatever happen not can there is

‘Whatever cannot be is there.’

(Apuleius, Metamorphoses II, 19, 2)
On the distribution of subject properties in formulaic presentational...

The underlying argument structure is thus basically identical to the one of the English and Danish copular presentational, i.e. it is ‘BE.LOCATED \((x_{TH}, y_{LOC})\)’. It seems to us that the motivation for the insertion of a locative expletive in (late) Latin is strikingly different, however. In English and Danish, the expletive was probably inserted for syntactic reasons, namely to fill the obligatory preverbal slot, i.e. the Forefield (in V2-syntax) or the subject position (when English had turned SVO). In Latin, by contrast, there was no such obligatory preverbal slot, and as is illustrated by the Romanian case (cf. §5.2.1), Italian could have preserved a presentational of the form ‘est/sunt NP’ (such forms were actually widespread in early Italo-Romance vernaculars, e.g. 13th cent. Tuscan; cf. Ciconte 2009). The expletive probably established itself as part of the existential formula for reasons of ambiguity avoidance, not as a result of syntactic pressure. In contemporary Italian, \(ci\) can block a referential interpretation of the (phonologically empty) subject. The contrast between copula sentences with and without expletive is illustrated in (53).

(53) Italian

a. È un uomo che sa fare qualunque lavoro.
   ‘He is a man who knows (how) to do any type of work.’

b. C’è un uomo che sa fare qualunque lavoro.
   ‘There is a man who knows (how) to do any type of work.’

   (Federica da Milano, p.c.)

In (53a), the absence of an overt subject is (by default) interpreted as the presence of a phonologically empty subject, which is necessarily topical and, hence, typically referential. In examples of the type of (53b), the presence of a locative adverbial (\(ci\)) triggers a locative (rather than equative or predicative) reading of the copula (‘\(x\) is located at \(p\)’ instead of ‘\(x = y\)’ or ‘\(x \in P\)’). While this locative element may have been optional at a certain stage, it was later reanalyzed as an obligatory component of the existential formula of Italian.

The difference in the historical developments of locative expletives in Romance and Germanic languages is also reflected in their synchronic behaviour. We will discuss this point with respect to Italian \(ci\), assuming that similar observations can be made about Sardinian \(bi\) (cf. Remberger 2009 for relevant data). \(Ci\) does not display as many subject properties as English \(there\), and we will in fact argue that it is not a subject at all. Note first that Italian \(ci\) appears to be susceptible to raising, just like English \(there\).

(54) \(Ci\) sembra essere un certo numero di problemi.
   ‘There seems to be a number of problems.’

   (Federica da Milano, p.c.)
However, (54) should probably not be analyzed as the result of raising *ci from the lower clause to the higher clause. Unlike in (spoken) English, the raising verb *sembre obligatorily agrees with the novel NP (cf. (55)).

(55) *Ci *sembra/*sembra essere molti studenti. there seem.pl/seem.sg to.be many students
‘There seem to be many students.’ (Federica da Milano, p.c.)

Given that Italian is a thetic-V1 language, (55) is thus best analyzed along the lines of the Spanish example (29) above, and molti studenti can be regarded as the (postverbal) raised subject of *sembra.

A second fact (presumably) showing that Italian *ci has subject status is that the formula c’è, like Engl. *there’s, allows singular agreement in specific dialects/registers, as in (12a) above. However, in our view an analysis of *ci as a subject is not very likely. As Koch (2003:158) has pointed out, agreement and linear order are the only indicators pointing to a subject status of *ci. The criterion of cliticization shows that the postverbal NP is certainly not an object, and that it is most probably a subject. The (indefinite) clitic *ne can be used with both postverbal subjects and direct objects, whereas the *lo/la-series is used only with objects. While ne-cliticization is possible for contadini in (56a+b), *lo/la is ungrammatical in the relevant varieties (cf. (56c), from Koch 2003:158–159; note that *ci is realized as *ce in this context). This shows that contadini is not treated as an object, as far as ne-cliticization is concerned.

(56)
a. C’-era dei contadini → Ce n’-era
there-was INDEF.PL farmers there INDEF-was
b. C’-erano dei contadini → Ce n’erano.
there-were INDEF farmers there INDEF-were
c. C’-era i contadini → *Non ce li era.
there-was the farmers not there them was

Given that we do not regard the postverbal position of the novel NP as an indicator of its status as a non-subject in thetic-V1 languages, the only (assumed) subject property of *ci that remains thus is the singular inflection of the copula in specific varieties of Italian. This is not necessarily an indication that *ci has become a subject, however. It seems more likely to us that (in the relevant varieties) the entire formula c’è has been reanalyzed as an existential (one-place) predicate, more or less like Romanian există. As many other existential predicates, it has become number invariant, i.e. it does not agree at all. Note that this explanation might seem somewhat *ad hoc, given that we have regarded Engl. *there (as well as Danish *der) as a subject, and the two cases appear to be largely parallel. Remember, however, that there is a clear difference in the raising behaviour of *there and *ci: While *there can control agreement in raising structures (in the relevant registers), *ci cannot do so.
We consequently assume that in Italian it is the novel NP that functions as a subject. The expletive *ci* can be regarded as a cataphoric copy of the Location (cf. Schwarze 1995:329) and is used for ambiguity avoidance. (57) shows the argument structure of the construction found in the standard language, where the copula agrees with the postverbal subject.

(57) Copular presentational with a locative expletive:

\[
\begin{array}{c}
\text{BE.LOCATED} \\
\downarrow \\
\text{ci} \\
\downarrow \\
\text{essere} \\
\downarrow \\
\text{NP}_{\text{SUBJ}} \\
\downarrow \\
\text{XP}_{\text{COMP}} \\
\downarrow \\
\text{cataphoric copy}
\end{array}
\]

Once the combination of *ci* and *è* (i.e. *c'è*) is reanalyzed as a single existential formula, it acquires the argument structure of a one-place presentational (cf. §5.1). This is shown in (58).

(58) Copular presentational with a locative expletive reanalyzed as one-place presentational

\[
\begin{array}{c}
\text{EXIST} \\
\downarrow \\
\text{c'è} \\
\downarrow \\
\text{NP}_{\text{SUBJ}} \\
\downarrow \\
\text{XP}_{\text{LOC}}
\end{array}
\]

5.2.4 …with weak pronominal expletives: Scandinavian and Dutch

Copular existentials with weak pronominal expletives are found in some Germanic languages, most notably in Scandinavian ones (Icel. *það er*, Norw. *det er*, Swed. *det är*) and in Dutch (*er is*). We will consider the Scandinavian cases in some detail. Old Norse did not (obligatorily) employ expletives in existential constructions (cf. (59)) and used the copula (more or less) like Latin as an existential predicate.

(59) Old Norse

\[
\text{kastali var fyrir austan sundit} \\
\text{castle.NOM was for east.of the.sound}
\]

‘There was a castle to the east of the sound.’ (Faarlund 1989:84)

In Modern Scandinavian languages, expletives have become obligatory and “behave like subjects syntactically” (Lødrup 1999:206; but see Börjars & Vincent 2005:9 for a
qualification of this statement). They can be raised (cf. (60a)) and they are used in tag questions in a slot that is generally reserved for subjects (cf. (60b)).

(60) Norwegian

a. *Det synes å være mange tilhørere.*
   
   EXPL seem to be many hearers
   ‘There seem to be many listeners.’

b. *Det kom mange kundar, gjorde det ikkje?*
   it came many customers did it not
   ‘There came many customers, didn’t there?’ (Faarlund et al. 1999: 833)

The novel NP, by contrast, has been claimed to “unequivocally” have the status of a direct object (Askedal 1986: 31; but see, again, Börjars & Vincent 2005). Or, as Faarlund (2001: 1158) puts it: “The NP argument is now in object position, following the non-finite verb. The only possible subject in the sentence is therefore det.”

We will followAskedal (1986), Lødrup (1999) and Faarlund (1989, 2001) in assuming that the expletive functions as a subject (with respect to its syntactic properties). However, the syntactic function of the postverbal NP seems to be characterized basically in negative terms by these authors, i.e. it is primarily a non-subject rather than an object (see e.g. Lødrup 1999: 206–208). We will assume that the grammatical relation of the novel NP is construction-specific, but that does not prevent it from being similar to objects in common transitive clauses.

According to Faarlund (1989, 2001), Old Norse allowed verb-initial clause structure in declarative clauses. The first instances of weak pronominal expletives among the Scandinavian languages are attested in the 15th century (cf. Faarlund 2001: 1157, referring to Falk & Torp 1900). (61) is an example from 15th cent. Danish.

(61) Danish (15th cent.)

*thet war een man, hwilken som haffde twenne sønær.*

it was a man who REL had two sons
‘There was a man who had two sons.’

(Faarlund 2001: 1157, referring to Falk & Torp 1900)

It is possible that weak pronominal expletives were void of any meaning at the time they came to be used as expletives and merely had a syntactic function. However, they may also have had a cataphoric function, which is still recoverable in the relevant structures of Modern Scandinavian languages. Copular presentational clauses with weak pronominal expletives seem to be structurally parallel to nominal predicative constructions of the form ‘A_{SUBJ} is B_{PRED},’ with the expletive taking the subject slot. Consider (62).
(62) Norwegian
   a. Hvem er det?
      who is it
      ‘Who is it?’
   b. Det er Jan.
      it is Jan
      ‘It’s Jan.’

In (62b), det refers back to the pronoun from the question in (62a); but given its referential identity with the predicative nominal Jan, it can easily be reanalyzed as a cataphoric copy of the latter. The cataphoric nature of weak pronominal expletives in Scandinavian languages has also been noticed by Faarlund (2001:1158), who points out that “[t]he use of the neuter pronoun ‘it’ as an expletive subject may have been reinforced by its use in extraposition constructions, where it originally was an anticipating pronoun referring to the sentential argument” (cf. the Old Norse example (63)).

(63) Old Norse
   er þat minn vili, at svá gøri vér allir
   is it my will that so do we all
   ‘It is my wish that we all do so.’ (Faarlund 2001:1158)

We will thus assume that at an earlier stage, det functioned as a cataphoric copy of the novel NP in presentationals, and that the argument structure of the construction corresponded to a nominal predication of the form (‘A_{SUBJ} is B_{PRED}’). The predicate of the construction is a copula with an equative function, represented as ‘be’ in (64). There is, thus, a Theme NP which is stated to be (identical to) a (referential) noun phrase functioning as a complement of the copula (cf. as in (62) above).

(64) Copular presentationals with weak pronominal expletives:

Source structure

\[
\begin{array}{c}
\text{BE} \\
\downarrow \\
\text{x_{TH}} \quad y_{PRED} \\
\downarrow \\
\text{det}_{SUBJ} \quad \text{er} \\
\downarrow \\
\text{NP}_{NOV} \\
\end{array}
\]

\text{cataphoric copy}

In modern Scandinavian languages, the expletive probably does not perform any semantic function. It only has a syntactic function, and the construction qualifies as impersonal, according to the criteria adopted in the present study (cf. Siewierska & Malchukov this volume). The verb can be regarded as functioning as an existential
predicate, taking only one (postverbal) argument. As pointed out above, we will assume that the grammatical relation of this argument is basically construction-specific.

(65) Copular presentational with weak pronominal expletives:

Target structure

\[
\begin{align*}
\text{EXIST} & \quad x_{\text{TH}} \\
\text{det}_{\text{SUBJ}} & \quad \text{er} & \quad \text{NP}_{\text{NOV}}
\end{align*}
\]

Assuming that the main presentational strategy of Old Norse was a copular one without an expletive where the postverbal NP functioned as a subject (cf. (59)), subject properties must have been transferred from the postverbal NP to the expletive in Modern Scandinavian languages. As far as we can tell, this change was causally related to the introduction of an expletive, which, in turn, acquired subject properties basically because it occupied the preverbal slot (cf. also §5.2.2 on Danish). There is thus, again, no evidence that the novel NP has shed subject properties, as is predicted by Lambrecht’s (2000) ‘(global) repulsion hypothesis’. Rather, subject properties seem to have been attracted by the expletive (our ‘attraction hypothesis’).

5.3 Transitive presentational

Transitive presentational based on possessive predicates are attested from the earliest records of Indo-European onwards. Latin *habet* (‘it has’) – often accompanied by the deictic pronoun *ibi* ‘there’ – is the historical source of Span. *hay*, Fr. *il y’a* and Catalan *hi* *ha*, among other existential formulas. Possessive predicates obviously lend themselves to being used as existential predicates, considering that they can be analyzed as predicting existence at a specific location (cf. Lambrecht 1986, 2000; Freeze 1992, among others). In other words, ‘*x* has *y*’ can be interpreted as ‘*y* is located at *x*’. The two arguments of a possessive predicate can thus be regarded as a Location (the first argument) and a Theme (the second argument). Given that the Theme functions as the second argument of a predicate, it will normally occur late in the sentence, thus complying with one of the requirements of presentative utterances pointed out in §3.2.

5.3.1 … without an expletive

Transitive presentational without an expletive can be found in Portuguese (cf. (66)).

(66) Portuguese

\[
\begin{align*}
\text{Havia} & \quad \text{muitos problemas} \\
\text{EX.PAST} & \quad \text{many problems}
\end{align*}
\]

‘There were many problems.’
The Portuguese construction corresponds quite closely to the underlying Latin construction, where the novel NP functioned as an object. The subject position can be assumed to have been realized by an empty pronominal element with an impersonal or generic interpretation in Latin (‘people had many problems’). The argument structure underlying this type of presentational can be represented as shown in (67).

(67) Transitive presentational without expletive

Source structure

\[
\begin{array}{c}
\text{HAVE} \\
\downarrow \\
\text{habere} \\
\end{array}
\begin{array}{c}
\text{LOC} \\
\downarrow \\
\text{SUBJ} \\
\end{array}
\begin{array}{c}
y_{\text{TH}} \\
\downarrow \\
\text{NP}_{\text{OBJ}} \\
\end{array}
\]

Given that in Romance languages (unlike in Latin) empty pronominal elements are necessarily associated with highly topical and referential arguments, the generic interpretation of the empty subject must have been lost at some point, and the empty subject was reinterpreted as no subject at all. Accordingly, the originally transitive construction changed into an intransitive one. Given that there was only one argument left – a Theme argument – the predicate could be reanalyzed as a plain existential predicate. As a result, the postverbal argument could acquire subject status, as is witnessed by agreement between the postverbal NP and haver in specific varieties of Portuguese. Interestingly, agreeing postverbal NPs appear to be particularly widespread among educated speakers (S. Perreira, p.c.). This seems to point in the direction of normative influence. (68) is a relevant example.

(68) Spoken Portuguese

\[
\%\text{Haviam muitos problemas.}
\]\n
EX.PAST.PL many problems

‘There were many problems.’

The argument structure of transitive presentational that have been reanalyzed as one-place presentational is thus identical to the one assumed for Romanian există. It is shown in (69).

(69) One-place presentational

Reanalysis of transitive presentational

\[
\begin{array}{c}
\text{EXIST} \\
\downarrow \\
\text{haver-} \\
\end{array}
\begin{array}{c}
x_{\text{TH}} \\
\downarrow \\
\text{NP}_{\text{SUBJ}} \\
\end{array}
\]

We may note that the Spanish existential formula based on haber could also have been included in this section. In the past tense, this formula is completely parallel to the
Portuguese case, being based on a plain predicate deriving from Lat. *habere* (sg. *había*, pl. *habían*). However, as pointed out in §2.1, the (number invariant) present tense form *hay* incorporates remnants of a former locative element *ibi*. We will therefore treat *hay* as an instance of a transitive presentational with a locative expletive and return to it in §5.3.3.

5.3.2 ... with a weak pronominal expletive: German and French

Weak pronominal expletives are found in *es hat*-presentational of (Southern) German as illustrated in (70):

(70) German

\[
\text{Es hat viele Menschen hier.} \\
\text{it has many people here} \\
\text{‘There are many people here.’}
\]

As pointed out by Lazard (1994), German provides a particularly clear case of a language whose expletive exhibits most or all of the properties commonly associated with subjecthood, whereas the novel NP is clearly an object. Note that, unlike in most other languages under consideration, in German the syntactic function of the constituents involved is also reflected in morphological case, especially on the object (the expletive *es* is case-invariant; cf. (15) above).

A weak pronominal expletive is also used in French, together with the locative element *y*. As pointed out in §5.1, *il* can *bona fide* be analyzed as a subject, as it displays most or all of the relevant properties. (71) illustrates that it undergoes subject-to-subject raising.

(71) French

\[
\text{Il semble y avoir beaucoup de problèmes.} \\
\text{it seems there have many part problems} \\
\text{‘There seem to be many problems.’}
\]

The French case is particularly interesting because it tells us something about the relationship between different types of expletives. It is the pronominal, not the locative, expletive that attracts subject properties. In fact, it seems to us that locative expletives may acquire subject status only under rather specific circumstances, e.g. in English, where weak pronominal expletives are not used in presentational.

Transitive presentational with weak pronominal expletives seem to have emerged from personal constructions in cases where coreference with a nearby referential element is possible but not compelling (cf. Behaghel 1923: 319; see also Wegener 2008). Such ambiguities often arise in combination with uncontrollable events or natural forces like weather conditions. Behaghel (1924: 137) provides the example in (72) to illustrate this type of ambiguity.
(72) German

Das Wetter ist sehr günstig: das/es gibt gute Ernte
the weather is very favourable that/it gives good crop
'The weather is very favourable: it will yield/there will be a large crop.'

Assuming that transitive presentational constructions have emerged from personal constructions, their original argument structure can be assumed to have been as shown in (73). There is a ‘genuine’ (morphosyntactic) subject es and also a ‘genuine’ object NP.

(73) Transitive presentational with weak pronominal expletives:

Source structure

\[
\text{HAVE} \quad x_{\text{LOC}} \quad y_{\text{TH}}
\]
\[
es_{\text{SUBJ}} \quad \text{hat} \quad \text{NP}_{\text{OBJ}}
\]

In the target structure, the expletive is no longer interpreted as a Location argument and does not have any semantic role (cf. (74)). The construction has thus become impersonal. The novel NP retains its morphosyntactic status as an object.

(74) Transitive presentational with weak pronominal expletives:

Target structure (impersonal)

\[
\text{HAVE} \quad x_{\text{TH}} \quad y_{\text{LOC}}
\]
\[
es_{\text{SUBJ}} \quad \text{has} \quad \text{NP}_{\text{OBJ}} \quad \text{XP}_{\text{LOC}}
\]

5.3.3 … with a locative expletive

A transitive presentational containing with a locative expletive can be found in Catalan (cf. (75)). Remember that Spanish hay originally also incorporated a locative element (hay < habet ibi), and thus belongs in the class of constructions dealt with in this section.

(75) Catalan

Hi ha aigua...
there has water
'There is water.'

(Alonso Capdevila & Suïls Subirà 1998: 7)

Just as in Spanish and Portuguese (but unlike in French), the Catalan construction is also used with plural agreement. A relevant example is given in (76).

(76) Hi han dos homes.

here have.pl two men
'There are two men here.'

(Alonso Capdevila & Suïls Subirà 1998: 7)
The fact that the postverbal NP has attracted agreement confirms our hypothesis that subject properties are attracted by expletives only in thetic-XV languages. The use of a locative adverb in the Catalan presentational can be motivated just like the use of *ci* in Italian. In all likelihood, it was used to prevent a referential interpretation of the (empty) subject argument. Remember that the locative adverb *ibi* was added to transitive presentationalas as early as in postclassical Latin, at that time probably being motivated pragmatically. The argument structure of the Latin source construction can be represented as in (77). The Location argument is realized twice, once in the empty subject and once overtly in the adverbial.

(77) Transitive presentationalas with a locative expletive:

Source structure

\[ \text{HAVE} \quad x_{\text{LOC}} \quad y_{\text{TH}} \]

\[ \begin{array}{c}
\text{hic} \\
\emptyset_{\text{SUBJ}} \\
habet \\
\text{NP}_{\text{OBJ}}
\end{array} \]

In Catalan, the combination of a locative expletive with a predicate of possession seems to have been reanalyzed as an existential predicate, similar to It. *cè*. As a result, the novel NP – formerly an object, now the only argument of the sentence – has acquired subject status. This is illustrated in (78).

(78) Transitive presentationalas with a locative expletive:

Target structure

\[ \text{EXIST} \quad x_{\text{TH}} \quad y_{\text{LOC}} \]

\[ \begin{array}{c}
\text{hi} \\
\text{ha} \\
\text{NP}_{\text{SUBJ}} \\
(\text{XP}_{\text{LOC}})
\end{array} \]

6. Summary

The major changes observed in the nine (diachronic) types of presentationalas are summarized in Table 3 (for thetic-VX languages) and Table 4 (for thetic-V1 languages). The first two columns represent the type of predicate and the type of expletive, respectively. In the third and fourth columns the argument structure of each source construction is indicated, and in the sixth and seventh columns changes in argument structure, if any, are shown. In those cases where no relevant change can be observed, the cells are shaded. The acquisition of subject properties is indicated by ‘+ subj’ and the loss of subject properties by ‘– subj’. The rightmost column provides examples.
Let us first consider thetic-XV languages. In one-place and copular presentationals, there has been a shift of subject properties from the postverbal NP to the expletive. In German, however, the original argument structure has remained unchanged and seems to be stable. A transfer of subject properties in the other direction (towards the novel NP) cannot be observed. All constructions surveyed in Table 3 are impersonal, insofar as the subject does not have a referent or semantic role.

Table 3. Historical changes in presentationals of thetic-XV languages

<table>
<thead>
<tr>
<th>pred</th>
<th>expl</th>
<th>underlying arg struc</th>
<th>changes in arg struc</th>
<th>examples</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>expl</td>
<td>novel NP</td>
<td>expl</td>
</tr>
<tr>
<td>one-place</td>
<td>weak pron</td>
<td>–</td>
<td>subj</td>
<td>+ subj</td>
</tr>
<tr>
<td></td>
<td>loc expl</td>
<td>loc</td>
<td>subj</td>
<td>+ subj</td>
</tr>
<tr>
<td>copular</td>
<td>weak pron</td>
<td>–</td>
<td>subj</td>
<td>+ subj</td>
</tr>
<tr>
<td></td>
<td>loc expl</td>
<td>loc</td>
<td>subj</td>
<td>+ subj</td>
</tr>
<tr>
<td>transitive</td>
<td>weak pron</td>
<td>subj</td>
<td>obj</td>
<td>–</td>
</tr>
</tbody>
</table>

Table 4 almost represents the mirror image of Table 3. There is no case in which the novel NP has lost subject properties, but there are two cases where subject properties have been acquired by the postverbal NP (transitive presentationals without expletives and with locative expletives). All constructions surveyed in Table 4 are personal, insofar as the novel NP either is a subject in the source construction already, or else acquires subject properties in the course of its historical development.

Table 4. Historical changes in presentationals of thetic-V1 languages

<table>
<thead>
<tr>
<th>pred</th>
<th>expl</th>
<th>underlying arg struc</th>
<th>changes in arg struc</th>
<th>examples</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>expl</td>
<td>novel NP</td>
<td>expl</td>
</tr>
<tr>
<td>one-place</td>
<td>no expl</td>
<td>–</td>
<td>subj</td>
<td>–</td>
</tr>
<tr>
<td>copular</td>
<td>no expl</td>
<td>–</td>
<td>subj</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>loc expl</td>
<td>loc</td>
<td>subj</td>
<td>–</td>
</tr>
<tr>
<td>transitive</td>
<td>no expl</td>
<td>adj</td>
<td>obj</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>loc expl</td>
<td>adj</td>
<td>obj</td>
<td>–</td>
</tr>
</tbody>
</table>

The results of our study can be summarized in the form of the following generalizations:

- A loss of subject properties on the part of the novel NP can only be observed in thetic-XV languages. Moreover, this type of development invariably seems to be associated with the simultaneous acquisition of subject properties by an expletive. This strongly suggests that it is in fact the expletive – or, to be more precise, the structural position occupied by the expletive – that attracts subject properties.
Given that expletives attract subject properties in the relevant constructions, the presentationals of thetic-XV languages show a strong tendency to be impersonal.

- The acquisition of subject properties by the novel NP can only be observed in thetic-V1 languages. It takes place when a transitive predicate of possession is reanalyzed as a predicate of existence due to the loss of the first (Locative) argument slot. Presentationals of thetic-V1 languages do not show a tendency towards 'impersonalization', insofar as the novel NP seems to function as a subject in the relevant constructions.

Another observation that has emerged concerns the role of expletives in thetic-XV as opposed to thetic-V1 languages: While in the former languages, expletives appear to be motivated syntactically – the obligatory preverbal slot must not be empty – in thetic-V1 languages there is no such syntactic pressure. On the basis of data from Italian, we have suggested that ambiguity avoidance might be the driving force behind the insertion of locative expletives in V1-languages. The function of the expletive is to block an equative or predicative function of the copula, thus leading to a locative or existential reading.

If our hypothesis is correct, we would expect thetic-V1 languages to invariably use locative, rather than weak pronominal, expletives. As far as we can see, this generalization is robust. While there are in fact pronominal elements in Romance languages that have been analyzed as expletives – most notably, (European) Portuguese ele – under closer scrutiny these elements turn out to perform a discourse-functional (e.g. cohesive), rather than purely syntactic, function (see Hinzelin 2009 for an overview and Carrilho 2005 specifically on European Portuguese). Given the very limited range of languages investigated in the present study this hypothesis is of course open to challenge.

**Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACC</td>
<td>accusative</td>
</tr>
<tr>
<td>COMP</td>
<td>complement</td>
</tr>
<tr>
<td>COP</td>
<td>copula</td>
</tr>
<tr>
<td>EX</td>
<td>existential</td>
</tr>
<tr>
<td>EXPL</td>
<td>expletive</td>
</tr>
<tr>
<td>INDEF</td>
<td>indefinite</td>
</tr>
<tr>
<td>LOC</td>
<td>locative</td>
</tr>
<tr>
<td>MID</td>
<td>middle</td>
</tr>
<tr>
<td>NOM</td>
<td>nominative</td>
</tr>
<tr>
<td>NOV</td>
<td>novel NP</td>
</tr>
<tr>
<td>OBJ</td>
<td>object</td>
</tr>
<tr>
<td>PART</td>
<td>partitive</td>
</tr>
<tr>
<td>PAST</td>
<td>past tense</td>
</tr>
<tr>
<td>PL</td>
<td>plural</td>
</tr>
<tr>
<td>PRED</td>
<td>predicate</td>
</tr>
<tr>
<td>PRO</td>
<td>pronoun</td>
</tr>
<tr>
<td>PSV</td>
<td>passive</td>
</tr>
<tr>
<td>REL</td>
<td>relative pronoun/particle</td>
</tr>
<tr>
<td>SBJ</td>
<td>subjunctive</td>
</tr>
<tr>
<td>SUBJ</td>
<td>subject</td>
</tr>
<tr>
<td>SG</td>
<td>singular</td>
</tr>
<tr>
<td>TH</td>
<td>Theme</td>
</tr>
</tbody>
</table>
References


PART II

Impersonal constructions

Diachronic studies
Impersonal constructions and accusative subjects in Late Latin

Michela Cennamo
University of Naples Federico II

This paper explores the role played by impersonal constructions in the rise of active coding systems, with evidence from Late Latin. It is argued that the spread of accusative arguments from impersonal to personal structures (initially unaccusative, subsequently unergative and transitive) might stem from the ambiguity of voice forms resulting from the restructuring of the grammatical dimension of voice in the transition to Romance. This might have led to the personal reinterpretation of an impersonal pattern, with the original O argument reinterpreted as S_O.

Keywords: active systems; transimpersonals; accusative subjects

1. Introduction*

In this paper I investigate the role played by impersonal constructions in the use of the accusative in ‘subject’ function with one-argument verbs in Late Latin – so-called ‘extended accusative’ (Plank 1985) (e.g. crepitavit panem.acc ‘The bread crackled’) – attested initially with unaccusatives (nascitur contractionem.acc ‘A spasm arises’)

* I wish to thank Werner Abraham, Bernard Comrie, Romano Lazzeroni and Mair Parry for discussing some of the issues investigated in the paper. I am also grateful to the editors, Andrej Malchukov and Anna Siewierska, as well as to an anonymous referee, for their very interesting and enlightening remarks. All shortcomings are, of course, my own responsibility.

1. The following abbreviations are used:

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABL</td>
<td>ablative (case)</td>
</tr>
<tr>
<td>ACC</td>
<td>accusative (case)</td>
</tr>
<tr>
<td>ACT</td>
<td>active</td>
</tr>
<tr>
<td>AGR</td>
<td>agreement</td>
</tr>
<tr>
<td>AN</td>
<td>animate</td>
</tr>
<tr>
<td>DAT</td>
<td>dative (case)</td>
</tr>
<tr>
<td>F</td>
<td>feminine</td>
</tr>
<tr>
<td>FUT</td>
<td>future</td>
</tr>
<tr>
<td>GEN</td>
<td>genitive</td>
</tr>
<tr>
<td>GER</td>
<td>gerundive</td>
</tr>
<tr>
<td>HUM</td>
<td>human</td>
</tr>
<tr>
<td>IMPF</td>
<td>imperfect (tense)</td>
</tr>
<tr>
<td>IMPER</td>
<td>imperative</td>
</tr>
<tr>
<td>IMPERS</td>
<td>impersonal</td>
</tr>
<tr>
<td>IND</td>
<td>indicative</td>
</tr>
<tr>
<td>INF</td>
<td>Infinitive</td>
</tr>
<tr>
<td>INTR</td>
<td>intransitive</td>
</tr>
<tr>
<td>M</td>
<td>masculine</td>
</tr>
<tr>
<td>MPASS</td>
<td>(medio-)passive</td>
</tr>
<tr>
<td>NOM</td>
<td>nominative (case)</td>
</tr>
<tr>
<td>NEUT</td>
<td>neuter</td>
</tr>
<tr>
<td>-R</td>
<td>marker -R (in middle, passive or impersonal function)</td>
</tr>
</tbody>
</table>
and subsequently with unergatives (*ipsum currit* ‘He.acc runs’) as well as transitives (*fontem.acc colorem.acc mutat* ‘The spring changes its colour’) (Cennamo 2001, 2009, among others). I show that the (Late) Latin data offer interesting insights into possible paths through which active coding systems may arise, in the light of the current debate on the role played by impersonal constructions in this type of change (Creissels 2007; Donohue 2008; Malchukov 2008 and contributions in Donohue & Wichmann 2008).

This chapter is organized as follows. §2 illustrates the use of the accusative for non-O core arguments in Late Latin. §3 discusses how this interacts with the concomitant reorganization of voice distinctions. §4 explores the contribution of impersonal passives with an accusative argument to the use of the accusative in subject function in Late Latin. This is shown to spread from impersonal constructions to passive/intransitive clauses, probably owing to the interchangeability among voice forms (e.g. between passives and impersonals), consequent to the loss of the grammatical dimension of voice in the transition to Romance. Finally, §5 summarizes the conclusions.

2. ‘Quirky’ accusatives in Late Latin

A characteristic feature of Late Latin morphosyntax is the use of the accusative to mark core, non O arguments, as part and parcel of more general changes affecting the encoding of the argument structure of the clause in the transition to Romance, resulting in a deep reshaping of the marking of arguments and of the voice system. The present discussion focuses on two aspects of the restructuring of actancy and voice in Late Latin, namely the use of the accusative in subject function (§2.1) and the occurrence of accusative arguments in impersonal constructions (§2.2), i.e. constructions with a logically implied but unexpressed (A/S) argument.

2.1 Accusative subjects

A well-known change taking place in the domain of transitivity in Late Latin, as part of the loss of the case-system, is the use of the accusative in subject function,
well-attested by the 4th–5th century A.D. in various areas of the România. Early examples of this ‘quirky’ use of the accusative are reported for African execration tablets, dating back to the end of the 2nd–3rd century A.D. (Audollent 1967:387; Herman 1987:103–105, 1997:25), with intransitive verbs denoting change of state/location (cadere ‘fall’) (1a) and endoreflexives (vertere ‘turn’) (1b), featuring animate and human or nonhuman nouns of the second (Superstianus) (1b) and third (Victor) declensions (1a), alternating with canonical nominative subjects (1c).4 In these early attestations, however, the accusative argument might

as (i) **nominative-accusative** for the singular/plural of neuters (e.g. templum ‘temple’, cubile ‘bed’ ~ templa, cubilia), and for the plural of 3rd, 4th and 5th declension nouns (e.g. gentes ‘people’, manus ‘hand’, dies ‘day’), (ii) **genitive-dative** for 1st and 5th declension nouns in the singular (e.g. roae ‘rose’/diei), (iii) **dative/ablative** for the 2nd declension in the singular (e.g. domino ‘master’, templo) and for the 3rd declension in the singular with some nouns (e.g. cubili -DAT/ABL vs regi-DAT ~ rege-ABL), with formal identity between the two cases for all declensions in the plural (e.g. rosis, dominis, templis, cubilibus, fructibus ‘fruit’, diebus). Neuter nouns, therefore, only had a three-case system, namely (i) **nominative-accusative** (e.g. templum/templa), (ii) **genitive** (templi/templorum), (iii) **dative/ablative** (templo/templis) (examples and discussion from Haverling forthc). In Late Latin, owing to the conflation of phonetic/phonological and (morpho)syntactic factors (e.g. phonological mergers resulting from phonetic erosion of final, unstressed vowels (Barðdal & Kulikov 2009:471–472; Pei 1932; Haverling forthc, among others), and the rise of head-marked patterns of active alignment (Cennamo 2009), the case-system was gradually lost. It gave way to either a no-case system, as in southern areas (e.g. Italy) or a two-case system, with a nominative-accusative/oblique distinction, as in Gaul and Rhetia, and a nominative/accusative – dative/genitive opposition in Dacia, probably resulting from a preceding three-case system (nominative-accusative-oblique), with evidence from some pronominal systems (La Fauci 1997:37–53; Zamboni 1998:137–42, 2000:110–15; Ledgeway, 2011:462, note 69). In the intermediate stages of these two areally distributed case patterns, there seems to be evidence, however, for the spread of the accusative in nuclear and non-nuclear (oblique) function in all the România by the 6th century A.D. (Cennamo 2009), with different coexisting subsystems and quirky uses of case-endings.

3. The term refers to an intransitive pattern, derived from an originally transitive one, with the original animate O argument, the object, in S (subject) function, very frequent with body motion verbs (e.g. move, turn). In some languages (e.g. German, Turkish), endoreflexives show the same marker as anticausatives (see Note 5), the difference between the two structures consisting in the animacy/agentivity of the subject, whereby they are also referred to as ‘agentive anticausatives’ (Haspelmath 1987:27–29).

4. The use of the accusative with proper names in core and non-core function, is apparently a characteristic feature of African Latin (Pei 1932:209, and references therein), for which one cannot exclude Arabic influence (V. Bubenik, p.c.), an issue that needs further investigation.
also be extranuclear (Herman 1987: 103), conveying the pragmatic function of topic (Cennamo 2009):

(1) a. *Epafu* victore cadant … (intransitive)

\[
\begin{align*}
\text{Epafus.ACC} & \quad \text{Victor.ACC} \quad \text{fall.PRES.SUBJ.3PL} \\
\text{‘Let Epafus, Victor…fall…’} & \quad \text{(Def. Tab. 278A 3–6; Herman 1997: 25)}
\end{align*}
\]

b. *Superstianus*…. cadat, vertat… (agentive anticausative)

\[
\begin{align*}
\text{Superstianus.ACC} & \quad \text{fall.PRES.SUBJ.3SG} \quad \text{turn.PRES.SUBJ.3SG} \\
\text{servu} & \quad \text{cadat} \\
\text{servant.ACC} & \quad \text{fall.PRES.SUBJ.3SG} \\
\text{‘Let Superstianus …fall, turn, the servant fall’}
\end{align*}
\]

c. *Blandus* gemmatus … cadat

\[
\begin{align*}
\text{Blandus.NOM} & \quad \text{Gemmatus.NOM} \quad \text{fall.PRES.SUBJ.3SG} \\
\text{‘Let Blandus Gemmatus fall’} & \quad \text{(Def. Tab. 272, 9–11; Herman 1987: 104)}
\end{align*}
\]

In late subsequent (4th–8th–11th century) texts from other areas (e.g. Italy, Spain and even Gaul at a later period (5th–8th century A.D.)), quirky accusatives are well-attested in different types of intransitive structures, namely *equative clauses* (2), *fientives* (3) (patterns generally derived from an adjective, denoting the transition from a state to a new state (Haspelmath 1993: 34)), *anticausatives* (4a–b), *passives* (5), as well as intransitive one-argument verbs denoting telic change of state (6a–b)/location (6c), state (6d), non-telic change of state (e.g. the coming into being of an entity/event, as with *happen*, *arise*) (6e), that is, with unaccusative constructions. Only at a later stage (e.g. in 7th–9th century texts) accusative subjects also occur with unergatives, as illustrated in (6f–g), for non-agentive and agentive activity verbs, respectively (see Cennamo 2009 for details of the progression of the change along lexical-aspectual paths consistent with Sorace’s 2000, 2004 gradient approach to split intransitivity):

(2) *ustionem necessaria* res est (equative) \hspace{1cm} (Chiron 153) (IV A.D.)

\[
\begin{align*}
\text{burn.ACC} & \quad \text{necessary.NOM} \quad \text{thing.NOM} \quad \text{be.PRES.IND.3SG} \\
\text{‘A burn is necessary’}
\end{align*}
\]

(3) *et duriorem ventrem facit (=fit)* \hspace{1cm} (Chiron 890) (fientive)

\[
\begin{align*}
\text{and} & \quad \text{harder.ACC} \quad \text{belly.ACC} \quad \text{make.PRES.IND.3SG} \\
\text{‘And its abdomen becomes harder’}
\end{align*}
\]

\[5\] Anticausatives are intransitive structures whose inanimate subject is a derived S, corresponding to an original object (O); the process is presented as occurring spontaneously, without the intervention of a wilful animate causer (Siewierska 1984: 77; Haspelmath 1987: 15, among others).
(4) a. *multos languores sanantur* …
   many.acc illnesses.pl.acc/nom heal.mpass.pres.ind.3pl
   ‘Several illnesses heal …’ (Anton. Plac. *Itin*. 9; Corp. Christ. 165, 16)
   (anticausative)

b. *equae aspretudinem et ragadia in suffraginibus*
   mare.gen hardness.acc and rhagadis.pl.neut in knees.abl
   se ostendent (Chiron 609)
   ‘If the mare develops hardness and rhagades in its knees’

(5) *cutem … esocis … non manducetur* (Anthim. 41)
   skin.acc pike.gen not eat.pres.subj.3sg
   ‘The pike’s skin should not be eaten/one should not eat the pike’s skin’

(6) a. *si nascitur ei*
   if be-born.pres.ind.3sg he.dat
   genuorum contractionem (telic change of state)
   knee.gen.pl spasm.acc
   ‘If there arises a spasm of its knees’ (Chiron 516) (intransitive)

b. *clavum morticinum … si natum fuerit* (Chiron 615)
   corn.acc dead.acc if born.pp.m.sg.acc be.fut.perf.3sg
   ‘If a corn … arises’

c. *ut sanguinem in-or-der-to blood.acc*
   exeat copiosum (Chiron 618) (telic change of location)
   come-out.pres.subj.3sg abundant.acc
   ‘So that the blood comes out abundantly’

d. *lucem… caruit* (CIL VIII, 5372; Herman 1997: 323) (state)
   light.acc lack.perf.ind.3sg
   ‘The light failed’

e. *si inter eos …*
   if between they.acc
   causam advenirit (Lex Cur. II 1, 2) (coming into existence)
   dispute.acc happen.suby.perf.3sg
   ‘If a dispute arose between them’

f. *crepitavit panem in furno* (Agnell.175) (non-agentive activity)
   crackle.perf.ind.3sg bread.acc in oven.abl
   ‘The bread crackled in the oven’

g. *si… ipsum currit* (Lex Alam. XCIV codd. A) (agentive activity)
   if he.acc run.pres.ind.3sg
   ‘If he runs’
Interestingly, when the verb in the 3rd person singular passive voice agrees with the accusative argument, as in (5), the clause may be ambiguous between a passive (‘The pike’s skin should not be eaten’) and an impersonal (‘One should not eat the pike’s skin’) interpretation (see further discussion in §2.3, also Cennamo 2009).

In the case of one-argument verbs, accusative subjects are attested initially with core unaccusatives (e.g. telic change of state/location such as *nasci* ’be born’, *exire* ’come out’) (6a, c), and only at a later stage with unergatives (e.g. *currere* ’run’) (6g) (see further discussion in Cennamo 2009).

In late, 8th–9th century texts the accusative may also figure, albeit rarely, on the subject of a passive construction with overt expression of the Agent, which is realized through a prepositional phrase introduced by the prepositions *a/ab* + ablative, as exemplified by *ecclesiam* in (7a), and *nullam licentiam* in (7b). This pattern, in the *inf ectum*, when the argument is in the accusative case and the verb in the 3rd person singular (synthetic) passive voice, as in (7b), is partially reminiscent of analogous constructions with non-promoted Os and default verbal agreement in Slavic and Indo-Aryan⁶ (see Siewierska 1984: 106, among others):

(7) a. *ut ecclesiam Beati Apolinaris ab Iuliano in-order-to church.ACC blessed.GEN Apolinarius.GEN by Julian.ABL Argentario fundata et Argentarius.ABL found.PP.SG.F.NOM and consummata fuisset (Agnell. 63)

‘For the church of Saint Apolinaris to be founded and completed by Julianus’

---

6. In these languages, in fact, in the passive there occurs default verbal agreement if the O argument retains its original marking (e.g. accusative case in Ukrainian (Fici Giusti 1994:123) and the -ko suffix in Hindi (Siewierska 1984:106, among others):

(i) *cervku bulo zbudovano Lesevym* (Fici Giusti 1994:123)

church-ACC be-PST.NEUT build.PP.NEUT Lesev.INSTR

‘The church was built by Lesev’

(ii) *Rām-se rōti ko khā-yā gayā*

Ram by bread.F.3SG ko eat.PP.M. go.PFT.3SG

‘The bread was eaten by Ram’

Interestingly, in Marwari (Magier 1990:217) if the Agent is unexpressed, in the passive the non promoted object optionally agrees with the verb, as in (iii), which is analogous to the Latin example (7a), without the agentive phrase:

(iii) *darfī-ne māciyo giyo/mārijyō*

tailor-ACC killed went/kill-PASS-PST

‘The tailor was killed.’ (i.e. murdered)
Impersonal constructions and accusative subjects in Late Latin

b. *a nullo principe ei nullam licentiam detur*
   by no.ABL prince.ABL he.DAT no.F.ACC
   permission.ACC give.M.PASS.PRES.SUBJ.3SG
   ‘That no permission ought to be given him by a prince’

In 4th–5th century texts the accusative may also occur on the A argument of transitive clauses. This is attested initially with inanimate nouns (8a) and subsequently with animate ones (8b). This type of accusative marking occurs also in the Gallic inscriptions (6th century) and in Merovingian royal charters (8th century) (Pirson 1901: 189; Pei 1932: 214):

(8) a. *fontem vero … quater in anno*
   spring.ACC in-fact four-times in year.ABL
   *colorem mutat* (Per. Aeth. Excerpta; Rovai 2005)
   colour.ACC change.PRES.IND.3SG
   ‘The spring- in fact changes colour four times a year’

b. *filios et nepotes… fecerunt* (ILC 3052.B) (transitive)
   child.PL.ACC and nephew.PL.ACC make.PERF.IND.3PL
   [‘For Iulia Crescentia, for whom] her children and nephews
   made [the tomb]’

By the 4th–5th century A.D., therefore, there appears to have developed an active coding system, indicated through case-marking and, to a lesser extent, agreement, with adjectives and participles sometimes agreeing with the subject in the accusative case, like *duriorem* in (3) and *copiosum* in (6c), *natum* in (6b) – sensitive to the notions of animacy and control, whereby the inactive, inanimate S argument of unaccusative constructions (S_o) may be marked by the accusative case like the O argument of canonical transitive clauses. At this stage unergatives continue to pattern nominatively, with the S_A argument occurring in the canonical nominative case and agreement consistently patterning nominatively (Cennamo 2009: 329–331). The alignment of verbal arguments, however, is already moving towards a neutral encoding, as exemplified by the accusative marking of both inanimate A arguments and unaccusative S_o arguments, in the same texts, for example, the *Mulomedicina Chironis*, a veterinary treatise of the second half of the 4th century A.D. (see further discussion in Cennamo 2009 and references therein).

---

7. This notion, reflecting the degree of primary responsibility of a participant over the verbal process, is a multifactorial dimension determined by the interplay of a number of transitivity features such as the inherent characteristics of the participant(s), their relation to the predicate (i.e. thematic relations) and inherent lexical aspect (Lehmann 1988; Comrie 1989: 59–62).
This pattern is known in the typological literature as the extended accusative, following a term originally introduced by Moravcsik (1978) to refer to the extension of a morphological marker (the accusative case), generally marking the object (O) of a transitive verb, to encode the sole argument (S) of some intransitive verbs, mainly denoting mental process, involuntary actions and existence, well-attested cross-linguistically, both synchronically (Moravcsik 1978:241–54) and diachronically (Plank 1985, 1995; Burridge 1994:152–160, for middle Dutch). As first proposed by Plank (1985), the extended accusative in Late Latin seems to represent a transitional stage in the encoding of grammatical relations in the transition to Romance, one during which they are no longer identified on a nominative-accusative basis, but on an active-inactive one, with case-marking as the overt manifestation of their realignment: the patient/theme subjects of passives and of some intransitives are encoded like the object (O) of transitive verbs. This type of intransitive-centred change often takes place in the transition from accusative to ergative coding systems and vice versa, with intermediate stages conforming to an active-inactive patterning, and extension of the accusative/regression of the nominative (so-called extended accusative/restricted nominative alignment (Plank 1985)). The accusative generally occurs with inactive, patient (and/or theme) subjects and the nominative recedes to the coding of active, dynamic arguments of intransitive situations, that resist the realignment (cf. Plank 1985, 1995; Harris 1990:85–88; Dixon 1994:187–92; Harris & Campbell 1995:273–81). This picture, however, is only partially confirmed by the Late Latin and Romance data. In point of fact, although the basic lexical form of Romance nouns seems to support this interpretation, with inanimate nouns entering the lexicon in the accusative form (e.g. Italian cicatrice ‘scar’ (< Latin cicatricem.ACC), animate nouns enter the lexicon in either the nominative or the accusative forms, sometimes showing doublets (e.g. Italian ladro ‘thief’ < Lat. latro.NOM ~ ladrone (‘Latin latronem.ACC, French compain/compagnon (‘boy)friend/companion’)) (see Zamboni 2000; Lazzeroni 2002 2006; Ledgeway 2011 al.). Indeed, in Late Latin the nominative may also occur in O function, with both animate and inanimate nouns, as in (9):

(9) a. pulvis superaspargis
    powder.NOM sprinkle.PRES.IND.2SG
    ‘You sprinkle the powder’ (Orib. 860.1; Mørland 1932: 104)

   b. si pater percuserit
    if father.NOM beat.SUBJ.PERF.3SG
    ‘If he beats his father’ (Ed. Roth. 169: II; Löfstedt 1977: 215)

The patterns of accusative and nominative marking do not thus qualify as straightforward instances of active alignment. Rather, the distribution of accusative and nominative cases for core arguments in Late Latin seems to reveal the existence of different coding subsystems, an initial active one, a subsequent neutral one and possibly even
an ergative alignment system, the last emerging once all S arguments are marked by
the accusative (Zamboni 1998, 2000: 114–115, Cennamo 2009 for a more detailed
discussion and further references).

2.2 Other patterns with accusative arguments in Late Latin
and their Early Latin antecedents

2.2.1 Impersonal constructions
A characteristic morphosyntactic development in Late Latin, attested in 4th–5th cen-
tury texts, is the use of the accusative for the O/SO argument of impersonal, existential-
impersonal constructions. In such patterns the verb is in the unmarked 3rd singular,
either of the active (10a,g) or of the passive voice (10b–e). In the latter voice one finds
the synthetic -\(R\) form\(^8\) in the infectum, with its endings added to the verb stem, as in
(10b–d), and a form of the verb ‘be’ (esse) + neuter pp in the perfectum, as in (10e).
Also in the gerundive the verb ‘be’ (esse) is in the default 3rd singular, and the verbal
adjective in the neuter singular form, as in (10f). In these constructions the verb often
figures with a pre/postverbal non-agreeing argument (in case and/or number), con-
veying either given (10b, f) or new information (10a, c), although this is not always the
case, as in (10d, e), where the postverbal argument conveys given information:

(10)  a. facit (=fit) hiemen
take.pres.ind.3sg winter/cold.acc
‘It is cold’ (lit. makes cold/winter. acc)
b. ipsos ficos imponatur
these.acc figs.acc gather.pres.subj.mpas.3sg
‘One should gather these figs’
c. faciatur... tricinia
make.mpas.pres.subj.3sg banquet.pl.acc
‘So that banquets are made’
d. ne baptizetur cos
not baptize.mpas.pres.subj.3sg they.pl.acc
‘That one should not baptize them’
e. cum factum fuerit missam (impersonal)
when make.pp.neut.sg be.fut.perf.3sg mass.acc
‘When the Mass is over’ (Per. Aeth.32, 2)

---

\(^8\) The -\(R\) suffix in Latin may be viewed as the marker of the ‘non-active’ voice (Cennamo
1998: 78), covering different types of (in)transitive patterns, such as middles, anticausatives,
passive and impersonal. It is traditionally regarded in the literature as being originally either
an impersonal (Ernout 1908–1909: 273–279, among others), or a (medio)-passive suffix,
having primarily a ‘passive’ function in a synchronic grammar of Latin (Baldi 1977; Kurzová
1993 and references therein).
f. *de carne* *putras*
   about *meat.pl abl stew.pp.pl.f.acc*
   
   factas … *utendum*
   make.pp.pl.f.acc *use.ger.neut*
   ‘As for meat, it should be be eaten stewed’

g. *in Hebraeo* … *non habet*
   in *Hebrew abl not have.pres.ind.3sg*
   *hunc numerum…*
   *this.acc number.acc*
   ‘In Hebrew this number does not exist’ (lit. not has this number.acc)

These structures further develop already existing, albeit rare (11a) and sometimes philologically uncertain/controversial early Latin forms (11b–c) (Rovai 2007), well-attested instead in the gerundive (11d):

(11) a. *vita vitatur*
   *life.acc live.mpass.pres.ind.3sg*
   ‘One lives life’

b. *me … despicatur*
   *1.acc despise.mpass.pres.ind.3sg*
   ‘I am despised/one despises me (lit. (it) is despised me)’

c. *nilne te populi vereretur?*
   *not at all you.acc people.gen fear.mpass.subj.impf.3sg*
   ‘Don’t you have any respect for the audience?’
   (lit. ‘does it not respect you of people?’)

(Atta, *com. 7*)

d. *(ut) vasa vinearia et*
   *in-order-to containers for-wine.pl.neut and oil.pl.neut make.ger.neut.sg*
   ‘In order to make containers for wine and oil’

(Varro. *XIII, 1*)

As already pointed out (§2.1), when the verb in the impersonal passive form agrees in number/gender with the argument in the singular accusative case, as in (11a), the pattern may be ambiguous between a passive reading, ‘life is lived’ and an impersonal (generic, existential) interpretation, ‘one lives life’ (see further discussion in Pinkster 1992; Cennamo 2008b; Napoli 2009 and references therein). These constructions are quite common in Late Latin, with the argument occurring in pre/postverbal position, conveying given/new information, the preverbal position often being associated with a nominal conveying given information and favoring a passive interpretation, as in (12):
2.2.2 Accusative in nominal clauses

In Late Latin the accusative often occurs in exclamations, listings, recipes, appositions and commands, i.e. in nominal clauses, sometimes with omission of an implied predicate, as in (13) (cf. Svennung 1935: 187–88; Gerola 1950: 212–13; Norberg 1943: 93–95; 1944: 31–32):

(13) a. *aquam foras, vinum intro* (Petron. 52)
   water.ACC out wine.ACC in
   ‘(Take the) water out (but take the) wine in’

b. *potionem ad eos, qui sanguinem meient* (Chiron 822)
   potion.ACC to them.ACC who.NOM blood.ACC urinate.FUT.IND.3PL
   ‘(One should give) a potion to those who urinate blood’

This usage develops already existing uses of the accusative in this low transitivity domain, in which it alternates with the nominative in the case of listings and appositions. These patterns are interpretable at times as equative structures with omission of the copula, as in (14) (see further discussion in Cennamo 2009: 308–309):

(14) a. *me infeliciem et scelestam* (Plaut. Cist. 685)
   I.ACC unhappy.ACC and dreadful.ACC
   ‘What an unfortunate scoundrel I am’

b. *manum de tabula!* (Cic. fam. 7, 25, 1)
   hand.ACC from table.ABL
   ‘Take your hand off the table!’

2.2.3 Accusative in topic function

In 4th and 5th century technical texts, and even more frequently at a later time, there also occur several examples of topics in the accusative, at times ambiguous between a purely pragmatic and a syntactico-pragmatic function, as in (15):

(15) *Herbam quae Gallice dicitur blutthagio* (Marcell. 9, 132; Svennung 1935: 186)
   herb.ACC which.NOM Celtic say.PRES.IND.MPASS.3SG blutthagio
   nascentur locis umidis eam teres born.PRES.IND.MPASS.3SG place.ABL humid.ABL it.ACC tear.PRES.IND.2SG
   ‘The herb which is called blutthagio in Celtic, grows in humid places, you tear it’
In (15) *herbam* is both the topic of the sentence, occurring as such in initial position, and the subject of the verb *nascitur*, as shown by verb agreement, and is cross-referenced by means of the resumptive pronoun *am* ‘it’, which refers back to it. Indeed, it might be claimed that the grammaticalization of a topic argument in the accusative might be the possible channel of the spread of the accusative to an argument in A function (see further discussion in Cennamo 2009: 328–329).

2.3 Interim summary

In Late Latin the accusative gradually comes to mark non-O core arguments, initially S_O, subsequently S_A and later A. In particular, it appears to spread from peripheral transitivity domains such as nominal clauses and impersonal constructions with an expressed argument, to more canonical intransitive and then transitive structures: equative clauses with expression of the copula, other intransitive patterns such as fientives, anticausatives, passives, unaccusative verbs, as well as unergatives and transitives (later development). It is not clear, however, what triggers this change, and how accusative subjects arise. I argue that impersonal constructions might have played a role in the change, interacting, in the course of time, with the reorganization of voice distinctions, well-attested by the 4th century A.D. In point of fact, the ambiguity resulting from the interchangeability among voice forms (e.g. between passives and impersonals), might have led to the personal reinterpretation of an impersonal structure, with the original O argument reinterpreted as S_O (see discussion in §4).

3. Voice and accusative subjects in Late Latin

The use of the accusative in subject function is usually viewed in the literature as resulting from confusions in the planning of discourse, with ensuing contaminations between active and passive clauses: an initially active transitive clause with the object (O) in the accusative case may end – often interspersed with subordinate clauses – as a passive construction, with the original object in “subject” function (Löfstedt 1911: 290–293; Norberg 1944: 22–24; Plank 1985: 288–289), a phenomenon already occurring in classical authors (e.g. in Cicero’s philosophical writings (Norberg 1944: 21–22));

9. In the discussion I do not consider deponents, to which the canonical rules of Latin do not apply, since the ‘passive’ form always signals an A, S and sometimes O argument as subject (see Cennamo 1998, among others).
Impersonal constructions and accusative subjects in Late Latin

(16) a. stupeo vos … Gesalecum… in vestram defensionem
   amaze.pres.ind.1sg you.acc Gesalecus.acc in your.acc defence.acc
   sic fuisse susceputum (vs suscepisse)
   thus be.past.inf take.pp.m.sg.acc take.past.inf.act
   (Cassiod. var. 5, 43, 2; Norberg 1944:22)
   ‘I am amazed by the fact that you have taken Gesalecus in your defence’

In (16a), for instance, in the subordinate clause the verb is in the passive voice, fuisse susceptum, rather than the active form, suscepisse. Also the converse phenomenon is attested, as in (16b), where an initially passive clause with the Agent (A) expressed as a prepositional phrase, a decessoribus suis, ends with an active infinitive, dedisse, rather than the passive data esse (see also Plank 1985; Cennamo 2009:331–332):

(16) b. Iocundus presbyter… nobis… petitione suggestit,
   locundus.nom priest.nom we.dat request.perm.ind.3sg
   a decessoribus suis… sacra ministeria …
   by predecessor.pl.abl his.pl.abl sacred.pl.acc function.pl.acc
   Albinus cuidam … dedisse (vs data esse)
   Albinus.dat certain.dat give.act.perm.inf (vs give.pass.perm.inf)
   ‘The priest Iocundus submitted to us the request that the sacred services be performed by a certain Albinus’ (Pelag. Epist. Pontif.. 39: Norberg 1944:22)

Contrary to the above, my contention is that the rise of accusative subjects does not reflect uncertainties in the planning of discourse, but is related to the reorganization of voice distinctions taking place in Late Latin, already perceivable in technical texts from the second half/end of the 4th century A.D. (e.g. veterinary treatises such as the Mulomedicina Chironis) (Cennamo 1998, 2006). In particular, the functional domains of voice forms and the dividing lines among them are no longer clearcut. In point of fact the active intransitive, the passive and the reflexive pattern are often used interchangeably, although continuing to occur in their canonical functions. In (17), for instance, fieri ‘become/be done/made’, the lexical passive of the verb facere ‘do/make’, equals the active reflexive pattern in anticausative function, facere se ‘become’, as well as the plain active form facere:

(17) a. tumor fit = se facit
   swelling be-made/arise.pres.ind.3sg refl make.pres.ind.3sg
   = facit
   make.pres.ind.3sg
   ‘A swelling occurs/arises’

The equivalences and confusions among voice forms involve all their domains, and in 5th–6th century texts there also occur ‘quirky’ uses of the active in truly
passive function, sometimes with overt expression of the A argument, as illustrated in (18), where the active form vexaverit ‘will have oppressed/burdened’ stands for the (expected) passive form vexatus erit ‘will have been oppressed/burdened’ (cf. Cennamo 1998, 2006):

(18) \textit{item si a rota vexaverit (sc. equus)} \hspace{1cm} (Pelag. 233; Feltenius 1977:137) 
then if by wheel.ABL trouble.PERF.FUT.3SG (horse) 
‘If it (= the horse) will be troubled by the wheel’

There also occurs the converse phenomenon, whereby the passive voice (i.e. the -R suffix in the infectum (19a) and esse + pp in the perfectum (19b)) figures in active function, so-called Deponentization (Flobert 1975), viewed traditionally as a sign of the vitality of the category of deponents, that in Late Latin spreads to all verbs (both intransitive and transitive) (see further discussion in Cennamo 1998 and references therein):

(19) a. \textit{lumbi ei vacillantur} \hspace{1cm} (Chiron 478) 
\hspace{1cm} loin.PL.NOM he.DAT toter.MPASS.PRES.IND.3PL 
‘His loins vacillate’

b. \textit{optati sumus ire} \hspace{1cm} (Per. Aeth. 10, 9) 
\hspace{1cm} ask.PP.M.PL.NOM be.PRES.IND.1PL go.PRES.INF 
‘We have asked to go’

The ‘wrong’ uses of voice forms illustrated in (18)–(19) above, signal the temporary loss of the grammatical dimension of voice in the transition to Romance, that culminates in the abandonment of the passive as a strategy, often replaced by the active, as attested in Merovingian Latin (Herman 2002, also Svennung 1935:460, for earlier periods). Only occasionally there occur alternative strategies such as the analytic passive forms esse + pp (both in the infectum and the perfectum), other verbal periphrases (e.g. fieri ‘become’/venire ‘come’ + pp) and the reflexive passive, that by the 8th century, however, do not appear to be organized in clearcut paradigms yet (Herman 2002). Up until the beginning of the 8th century, in fact, very few examples of the esse + pp pattern in imperfective function are attested, in place of the canonical synthetic passive realized by the -R form. Thus the form laudor ‘I am (being) praised’ is only rarely replaced with laudatus sum (lit. ‘praised.PP.M.SG I-am’). In early and Classical Latin the latter could have either a resultative stative ‘I am (a) praised (man)’ or a perfective, eventive-resultative reading ‘I was praised/I have been praised’ (see discussion in Cennamo 2005, 2006, among others).

The interchangeability among voice forms discussed above may be regarded as the surface manifestation of changes in argument linking, and more generally in the encoding of the argument structure of the clause in the passage to Romance (see Cennamo 2009). In particular, the use of the passive in active function and of the
active in passive function signals a violation of the canonical rules for the assignment of grammatical functions to the arguments of a verb, so called linking rules. There occurs in fact an exchange in the markedness relationship between clauses marked with the active and the passive voice, so that the active no longer consistently signals an A participant as subject and vice versa the passive morphology no longer univocally indicates an O participant in subject function, as in canonical active and passive clauses respectively. It is still possible, however, to identify the syntactic status of verbal arguments through case-marking and agreement. Once case-marking (and later and to a lesser extent agreement too) also operate on an active-inactive and then neutral basis – as testified by the accusative in subject function, initially with unaccusatives, and subsequently with transitives and unergatives – the balance existing within the system gets altered. The restructuring, in fact, affects the identification of the role of the verbal arguments (i.e. their A/O status) and no longer only the rules linking arguments to their syntactic function. This state of affairs is exemplified in (20):

(20) a. calyculos… sale aspersos
    mushrooms.acc salt.abl sprinkled.acc
    inferunt (= inferuntur)
    serve.pres.ind.3pl serve.mpas.pres.ind.3pl
    ‘Mushrooms are served sprinkled with their own juice or with salt’
    (Apic. VII, 13, 5; Rovai 2005)

b. cataplasmabis eum (= tumorem), donec
    smear.pres.fut.2sg it.acc (= swelling) until
    maturum faciat
    soft.acc make.pres.subj.3sg
    ‘Smear it (=the swelling) until it becomes soft’
    (Chiron 91)

c. si a rota vexaverit (sc. equus) (Pelag. 233; Feltenius 1977: 137)
    if by wheel.abl oppress.fut.ind.3sg
    ‘If it (sc. the horse) will be oppressed by the wheel’

In (20a), for instance, the verb is in the 3rd person plural of the present indicative, and the argument is in the accusative case. Calyculos, however, is not the object of the verb inferunt, but its subject. The interpretation of the pattern, however, can only be inferred from the wider context. The same is true of (20b), where the active facit ‘(it) makes’ stands for the (lexical) passive fit ‘(it) is made/it becomes’. Out of context, the clause may be interpreted as involving an unexpressed 3rd person singular subject softening a swelling (active transitive reading ‘it/he makes it ripe’), as well as an intransitive active (fientive) reading, with the argument in the accusative functioning as subject (‘it becomes ripe’), and even a passive interpretation, ‘it is made ripe’ (see also discussion in Cennamo 2006, 2009).
The examples illustrated in (18)–(20) above show the functional ambiguity of voice forms, stemming from the opacity of voice morphology and of the syntactic structures encoding the linking of arguments to their grammatical function.

The ambiguity of interpretation of a passive or active form, therefore, comes to involve both the assignment of grammatical functions and the identification of the status of verbal arguments. The same argument, then, may be interpreted as either O or S_O, depending on the context. This ambiguity might have facilitated the reinterpretation of a grammatical object (O) as an S_O argument.

4. Impersonal constructions and the rise of split S systems

Recent work on semantically aligned systems has put forward the hypothesis that agentive systems might arise from the reanalysis of transitive impersonal clauses as intransitive, with the original objects reinterpreted as “grammatical patients” (Mithun 2008: 319, 332; Holton 2008; Malchukov 2008) along the following scheme (from Mithun 2008: 309):

\[
\begin{align*}
\text{(S UBJECT)} & \quad \text{OBJECT} & \quad \text{TRANSITIVE VERB} & \quad \text{‘It scared me (object)’} \\
\text{me} & \quad \text{scared} & \quad \text{P A TIENT} & \quad \text{INTRANSITIVE VERB} & \quad \text{‘I (PATIENT) was scared’}
\end{align*}
\]

The precondition for the reanalysis is “zero 3rd person agent affixes”, i.e. the lack/non expression of topical 3rd person agents and the formal identity between transitive and intransitive verbs, that is, the lack of an overt marker of transitivity (Malchukov 2008: 83, 86; Mithun 2008: 331–332).

The Late Latin data seem to give evidence for the path of development proposed in (21). The chronology of accusative subjects, in fact, appears to support the hypothesis that impersonal patterns with an accusative argument, already attested in Early Latin (though rarely in non-gerundive forms) might have been one of the paths along which an active-inactive system marked though case-marking arose in Late Latin. In point of fact, owing to the (temporary) loss of the grammatical dimension of voice, whereby voice forms become interchangeable and no consistent formal differentiation occurs between transitive and intransitive clauses, either active or passive, a pattern with the verb in the passive voice and an accusative argument may be ambiguous between an active transitive/intransitive (anticausative), an impersonal or a passive interpretation. The argument in the accusative case collocating with the verb in the ‘passive’ morphology and no overt subject, therefore, may be interpreted as either S or O, depending on the context. Thus a pattern such as (22), lapide(m) non revolvatur, from a 6th century Gallic inscription (Pirson 1901: 189), may have four different readings out of context:
Impersonal constructions and accusative subjects in Late Latin

(22) lapide(m) non revolvatur
the-gravestone =O not turn-over.MPASS.PRES.SUBJ.3SG
verb: 3rd sing. impersonal passive form + accusative argument
a. impersonal: 'one should not turn the gravestone over', lapidem = O, object
b. passive: 'the gravestone should not be turned over', lapidem = SO, subject
c. intransitive (anticausative): 'the gravestone should not turn over', lapidem = SO, subject
d. transitive: 'He should not turn the gravestone over', lapidem = O, object.

The ambiguity resulting from the interchangeability among voice forms (e.g. between passives and impersonals of transitive verbs) might have facilitated the personal reinterpretation of an impersonal structure, with the original O argument, the grammatical object, reinterpreted as an SO. Patterns such as (22), therefore, might have been the hinge between the accusative in impersonal patterns, already attested in early Latin (e.g. vitam vivitur) and the Late Latin usage of the accusative in intransitive (personal) structures (e.g. nascitur contractionem).

5. Conclusions

The analysis of some aspects of the restructuring of the encoding of the argument structure of the clause in the passage to Romance has revealed the existence of a semantically based alignment in Late Latin, sensitive, initially, to the notions of animacy and control, manifested by the occurrence of the accusative in subject function with the inactive arguments of unaccusative structures.

Impersonal constructions with accusative arguments might have played a role in the use of the accusative in SO function, interacting, in the course of time, with the reorganization of the voice system, determining the spread of the accusative from impersonal structures to passives and intransitives.

Indeed, the temporary loss of the grammatical dimension of voice, with the lack of a consistent formal marking of transitivity, might have been one of the paths through which a semantically aligned system coded through case-marking arose in Late Latin.

References


Ledgeway A. 2011. Syntactic and morphosyntactic typology and change in Latin and Romance.


Malchukov, A. 2008. Split intransitives, experiencer objects, and ’transimpersonal’ constructions: (Re-)establishing the connection. In Donohue & Wichmann (eds), 76–100.


Pirson, J. 1901. La Langue des Inscriptions de la Gaule. Bruxelles: Bibliothèque de la Faculté de Philosophie de l’Université de Liége, XI.


From passive to impersonal
A case study from Italian and its implications*

Anna Giacalone Ramat¹ & Andrea Sansò²
¹Università di Pavia/²Università dell’Insubria, Como

In contemporary Italian, there is a passive *-construction in which the patient governs agreement on the verb and an impersonal *-construction in which either the verb is intransitive or the patient is not promoted to subject. The coexistence of the two constructions is the result of a long-lasting process by which an impersonal has developed out of (and has progressively differentiated itself from) an original passive. In this paper we focus on the initial stage of this process, namely the extension of the *-construction to intransitive verbs, and the emergence of the non-agreeing pattern with transitive verbs. Based on a large corpus of literary and non-literary documents, we argue that both these phenomena require a reanalysis of * as a marker of generic human agency as a necessary precondition, and that such a reanalysis starts with patients that are unlikely candidates for subjecehood.

Keywords: reanalysis; impersonal passive; non-promotional passive; impersonalization of passive constructions

1. Introduction: From passive to impersonal

Impersonal constructions derived from passive constructions are known from a number of languages (Polish, Frajzyngier 1982:272–275, Siewierska 1988; Icelandic, Maling & Sigurjónsdóttir 2002; German, Abraham & Leiss 2006; Vogel 2006). The following examples from Italian, German, and Polish all exemplify constructions displaying

* This article is the result of joint work by the two authors. Andrea Sansò is responsible for §1 and 3, while Anna Giacalone Ramat is responsible for §2 and 4. §5 has been written jointly by the two authors. Both authors, however, subscribe to the general ideas presented in the article. We wish to thank Anna Siewierska, Andrej Malchukov, Elisa Roma and an anonymous referee for valuable discussions of the ideas presented in this paper. The usual disclaimers apply.
typically passive morphology (either a past participle – as in (2) and (3)\(^1\) – or a reflexive marker, as in (1)), but in which the patient remains unpromoted to subject:

(1) Italian

\[
\text{In Italia si mangia spaghetti} \\
\text{in Italy si[REFL] eats spaghetti[PL]} \\
\text{‘In Italy, people eat spaghetti (it is usual to eat spaghetti)’}
\]

(2) German

\[
\text{Heute abend wird richtig Zähne geputzt} \\
\text{tonight becomes properly teeth brushed} \\
\text{‘Tonight you should brush your teeth properly’}
\]

(3) Polish

\[
\text{Dano mu ksiazke} \\
given:NEUT he:DAT book:ACC.SG.F \\
\text{‘He was given a book’}
\]

All these constructions are also possible with intransitive (or intransitively-used transitive) predicates, as in the following examples:

(4) Italian

\[
\text{Qui si lavora troppo!} \\
\text{here si[REFL] works too much} \\
\text{‘Here people work too much!’}
\]

(5) German

\[
\text{Es wurde getanzt} \\
\text{it became danced} \\
\text{‘There was dancing (going on)’}
\]

(6) Polish

\[
\text{Tutaj tanzono} \\
\text{here danced} \\
\text{‘There was dancing here.’/ ‘People danced here.’}
\]

Both the constructions in (1)–(3) and the constructions in (4)–(6) are generally subsumed under the same rubric of “impersonal passives”, to be intended as constructions in which “the predicate is associated with passive morphology” (past participle, reflexive marker/suffix, etc.; Abraham & Leiss 2006: 509), but in which either there is no patient, i.e. the predicate is intransitive, or the patient does not fill the subject position. Under the subject-based view of impersonality proposed by Siewierska (2008: 116ff.),

\(^1\) -no and -to are petrified neuter singular forms of the so-called nominal declension of the past participle. The neuter form of the past participle in Polish is now -ne/-te.
however, these constructions are all instances of impersonal constructions to full right. Following Siewierska (2008:116), indeed, an impersonal construction can be defined as one which lacks a canonical subject, i.e. “a verbal argument which is fully referential and manifests the morphosyntactic properties of subjects in a language”. This definition includes (i) constructions in which the subject is not fully referential (e.g. man-constructions in German, vague they in English, and comparable constructions in pro-drop languages); (ii) constructions in which the subject is identified (in one way or another) by means of special morphology, e.g. by means of an originally reflexive marker (as in the case of the so-called Romance and Slavic reflexive impersonals; Siewierska 2008:117); (iii) constructions with overt expletive subjects (as the so-called “impersonal passive” of German and Dutch), (iv) constructions that lack an overt subject at all (as the no/to construction in Polish).  

There appears to be little known about the emergence of the impersonal constructions exemplified in (1)–(3), and their relationship with both “promotional” passives and the “impersonal passive” of intransitive verbs exemplified in (4)–(6). Although the nature of this process remains “an open question” (Abraham & Leiss 2006:511), it is generally stated that the constructions in (1)–(3) represent a rather late development, and the diachronic mechanisms involved in this development are generally described as including two steps:

i. a promotional passive construction/marker starts being used with intransitive verbs, or with intransitively-used transitive predicates;

ii. on the basis of analogy with these intransitive counterparts, the construction/marker ceases to be perceived as passive even when the verb is transitive, and acquires impersonal traits.  

2. Passive constructions, on the other hand, are minimally defined as constructions in which a non-agent argument has been promoted to subject (or, at least, has taken on a subset of the morphosyntactic properties of subjects in a given language, e.g. it controls agreement with the verb) and the verb is morphosyntactically derived in some way from the form used in the unmarked active voice construction (Haspelmath 1990). The motivation for adopting a purely structural definition of impersonal and passive constructions has to do with the nature of the process of change described in this paper, which involves the emergence of an impersonal construction out of an originally passive construction. Under a different, functionally-driven view, both passive and impersonal constructions serve the function of defocusing the main initiator of the event (Siewierska 2008:122), and although it is possible to distinguish between different types of agent defocusing (Myhill 1997; Sansò 2006), these distinctions only loosely correspond to the structural divide between passive and impersonal constructions (see the discussion in Sansò 2006:265–267).

3. This diachronic sequentiality (impersonal passives of intransitive verbs precede impersonal passives of transitive verbs) is also mirrored by the cross-linguistic synchronic
This process may result in a complete “impersonalization” of the formerly passive construction, which no longer exists as such, or it may be the case that the passive construction continues to exist along with its impersonal “offspring”.

Semantically, this process may involve a change in the meaning of the passive marker/construction: whereas in promotional passives the passive marker usually signals a switch in the relative prominence of agent and patient (triggered, e.g. by the topicality/discourse-relevance of the patient), when the same marker is used with intransitive verbs it develops a different function, namely that of expressing that the action denoted by the verb has a loosely specified human agent. The same function can be assigned to the formerly passive marker in the impersonal constructions exemplified in (1)–(3), in which there is a patient that remains unpromoted to subject. This semantic change sets the so-called impersonal passives apart from passive constructions: while the function of passive markers (and their various diachronic sources: reflexives, anticausatives, resultatives, etc.) has generally to do with the affectedness of the patient (signalling, e.g. the resulting state of the patient, as in resultatives, or its being affected by the verbal action, as in anticausatives and middles), the function of impersonal markers derived from passive markers is simply to signal that the agent is generic and human, without any implications of patient affectedness.

From a general point of view, step (i) is usually thought of as an instance of extension of a construction to new contexts, while step (ii) can be characterized as a process of reanalysis without grammaticalization, as “no particular element … become(s) more grammatical(ized) as a result of the change, i.e. no element changes from a lexical item to a grammatical item, and the whole construction does not necessarily become tighter” (Haspelmath 1998: 325, adapted).

Frajzyngier (1982:274ff.) posits these two steps to explain the emergence of the impersonal no/to construction in Polish, which was only possible with intransitive verbs in the oldest Polish texts (Frajzyngier 1982:275, quoting Brajerski 1979). Siewierska (1988:266), on the other hand, invokes such a path to account for the distribution of these two construction types: as Kazenin (2001:905) states,“(1) no language has impersonal passive of transitives without having (impersonal) passives of intransitives; (2) no language has (impersonal) passive of intransitives without having some type of passive of transitives”.

4. It should be remarked, however, that according to Siewierska’s (1988:271, our emphasis) reconstruction, “impersonal passives with the no/to participle were formed [in Old Polish, AGR-AS] mainly, but not exclusively from transitive verbs, in Russian (sic) predominantly from perfective verbs. Intransitive verbs stopped being used in this construction in literary Russian […] around the eighteenth century, while in Polish they began to be more widely used, and nowadays occur with the same frequency as transitive verbs”.

development of the reflexive impersonal construction in Polish. In this construction, illustrated in (7), the patient appears in the accusative rather than in the nominative case and the verb is invariably third singular:

(7) Polish

Traktuje się go jako malarza autentycznej natury ludzkiej

\[\text{treats:3SG REFL he:ACC as painter authentic nature human}\]

‘He is treated (one treats him) as a painter of the real people’

Similarly, Maling & Sigurjónsdóttir (2002:101; see also Maling 2006; Eythórsson 2008) consider the “new impersonal construction”\(^5\) of present-day Icelandic as representing a case of reanalysis of the canonical passive morphology from passive to syntactically active (i.e. impersonal). Unlike the canonical passive in which the direct object is moved to subject position (see (8a)), in the “new impersonal construction” exemplified in (8b) “the null pro […] is an external argument which gets interpreted as an ‘unspecified human subject’”, and the direct object retains object marking:\(^6\)

(8) Icelandic

\(a.\) Stúlkan var lamin í klessu

the.girl:NOM was hit:FSG.NOM in a.mess

“The girl was badly beaten”

\(b.\) Það var lamíð stúlkuna í klessu

it[EXPL] was hit:NEUT.SG the.girl:FSG.ACC in a.mess

“The girl was badly beaten”

The explanation proposed by Maling & Sigurjónsdóttir for this reanalysis, however, is somewhat different from the two-step process described above. In their view, the process has been fostered by the fact that only in Icelandic among Scandinavian languages

---

5. This construction is rapidly spreading in the language of young Icelanders, and its acceptance rate is significantly lower in Inner Reykjavik than in other parts of the country (Maling & Sigurjónsdóttir 2002:109ff.), possibly due to normative pressure. The German construction *Es wird Bücher gelesen* is often cited as a parallel to the new Icelandic construction, but differs from it in several crucial ways. For instance, the German construction seems to allow mainly determinerless, and more generally indefinite and inanimate nouns as patient arguments, while the Icelandic construction also allows animate and definite nouns.

6. The expletive subject Það is not a grammatical subject, but serves only to satisfy the V2 constraint (Maling 2006:198). As such, it does not appear, for instance, in yes-no questions:

\(i.\) Var [*Það] beðið þig að vaska upp?

was [it] asked you:ACC to wash up

‘Did they ask you to do the dishes?’
the promotional periphrastic passive is restricted to [+human] agents, and this feature has set the stage for the reanalysis of a thematically empty null subject as a generic human agent. Once reanalyzed, the construction spread first to intransitive verbs and then to transitive verbs via inherently reflexive predicates.\footnote{Vogel (2006:96ff.) puts forward a model for the development “passive > impersonal” in which the emergence of impersonal passives of intransitives is kept apart from the emergence of impersonal passives of transitives. In her view, the emergence of impersonal passives of intransitive verbs (es wurde getanzt) presupposes the existence of passives in which there is a generalized patient (“generalisiertes Subjekt”, as in German gerade wird gegessen, ‘people are eating right now/there is eating going on’, in which a transitive predicate is used intransitively and the patient remains unspecified). The emergence of impersonal passives of transitive verbs presupposes the existence of passive constructions with “rhematic” subjects, which are typically indefinite (es wurden viele Grundsteine gelegt). Both passive constructions with subjects low in topicality and passive constructions with a generalized patient focus on the event rather than on the subject (in Vogel’s terms, they are “event-zentral” rather than “entity-zentral”). As will be discussed later, however, the two structures from which impersonal passives are supposed to arise (passives with subjects low in topicality and passives with generalized patients) share a significant property, i.e. they have patient arguments that represent unlikely candidates for subjecthood, so that it is not strictly necessary to posit two different paths. Moreover, if the two paths of development were independent from one another, there would be no way to account for the cross-linguistic distribution of impersonal passives: given the fact that impersonal passives of transitive verbs always imply the existence of impersonal passives of intransitive verbs across languages, it is more economic to consider the two developments as subsequent steps of a single process.}

Those who postulate a two-step evolution from passive to impersonal have never discussed why passive morphology should extend to intransitive verbs. As discussed above, passives and impersonals are semantically so markedly different that we must be cautious when postulating a mere extension of a construction to new contexts: impersonal constructions typically have a generic human agent, whereas passive constructions are in general not characterized by any semantic restriction concerning the agent. As Abraham & Leiss (2006:502, adapted) correctly observe, impersonal passives “do not involve any passive semantics … [and] the demoted subjects of these constructions carry the features [+AGENT], [+HUMAN]”. In other words, while in (9b) the intended agent may be only a generic person (= one) or a loosely specified set of humans, in (9a) it can coincide with a specific person the speaker does not want to mention, with a generic person, and even with a natural force.

(9) Dutch (Kirsner 1976: 387–388)

a. De huizen werden verwoest
    the houses became destroyed
    ‘The houses were destroyed’ (by the enemies/by the hurricane etc.)
Following Comrie (1977; see also Kirsner 1976; Shibatani 1985), one could explain the extension of passive constructions to intransitive verbs in terms of “spontaneous demotion”, i.e. as an instance of defocusing not caused by the promotion of some other argument to subject. Such a hypothesis enjoys wide currency in the literature on “impersonal passives”. If this hypothesis is accepted, a passive marker would have a very abstract function, that of defocusing the main participant (i.e. the agent), and the meaning of this marker when used with different verb types would be the result of a process of accommodation: if there is another participant, the defocusing of the agent is only relative, i.e. it depends on the higher topicality of the patient; on the contrary, in impersonal passives the defocusing of the agent is absolute, which is tantamount to saying that the agent is not particularly noteworthy, representing virtually all humanity, or a subgroup thereof (e.g. people in a given location).

Based on the history of the Italian *si*-construction, in this paper we will argue in favour of a context-based reinterpretation of a passive marker as a marker of generic human agency. This reinterpretation precedes and motivates its use with intransitive verbs, which cannot be simply thought of as a mere extension of a passive marker to new contexts. In other words, we will not postulate a general abstract function of the passive marker that is “accommodated” when this marker is extended to intransitive verbs, and we will propose a different view of the passive-to-impersonal reanalysis, which is at the same time more gradual (i.e. the extension of the passive marker to intransitive verbs does not affect all the verbs at one time) and more “local”, i.e. triggered by a bundle of ambiguous contexts which facilitate the reinterpretation of the passive marker as a marker of generic human agency even in the presence of a patient.

In what follows, we will describe the initial stage of the history of the *si*-construction in Old Italian (§2): it will be shown that the prerequisites for the reinterpretation of the passive marker *si* as a marker of generic human agency were already present in the earliest stage of the Italian language. In §3, the Old Italian situation will be compared with the situation of present-day Italian, with a view to underscoring the differences between the two stages. In present-day Italian two different *si*-constructions exist, a passive and an impersonal one; besides them, there is an “inclusive” *si*-construction, structurally indistinguishable from passive and impersonal *si*-constructions, which roughly corresponds to a 1st person plural form of the verb. §4 will sketch the main steps of the passive-to-impersonal development, which is by no means an abrupt process, as testified by the co-existence of two different *si*-constructions in contemporary Italian. Finally, in §5, the main stages of
this process will be briefly recapitulated, and the general implications of the present analysis will be discussed.

2. The si-construction in Old Italian

In this section, the syntactic and semantic features of the si-construction in Old Italian will be discussed in detail, with a view to identifying the characteristics that have favoured its reinterpretation as an impersonal construction. Old Italian is to be intended here as a rather arbitrary term covering literary and non-literary Tuscan documents written in the 13th century (more details on the corpus used in the present study are provided in the Appendix).8

The construction in question involves the use of the 3rd person reflexive marker si, which had grammaticalized into an anticausative and a passive marker already in Late Latin (Cennamo 1998). In more than 95% of the cases in our Old Italian corpus, the si-construction behaves syntactically as a passive with respect to agreement.9 This means that the patients in this construction almost always control agreement on the verb, as in example (10). This example also nicely shows that the si-construction is a functional equivalent of another passive construction of Old Italian, the so-called periphrastic passive (fu coronato = corono-ssi), used immediately after the si-construction to refer to the same event, namely the coronation of Frederick the First. Moreover, Frederick the First, the subject of corono-ssi (‘crowned-si’), is also the main topic of the portion of text from which this passage is taken, which means that both the si-construction and the periphrastic passive, in Givónian terms, can be used to introduce

8. Impersonal constructions in non-Tuscan texts will not be discussed in this paper. In Old Venetian the si-construction was limited to transitive and intransitively-used transitive predicates (chom se leze in la ystoria,’as one can read in the history’), while in Old Neapolitan there are examples of si-constructions with unergative and unaccusative intransitive verbs (non se pò plu andare,’one cannot go further’). The reader is referred to Cennamo (2000:94, 98) for a survey of passive and impersonal si-constructions in these two vernaculars.

9. This percentage has been obtained by counting only si-constructions with plural patient arguments. In the case of si-constructions with singular patient arguments, there is no way to establish whether the singular form of the verb agrees with the patient or is simply the default form (see below). This fact advises us to remove si-constructions with singular patient arguments from the statistical count at issue in order to obtain a more realistic picture. The reader should be aware, however, that si-constructions such as those exemplified in (10)–(11) are undoubtedly genuine instances of passive constructions, given that the singular patient appears before the verb in these cases (a behavioural property of subjects in Italian).
From passive to impersonal 197

a non-agent topic (and to maintain it discourse-central as the narration proceeds). An example of si-construction with the same function is provided in (11):

(10) Federigo primo deceto Barbarossa imperò anni xxxvij, et corono-ssi
F. first called B. reigned years 37 and crowned-si
ne-lla chiesa di Sam Pietro ad Roma … Elli fu coronato da
in-the church of Saint Peter at Rome he was crowned by
papa Adriano il di medesimo ch’elli giunse ad Roma
pope A. the day same that he arrived to Rome

(Cronica Fiorentina, 102, 21, 1; 13th century)

'Frederick the First, called Barbarossa, was emperor for 37 years, and was
crowned in the church of St Peter in Rome … He was crowned by pope Adrian
on the very same day he arrived in Rome'

(11) ma il suo figluolo, il qual era allora co llui, si llo
but the his son who was then with him EMPH him
inbalsimò e reco-llo infino a Tiro, e quivi si soppellio
enbalmed and took-him as far as T. and here si buried

(Cronica Fiorentina, p. 106; 13th century)

'But his son, who was with him at that time, enbalmed him and took him
as far as Tyre, and he was buried there'

The demoted agent in the si-construction is often a specific person or group, and
can be overtly encoded by means of a PP headed by per or da, a possibility definitely
excluded in present-day Italian (Sansò 2011), as in (12). This fact shows that the si-
construction in Old Italian directly concurred with the periphrastic passive in many, if
not all, of its contexts of usage:

(12) Anche fue ordinato la seççaia domenica di giennaio, per
also was ordered the last Sunday of January by
Ghese e per Bonaguida capitani e per li loro consiglieri, che
G. and by B. captains and by the their councillors that
si dovesse bandire la nostra processione la primaia domenica di
si should announce the our procession the first Sunday of
ciascheuno mese per Angello banditore
each month by A. town-crier

(Carmine, §26; 1280–1298)

'The last Sunday of January the Captains Ghese and Bonaguida and their
councillors ordered that our procession should be announced publicly by
Agnello, the town-crier, the first Sunday of each month'

On the other hand, the patient NP promoted to subject acquires not the complete
set of subject properties, but only a subset thereof. Indeed, in approximately half of
the occurrences (55.94%) of the *si* construction in the 13th century, the patient NP maintains a **behavioural property typical of objects**, i.e. it appears postverbally (*si* + V + N, as in examples (13)–(15)). From a semantic point of view, the preverbal vs. postverbal position of the patient in the *si*-construction tends to correlate with the referentiality/topicality of that participant. In other words, there is a tendency for postverbal patients to be non-topical and non-referential and for preverbal patients to be topical and referential:

(13) A *Roma* *si vende* ogne cosa e a pochi è più cara la
    in Rome *si* sells each thing and to few:pl is more valued the
    [fede] that *la* pecunia
    faith than the money

    *(Fiori e vita di filosafi e d’altri savi e d’imperadori, p. 170; 1271–1275)*

    ’In Rome everything is sold, and few care about faith more than about money’

(14) *Fue trovato che* in *Roma* *si trattava* tradimento
    was found that in Rome *si* plotted betrayal

    *(Brunetto Latini, Rettorica, p. 90; 1260–1261)*

    ’It was found that in Rome people were plotting betrayal’

(15) *Et lo camarlenge* *si debbia* chiamare comunemente per tutta
    and the chamberlain *si* should call communally by all
    l’ Arte, *là unque* *si trova* lo milliore
    the Art wherever *si* finds the best

    *(Statuto dell’Università ed Arte della lana di Siena, 1, 9; 1298)*

    ’And the chamberlain should be elected communally by the whole Art
    (a professional association, AGR-AS), wherever one can find the best one’

The passages in (13)–(15) exemplify the kind of contexts that are associated with the reinterpretation of the *si*-construction as an impersonal construction. These contexts are highly ambiguous, both in structure and meaning:

- from a semantic-pragmatic point of view, while preverbal patients are likely candidates for subjecthood, being topical and persistent referents, those following the verb generally lack the semantic-pragmatic characteristics of subjects, though being syntactically subjects;
- from a structural point of view, when a singular patient NP appears in postverbal position the construction is ambiguous between a passive and an impersonal interpretation, as the 3rd person singular agreement on the verb might equally well be triggered by the singular patient or be the default choice in the case of lack of agreement.
It must also be added that a generic human agent interpretation is also favoured in many cases in which the *si*-construction is accompanied by a temporal or spatial specification, as in examples (13) and (14) above (*in Roma*, ‘in Rome’), in which the spatial or temporal specification induces an interpretation of the agent as a loosely specified set of individuals (e.g. ‘people living in Rome’).

These characteristics might be considered as the necessary preconditions for the reanalysis of the construction as impersonal. In other words, the reanalysis starts with postverbal patients that are particularly low on the referentiality/topicality scale, as these are unlikely candidates for subjecthood from a semantic point of view, and is favoured by clusters of contextual features that invite a generic human agent reading of the construction (e.g. a spatial/temporal specification, the atemporal present tense, or the imperfective/unbounded aspect). Contexts such as those exemplified in (13)–(15) thus represent the critical contexts (in the sense of Diewald 2002:109) in which semantic and syntactic ambiguity provides “several options for interpretations”, among which the innovative impersonal reading appears to be favoured. In contexts such as those exemplified by (13)–(15) the reanalysis of the construction is covert, i.e. there are no formal clues that *si* has been reinterpreted as a marker of generic human agency. During a narrow time span in the second half of the 13th century, however, we find the first formal symptoms of such a reanalysis. These include both the extension of the construction to intransitive verbs (and intransitively-used predicates) and the emergence of a new pattern with transitive verbs in which the patient does not agree with the verb. These two new syntactic environments in which the *si*-construction can be found will be discussed in detail in the next two subsections.

---

10. An anonymous referee points out to us the possibility that the reflexive pronoun *se* had already become an indefinite pronoun in Late Latin, and quotes examples such as (ii) and (iii) (from Cennamo 1993:58):  

(ii)  *attendite se ipsis*  
*pay.attention:IMP.2PL se self:DAT.PL*  
‘pay attention to yourselves’  

(*Itala*, act. 20, 28 (cod. e))

(iii)  *speramus beneficia uber-ior-a pos se consequituros*  
*hope:1PL benefits abundant-comp-neut:PL after se obtain:PTCP:FUT*  
‘we hope to obtain more benefits’  

(*CIL IX* 1681; Benevento)

In these examples, however, *se* cannot be considered as an indefinite pronoun under any definition of this term: it is simply a generalized reflexive pronoun, used for all persons, and there are no examples of *se* appearing in subject position as an equivalent of ‘(some/any)one’.
2.1 The extension to intransitive verbs

As to the extension to intransitive verbs, Salvi (2008: 135) states that “of the intransitive verbs, only the unergatives were possible in this construction” in Old Italian and that “in the works of Dante the construction begins to be used with […] unaccusative verbs too, mainly with verbs of movement”. The following examples show, contra Salvi, that this construction was equally possible with both unergative (questionare ‘discuss, question,’ contrastare ‘fight with,’ etc.) and unaccusative (andare ‘go,’ venire ‘come,’ entrare ‘enter’) predicates within the same rather narrow temporal span (and that unaccusative verbs were possible well before Dante’s works, as examples (20) and (21) show). If our analysis is correct, this is exactly what we expect to find, given that at least a sub-part of unaccusative verbs (especially motion verbs) imply a volitional agent that can be conceptualized as generic.

(16) se ne l paese di Francia si guadagniase melgio
    if in-the country of France si earn:SBVFIMP.3SG better
    che no vi si può guadagniare oggi
    than NEG there si can earn:INF today

    (Lettera di Vincenti di Aldobrandino Vincenti e compagni, da Siena, a Iacomo di Guido Cacciaconti, 1260)

    ‘if in France one could earn more (money) than one can earn today’

(17) quando i farai el pagamento, si ne fa
    when to.him make:FUT.2SG the payment EMPH of.it make:IMP
    fare la sc[r]ipta ne libro di Signori de’ mercatanti,
    make:INF the registration in book of Lords of merchants
    chome si chustuma di fare
    as si use:3SG of do:INF

    (Lettera di Vincenti di Aldobrandino Vincenti e compagni, da Siena, a Iacomo di Guido Cacciaconti, 1260)

    ‘when you will pay him, make sure that the registration [of payment] be made in the book of the Lords of merchants, as it is customary’

(18) La iudiciale constituzione è quella ne-lla quale … si questiona
    the iudicial constitution is that in-which si questions
    sopra la quantitate o sopra la comparazione o sopra la
    about the quantity or about the comparison or about the
    qualitade d’ un fatto
    quality of a fact

    (Brunetto Latini, Rettorica, p. 103; 1261)

    ‘The iudicial constitution is the one in which one questions about the quantity or the comparison or the quality of a fact’
(19) *i fermi argumenti ne-l principio, i deboli ne-l mezzo, i*
the strong arguments in-the beginning the weak in-the middle the
fermissimi, *co’ quali non si possa contrastare lievemente,*
strongest with which *NEG si can:SBV.3SG contrast easily*
ne-l la fine
* in-the end* (Brunetto Latini, *Rettorica*, p. 75; 1261)
‘the strong arguments (should be put) in the beginning, the weak ones in the
middle, and the strongest ones, against which one cannot contrast easily,
in the end’

(20) *tutte le cupidità [sono porte] de-l ninferno per le quali si*
all the pleasures are gates of-the hell through which *si va a la morte*
goes to the death (Andrea da Grosseto, 2, 17; 1268)
‘all the pleasures are gates of the hell, through which *one goes* to death’

(21) *Et per la [necessità] si viene a la povertà*
and through the necessity *si comes to the poverty*
* (Andrea da Grosseto, 4, 9)
‘and through necessity one comes to poverty’

(22) *Sappi che cinque sono le porti per le quali s’ entra,*
know:IMP that five are the gates through which *si enters*
anzi che *andare si possa in paradiso*
before *go:INF si can:SBV.3SG to paradise*
* (Bono Giamboni, *Libro de’ Vizi e delle Virtudi*, 69,2; 1292)
‘Be aware that there are five gates through which *one can enter,*
before *one can go* to heaven’

(23) *Io tenni li piedi in quella parte de la vita di là da*
I kept the feet in that part of the life beyond
*la quale non si puote ire più per intendimento*
which *NEG si can go further through intention*
*di ritornare*
of *come.back:INF*
* (Dante, *Vita Nuova*, 14, 1; 1292–1293)
‘I have just set foot on that boundary of life beyond which
*no one can go,* hoping to return’

(24) *Legge-si de-l la bontà de-l re giovane…*
read-si about-the kindness of-the king young
* (Il Novellino, 18, 3; end of 13th century)
‘*One can read* about the kindness of the young king’
and in-the welcome:INF any novice si proceed:sbjv.3sg in this way

(Capitoli della Compagnia dei Disciplinati di Siena, 11; 1295)

‘and when welcoming a novice, one should proceed in this way’

There is no doubt a relative difference in terms of types (though not in terms of tokens) of unergative vs. unaccusative intransitive verbs that can be used in the si-construction, but this difference has possibly an independent explanation: most unaccusative verbs had a si-marked counterpart in Old Italian (Jezek 2010); this is especially true of verbs of change of state (morire and morir-si ‘die’) and location (andare and andar-si ‘go’, partire and partir-si ‘leave’), in which si goes back to the Late Latin dative reflexive pronoun sibi, used to mark change-of-state verbs and spontaneous processes (Cennamo 1999:122). The alleged incompatibility of unaccusative predicates with the si-construction would thus be motivated by the necessity of avoiding ambiguities between an impersonal (si va ‘one goes’) and a purely intransitive ([egli] si va ‘he goes’) interpretation of these verbs.

More importantly, the earliest examples of this construction with intransitive verbs are exceptionless in the (omnitemporal) present tense or in other tenses/moods that are most compatible with the genericity of the human subject (as in examples (16)–(25)). Salvi (2008:136) only talks of a restriction to non-compound tenses, and seems to imply that this restriction only characterizes the use of the si-construction with unaccusative intransitives. This is clearly incorrect, as in the 13th century the restriction to non-compound tenses characterizes si-constructions with unergative intransitive and transitive verbs as well (see §4.1). The crucial point is that while passive si-constructions are equally possible with specific and generic time reference, the first examples of the impersonal si-construction with intransitive verbs all have generic time reference. The fact that the first contexts in which the si-construction is used with intransitive verbs are those in which the genericity of the agent is also independently triggered by other contextual features corroborates our hypothesis that the reinterpretation of si as a marker of a generic human agent precedes and motivates its extension to intransitive verbs.11

---

11. Another argument in favour of an early reinterpretation of the si-construction as an impersonal construction implying a generic human agent is the early attestation of the inclusive reading of this construction, that will be discussed in §4.3: an impersonal/generic reading of the construction is indeed necessary for an inclusive interpretation to emerge.
2.2 The emergence of the impersonal *si*-construction with transitive verbs

During the same temporal interval, we also see the emergence of another impersonal trait, i.e. the **optionality of agreement** with patient NPs. This typically happens in four cases:

- when the patient is a **quantified noun** (example (26)–(28))
- when the patient is a **determinerless plural noun** (example (29)–(30))
- when the patient is a **coordinated noun phrase** (N and N); (example (31) and (32))
- with **nominals in light-verb constructions**, formed by a semantically bleached verb (mainly *fare*, ‘do’, ‘make’) and a nominal contributing a large part of the predicative content (example (33))

(26) *E ancho vi si rameta tutti i miei denari propi, ch’io and also there *si* recalls all the my money[PL] own that I
debo avere di chostà
must have from there

(Lettera di Iacomo de’ Sansedoni da Siena a Goro e Gonteruccio de’ Sansedoni, in Parigi, 1294)

‘and all my money that I must have back from there is mentioned there
(i.e. in that letter)’

(27) *andando su per questo fiume XII giornate, si trova
going up along this river 12 days *si* find.3sg
città e castella assai
towns and castles many

(Marco Polo, *Il Milione*, 126, 4; beginning of the
14th century)

‘travelling along this river for 12 days, one meets with a good number of
towns and castles’

(28) *in una cosa si trova quattro cause, cioè la cagione materiale
in one thing *si* finds four causes that.is the cause material
e la cagione formale e la efficiente e la finale
and the cause formal and the efficient and the final

(Andrea da Grosseto, *Trattati morali di Albertano da Brescia volgarizzati*, 1, 5; 1268)

‘in each thing one finds four causes, the material one, the formal one,
the efficient one, and the final one’
(29) **uno paramento da prete, co-l quale vi si debia dicere**

one paraments[sg] for priest with-which with.it si should[sg] say

**messe per anima del conte Guido Guerra, mio figluolo**

masses for soul of.the count G. G. my son

(Testamento della contessa Beatrice da Capraia, 1279)

‘priest vestments, with which masses should be said for the soul of Count Guido Guerra, my son’

(30) **In questa città si fa giambellotti di pelo di camello, li più belli del mondo**

in this city si make:3sg camlet:pl of hair of camel

the most beautiful of.the world

(Marco Polo, *Il Milione*, 72, 1)

‘They make in this city great quantities of camlets of camel’s wool, the finest in the world’

(31) **Et certo la colpa e la cagione si mette sopra altrui**

and surely the fault and the cause si puts on others

(Brunetto Latini, *Rettorica*, p. 113; 1261)

‘and surely the blame and the fault are put on others’

(32) **li argomenti e le legagioni [...] per li quali si fa la confermazione e la risponsione**

the arguments and the legations through which si makes the affirmation and the answer

(Fiore di rett., red. gamma, p. 140; 1292)

‘the arguments and the legations through which confirmations and replies are made’

(33) **tutto l’ arde infino che l’ uccide, se non vi si fa argomenti**

whole him burns until him kills if not against.it si does arguments

(Tesoro di Brunetto Latini volgarizzato [Libri III, IV e V], p. 76; 2nd half of the 13th century)

‘[the poison] burns him entirely and kills him if one does not take an antidote for it’

According to Salvi (2008: 134–135), cases such as (26)–(33) are not indicative of the reinterpretation of the construction as impersonal. In his view, these non-agreeing postverbal patients are not the direct objects of the verb but the subjects, lack of agreement being “normal in Old Italian with a postverbal Subject in unaccusative constructions” (Salvi 2008: 134), given examples such as (34)–(37):
From passive to impersonal

(34) *Iddio [...] da-l quale nasce tutti e' beni e tutte le grazie* God from-whom is.born all the goods and all the graces (Anonimo, Navigatio Sancti Brendani, volgarizzamento toscano, p. 41; 1300) ‘God, from whom all the goods and all the graces derive’

(35) *E appresso viene isperanza e disperanza, che vanno innanzi a* and after comes hope and despair that go before *paura e ardimento* fear and courage Egidio Romano volg., 1, 3, 2; 1288) ‘and after come hope and despair, which go before fear and courage’

(36) *De-lia cattiva e rea volontà, di cui nasce i sette* from-the bad and guilty will from which is.born the seven *vizi capitali* sins deadly (Bono Giamboni, Trattato di Virtù e di Vizi, Cap. 3, rubr.) ‘from the bad and guilty will, from which the seven deadly sins derive’

(37) *per questa guerra nacque grandissime battaglie e mortali[ta]te* for this war was.born great battles and mortality *tra Guascongesi e Normandi e France[sc]hi* among Gasconian.PL and Norman.PL and French.PL (Cronica Fiorentina, p. 140) ‘because of this war many great battles arose among Gasconians, Normans and the French, which caused many victims’

Salvi fails to acknowledge that the types of patients that do not trigger agreement on the verb in the *si*-construction share a crucial property, i.e. they rank low on the scale of referentiality.topicality. Moreover, the statement that lack of agreement is “normal” with postverbal arguments of unaccusatives invites the inference that we have to do with a rule. On the contrary, this is just a possibility, and examples of postverbal arguments that trigger agreement are easy to find:

(38) *la delettanza de-l corpo [...] de la quale nascono tradimenti* the delight of-the body from which are.born betrayals *de le Terre* of the lands (Andrea da Grosseto, 2, 17) ‘bodily delights, from which betrayals of the lands derive’

(39) *Da questo peccato vengono furti, simonia, inganni, usura,* from this sin come:3PL thefts simony deceits usury *tradimenti (et) deceptioni* betrayals and deceptions (Questioni filosofiche, p. 160; 1298) ‘thefts, simony, deceits, usury, betrayals and deceptions derive from this sin’
Moreover, lack of agreement in the *si*-construction is also attested in a few cases ((40)–(41)) in which the patient appears in preverbal position as the head of a relative clause (but also in other cases: recall example (31)):

\[(40)\]  
\[
\begin{array}{l}
\text{i quali danari si diedero ne-le spese che si fece p(er)} \\
\text{which money si gave:3pl in-the expenses that si made:3sg for} \\
\text{Baldovino il die che si supelio} \\
\text{B. the day that si buried:3sg} \\
\end{array}
\]

\[(Libro d'amministrazione dell'eredità di Baldovino Iacopi Riccomanni, 1278)\]

'this money was invested in the expenses that were made for Baldovino the day he was buried'  

\[(41)\]  
\[
\begin{array}{l}
\text{VII. [soldi] che si diè ne la soprasberga di Matasala} \\
\text{seven coins that si gave:3sg in the soprasberga of M.} \\
\end{array}
\]

\[(Libro di conti di Matasala di Spinello, 1233–1243; quoted after Wehr 1995: 112)\]

'seven coins that were paid for Matasala's soprasberga [a kind of military vestment]'  

To sum up, as a result of the availability of a generic human agency interpretation, the *si*-construction started developing impersonal features already in Old Italian. As will be described in the next section, this process did not lead to the transformation of a passive construction into an impersonal one (as in Polish), and in present-day Italian a passive and an impersonal *si*-construction co-exist. The stages of the process that resulted in the progressive differentiation and divergence between an impersonal and a passive construction will be sketched in §4.

3. The situation in present-day Italian

In present-day Italian, two types of *si*-constructions can be distinguished on structural grounds. One of them can be considered as an impersonal construction, following the definition introduced in §1. The other one can be defined as a passive construction, although it maintains some hybrid syntactic and semantic features that distinguish it from the other passive construction of present-day Italian, formed with the auxiliaries *essere* 'be' and *venire* 'come' + the past participle. A third construction type, formally indistinguishable from the two other constructions, will be labelled 'inclusive *si*-construction': in this construction *si* + 3rd person verb is interpreted as equivalent to a 1st person plural form of the verb. The three constructions will be examined in turn in the
next paragraphs on the basis of two corpora of present-day spoken and written Italian (see Appendix).12

3.1 The passive *si*-construction

In the passive *si*-construction the verb agrees with the patient, in gender and number but in the overwhelming majority of cases (more than 70% in our corpus) it appears postverbally, and tends to be inanimate and non-topical. From a semantic point of view, the agent in the passive *si*-construction is typically generic, the construction tends to occur in the present tense and with imperfective aspect, and is often characterized by modal overtones (see e.g. example (42) and (46)), even in the absence of overt modal operators or verbs:

(42) *si annota il trasferimento solo su-* l foglio complementare

> the transfer is recorded only on the additional sheet

(Lessico dell’Italiano Parlato, Naples, 55, 89)

‘The (property) transfer is recorded only on the additional sheet’

‘The property transfer must be recorded only on the additional sheet’

(43) *La sicurezza di tutti […] potrà essere ottenuta […] se * si

> the security of everybody can: fut.3sg be obtained if *

*accetteranno efficaci sistemi di verifica. *

> accept: fut.3pl effective systems of checking

‘the general security can be obtained only if effective checking systems will be accepted’

(44) *Ne-l’ ona si sono fatte, a-* l’inizio de-gli anni

> in-the area si are[aux.3pl] made at-the beginning of-the years

70, *battaglie famose per la tutela de-ll’ ambiente

> seventies battles famous for the protection of-the environment

‘in this area many famous battles for the protection of the environment were made in the early Seventies’

(45) *bruciata come * si bruciavano le streghe e gli eretici

> burned as si burnt:3pl the witches and the heretics

‘burnt just as witches and heretics used to be burnt’

12. Unless otherwise specified, the examples in the following subsections are taken from “La Repubblica” corpus.
(46) *Non si regolano in quarantotte ore [...] i problemi accumulatisi in cinque anni*  

in five years  

‘The problems piled up in five years cannot be resolved in 48 hours’

3.2 The impersonal *si*-construction

The impersonal *si*-construction in present-day Italian appears with every kind of intransitive (or intransitively-used) predicates, including reflexives (example (47)), reciprocals (example (48)), the copula (example (49), and the periphrastic passive (example (50)):

(47) *per non sputar-si in faccia quando ci si guarda al-lo*  

in face when oneself *si* watches at-the  

specchio, a volte *bisogna saper-si accontentare*  

mirror sometimes it.is.necessary be.able.to-REFL content:INF  

‘in order not to spit on your face when you look at yourself in the mirror sometimes you must content yourself’

(48) *In famiglia e tra amici ci si consola de-lle privazioni e de-lle limitazioni*  

in family and among friends *recip* *si* consoles of-the deprivations and of-the limitations  

‘within families and among friends people console each other of deprivations and limitations’

(49) *La vita de-lla pendolare, soprattutto quando si è giovani, belli e futuri “dottori” appare crudele*  

the life of-the commuter mostly when *si* is young:M.PL beautiful:M.PL and future:M.PL doctors appears cruel  

‘living as a commuter, particularly when one is young, beautiful and is going to graduate, looks cruel’

(50) *Poi, a–l terzo giorno, si spiega come comportarsi quando si viene presi in ostaggio*  

then on-the third day *si* explains how behave:INF when *si* comes[AUX] taken:M.PL in hostage  

‘then, on the third day, they explain how to behave when one is taken hostage’

Unergative and unaccusative intransitive verbs behave differently in compound tenses: whereas the past participle of unergative verbs displays a singular ending (*si è telefoniato, si is[AUX] phoned, ‘people called’), in unaccusative verbs the past participle has the plural ending (*si è arrivati, si is[AUX] arrived:M.PL, ‘one/we arrived’).
The impersonal *si*-construction is also possible with transitive verbs. Unlike Old Italian, the impersonal construction is possible with patients of any kind, including specific and definite patients:

(51) *io spero* *si tocca* *anche argomenti* *anche piuttosto*

I hope given that *si* touches upon also issues also rather
difficoli *spero* *ci* *sarà una partecipazione anche*
difficult I hope there will be a participation also
*da parte de-i professori*
by-the professors

'(I hope ... because also some rather difficult issues will be touched upon, I hope that also teachers will take part in it)'

(52) *si è agitato* *questa questione* *de-lla fiscalizzazione*

*si* is*(AUX)* raised.M.SG this.F.SG question.F.SG of the fiscalization
de-gli oneri sociali
of the burdens social

'(This question of the fiscalization of social burdens has been raised)'

(53) *io volevo fare* *delle liste* *tanto per avere un’ idea*

I wanted make INF some lists just in order to have INF an idea
*su come come si svolgerà i lavori*
on how how *si* carry out FUT.3SG the PL works

'(I wanted to make some lists, just to figure out how the activities will be carried out)'

With transitive predicates, the impersonal *si*-construction is the only option available when the patient is pronominal:

---

13. The context appears to exclude that the *si*-constructions in examples (51)–(53) are instances of inclusive *si*-constructions.

14. With third person plural pronouns a further possibility is mentioned in the literature, namely a construction in which the verb agrees with the patient, although the pronoun itself remains in the accusative form (Lepschy 1989:112):

(iv) *Le si considerano e studiano*

them:F.PL *si* consider:3PL and study:3PL

'They are considered and studied'

This possibility appears to be syntactically a blend between the passive and the impersonal *si*-constructions and is quite rare in the corpora we made use of.
(54) **un po’ come si fa con le “Sonate” quando le si suona su-l fortepiano**

‘a bit like one does with Sonate when one plays them on the fortepiano’

(55) **spero proprio che non mi si accusi di falso moralismo**

‘I hope really that NEG me si accuses of false moralism’

According to D’Alessandro (2007: 37, and *passim*), lack of agreement between the patient and the transitive verb reflects “an aspectual difference”: in particular, si-constructions with agreement (passive si-constructions in our terminology) encode accomplishments (to be intended as bounded events with a duration and an endpoint), whereas si-constructions without agreement encode activities (i.e. unbounded events with a duration but no endpoint). While there may be some correlation between agreement (or lack thereof) and the aspectual properties of the predicate, it must be admitted that the passive si-construction is largely more frequent than its impersonal counterpart in the written language, so that it is not difficult to find cases of activity predicates with verb-patient agreement (see, for instance, example (68)). At the same time, the impersonal construction appears to be favoured with a special class of bounded predicates involving the phasal verbs *cominciare/iniziare a*, ‘start’ and *finire di*, ‘finish’. With these phasal verbs both the passive and the impersonal si-construction are allowed (as shown by the contrast between (56)–(57), and (58)–(59) respectively); yet, the rate of impersonal si-constructions is somewhat higher (e.g. 15.17% with *cominciare/iniziare*; 26.66% with *finire* in the two corpora of spoken and written present-day Italian, see the Appendix) than in other syntactic configurations (*si* + *verb* + *NP*; *si* + modal auxiliary + *verb* + *NP*). This fact alone shows that the aspectual properties of the predicate are not the only factor at play:15

15. An alternative analysis should be mentioned with regard to what constitutes the subject of phasal and modal verbs. If we take a couple of sentences such as (iv) and (v), we might say that there is agreement in (iv) because in this case the lexical verb (*fare*) is considered to be part of the predicate, whereas it is felt as a phrasal subject together with its object (*fare paragoni*) in (v) (therefore triggering singular agreement):

(iv) **[si possono fare] [paragoni]**

si can:3pl make:inf comparisons

‘comparisons can be made’

(v) **[si può] [fare paragoni]**

si can:3sg make:inf comparisons

‘comparisons can be made’
su consiglio de-l segretario di Stato Shultz [...] si comincia

on advice of-the Secretary of State S. si starts

a frenare le aspettative

at repress the expectations

‘on advice of the Secretary of State Shultz, they are starting to repress expectations’

per il cui pagamento si cominciano a nutrire preoccupazioni

for whose payment si start:3PL at nurture concerns

‘about whose payment one starts to nurture some concerns’

finché non si finisce di contare tutte le schede e di

until neg si finish:3SG of count:INF all the votes and of

riempire tutti i moduli

fill:inf:INF all the forms

‘until they finish to count all the votes and to fill in all the forms’

mentre si finiscono di spogliare i voti de-l Senato

while si finish:3PL of sort:out:INF the votes of-the Senate

‘while they finish to sort out the votes of the Senate’

Moreover, our corpus data show that the crucial property favouring lack of agreement is one that is superordinate to the aspectual properties of the predicate, namely the

That there are two alternative constituency relations at play is testified by the impossibility of (vi–c) and (vi–e) below:

(vi) a. si può mangiare le caramelle adesso?

‘can we eat candies now?’/‘is it possible to eat candies now?’

b. si, si può

‘yes si can:3SG’

eyes, you can/‘yes, it is’

c. *si, si può mangiare

‘yes si can:3SG eat:INF’

‘yes, you can’/‘yes, it is’

d. si possono mangiare le caramelle adesso?

‘can:3PL eat:INF the candies now’

e. *si, si possono

‘yes si can:3PL’

‘yes, you can’

f. si, si possono mangiare

‘yes si can:3PL eat:INF’
genericity of the event, which subsumes other temporal, aspectual and modal properties. In particular, the impersonal *si*-construction appears to be favoured with modal verbs (*dovere* ‘must’, *potere* ‘can/may/be able to’; *volere* ‘want’), which significantly correlate with generic human agency. The three modal verbs *dovere* ‘must’, *potere* ‘can/may’, and *volere* ‘want’ behave differently with respect to their preference for the passive vs. the impersonal *si*-construction. While *dovere* and *potere* generally appear with the passive *si*-construction (only 10.35% of the cases of *si* + *potere* + V + N are impersonal; even less so – 6.42% – for *dovere*), *volere* appears to admit the impersonal *si*-construction more easily (20.95%):\(^{16}\)

(60) non *si* può *fare* paragoni co-i giocatori di cinquanta anni fa

‘One cannot compare (them) with the (football) players of fifty years ago’

(61) un risvolto tecnologico [...] di fronte a-l quale *non si può* chiudere gli occhi

‘a technological side-effect in front of which one cannot close one’s eyes’

(62) Ma quand’anche [...] *si volesse* ridare a-lle facoltà quelle competenze [...] *non si potrebbe fare* lo lasciando in piedi la struttura dipartimentale

‘but even if one wished to give back those competences to the faculties, one could not do that leaving the departmental structure intact’

3.3 The inclusive *si*-construction

The third construction in question will be labelled inclusive *si*-construction: in this construction, which is structurally indistinguishable from passive and impersonal

---

16. From these counts two cases have been removed in which the impersonal *si*-construction is the only possibility in contemporary Italian:

- intransitive (both unergative and unaccusative) verbs (*si può andare*, ‘one can go’)
- verbs with a pronominal argument (*li si può vedere*, ‘they can be seen/one can see them’).
si-constructions, the agent is identified as an “unspecified set of people including the speaker” (Cinque 1988: 542):

\[(63)\] In 20 anni in questo paese non si è andati a fondo
in twenty years in this country NEG si [aux] gone to bottom
a uno solo de-i casi scottanti
to one only of-the cases hot
‘In this country in twenty years we didn’t manage to discover the truth about any of the hot cases’

\[(64)\] Quindi, secondo lei, si è stati troppo precipitosi?
then according to You si is [aux] been too rash
‘Then, according to You, were we too rash?’

According to Cinque (1988: 542), in independent sentences with specific time reference si acquires this new interpretation only when it occurs with unaccusative, psych-, copular and passive verbs.\(^{17}\) In the following examples (drawn from Cinque 1988: 542), the choice of a predicate that is incompatible with the inclusion of the speaker results in a pragmatically odd sentence:

\[(65)\] %Oggi, a Beirut, si è nati senza assistenza medica
‘Today, in Beirut, we were born with no medical assistance’

\[(66)\] %Oggi, a Beirut, si è stati uccisi inutilmente
‘Today, in Beirut, we have been killed in vain’

An inclusive interpretation is not excluded also when the verb is transitive or unergative (as in examples (67)–(68); see D’Alessandro 2007: 144). D’Alessandro (2007: 152ff.), elaborating on Cinque’s (1988) seminal proposal, argues that the crucial factor triggering an inclusive reading of a si-construction is not time-specificity per se but rather the boundedness of the event. The following examples, however, show that the inclusive si-construction can also refer to unbounded events, provided that the context facilitates such a reading.

\(^{17}\) As Cinque (1988: 544) observes, a pure impersonal interpretation of the si-construction is possible if the unaccusative, psych-, copular, or passive verb “is embedded in a context that suspends the specificity of the time reference”, as in the following examples:

\[(vii)\] Mi domando se a Beirut si sia nati senza assistenza medica anche oggi
‘I wonder whether in Beirut one was born with no medical assistance today still’

\[(viii)\] Quando si è sfiniti dalla fame non si ragiona
‘When one is worn out by hunger one cannot reason properly’
(67) Ne ho fatti anche sette contemporaneamente, si lavorava a casa mia, in ogni stanza c’era un gruppo che scriveva un film diverso.

(Context: a movie director talking about his activity)

‘I worked on even seven movies at the same time, we used to work at my place, (and) in every room there was a group working on a different movie’

(68) L’indicazione decisiva è arrivata da Craxi: “non si fanno questioni di uomini […] ma si affrontano problemi politici e si tiene conto de-i bisogni […] che ha Milano”

(Context: Craxi is giving a speech at a political meeting of his party)

‘The final indication has been given by Craxi: “we are not questioning about specific persons; rather, we are dealing with political issues and we are taking into account the needs of Milan’

Although the inclusive si-construction is indistinguishable from the passive and impersonal si-constructions, there may be some unambiguous clues pointing to an inclusive interpretation such as, e.g. the use of the 1st person plural possessive pronoun nostro, ‘our’ (when an inclusive interpretation is excluded, the possessive pronoun triggered by si is proprio, ‘own’), or the presence of the independent 1st person plural pronoun noi before si. The reader is referred to D’Alessandro (2007: 140–143) for a detailed discussion of the tests for an inclusive reading of the si-construction.

To sum up, except for the inclusive si-construction, both passive and impersonal si-constructions in present-day Italian share an important semantic feature, namely the genericity of the agent. In particular, the passive si-construction appears to be much more infused with generic nuances than it was in Old Italian (recall examples (10) and (11)). As will be argued in §5, this is the result of a long-lasting process in which (i) the impersonal si-construction has progressively emerged and spread, and (ii) the existence of an impersonal construction formed with the same building blocks has fostered to some extent the (semantic, if not syntactic) “impersonalization” of the passive si-construction. In the next section, the most important stages of this long-term development will be discussed in detail.
4. What happened in the meantime

As discussed in the previous section, the impersonal *si*-construction in Italian has not replaced the passive *si*-construction and is less widespread than it. From a semantic point of view, however, the reinterpretation of *si* as a marker of generic human agency has gone a step further. Such an interpretation holds in most cases in which there is agreement between the patient and the verb, as in examples (42)–(46).

In this section we will sketch the main stages of the diachronic processes leading to the present-day situation. We will mainly focus on three different evolutionary paths, which will be the object of the next three subsections:

i. the extension of the *si*-construction to intransitive verbs (§ 4.1);
ii. the development of impersonal *si*-constructions with transitive verbs (§ 4.2);
iii. the emergence and establishing of the inclusive *si*-construction (§ 4.3).

4.1 The extension to intransitive verbs

As discussed above, the earliest examples of the *si*-construction with intransitive verbs all have generic time reference. The first examples of the impersonal *si*-construction with specific time reference are as early as the 14th century, and there are no differences between unergative and unaccusative predicates. In (69) and (70) the verbs are unergative, while in (71)–(73) they are unaccusative:18

(69) *cioè sacrifici de’ quali si raccontò di sopra*
that is sacrifices of which *si* told:3sg above

(Anonimo, Volgarizzamento B del secondo libro di Valerio Massimo, par. 31, glossa x; 1326)

‘that is, the sacrifices discussed above’ (lit: about which it was told above)

(70) *Ma molto si ragionò a quello mangiare de-l’ opere*
but much *si* talked about at that dinner of the deeds

*di Merlino*

of M. (Paolino Pieri, La Storia di Merlino, 34, 15; between 1310 and 1330)

‘at that dinner people talked a lot about Merlin’s deeds’

---

18. *Si raccontò* in (69) is not to be intended as an instance of an inclusive *si*-construction; the example is drawn from a commentary added by a person different from the author on the margins of the manuscript.
(71) *e co-l gonfalone de-l popolo in furia si corse a casa i* and with-the gonfalon of-the people quickly *si* ran to house the *Cavalcanti, e mise-vi-si fuoco, e da capo furono cacciati* C. and put:*pst.3sg-there-si* fire and again were chased *di Firenze i Cavalcanti* from Florence the C. (Villani, *Cronica*, 9, 33; 1348) ‘and the crowd ran quickly to the Cavalcanti’s house with the people’s gonfalon, and set it on fire, and the Cavalcanti were exiled from Florence again’

(72) *quando si venne al-la battaglia, fu quell’ asprezza utile* when *si* came to-the battle was that harshness useful *(Deca prima di Tito Livio volgarizzata, 8, 8; 1350)* ‘when people started to fight, that harshness turned out to be useful’

(73) *Come si sapesse, vi si andò, e trovo-ssi, e* as *si* know:*sbjv.pst.3sg* there *si* went and found:*3sg-si* and *fu preso* was taken (Marchionne di Coppo Stefani, *Cronaca fiorentina*, 830; 1385) ‘as soon as the news spread, people went there, and he was found and captured’

On the contrary, there is an important difference between unergative and unaccusative intransitive verbs with regard to the possibility of appearing with compound tenses in the *si*-construction. While unergative verbs appear quite early with compound tenses (see examples (74)–(75); (74) has an inclusive interpretation), the earliest instances of the compound tenses of the *si*-construction with unaccusative verbs are as late as the 16th century (see examples (76)–(78)). In all these instances, unlike in present-day Italian, the past participle displays singular agreement:

---

19. The passive *si*-construction could be used with compound tenses already at the beginning of the 14th century, given examples such as (ix):

(ix) *e ben si sarebbe auta la terra* and well *si* be:*cond.3sg* have:*ptcp.f.sg* the land *(Cronaca fiorentina*, p. 135; beginning of the 14th century) ‘and that land would have been easily conquered’
(74) *benché di lui e de’ suoi compagni assai ragionato*  
although of him and of his fellows much discussed

si sia  
*Boccaccio, Decameron, IX, 3; 1370*

‘although we have (already) discussed a lot about him and his fellows’

(75) *e sotto lo stecatto medesimo si sarebbe combattuto*  
and under the fence same si be:COND.3SG fought

*Deca terza di Tito Livio volgarizzata, p. 324; 1400*

‘and they would have engaged battle under the same fence’

(76) *dove quando l’ esercito era stato meno potente non si*  
where when the army was[AUX] been less mighty neg si

era potuto andare senza pericolo  
was[AUX] can:PPT go:INF without danger

*Guicciardini, Storia d’Italia, 6, 11; 1537–1540*

‘where one could not go with no risk when the army was less mighty’

(77) *Stanno-si a–l presente pacifiche e servon […] a-gli*  
stay:3PL-si at-the present peaceful and serve:3PL to-the

*Spagnuoli […] se ben per mancar la gente non si è*  
Spanish:M.PL although for lack:INF the people neg si is[AUX]

tornato ad abitar la terra  
come.back:PPT to inhabit:INF the land

*Ramusio, Relazioni di Cortés su Nuova Spagna, 4, 3; 1550–1559*

‘they are presently peaceful and serve the Spanish Empire, although nobody came back to inhabit (those lands) when the populace decreased in size’

(78) *immediate si è venuto a conoscere che ’l detto*  
immediately si is[AUX] come:PPT to know:INF that the mentioned

mappamondo fu senza alcuno dubbio cavato da quello di  
world.map was without any doubt drawn from that of

*messer Marco Polo*  
Sir M. P.

*Ramusio, I viaggi di Marco Polo, Dichiarazione di Ramusio*

‘Immediately people came to know that the aforementioned world map was no doubt drawn from the one sketched by Sir Marco Polo’
In the 16th century, the earliest examples of the si-construction with the verb essere ‘be’ are also attested, apparently only with prepositional phrases, as in (79)–(80):

(79) *Si era* insino a ora stato in ambiguo quale dovesse essere il cammino de’ tedeschi

‘Up to now people had been in doubt about which path the Germans would have followed’

(80) *mentre si era* su-l maggior furor de-l’ala battaglia un’ ala
de cavalaria polona secretamente passò a-lle trinciere nemiche

‘while the battle raged most (lit.: while one was on the biggest rage of the battle), a wing of Polish cavalrymen secretly slipped into the enemy’s trenches’

The earliest attestations of the si-construction with essere + an adjective or a noun in our corpus date to the end of the 18th century. In example (81), drawn from Vittorio Alfieri’s autobiography (1790–1803), the adjective displays plural agreement (and the construction has an inclusive meaning). The singular agreement, on the other hand, continues to be attested after the emergence of the plural agreement pattern, as example (82) shows. In this period, the coexistence of singular vs. plural agreement also characterizes the si-construction when it occurs with unaccusative intransitive verbs (as shown by the contrast between (83) and (84)):

(81) *si era certi* di non poter più partire

‘we were sure we couldn’t leave any longer’

(82) *Ma, se vi vuole del coraggio per dar-si la morte*,

‘but if there wants some courage for give:INF-refl the death

non se ne richiede uno minore per non dar-se-la, quando

NEG si than.it requires one lesser for NEG give:INF-si-it when

si è certo di aver-la da altri

‘but, if some courage is necessary in order to commit suicide, no less courage is required in order not to kill oneself when one is sure to be killed by someone else’
Finally, in the early 19th century the earliest examples of the *si*-construction with the periphrastic passive are attested:

(85) _Quando si tratta d’ uomini che camminano in cadenza e_ 
_When si treats of men who walk in rhythm and_ 

gestiscono invece di parlare, _si è portati in_ 
gesture:3pl instead of talk:inf _si is[aux] brought:m.pl in_ 

_un altro mondo_ 
_another world_ 

_(Il Conciliatore, n.28 [Sulla poesia romantica]; 1818–1819)_ 

‘when we have to do with men who walk rhythmically and gesture instead of talking, we are brought to another world’

4.2 The development of impersonal *si*-constructions with transitive verbs

As discussed above, in the 13th century optionality of agreement with patient NPs in the *si*-construction was possible with a limited number of patients, including quantified and non-referential patients, and nominals in light verb constructions. The extension to other kinds of patients is quite precocious: starting from the very beginning of the 14th century, lack of agreement is attested also with definite patients (both
uniquely identifiable referents as in (87) and patients that have been mentioned before, as in (86) and (88)):

(86) *Ne la terza parte* *si pone l’ armi le quali il Segnore prese, e* in the third part *si posits the arms which the God took and* 
*co le quali egli s’ armò e combatté e vinse* 
*with which he refl armed and fought and won* 

(Giordano da Pisa, *Quaresimale fiorentino*, p. 41; 1306) 

‘in the third part one posits the arms that our Lord took, of which he armed Himself and with which he fought and won’

(87) *Sopra la detta porta si è lavorato, d’ opera musica,* over the said door *si is*[^aux] worked:M.SG of work mosaic 
*santa Maria co-l suo Figliuolo in braccio* 
*holy Mary with-the her Son in arm* 

(Niccolò da Poggibonsi, *Libro d’oltramare*, Cap. 211; 1345) 

‘over the aforementioned door someone represented the Virgin Mary with her Son in her arms in a mosaic work’

(88) *Ne-lla detta chiesa si è figurato per ordine tutta* in-the said church *si is*[^aux] represented:M.SG through order all:F.SG 
*questa storia di Moisè* 
*this:F.SG story:F.SG of Moyses* 


‘in the aforementioned church someone represented this entire story of Moyses in order’

The earliest examples of the impersonal construction with 3rd person accusative clitics are also found in the early 14th century (see Wehr 1995: 116):

(89) *ke l mi debbie sotterare, si che no-l* *si sappia* 
*that he me should bury so that neg-it:obj si know:sbjv.3s* 

(Volgarezzamento di un frammento della *Disciplina Clericalis di Pietro di Alfonso*; 1300) 

‘He should bury me so that nobody will know it’

4.3 The emergence and establishing of the inclusive *si*-construction

As discussed before, *si*-constructions with an inclusive meaning are structurally indistinguishable from passive and impersonal *si*-constructions. An inclusive flavour
From passive to impersonal

of the *si*-construction, however, can be safely reconstructed on the basis of the context already in some passages from texts of the 14th and 15th century, as the following examples show. In (90), (91) and (92), for instance, the *si*-construction alternates, respectively, with a 1st person plural pronoun (*noi*), with a 1st person plural possessive (*nostra eletta*), and with a first person plural verb form (*abiamo ricevuto*), whereas in (93) and (94) it is clear from the context (and from previous letters belonging to the same correspondence) that the author is talking about a group of people including herself:

(90) *quest’ ultima preghiera, segnor caro, già non si fa per noi,*

this last prayer Lord dear now NEG si makes for us

*ché non bisogna, ma per color che dietro a noi restar*

because NEG needs but for those who behind us remained

(Dante, Purgatorio, 11, 23)

‘this last request we now address to You, dear Lord, not for ourselves – who have no need – but for the ones whom we have left behind’

(91) “*Se qui per dimandar gente s’ aspetta*,” ragionava il

if here for inquire:INF people si waits argued the

*poeta, “io temo forse che troppo avrà d’ indugio*

poet I fear perhaps that too.much will.have of delay

*nostra eletta”*

our choice

(Dante, Purgatorio, 13, 10)

“If we wait here in order to inquire of those who pass,” the poet said, “I fear our choice of path may be delayed too long.”

(92) *e se meno si ne fosse ricevuto di fior.* 3200 d’

and if less si of.it be:SBV:3SG received than *fiorini* 3200 of

*oro per anno si gli ci dèe fare aconpiere d’*

gold per year EMPH to.him to.us must:3SG make:INF fulfil:INF from

*altre sue rendite [...] sicchè in capo de l’ anno [...] abiamo*

other his incomes so.that in the end of the year we.have

*ricevuto interamente la sopradetta pagha di fior.*

received entirely the aforementioned pay of *fiorini*

3200 d’ oro

3200 of gold

(Libro delle rede di messer Niccholò Gianfigliazzi, p. 96; 1325)

‘and should we receive less than 3200 golden *fiorini*, he should compel him to satisfy us (by drawing money) from other incomes of his, so that at the end of the year we have received the aforementioned pay of 3200 golden *fiorini* entirely’
(93) *Questo di si comperorno gli occhiali, e ti si mandano* this day si bought:3pl. the glasses and to.you si send:3pl. sotto lettere di Niccolò Strozzi under letters of N. S. (A. Macinghi Strozzi, Lettere, 22 marzo 1463) ‘today we bought the glasses (for you), and we are now sending them to you along with letters by Niccolò Strozzi’

(94) *De-lla dota ti si scrisse, che’ ella non si poteva avere, se non si pagava el Comune* of-the dowry to.you si wrote that it neg si could have:inf if neg si paid:3.sg the municipality

(A. Macinghi Strozzi, Lettere, 31 agosto 1465) ‘as far as the dowry is concerned, we wrote to you that we couldn’t have it if we hadn’t pay the municipality’

It could be remarked that in these examples specific time reference is not a necessary condition for an inclusive interpretation to hold: while there is specific time reference in (90) and (93)–(94), in examples (91) and (92) the *si*-construction is under the scope of the conditional operator *se* (‘if’) and thus has generic time reference. The relatively early occurrence of inclusive *si*-constructions is another piece of evidence in favour of the hypothesis discussed in this paper that the reinterpretation of *si* as a marker of generic human agency occurs at earlier times than assumed by Salvi (2008), as the impersonal/generic reading is the necessary precondition for an inclusive interpretation: both interpretations usually refer to groups of people, which vary in size and composition according to the context, and this semantic overlap is notoriously the reason why languages often use impersonal forms for 1st person plural reference.

Salvi (2008: 140) also tentatively proposes that the use of the construction to express 1st person plural might have influenced the plural agreement pattern of adjectives and past participles discussed above and exemplified in (81) and (84). In our view, the availability of the inclusive reading is not directly responsible for the emergence of the plural agreement pattern in the 18th century: although plural adjectives and participles are a necessary condition for the inclusive interpretation by this time, the earliest examples of this pattern more often than not have a generic rather than an inclusive reading:

(95) *Confrontando una pittura ed un ballo, questo ha lo svantaggio di non offrire forme ideali, giacché si è costretti a* comparing a painting and a dance this has the disadvantage di non offer:inf forms ideal because si is forced:m.pl to
servirsì di ballerini tali quali sono
make.use of dancers just like they.are

(Il Conciliatore, n. 28; 1818–1819)
‘comparing a painting and a dance, the latter has the disadvantage that it does
not offer ideal forms, because one is forced to make use of dancers just like
they are’

(96) quanto più si è lontani da-llo stato naturale, cioè
the more si is far:M.PL from-the state natural that.is
quanto più si sa
the more si knows
(Leopardi, Zibaldone, 22 dicembre 1820)
‘the farther one is from the natural state, i.e. the more one knows’

(97) Ma ne-lla casetta di Lucia da-l momento che il
but in-the small.house of L. from-the moment that the
padre ne era partito non si era stati
father from.it was[AUX] left NEG si was[AUX] been:M.PL
in ozio
in idleness
(Manzoni, Fermo e Lucia, 1, 6; 1827)
‘but in Lucia’s house from the time her father passed away people had
not stayed idle’

We argue on the contrary that the plural agreement pattern in the examples above
is, so to speak, the “natural” result of the reanalysis of si as a generic human agent:
generic human agents often coincide with loosely defined groups of people, and plural
agreement functions as a marker of such a plurality.

5. Conclusions

On the syntactic side, the passive-to-impersonal development sketched in this paper
has led to the coexistence of two si-constructions in present-day Italian, a passive and
an impersonal one. On the semantic side, the process has gone somewhat further:
in both the passive and the impersonal si-constructions the agent is typically generic
in present-day Italian, and both constructions tend to be associated with a cluster of
morphosyntactic features that favour a generic interpretation, such as, e.g. the present
tense or the imperfective aspect. Table 1 summarizes the main stages of the passive-to-
impersonal development described in the previous sections.
Table 1. The main stages of the passive-to-impersonal construction reanalysis of the *si*-construction in Italian

<table>
<thead>
<tr>
<th>13th century</th>
<th>14th century</th>
<th>16th century</th>
<th>18th–19th centuries</th>
<th>present-day Italian</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Impersonal <em>si</em>-construction:</strong></td>
<td><strong>Impersonal <em>si</em>-construction:</strong></td>
<td><strong>Impersonal <em>si</em>-construction:</strong></td>
<td><strong>Impersonal <em>si</em>-construction:</strong></td>
<td><strong>Impersonal <em>si</em>-construction:</strong></td>
</tr>
<tr>
<td>– possible with unergative and unaccusative intransitive verbs (present tense, or other tenses/moods inducing generic-time reference)</td>
<td>– possible with unergative and unaccusative intransitive verbs (past tense and other tenses/moods inducing specific-time reference)</td>
<td>– possible with compound tenses of unaccusative intransitives</td>
<td>– possible with compound tenses of unaccusative intransitives (the past participle shows singular agreement)</td>
<td>– plural agreement of the past participle (with unaccusative verbs) and the N or Adj (with <em>essere</em> + AP/NP) is obligatory</td>
</tr>
<tr>
<td>– possible with a small subset of non-referential patients, generally indefinite and appearing postverbally</td>
<td>– possible with definite and referential patients</td>
<td>– possible with 3rd person accusative clitics</td>
<td>– possible with reflexive verbs and with the periphrastic passive (19th century)</td>
<td></td>
</tr>
</tbody>
</table>
As shown in §2, the prerequisites for such a reinterpretation were already available in the earliest written documents analyzed in this paper. This does not amount to saying, however, that these prerequisites represent a sufficient condition, as other factors might equally well have played a role in this development. One of these factors is possibly the existence of another construction, the periphrastic passive, which is functionally comparable to the *si*-construction, in that both can be considered as agent-defocusing strategies. In the earliest stage discussed in this paper, the area of overlap between the two constructions was larger than in Modern Italian: example (10) exemplifies the functional similarity between the periphrastic passive and the *si*-construction. The periphrastic construction was also possible with both unergative and unaccusative intransitives, a possibility definitely ruled out in present-day Italian:

(98)  
\[
\text{Veramente per diversi filosofi de-lla differenza de-lle nostre anime fue diversamente ragionato (Dante, Convivio, 4, 21, 2; 1304–1307)}
\]

'Different philosophers, it is true, have held different opinions regarding the difference of our souls'

(99)  
\[
\text{quella provincia nel-la quale prima fu andato da Romani (Deca terza di Tito Livio, 8, 12; early 14th century)}
\]

'that province where the Romans went first'

In present-day Italian the periphrastic construction has usually a specific reading, which is possibly connected with the resultative meaning of the past participle: the past participle contains an intrinsic feature of completion which renders it unsuitable for representing typically imperfective/atemporal states of affairs such as those connected to generic human agency. Throughout the history of Italian, the periphrastic passive has lost its “impersonal” features (ceasing to be employed with both unergative and unaccusative intransitive verbs quite early) and has reinforced its resultative features, resulting in a fully promotional passive construction, typically used when the patient is a discourse-salient entity. In other words, two concurrent and largely overlapping constructions belonging to the same functional space ended up being preferentially associated with the expression of two different situation types, as a result of the emergence of a formal/functional contrast between the two. We may speculate whether this functional/formal contrast was already present at an embryonic stage in 13th century Italian (as the data discussed in §2 seem to suggest; see also Sansò 2011), but what is crucial for the present purposes is that the process sketched in this paper is a clear instance of polarization, to be intended as the functionalization of
an opposition between concurrent and alternative grammatical strategies within the same functional domain.

The diachronic process described in this paper is also language-specific and idiosyncratic, and in other languages different processes of passive-to-impersonal development with different outcomes may have taken place. The present Italian case study, however, may be instructive as to the initial stage of the process, namely the extension of a passive construction to intransitive verbs: although passive and impersonal constructions are functionally similar as agent-defocusing strategies, the different type of agent defocusing that characterizes the two construction types requires us to posit a reanalysis of the passive marker as a marker of generic human agency as a necessary precondition motivating its extension to intransitive verbs.

### Abbreviations

| 1, 2, 3 | 1st, 2nd, 3rd person | M | masculine |
| ACC | accusative | NEG | negation |
| ADJ | adjective | NEUT | neuter |
| AUX | auxiliary | NOM | nominative |
| COMP | comparative | OBJ | object |
| COND | conditional | PL | plural |
| DAT | dative | PST | past |
| EMPH | emphatic particle | PPT | past participle |
| EXPL | expletive | PTCP | participle |
| F | feminine | RECIP | reciprocal |
| FUT | future | REFL | reflexive |
| IMP | imperative | SBJV | subjunctive |
| IMPF | imperfective | SG | singular |
| INF | infinitive |

### Appendix

**Corpora**

13th and 14th century Italian

**OVI (Opera del Vocabolario Italiano) Corpus** – The OVI Corpus is a large-scale corpus of early Italian containing 1849 vernacular texts (21.2 million words, 479,000 unique forms), the majority of which are dated prior to 1375, the year of Giovanni Boccaccio’s death. This corpus aims to collect both literary and non-literary texts in a single repository: besides early masters of
From passive to impersonal

Italian literature like Dante, Petrarch, and Boccaccio, lesser-known and obscure texts by poets, merchants, and medieval chroniclers, as well as non-literary texts such as private letters, Statuti (statutes) and Registri (official records) are well-represented in the database.

1200–1900 Italian

LIZ 4.0 (Letteratura Italiana Zanichelli) Corpus – The LIZ 4.0 Corpus contains approximately 1000 works of Italian literature. The works included in this resource span a chronological period of time beginning with Francesco d'Assisi’s Laudes Creaturarum (12th century) and ending with Italo Svevo’s Coscienza di Zeno (20th century).

Present-day Italian

LIP (Lessico dell’Italiano Parlato) Corpus – The LIP Corpus is a 500000 word corpus of contemporary spoken Italian, collected by a team of linguists coordinated by Tullio de Mauro. The corpus includes different types of everyday conversation, ranging from spontaneous face-to-face interactions to unidirectional speech such as broadcast news speech. The data have been collected in four major Italian cities (Rome, Milan, Naples and Florence).

La Repubblica Corpus – The “La Repubblica” corpus is a very large corpus of Italian newspaper texts (approximately 380M tokens) developed at the University of Bologna (Scuola Superiore di Lingue Moderne per Inter preti e Traduttori, Forlì) and available online at the following address: http://dev.sslmit.unibo.it/corpora/corpora.php. The corpus is tokenized, pos-tagged, lemmatized, and categorized in terms of genre and topic.

References


Passive to anticausative through impersonalization

The case of Vedic and Indo-European

Leonid Kulikov
Leiden University/Institute of Linguistics, Moscow

Vedic Sanskrit and other Indo-European languages attest a typologically remarkable change of passives to anticausatives. This semantic development is attested, foremost, for passives of several verbs of perception and knowledge (knowledge transfer) obviously, according to the scenario ‘Y is seen (known etc.) by smb.’ → ‘Y is seen (known etc.) [by smb.]’ → ‘Y is seen (known etc.) [by generic passive agent]’ → ‘Y is visible (famous, etc.)’. A special variety of this development is instantiated by the passive of a verb of speech, *ucyáte* ‘Y is pronounced’ → ‘Y [e.g. speech, musical instrument] sounds’. In addition, passive to anticausative transfer is attested for a small subgroup of verbs of caused motion. While in this latter case the rise of anticausative usages may be due to conceptualizing simple transitives as causatives (*throw* = ‘make fall, make fly’, etc.), in cases of verbs of perception and knowledge we observe the rise of the anticausative usages through the stage which is called ‘impersonalization’ in Siewierska 1984 and explained in terms of ‘objectivization of knowledge’, i.e. knowledge without a knowing subject. In connection with these verbs, I will briefly discuss the relationships between ‘agentless’, ‘impersonalized’ and ‘impersonal’ passives.

Keywords: passive; anticausative; impersonalization; impersonal passive; verbs of perception; objectivization; Vedic; Indo-European

1. Between passives and anticausatives: Introductory remarks

1.1 Preliminaries and definitions

The distinction between (agentless) passives and non-passive intransitives (foremost, anticausatives) is one of the most tricky issues with which a linguist may be confronted
when undertaking a syntactic study of the verb. On the border between these two categories, we find a large ‘grey zone’ where languages exhibit amazing versatility in the ways of conceptualizing events. Before entering into a discussion of this dichotomy, it will be helpful to give definitions of the basic terms and syntactic notions.

In what follows, I will be dealing with a number of voices and voice-related categories using the powerful theoretical framework developed by the Leningrad/St. Petersburg Typology Group. In accordance with this approach, voices are determined as grammaticalized diatheses, that is, patterns of mapping of semantic roles (or, rather, macroroles, designated with X = Actor and Y = Undergoer) onto syntactic functions, or grammatical relations (Subject [S], Direct Object [DO], Indirect Object [IO], Oblique Object [Obl]).

Thus, the canonical (‘full’) passive construction (diathesis) based on a simple transitive verb suggests the demotion of the initial subject of the base construction and the concomitant promotion of the direct object to the subject position, as schematized in (1) and illustrated in (2b):

\[
\begin{array}{ccc}
X & Y \\
S & DO \\
\end{array}
\Rightarrow
\begin{array}{ccc}
X & Y \\
Obl & S \\
\end{array}
\]

(2) Latin

a. Miles hostem occidit  
   warrior:nom enemy:acc kill:pres:3sg  
   ‘The warrior kills the enemy.’

b. A militē hostis occidi-tur  
   by warrior:abl enemy:nom kill:pres-3sg.pass  
   ‘The enemy is (being) killed by the warrior.’

c. Hostis occidi-tur  
   enemy:nom kill:pres-3sg.pass  
   ‘The enemy is (being) killed.’

The much more common agentless passive pattern (sometimes also called ‘reduced passive’; see e.g. Matthews 1997:311) differs in that the agent noun is removed from the structure, as shown in (3 and 2c):

\[
\begin{array}{ccc}
X & Y \\
S & DO \\
\end{array}
\Rightarrow
\begin{array}{ccc}
X & Y \\
- & S \\
\end{array}
\]

1. For a detailed description of this apparatus, see, for instance, Geniušienė (1987); Melčuk (1993); Kulikov (2011).
Obviously, this pattern is very similar to the anticausative (decausative) derivation, schematized in (4) and illustrated in (5b):²

(4) Anticausative

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>DO</td>
</tr>
</tbody>
</table>

⇒

<table>
<thead>
<tr>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
</tr>
</tbody>
</table>

Both anticausative and agentless passive derivations entail the promotion of the initial Direct Object (Patient) and the demotion (de-focusing) of the Agent. This functional similarity of the two categories, repeatedly mentioned in the typological literature (see, among others, Comrie 1985:328ff.; Haspelmath 1987:29ff.; Shibatani 1985; Myhill 1997), accounts for their similar morphological marking in many languages. Some languages, nevertheless, make a morphological distinction between these two categories, cf. (5b–c) in Russian:

(5) Russian

a. *Petr* sloma-l-Ø derevo
   Peter:NOM broke-PAST-SG.M tree:ACC
   ‘Peter broke the tree.’

b. Derevo sloma-l-o-s’ (*Petrom)
   tree: NOM broke-PAST-SG.N-REFL (Peter:INS)
   ‘The tree broke (*by Peter).’

c. Derevo by-l-o sloma-n-o (Petrom)
   tree: NOM be-PAST-SG.N broke-PART.PERC.PASS-SG.N (Peter:INS)
   ‘The tree was broken (by Peter).’

1.2 Agentless passive vs. anticausative: Semantic distinction and its conceptual basis

Apparently, the only difference between (3) and (4) resides in the fact that the agentless passive, while leaving the Agent (Actor) overtly unexpressed, preserves it in the semantic structure, whereas the anticausative entirely removes the Agent from the inventory of semantic roles. In cases where the markers of the passive and anticausative overlap, passives without an overtly expressed Agent can be distinguished from anticausatives only semantically. The standard formulation of this semantic dichotomy is as follows: “Passive and anticausative differ in that, even where the former has no agentive

². Terms used to refer to this category include, in particular, ‘inchoative’ (a rather confusing term, widely used after Haspelmath (1993)), ‘unaccusative’, ‘ergative (intransitive)’, ‘quasi-passive’, ‘middle passive’, ‘pseudo-passive’ (see Shibatani 1985), ‘fientive’ (now common in the Indo-European scholarship); see Kulikov (2001a: 888) for a survey.
phrase, the existence of some person or thing bringing about the situation is implied, whereas the anticausative is consistent with the situation coming about spontaneously” (Comrie 1985: 326). However, this definition oversimplifies the real picture. The use of the anticausative morphology implies conceptualizing the corresponding event as spontaneous, even in cases where the presence of an external agent is possible, and even quite probable. Thus, the sentence *The door is being opened* strongly imposes the existence of someone opening the door. By contrast, when uttering the sentence *The door is opening*, we present the event as coming about spontaneously, even though in the majority of cases, an agent (someone who is opening the door) would be involved. In other words, the Agent, even if possible, is considered much more irrelevant than in the case of agentless passives – so irrelevant that the corresponding verbs (*The door is opening*, *The chair is breaking* [*by Peter*], etc.) are grouped together with verbs that denote processes which are necessarily spontaneous (*The tomato is rotting*, *The tree is growing*, etc.).

The affinity between these two categories is reflected in their historical connections. The grammaticalization path from non-passive derived intransitives (in particular, from reflexives or anticausatives) to passives is indeed well-known and has been repeatedly discussed in the typological literature; see, e.g., Haspelmath (1990); Heine & Kuteva (2002: 252f).

Much less studied is the opposite transition, from passives to non-passive intransitives (anticausatives). Yet this development is not infrequent – in particular, in a number of Indo-European languages. Importantly, this transition starts with agentless passives3 and suggests a compulsory intermediate stage of impersonalized passive. An unfortunate corollary of the delicate character of the passive/anticausative opposition is the fact that the passive to anticausative transition is only rarely explicitly mentioned in grammars and has not received due attention in the literature on the diachrony of passive, voice and valency-changing categories in general.

A very clear instance of such development is observed, in particular, in Old Indo-Aryan (= Vedic Sanskrit). After a short summary of passive formations in Vedic Sanskrit, which will be given in §2, I will scrutinize the passive to anticausative transition in Vedic (§3). §4 gives a short overview of typological parallels attested in some other Indo-European languages. §5 offers a general description of this development, paying special attention to its epistemic roots. The concluding §6 concentrates on some terminological aspects of the phenomenon in question, in particular, on the distinction between impersonal and impersonalized passives.

3. That is, passives constructed without agents; for definitions, see, for instance, Chalker & Weiner (1994:17); Matthews (1997).
2. Chronological and grammatical notes on Vedic

Vedic Sanskrit, or Vedic, which can be roughly identified with Old Indo-Aryan, is the earliest attested language of the Indo-Aryan group of the Indo-European language family and one of the most ancient attested Indo-European languages. The language of the most ancient Vedic text, the Rgveda (RV), can approximately be dated to the second half of the second millennium BC. The language of the second most ancient text, the Atharvaveda (AV), resembles in many respects – and is essentially synchronic with – the language of the late RV. Early Vedic (i.e. the language of the RV and AV) is followed by middle and late Vedic (covering the middle of the first millennium BC) documented in the Brāhmaṇas, Āraṇyakas, the oldest Upaniṣads and Śūtras.

There are several verbal formations in Vedic which can be employed in passive constructions; for details, see Kulikov (2006). In the language of the RV, there are three sets of passive formations, used for the three main tense systems, present, aorist, and perfect. These include:

i. **present** passives with the accented suffix -yā- (e.g. yuj ‘yoke, join’: 3rd person singular form yujyāte ‘is (being) yoked, joined’, 3rd person plural form yujyānte ‘are (being) yoked, joined’, participle yujyāmāna- ‘being yoked, joined’, etc.);

ii. **medio-passive aorists** with a defective paradigm, which includes 3sg. in -i, 3pl. in -ran/-ram, and participle with the suffix -āna- (e.g. yuj ‘yoke, join’: 3sg. āyoji, 3pl. ayujran, part. yujānā-); and

iii. **stative**, which supplies passives in the system of **perfect** and also has a defective paradigm (3sg. in -e, 3pl. in -re, and participle with the suffix -āna-); statives can be derived either from present stems (e.g. for the root hi ‘impel’, from the stem of the class V present hinó-/hinu-: 3sg. hinvé ‘(it) is/has been impelled’, 3pl. hinviré ‘(they) are/have been impelled’, hinvānā- ‘impelled’), or from perfect stems, thus being formally identical with middle perfects (cf. for the root yuj ‘yoke, join’: 3pl.

4. Abbreviations used for Vedic texts (text sigla) are the following:

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>AV(ś)</td>
<td>Atharvaveda (Śaunakīya recension)</td>
</tr>
<tr>
<td>AVP</td>
<td>Atharvaveda (Paippalāda recension)</td>
</tr>
<tr>
<td>JB</td>
<td>Jaiminīya-Brāhmaṇa</td>
</tr>
<tr>
<td>Kps</td>
<td>Kapiṣṭhala-Kaṭha-Saṃhitā</td>
</tr>
<tr>
<td>KS</td>
<td>Kāṭhaka</td>
</tr>
<tr>
<td>RV</td>
<td>Rgveda</td>
</tr>
<tr>
<td>ŠB</td>
<td>Śatapatha-Brāhmaṇa</td>
</tr>
<tr>
<td>TB</td>
<td>Taittiriya-Brāhmaṇa</td>
</tr>
<tr>
<td>TS</td>
<td>Taittiriya-Saṃhitā</td>
</tr>
<tr>
<td>VS</td>
<td>Vājasaneyi-Saṃhitā</td>
</tr>
</tbody>
</table>

5. Finite verbal forms are normally unaccented except when appearing in a subordinate clause and/or at the beginning of a sentence or metrical unit (pāda), i.e. a verse which forms the minimal constituent of a stanza.
yuyujrē ‘(they) are/have been yoked, joined’, yuyujānā- ‘yoked, joined’).6 By the end of the early Vedic period, the category of stative disappears.

There are also isolated passive occurrences of non-characterized middle forms (i.e. middle forms that have no specific passive morphology).

3. Passive to anticausative transition in early Vedic: Patterns and semantic classes

The passive to anticausative transition is attested for a few compact semantic classes of verbs. These include: (i) verbs of perception and knowledge; (ii) a verb of speech (vac); and (iii) a few verbs of caused motion.

3.1 Passives of verbs of perception and knowledge

The non-passive usages of the passives derived from verbs of perception of the type ‘is seen’ → ‘is visible; appears’ represent the commonest instance of passive to anticausative transition, and can probably be found in most languages with passives.

In Vedic, this development is attested for passives of a few verbs such as ḍṛyāte ‘is seen’ → ‘is visible; appears’ and śṛuṣte ‘is heard’ → ‘is known, is famous’. In what follows, I will discuss at length the meanings of these passives and the corresponding syntactic patterns.

3.1.1 ḍṛṣ ‘see’: ḍṛṣyā-te etc.

The present ḍṛṣyā-te regularly occurs in the agentless passive. In fact, its meaning always wavers between (agentless) passive (‘is seen’, ‘can be seen’) and non-passive (‘appears, is visible’). Cf.:

(6) (AVŚ 7.101.1)

yát svápne ánnam aśnámi ná prātār
if dream:LOC food:ACC eat:PRES.1SG.ACT not in.the.morning
adhigam-yá-te sárvam tád astu me śiváṁ
find-PRES-PASS-3SG all that be:IMPF.3SG I:DAT propitious
nahí tád ḍṛṣ-yá-te dívā
for that see-PRES-PASS-3SG by.day

‘If I eat food in my dream, [and it] is not found in the morning, be all that propitious to me, for that is not seen by day.’

6. For the category of stative (which only recently has acquired full rights in the synchronic grammar of early Vedic), its paradigm and its status within the verbal system, see Kümmel 1996 and Kulikov 2006.
Passive to anticausative through impersonalization

(7) (AVŚ 10.8.25ab)

bālād ekam anīyaskām utā- ēkaḿ nā- iva drś-ya-te

child:ABL.SG one more.minute and one not like see-PRES-PASS-3SG

‘One [thing] is minuter than a child, and another one is as if it were invisible.’

Note also that the participle drśyāmāna- (and, with the negative prefix, ādṛśyamāna-) functions as an adjective, meaning ‘(in)visible’, as in (8):

(8) (ŚB 3.6.2.26)

té ha sma- etá ubhaye deva-manusyāḥ pitāraḥ

that:NOM.PL.M PRTCL these both god-man:NOM.PL father:NOM.PL

sām pibante, sā- esā sampā; té
together drink:PRES:3.PL.MED that this symposium that:NOM.PL.M

ha sma drś-ya-mānā eva purā sām

PRTCL see-PRES.PASS-PART:NOM.PL.M formerly together

pibanta, utāitārhy ā-drś-ya-mānāḥ
drink:PRES:3.PL.INJ.MED but now not-see-PRES.PASS-PART:NOM.PL.M

‘And, verily, these both, the gods and men, [as well as] the fathers used to drink together, that is this symposium; formerly they used to drink together [with us] quite visibly, but now [they do so] invisibly.’

Alongside with (i) present formations with the suffix -yā-, other morphological passives attested for the root drś include: (ii) passive aorists (3sg. ādṛśi ‘(he/she/it) has been seen; has been visible, has appeared’; 3pl. ādṛśran/-ram ‘(they) have been seen, visible etc.’, participles drśāna- RV1x and drśānā- RV2x); (iii) the 3pl. form of the sigmatic aorist adṛks.ata (which replaces the passive aorist ādṛśran/-ram after the RV; see Narten 1964: 146), and (iv) stative (?) 3sg. dādṛśe/dadrśe (traditionally regarded as middle perfects),8 as, for instance, in (9):

(9) (RV 8.82.8ab)

yöm apsū candrāmā iva sōmaś

who:NOM.SG. water:LOC.PL moon:NOM.SG like Soma:NOM.SG

camuṣu dādṛś-e
camu:LOC.PL see:PERF-3SG.STAT

‘Soma, who has appeared in the camū-vessels, like the moon in the waters…’

7. Cf. Delbrück’s (1888:502) translation: “diese beiden, Götter und Menschen sowohl wie Väter, pflegten zusammen zu trinken, das ist die sampā, und zwar tranken diese vormals sichtbar mit, jetzt aber unsichtbar”.

8. See Kümmel (2000:233f.) For the paradigmatic status of these formations, see Kulikov 2006.
3.1.2 śru ‘hear’: śrūyā-te etc.
Like the verb drś, śru attests a very rich and nearly complete passive paradigm, which includes: (i) present passive śrūyā-te (attested from the late RV onwards); (ii) passive aorist (3sg. injunctive) -śrāvi (RV 10.93.14); (iii) statives derived from the present stem śṛnō-/śṛnu- (3sg. śṛnvē, 2sg. śṛnvīṣē, 3pl. śṛnvīrē); and (iv) middle perfect (stative derived from perfect stem?) śuśruve (RV 8.66.9).

There are two types of usages attested for these formations: (α) an agentless passive usage, typical of the passives of verbs of perception: ‘be heard,’ and (β) a non-passive intransitive (anticausative) usage, which easily develops from (α): ‘is heard [by smb.]’ → ‘is audible’ → ‘is known’ → ‘is famous.’

In usage (α), passives are constructed with the subject of sound properly speaking, cf.:

(10) (RV 10.168.4c)

\[
\text{ghośā id asya śṛṇ-ire nā rūpām}
\]

voice:NOM.PL only his hear:PRES-3PL.STAT not form:NOM.SG

‘Only his (= Vāta’s) voices are heard (are audible), not the form.’

The parallel verse in the Atharvaveda (Paippalāda) has a present form instead of the more archaic stative (which virtually disappears from the verbal paradigm after the RV), thus being its secondary replacement.\(^9\)

(11) (AVP 1.107.3c)

\[
\text{ghoṣa id asya śṛ-ya-te na rūpam}
\]

voice:NOM.SG only his hear:PRES-PASS-3SG not form:NOM.SG

‘His (= Vāta’s) voice is heard, not the form.’

(12) (JB 2.1:4–5)

\[
yām imām śṛṣṭhī vāc-am
\]

which:ACC.SG.F this:ACC.SG.F chairman:NOM.SG speech:ACC.SG

vadati […] sā hi dūrāc chrū-ya-te pronounce:PRES:3SG.ACT that:NOM.SG.F indeed from afar hear:PRES-PASS-3SG

‘The speech which the chairman pronounces […] is indeed heard from afar.’

(13) (RV 9.41.3ab)

\[
\text{śṛṇ-é vrśṭér iva svanāh}
\]

hear:PRES-3SG.STAT rain:GEN.SG like sound:NOM.SG

pāvamānasya śuṣmīn-aḥ

\(^9\) The Atharvavedic form is qualified by Renou (1947:69, with Footnote 1) as “variante … linguistiquement « normalisante »” or as “modernisme en face de RV śṛṇvire” (Renou 1965:40).
Pavamāna: gen.sg  rushing-gen.sg.m

'The sound of the rushing Pavamāna is heard, like [that] of the rain.'

Usage (α) is attested only for pres. śrūyāte (after the RV) and statives.

In the other usage, (β), passives (śrūyate, śṛṇvire etc.) are constructed with the subject of a person or an abstract concept (see Cardona 1961: 339ff.; Kümmel 1996: 115ff.), who/which is the source of sound, thus being metonymically associated with the sound and therefore can be said to 'be heard'/'be famous.' The only RVic occurrence of pres. śrūyā-" belongs to this type:

(14) (RV 10.22.1ab)

kúha śrutá ṛndraḥ
where heard:nom.sg.m  Indra:nom
kásminn adyá jáne mitró ná śrū-ya-te
which:loc.sg today people:loc friend:nom.sg as hear-pres.pass-3sg

'Where is Indra famous? In which community is he known/famous today as a friend?'

The passive present śrūyate becomes more common in later texts, cf.:

(15) (TBm 2.5.1.3)

śrōtreṇa bhadrám utá śṛṇv-anti satyám […]
ear:ins.sg good:acc.sg and hear:pres-3pl.act truth:acc.sg
śrōtreṇa módaś ca máhaś ca śrū-ya-te
ear:ins.sg joy:nom.sg and greatness:nom.sg and hear-pres.pass-3sg

'With ear they hear good and truth; […] with ear the joy and greatness are heard.'

The same usage is also attested for the passive aorist (-śrāvi RV 10.93.14), statives (śṛṇvé, etc.) and middle perfect (stative?) śuśruve (RV 8.66.9),10 cf.:

(16) (RV 3.55.20c)

śṛṇv-é viró
hear:pres-3sg.stat hero:nom.sg
vindá-máno vásūni
find:pres-part.med:nom.sg.m goods:acc.pl

'He is known as the hero finding goods.'

10. On this attestation and on other occurrences of the middle perfect, in the AVP and PB (passive or absolute transitive?), see Kümmel 2000: 532.
The inventory of formations attested in the two types of passive usages is summarized in the table below:

**Table 1. Usages of the passive formations of śru**

<table>
<thead>
<tr>
<th>Formations</th>
<th>(α) agentless passive</th>
<th>(β) non-passive (anticausative)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>pres. śṛyāte (AV+)</td>
<td>pres. śṛyāte (RV+)</td>
</tr>
<tr>
<td></td>
<td>stative (3sg. śṛyāve,)</td>
<td>etc. (RV)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>passive aorist (-śrāvi) RV 10.93.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>stative/middle perf. śuśṛuve RV 8.66.9</td>
</tr>
</tbody>
</table>

3.1.3 khyā (kśā) ‘see, consider, know’: khyāyā-ṭe
The present passive khyāyā-ṭe ‘is seen, known, etc.’ attested from the Brāhmaṇaṇas onwards, occurs in agentless passive constructions. This present formation is attested almost exclusively in compounds with preverbs, foremost with ā, as in (17):

(17) (ŚB 10.5.4.4)

what:nom.sg.n thus and rays:nom.pl thus and ākhyā-yā-te tāl lokam-ṛṇā
consider-pres.pass-3sg that:nom.sg.n lokam-ṛṇā:nom.pl

‘That what is called both ‘regions’ and ‘rays’ is the lokam-ṛṇā’¹¹ [bricks].’

In earlier texts, before the period of Vedic prose, passive usages are only found for rare non-present middle forms of khyā, viz. for the thematic aorist -ākhyā-ṭa (only with the preverb sām). It is attested in an intransitive usage, meaning ‘appear together (with smb./smth.),’ in some contexts with the additional semantic nuance ‘appear together, and, by virtue of that, be considered/become associated (with smb./smth.).’ This intransitive usage can only be based on the original agentless passive (‘be seen/considered together (with smb./smth.),’), and this semantics still shimmers through the actually attested meanings. sam-ākhyā-ṭa typically denotes a particular spiritual (sacral) contact or connection between deities or between a deity and his/her adepts. Cf. the two earliest occurrences (for a detailed analysis of the passive usage of this aorist, see Kulikov 2008):

11. Lit.: ‘filling the space,’ i.e. the brick set up with the formula lokam ṛṇā ‘fill the world!’.
(18) (RV 9.61.7c)

`sá m `ādityēbhīr akhya-ta
together Āditya:INS.PL consider:AOR-3SG.MED

'[Soma] has appeared together (and, by virtue of that, has become associated) with the Ādityas (a group of gods).\textsuperscript{12}

(19) (VS 4.23)

`sá m akhye devy-ā dhy-ā
together consider:AOR:1SG.MED divine-INS.SG insight-INS.SG

`sá m dākṣīnayā- urū-cakṣas-ā
together Dakṣinā:INS.SG far-seeing-INS.SG

'I have been considered together/have appeared together/(= I have become associated) with the divine insight, with the far-seeing Dakṣinā.'

3.1.4 \textsuperscript{2}vid ‘know’: vidē etc. and \textsuperscript{1}vid ‘find’: vidyā-\textsuperscript{te}, avedi

There are two synchronically distinct, but etymologically related, roots \textsuperscript{1}vid ‘find’ and \textsuperscript{2}vid ‘know’. Both meanings can readily be traced back to the original semantics of the Common Indo-European root *\textsuperscript{u}e)d- ‘see’, reconstructable for this root on the basis of evidence from other branches (cf. Latin \textit{videre}, Old Church Slavonic \textit{viděti} etc.; see Mayrhofer 1986–96 [EWAia]: II, 579f. for details), which is not preserved in Indo-Iranian.

The verbal system of \textsuperscript{1}vid ‘find’ includes, among others, the aorist ávidat, present vindāti, and reduplicated perfect vivéda. By contrast, the paradigm of \textsuperscript{2}vid ‘know’ consists of the perfect without reduplication (3sg. vēda etc.) and causative vedāyati ‘makes know’.

The inventory of forms of \textsuperscript{2}vid ‘know’ attested in passive usages includes two third-person forms, 3sg. vidē and 3sg. vidrē,\textsuperscript{13} as well as the participle vidāna- (with the accent on the root),\textsuperscript{14} which should be taken as statives rather than middle perfects (see Kümmel 1996: 104 and Kulikov 2006:67f.). Cf.:


\textsuperscript{13} Cf. Kümmel 1996:101f.

\textsuperscript{14} By contrast, the participle vidāna- (with the accent on the suffix) appears in non-passive constructions, in particular, in reflexive (‘knowing oneself’) and reciprocal usages (see Geldner 1917:331, footnote 2; Kümmel 1996:103f.).
Leonid Kulikov

(20) (RV 8.93.32ab)

\[
\begin{array}{l}
\text{dvitā yó vrtra-hān-tamo} \\
\text{twice which:NOM.SG.M resistance-breaker-SUPERLAT:NOM.SG} \\
\text{vid-ā indraḥ śatā-kratuḥ} \\
\text{know:PERF-3SG.STAT Indra:NOM.SG hundred-power:NOM.SG.M}
\end{array}
\]

‘… Indra, who is known in two aspects, as the best resistance-breaker [and] as possessing hundred powers.’

By contrast, the inventory of passive formations of \( \text{vid ‘find’} \) includes the -yá-passive \( \text{vidyá-} \) and the passive aorist \( \text{avedi.} \)

The -yá-present \( \text{vidyá-} \) is employed in the agentless passive (‘is found, can be found’), especially with negation (‘is not found, cannot be found’), as in (21–22):

(21) (RV 5.44.9cd)

\[
\begin{array}{l}
\text{ātrā nā hárdi kravaṇāsya reja-te} \\
\text{here not heart:NOM.SG Kravaṇa:GEN.SG tremble:PRES-3SG.MED} \\
\text{yātrā matir vid-yá-te pūta-bándhanī} \\
\text{where prayer:NOM.SG find-PRES-PASS-3SG purified-connected:NOM.SG.F}
\end{array}
\]

‘The Kravaṇa’s heart does not tremble here, where the prayer is found that is connected with the purified [Soma].’

(22) (AVŚ 19.50.4 = AVP 14.9.4)

\[
\begin{array}{l}
\text{yāthā śāmyākah pra-pāta-nn […]} \\
\text{as particle.of.śāmī.tree:NOM.SG forth-fly:PRES-PART.ACT:NOM.SG.M} \\
\text{nā- anu-vid-yá-te} \\
\text{not along-find-PRES.PASS-3SG}
\end{array}
\]

‘As a (tiny) particle of śāmī-tree,\(^{15}\) flying forth […], cannot be found …’

Such agentless passive constructions can easily be de-agentivized, developing non-passive intransitive, or anticausative, usages. Thus, ‘is found, can be found’ (and, very frequently, with negation: ‘is not found, cannot be found’) easily transforms to ‘exists, is’ (‘does not exist, is not’), as in (23); the distinction between the meanings ‘is found’ and ‘exists’ cannot be drawn with accuracy in many cases:

\(^{15}\) Whitney and Roth’s edition of the AVŚ conjectures ‘śyāmākah (‘millet-seed’) against the manuscript reading ś(y)āmyākah, but this emendation may be unnecessary. The word śāmyākah (attested in the manuscripts) could perhaps refer to a tiny particle of śāmī-tree (used for producing fire) (A. Lubotsky, p.c.).
3.2 The case of a verb of speech: vac ‘speak; pronounce; call’: ucyá-ite

Less trivial and typologically both rarer and more interesting is the instance of passive-to-anticausative transition attested for the passive of a verb of speech, vac ‘speak; pronounce; call’: ucyá-ite.

There are two transitive usages attested for the verb vac: (α) with the accusative of speech (‘XNOM says, pronounces SACC’), as in (24); and (β) with accusatives denoting the object of nomination and his/her/its name (‘XNOM declares YACC to be ZACC’; ‘XNOM calls YACC ZACC’), as in (25):

(24) (RV 1.78.5ab)
á vocāma rāhūgaṇā agnāye
say:AOR:1PL.ACT R.:NOM.PL Agni:DAT.SG
mādhumad vácaḥ
honey:sweet:ACC.SG.N speech:ACC.SG
‘We, the Rahūgaṇas, have pronounced for Agni a honey-sweet speech.’

(25) (RV 3.54.19)
devānāṃ dūtāḥ [...] ánāgān
god:GEN.PL messenger:NOM.SG sinless:ACC.PL.M
no voca-tu
we:ACC say:AOR-3SG.IMPV.ACT
‘Let the messenger of the gods … declare us [to be] sinless.’

3.2.1 Agentless passive usages of ucyá-ite

The passive counterpart of the latter transitive pattern, (β pass.), is the agentless passive constructed with two nominatives (‘YNOM is called ZNOM’), as in (26–27):

(26) (RV 10.97.6cd)
vipraḥ sā ucy-ya-te bhiṣāg
poet:NOM.SG that:NOM.SG.M say-PRES-PASS-3SG healer:NOM.SG
‘That poet is called healer …’

In fact, this passive usage tends to be conceived as an anticausative, meaning ‘have the name N’, which, together with its transitive counterpart, can be considered as a causative pair ‘have the name N’ ~ ‘make have the name N, give the name N, call’.

3.2.2 Anticausative usages of uc-ya-te

The transitive usage α underlies two intransitive patterns: (α_pass.) the regular passive ‘SNOM is said, pronounced’; and (α_anticaus.) the non-passive intransitive (anticausative) pattern ‘SNOM sounds’. An example of the passive pattern is given under (28):

(28) (RV 1.114.6a)

\[
\begin{array}{llll}
\text{idám} & \text{pitr-é} & \text{marút-ām} \\
\text{this:nom.sg.n} & \text{father-dat.sg} & \text{Marut-gen.pl} \\
\text{uc-ya-te} & \text{vācaḥ} \\
\text{say-pass-3sg} & \text{speech:nom.sg} \\
\end{array}
\]

‘This speech is pronounced for the Maruts’ father.’

The latter, anticausative, pattern\(^{17}\) has undoubtedly developed from the ordinary passive α_pass. through the stage of the agentless passive (‘X is pronounced [by smb.]’ → ‘X sounds’). Subsequently, this usage could be expanded to a larger class of subjects, so as to include, alongside sounds proper (speech, songs, etc.), a variety of sources of sound, such as tongue, pressing-stone,\(^{18}\) etc. Such a semantic derivation is typical of verbs of sounding; see Padučeva (1998: 19).

This type of usage has never become productive, however. The anticausative type ‘sound’ is much rarer than passive (α_pass.), ‘be pronounced’. It only appears a few times in the RV and disappears in later texts. The following three Rigvedic occurrences are taken by most or all translators as examples of the anticausative type:

(29) (RV 10.64.15c = 10.100.8c)

\[
\begin{array}{llll}
\text{grāvā} & \text{yātra} & \text{madhu-śūd} \\
\text{pressing.stone:nom.sg} & \text{where} & \text{honey-sweet:nom.sg.m} \\
\text{uc-ya-te} & \text{bṛhát} \\
\text{say-pass-3sg} & \text{loudly} \\
\end{array}
\]

‘… where the honey-pressing stone sounds loudly.’

---

\(^{17}\) Qualified by Grassmann (1873: 1191) as “Passiv mit medialer Bedeutung”; cf. also Velankar (2003: 174): “This passive form is used in the active sense of ‘resound’.”

\(^{18}\) On pressing-stones (grāvan-) and the sounds made by them, see Wright (2008).
A remarkable feature shared by all these constructions is the presence of an adverb (brhát ‘loudly’) or an adjective (urúci ‘far-reaching’) syntactically connected with the verb or subject, respectively, which expresses the loud and far-reaching character of the sound: ‘SNOM Adv./Adj. loud(ly) sounds’.

Yet another occurrence of the passive ucýate that probably belongs to the same type is attested at RV 6.28.6 (see Kulikov 2001b: 164f.; contra the interpretations offered by most translators).

To sum up, both varieties of agentless passive usages attested for ucýá-te drift to the non-passive (anticausative) type. However, while for (β pass.), the two types of usages (‘YNOM is called Z NOM’/‘YNOM has the name Z NOM’) remain virtually indistinguishable, in the case of (α pass.) the semantic shift from ‘SNOM is said, pronounced’ to ‘SNOM sounds’ constitutes a drastic lexicalization.


20. Thus Geldner (1951:1, 405), Elizarenkova (1989:351 and 718). Renou (1958 [EVP IV]: 56) was hesitating in his comments on this passage (“ucýate, « erklingt » G[e]ld[ner], en conformité avec les passages où le sujet est grávan […] Mais l’emploi usuel mène au sens de: est appelée, est dite”), but has eventually adopted Geldner’s interpretation (“Ta langue […] qui se fait entendre chez les dieux’) in his translation (Renou 1959 [EVP V]: 18).
3.2.3 Verbs of speech and verbs of perception and knowledge: Systemic relations

In spite of the fact that we do not find other verbs of saying or sounding instantiating the same passive to anticausative transition as attested for ucyá-te ‘sound’ (← ‘be pronounced’)/‘have the name’ (← ‘be called’), this passive verb should not be considered isolated. In fact, vac ‘say, pronounce, call’, as a member of the class of verbs of speech, can be grouped together with verbs of perception (discussed in §3.1) for good reasons. Altogether, both verbs of perception and verbs of speech form a larger semantic class, which can be determined as verbs of knowledge transfer. By virtue of some particular semantic and epistemic features, verbs of this class display a particularly clear tendency to develop anticausative usages from their passives.

Note that yet another verbal root, 1vid ‘find’ (etymologically related to 2vid ‘know’; see 3.1.4), can also be considered as belonging to the semantic class of knowledge transfer, together with such verbs as ‘see’ or ‘know’.22 As a matter of fact, the process of finding an object has direct implications for our knowledge of this object. This accounts for the semantic development ‘see’/‘know’ → ‘find’ attested for this verbal root in Indo-Iranian.

This semantic feature of 1vid ‘find’ may be responsible for the agentless passive usage attested for its passives, the present vidyá-te and the aorist avedi. As mentioned above (§3.1.4), these passives readily develop into non-passive intransitives (anticausatives): ‘is found, can be found’ → ‘exists, is’, as illustrated in (21–22). Note that this semantic development is typical of intransitive (passive or reflexive) derivatives of the verb ‘find’; cf. Germ. sich finden, Fr. se trouver, Ital. trovarsi, Pol. znajdować się, Russ. naxodit’-sja, etc. Needless to say that the distinction between the meanings ‘is found’ and ‘exists’ cannot be drawn with accuracy in many cases.

3.3 Passive to anticausative transition in other semantic classes

Yet another semantic class which exhibits the passive to anticausative transition consists of a few verbs of causation of motion. These include, among others, the present passives -kíryá-te ‘be scattered; fall (down)’ (root kṛ), rudhyá.-te(ā) ‘be kept’; move, adhere’ (in some compounds) (root rudh), sicyá-te ‘be poured; pour (out)’ (root sic), srjyá-te ‘be set free, be emitted; run’ (root sic), and vacyá-te ‘move (waveringly)’ (root vañc). For some constructions, both passive and anticausative interpretations are possible (as in (33)). Cf.:

21. See also Rogers (1971, 1972) on the semantic affinity of verbs of perception and such predicates as sound.

22. See, for instance, Sibley (1955) on relationships between seeking, finding and seeing.
In example (34), it is particularly clear that the demons are not thrown but fall; that is, we are confronted with a spontaneous event, indirectly triggered by pulling out a brick.

Again, the origin of such non-passive usages must lie in their semantics, but the scenario of the passive to anticausative transition is different from the one attested for verbs of perception. Verbs belonging to the class of caused motion, such as throw (= ‘make fall, make fly’), send (= ‘make go, make move’), etc., can easily be conceptualized as causatives. Since for many such verbs, the present passive with the suffix -ya- is the only regular intransitive derivative, it could occasionally take over the anticausative function. This secondary function could further be supported by the influence of the middle non-passive presents with the suffix -ya- and root accentuation (class IV presents in the traditional Indian classification) derived from some verbs of motion,
such as pádyate ‘falls’ or réyate ‘whirls, swirls’ (see, for instance, Kulikov 1997). Note that these anticausative -ya-presents are typically opposed to morphological causatives (cf. pādāyati ‘makes fall’, rīnāti ‘makes whirl, makes swirl’). This type of passive to anticausative transition deserves a separate study and will not be discussed in detail in this paper.

4. Typological parallels from other Indo-European languages

The cases of passive to anticausative transition attested in early Vedic and discussed at length in the preceding section do not represent of course a unique typological feature of the Old Indo-Aryan linguistic system. Similar phenomena can also be found both within the Indo-European linguistic family and beyond.

Thus, Latin grammatical studies repeatedly notice the fact that some morphological passives may be employed in usages which do not instantiate a ‘canonical passive type’, but rather should be qualified as ‘medio-passives’. In such usages, the activity in question is conceived as ‘internal’, rather than originating from some external agentive force – whence the term ‘passif intrinsèque’ (used in the French literature on Latin passives; see Flobert 1975: 37; Touratier 1994: 175f.), as opposed to the canonical ‘passif extrinsèque’. This is, in particular, the case with the Latin passive videri ‘be seen’ → ‘appear, seem’ (which can be constructed with the dative of Experiencer), as in (35):

(35) (Cic., Lael. 86) (Touratier 1994: 562f.)

\[
\begin{align*}
\ldots & \text{cetera} & \text{quae} & \text{quibusdam} \\
\text{other: NOM.PL.N} & \text{that: NOM.PL.N} & \text{some: DAT.PL} \\
\text{admirabilia} & \text{videntur} & \ldots \\
\text{wonderful: NOM.PL.N} & \text{see: PRES.PASS:3PL} \\
\text{‘… other things that seem wonderful to some people …’}
\end{align*}
\]

Similar semantic shifts are attested for a number of other Latin passives, in particular, for some verbs of caused motion, cf.:

(36) (Caes., B.G. IV, 10, 3) (Claflin 1946: 205)

\[
\begin{align*}
\text{Rhenus [...] citatus fertur} \\
\text{Rhine quick carry: PRES.PASS: 3SG} \\
\text{‘The quick Rhine rushes …’}
\end{align*}
\]

23. See also Claflin (1942: 1943) on ‘middle’ usages of Latin passives. For a discussion of the terminology, see Flobert (1975: 36f., Footnote 7).
Obviously, such usages instantiate a development similar to what we observe in the case of Vedic verbs of perception and knowledge transfer. Yet, there is an important difference between these seemingly identical diachronic processes attested in two genetically related languages, Vedic and Latin. Latin morphological passives of the type amatur are historically related to the Proto-Indo-European middle (see, e.g., Claflin 1927:160ff. and Sihler 1995:470ff. for a general survey), which is at least one of the diachronic sources of the Latin passive. Accordingly, its non-passive (‘medio-passive’) usages may instantiate archaic traces of the earlier linguistic situation, being the vestige of (some) non-passive functions of the middle forms in the proto-language.

By contrast, Vedic passive presents with the suffix -yá- and passive aorists in -i/-ran (-ram) represent an Indo-Iranian innovation (see, e.g. Szemerényi 1990:271; Kümmel 1996), which was not based on the Proto-Indo-European middle. It was specially created to express the passive function, taking it over from the middle of the proto-language. Its non-passive usages cannot therefore be explained as archaisms.

Similar phenomena are attested for passives of verbs of perception and knowledge transfer in many other languages, both within the Indo-European language family and beyond. Thus, in Modern Greek, where the passive function is regularly rendered by middle forms, passives of verbs of perception are commonly employed in agentless, or ‘generic’ (in terms of Manney 2000), usages, as in (37):

(37) Mod. Greek

\[ \text{akústike óti óli i ipálili} \]
\[ \text{hear:3sg:mid/a comp all-the-employees:nom} \]
\[ \text{tu ipáryiu laðóðikan} \]
\[ \text{the-ministry bribe:3pl:mid/a} \]

‘It was rumored that all the employees of the ministry were bribed.’

Note also the Slavic reflexes of the Indo-European root for ‘see’ and ‘know’, *\( \text{uèi} \)èd-, which display similar features. Thus, the Old Church Slavonic present passive participle made from this root, vidimъ, regularly shows up with the meaning ‘visible’, rather than ‘being seen’ (cf. lexicographic marks such as ‘used as adjective’ in Blagova et al. 1994:115), even in spite of the presence of an instrumental agent, as in (38):

24. Flobert’s classification of the Latin ‘passifs intrinsèques’ (= anticausatives based on passives) is quite different, however, and in some respects, appears not quite adequate. Thus, the passive videri ‘be seen, appear’ is grouped with such verbs as gigni ‘be born’, frangi ‘break’ etc. under the cover term “mutatifs (être, paraître, situation)”.

25. This is the claim advocated, in particular, by Claflin (1927, 1942).

26. The only passive formation which must be historically related to the Proto-Indo-European middle is the stative, virtually limited to the language of the RV.
The same semantic shift is attested for the cognates of these forms in many other Slavic languages – in particular, in Russian (cf. видимый ‘visible’ etc.).

5. The de-agentivization of passives: A general scenario

5.1 From passive to anticausative through impersonalization

The scenario of evolution within the system of voices discussed above is represented below within the framework of the Leningrad/St. Petersburg Typology Group:

![Diagram of passive to anticausative transition]

Figure 1. General scenario of passive to anticausative transition

For the sake of convenience, this schema opens with the pattern representing the canonical passive (which is directly related to the corresponding transitive pattern). Most importantly, between the stages of agentless passive and anticausative we observe a compulsory transitional stage of impersonalization. While the standard agentless passive can be conceived as mere omission of the agentive noun phrase from the structure,
its impersonalization implies a much more drastic change to the semantics of the sentence – specifically, to the referential status of the Actor. The Actor is conceptualized as generic, or non-referential (cf. (III)). This type can readily evolve further into a pattern where the Actor is entirely lacking from the structure – that is, into the anticausative.

Such a diachronic scenario is theoretically possible for all agentless passives, but, in fact, it is particularly common for a few compact, relatively small semantic classes of verbs, such as verbs of perception and knowledge transfer. The epistemic roots of this development must reside in the semantic nature of these classes.

5.2 Epistemic roots of impersonalization

To begin with, let us take a closer look at the semantic evolution of the passive of the Indo-European verbal root *uèièd- 'see'/’know’, which instantiates a typical passive to anticausative transition. As already mentioned, this scenario can be outlined as follows:

(39) ‘Y is seen (known etc.) by smb.’
    \↓
‘Y is seen (known etc.) [by smb.]’
    \↓
‘Y is seen (known etc.) [by Øgeneric]’
    \↓
‘Y is visible (famous, etc.)’.

In all such cases we observe the rise of the anticausative usages through the stage which is called 'impersonalization' by Siewierska (1984: 241 et passim). The epistemic roots of such impersonalization of passives must lie in the phenomenon of 'objectivization of knowledge'. In other words, knowledge is easily conceptualized as having no subject – i.e. as 'knowledge without a knowing subject'; see, for instance, Lyons 1979: 129 (referring to Popper 1972; see especially p. 73f. and Chapters 3–4) and Ziff 1984: 12ff. For a detailed discussion of the objectivization of knowledge, see Rojszczak 2005: 146ff., where this process is explained as based on several conditions: (1) “making knowledge independent of time”, which suggests, in particular, that (1.1) “we do not need any knowing subject at all to perform an act at a particular point in time, because knowledge exists without any empirical subject”; (1.2) “absolutization of knowledge with regard to the act of knowledge”; … (2) making knowledge independent (objective) with regard to space; etc.

It is not necessary to enter here into a detailed discussion of the philosophical and epistemic aspects of objective knowledge. For our purposes, suffice it to mention that this process easily accounts for the impersonalization of agentless passives. Accordingly, the agentive force can be entirely removed from the scene, which results in a complete ‘de-agentivization’ of the original (agentless) passive. Subsequently, such de-agentivized passives evolve into anticausatives, as shown in the scheme above.
6. Concluding terminological remarks: 
Impersonal vs. impersonalized passive

To conclude this short discussion of impersonalized passives, it is worth pointing out an important terminological distinction, closely related to the subject of the present paper.

Impersonalized passives, often nearly indistinguishable from agentless passives, are sometimes (quite understandably) labelled ‘impersonal passives’. Thus, on the German web-site http://www.englisch-hilfen.de, which provides grammatical help for German students learning English, the phrases (40–41) (see http://www.englisch-hilfen.de/en/grammar/personal_passive.htm) are qualified as ‘impersonal passives’ – in spite of the presence of overt subjects in both cases:

(40) *It is said that children are afraid of ghosts.*

(41) *Children are said to be afraid of ghosts.*

Yet, in the literature on passives, the term ‘impersonal passives’ (also known as back-grounding passive) is typically employed to refer to quite a different phenomenon: passives with no direct object to subject promotion, where, accordingly, the subject position remains vacant. Handbook examples of impersonal passives are usually taken from such languages as Dutch, German or Polish, cf., for instance, examples quoted in Siewierska (1984: 96f). (Ch. 3 is dedicated to the phenomenon of impersonal passive):

(42) German 
*Es wurde getanzt.*

‘There is dancing.’

This frequent terminological confusion is due to the properties of the two phenomena in question. The affinity of the corresponding terms originates, quite naturally, in the structural affinity of the impersonalized and impersonal passive (see especially Siewierska 2008), as illustrated in the scheme below:

<table>
<thead>
<tr>
<th>Impersonal passive</th>
<th>Impersonalized passive</th>
</tr>
</thead>
<tbody>
<tr>
<td>(subjectless passive)</td>
<td>(agentless passive)</td>
</tr>
<tr>
<td>$X_{\text{generic}}$</td>
<td>$Y$</td>
</tr>
<tr>
<td>(Obl/) –</td>
<td>DO</td>
</tr>
<tr>
<td>$(X_{\text{generic}})$</td>
<td>$Y$</td>
</tr>
<tr>
<td>–</td>
<td>S</td>
</tr>
</tbody>
</table>

Figure 2. Impersonal vs. impersonalized passive

Impersonal passives do not promote the direct object to the subject position and therefore lack a subject. By contrast, impersonalized passives do promote the direct
object to the subject position. Thus, whilst impersonalized passives can be described as agentless, impersonal passives are subjectless – which, however, does not rule out the presence of an agent (surfacing as an oblique noun phrase). This constitutes a crucial difference between these two voice categories. Yet, in both cases the status of the main participant of the situation, the Actor, is similar; it ranks somewhere between unimportant and non-specified, non-referential or generic. In both cases, this feature leads to the removal of the corresponding noun from the structure.27

Another important difference between these two categories concerns their diachronic potential. Impersonalized passives can easily evolve into anticausatives, while for impersonal passives this path of evolution is closed.

This terminological confusion remained almost unnoticed in the literature till recently. Rare exceptions included Siewierska (1984) and Pinkster (1992:163), who rightly pointed out that “[t]he term ‘impersonal’ is also used, however, as some sort of synonym of ‘agentless,” mentioning a few standard grammars, where this terminological confusion can be found. Now, fortunately, more attention is paid to this important distinction; see Siewierska 2008, distinguishing between subject-defocusing and agent-defocusing approaches to the analysis of impersonal passives.

Future research should pay more attention to the study of the impersonalized passive and, particularly, to its diachronic aspects. Of particular relevance are the relationships between impersonalized passives and anticausatives as well as, in general, their position and possible paths of development within the system of valency-reducing categories.

Acknowledgments

I am much indebted to Michela Cennamo, Andrej Malchukov, and Anna Siewierska for suggestions and critical remarks on earlier drafts of this paper.

Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACC</td>
<td>accusative</td>
</tr>
<tr>
<td>ACT</td>
<td>active</td>
</tr>
<tr>
<td>AOR</td>
<td>aorist</td>
</tr>
<tr>
<td>FUT</td>
<td>future</td>
</tr>
<tr>
<td>DAT</td>
<td>dative</td>
</tr>
<tr>
<td>DO</td>
<td>direct object</td>
</tr>
<tr>
<td>F</td>
<td>feminine</td>
</tr>
<tr>
<td>OPT</td>
<td>optative</td>
</tr>
</tbody>
</table>

27. The ‘impersonal’ character of the agentless passive is noticed, for instance, by Luukka & Markkanen (1997:174ff).
References


PART III

Cross-linguistic variation in Impersonal constructions
Case studies
The Maa (Eastern Nilotic)
Impersonal construction

Doris L. Payne
University of Oregon & SIL International

The Maa (Maasai) Impersonal construction occurs with intransitive, (di)transitive, active and stative verb stems. It implies either no agent or at most unspecified non-referential “people” as a possible agent. Likely because of the possible implication of “people”, it has not been found with meteorological verbs. The Maa Impersonal may focus on an event or situation, and is sometimes functionally used where passive constructions are found in other languages. It can further combine with ata 'have' to render an existential function. Though structurally this existential is entirely the same as the Impersonal, the existential cannot so easily be said to always have an impersonal sense. Historically the key morphological element of the Impersonal derives from a plural affix, but it is not certain this was necessarily pronominal.

Keywords: Maasai; Nilotic; impersonal; passive; historical development

1. Introduction

Various typologies and definitions have been offered as to what an “impersonal construction” is or should be. The definitional issue throws us immediately into the typologist's dilemma: does one approach the domain from a functional or from a structural perspective? That is, do we assume that the proper definition a linguistic structure is grounded first and foremost in semantics or function, and only then go about determining which constructions have that communicative or semantic profile? Or do we assume that the proper definition is grounded first and foremost in a certain morphosyntactic profile, and then determine what communicative functions or concepts that structure fulfills? Analyses grounded solely in either extreme can lead to conclusions that don't always “feel quite right”. For example, an extreme function-first approach to transitivity may end up calling What if she already ate at home? an “anti-passive”; or saying that a word that lexicalizes the concept of BIG, or that modifies a noun, in whatever language is necessarily an “adjective”. On the other hand, an extreme
form-first approach could end up saying that [T]his struggle goes largely unnoticed by the children is an “ergative”.¹

With regard to “impersonal constructions”, Siewierska (2008) notes that both structural and communicative-functional characterizations exist in the literature. A function-first approach grounds the definition of an impersonal in agent defocusing. But a strict interpretation of an impersonal as any construction which defocuses an agent would end up including as impersonals what are commonly called personal passives, locative subject clauses, and (in some languages) constructions with predicates that express emotion, need, etc. On the other hand, a structure-first approach to impersonal constructions centers on the construction having either no subject or a non-canonical subject.² But a strict interpretation of this sort could end up including existential and locative-presentative constructions, and even transitive ergative constructions, as impersonals.

The results of both extreme functionalist and extreme structuralist definitions thus often seem wrong-headed, at least as a cross-linguistic definitional strategy. Most functionalists and cognitively-oriented typologists tend toward the function-first approach (cf. Croft 2001; Givón 1984/1990; Thompson 1989), arguing that (a sophisticated view of) function, conceptualization, or semantics can be the only grounds for meaningful cross-linguistic comparison. But other typologists eschew an extreme or reductionist function-first approach to determining linguistic categories, and adopt something of a hybrid approach that combines structural and functional criteria in identifying what is a “passive”, an “adjective”, an “impersonal”, etc., construction.

This paper investigates what I call the Maa (Eastern Nilotic) Impersonal construction.³ My use of the term impersonal for this Maa construction is based on the hybrid


2. This approach assumes that one can identify a homogenous “subject” category cross-linguistically – which is far from a consensus view. See also Malchukov and Ogawa (this volume) for discussion of R(eferential), T(opicality), and A(gentivity) functional properties of “subjects”.

3. Maa is the term used for the language or languages spoken by Maasai, Il-Chamus, Samburu, Il-Parakuyo and other associated ethnicities. Data for this paper come primarily from southern Kenyan Maasai, though some textual data is from the northern Chamus variety. Data were collected as part of a text and lexicography database project supported by a Fulbright Foundation grant (1993–1994), and NSF grants SBR-9616482 (1987–1999) and SBR-9809387 (1998-ff) to Doris Payne and the University of Oregon. This work has also crucially relied on logistical, technical, and other assistance from the Maasai Cultural Center, the University of Nairobi, the Nairobi Evangelical Graduate School of Theology, SIL International and
approach to defining categories: the primary construction under discussion has the functional effect of defocusing any agent in the sense that no particular agent may be expressed and any sense of an agent can at most be taken only as some unspecified and non-referential “people”. Consonant with the functional properties, the formal subject of the construction is non-canonical compared to the majority of clausal constructions in the language. However, my bias for determining the full scope of what is under discussion in this particular paper is admittedly language-specific and, ultimately, structural. That is, I identify a particular Maa constructional form which, with great frequency, involves an unspecified agent; but then I attempt to widely explore the full range of functions that this constructional form has – whether or not all of those, strictly speaking, involve an unspecified agent. This approach may then lead us to a “family” of constructions which have some conceptual or historical core.

§2 introduces the basic structural facts of the Maa Impersonal construction: there is a suffix -i (or -i) glossed as “Impersonal” (IMP), the verb is inflected with a third person bound pronominal prefix, and only Accusative full NPs or Accusative free pronouns may (optionally) occur to express one or more non-Agents involved in the event or situation. We will see that this Impersonal construction can occur with intransitive, (di)transitive, active and stative verb stems. However, the construction itself is at least semantically if not also syntactically detransitivized, disallowing a Nominative NP in the clause even though the verb is marked as if it had a third person subject.

§3 explores the functional dimensions of the Maa Impersonal construction. First, it is used when the communicative focus is on the event or situation, as in something like ‘(Unspecified) people make-noise’ or ‘Noise-making happens’. Second, it is used in a range of what I will call “functional passive” situations. Together these functions appear to span the active impersonal and passive impersonal range noted in other languages. Third, the suffix -i can be used with the stative possessive verb ata ‘have’ to render an existential function, as in something like ‘There is(n’t) a leader.’ Though the structure of this Existential construction is entirely the same as the Impersonal, it cannot so easily always be said to have an “impersonal” sense.

§4 explores whether the Accusative NP in an Impersonal construction has acquired any subject properties. The general conclusion is that it has not. Finally, §5 discusses historical origins. The suffix -i appears to derive from a plural affix, but it is not clear whether this was necessarily a (third person) pluralizer, or a true pronominal

numerous colleagues. I particularly acknowledge the patient help of Sara Tukuoo, Leonard Ole-Kotikash, and Keswe Ole-Mapena, among many other Maa speakers, and significant elicitation by Kent Rasmussen. This work could not have been done without them. However, they are not responsible for errors of data or interpretation.
element. In the contemporary language, -i still can occur as an optional plural marker in a non-impersonal construction. Finally, we note a related Impersonal Existential which lacks the -i suffix.

2. Basic morphosyntax of the Maa Impersonal construction

Maa sentences are dominantly verb-initial, though constituent order variation does occur. The language has abundant morphological resources, and an inflected verb may constitute a complete sentence. The language has a “marked nominative” case system (König 2006), with formal contrast between what Tucker & Mpaayei (1955) (hereafter “TM”) called “Nominative” versus “Accusative” forms. I follow TM’s terminology in this paper. Case is marked on nouns and their dependents solely by means of tone. There are multiple tone classes, but each singular or plural noun form has just two tonal possibilities.

Table 1 presents the basic bound pronominal prefixes on verb forms (see Payne, Hamaya & Jacobs 1994 for full discussion of the system). This table actually contains two sets of prefixes. The first set is inside the highlighted box, and is functionally determined by an A “low” on an animacy hierarchy acting on a first or second person singular O; thus we have used the term Inverse for this set (and in particular, for the highlighted kĩ- prefix). In basic clausal constructions, the Inverse set occurs only with (di)transitive stems. The second prefix set, presented outside the highlighted box in Table 1, is used for (di)transitive verbs with a 3rd person or plural object, any 1st plural subject, and for intransitive verbs. Note that there are two prefixes with the segmental form /ki/ – one inside and one outside the highlighted box. The highlighted Inverse kĩ-behaves in a tonally distinct manner from the first person plural kĩ- prefix. Often the first plural prefix has a HL tone while the Inverse is just H. (The tonal details depend on overall stem form and aspect, but Inverse versus first plural inflected verb forms are always tonally distinct.)

4. Some modern typological linguists may object to the “Nominative”/“Accusative” terminology for Maa, as both case forms are used for more than just coding the core-argument groupings of A + S/O. However, several linguists working on Maa have followed TM’s terminological tradition, and in order not to proliferate terminology for studies of this language, in at least this paper I will continue to follow this tradition.

5. Note that the distributions of the highlighted kĩ- and aa- forms do not yield themselves to a simple hierarchical system in which the higher argument is simply marked on the verb regardless of whether it is A or O. By themselves the aa- forms would suggest a hierarchy with 1sg as the highest position (i.e. 1sg>2sg>3); but the kĩ- forms suggest a hierarchy with 2sg as the highest position (i.e. 2sg>1sg>3).
Table 1. Basic bound pronominal prefixes on verbs

<table>
<thead>
<tr>
<th></th>
<th>1sgO</th>
<th>2sgO</th>
<th>3 or plural O</th>
<th>Intransitive (i.e. no O)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sgS/A</td>
<td>áá-</td>
<td>á-</td>
<td>á-</td>
<td></td>
</tr>
<tr>
<td>2S/A</td>
<td>kí-</td>
<td>í-</td>
<td>í-</td>
<td></td>
</tr>
<tr>
<td>3S/A</td>
<td>áá-</td>
<td>kí-</td>
<td>kí-</td>
<td>kí- + L tone</td>
</tr>
<tr>
<td>1plS/A</td>
<td>kí- + L tone</td>
<td>kí- + L tone</td>
<td>kí- + L tone</td>
<td></td>
</tr>
</tbody>
</table>

As noted earlier, the Maa Impersonal construction is formed with a suffix -i (IMP) on the verb. This suffix follows all other suffixes that the verb may carry, suggesting that it is a comparatively more recent morphological element. Compare the personal transitive construction in (1a–b) with the Impersonal in (1c). Note the tone differences on the optional free pronouns in (1a–b) and (1c), which mark case.6

(1) a. É-dúny (nínýe).
   3-cut 3SG.NOM
   ‘He will cut it.’

   b. É-dúny ə-l-morrání (nínýé).
   3-cut MSG-WARRIOR.NOM 3SG.ACC
   ‘The warrior will cut him.’

   c. E-duŋ-i (nínýé).
   3-cut-IMP 3SG.ACC
   ‘He/it will be cut.’

As the contrast between (1c) and (2) shows, the Impersonal construction cannot take a Nominative case free pronoun or full NP, but only an Accusative one.

(2) *E-duŋ-i nínýe.
   3-cut-IMP 3SG.NOM
   ‘It/He will be cut.’

The following set shows all the same facts for full NPs; compare the case forms for ‘God’ in the non-Impersonal clauses (3) and (4), versus (5).

(3) K-é-ışhū ənkuí.
   DSCN-3-live FSG-GOD.NOM
   ‘God is alive.’

6. Out of context, a 3rd person bound pronominal or (singular) free pronoun could be translated as ‘he/him, she/her, it’. For simplicity I usually just use one of these in translating particular examples. Maa is an aspect language. Out of context, verbs that are not marked for aspect can be translated into English as future or present, but this same form can also surface in past-time discourse contexts.
Example (6) now shows that in the Impersonal Construction, the form of ‘God’ is like that in (5), i.e. Accusative rather than Nominative. Example (7) demonstrates the impossibility of including a Nominative NP in the construction.

The Impersonal construction always takes a bound pronominal prefix on the verb (cf. Table 1). The allowable prefixes are all and only those that would correlate with a 3rd person S or A (i.e. from the row marked with an arrow in Table 1). Hence, in addition to ε- as illustrated in (1c) above, áa- and (the highlighted) Inverse kí- are also possible. In the Impersonal, the Inverse kí- cannot be interpreted as ‘you (sg.) act on me’ (an interpretation which is otherwise always possible with this prefix), but only as having some impersonal “people” as the A.

The Maa Impersonal is quite promiscuous in the sense that -i can occur with nearly every verb tested, whether intransitive, (di)transitive, active or stative. In (1c) above it occurs with a simple transitive active root. The acceptability of this suffix with active stems is explicitly revealed in the following, where a Progressive form occurs.
(10)  
\textit{E-ibel-i'ó-i}  
\textit{3-rock-PROG-IMP}  
'It is being rocked from side to side.'

In (11) it occurs on a basic active ditransitive stem; here we see two Accusative NPs in the clause.

(11)  
\textit{e-\text{ish}-\text{f}-\text{k-i} \text{en-kitók} \text{en-dáa}.}  
\textit{3-give-PF-IMP \text{FSG-woman.ACC} \text{FSG-food.ACC}}  
'The woman was given food.'

In (12) and (13) it occurs on derived ditransitive stems, formed with the Dative and Instrumental applicatives, respectively.

(12)  
\textit{Áa-tt-gil-ak-ak-f ol-páéki.}  
\textit{3>1SG-PF-break-DAT-PF-IMP \text{MSG-maize.ACC}}  
'The maize was broken for me.'

(13)  
\textit{E-idony-iék-i o-rinká shímpi.}  
\textit{3-pound-INST-IMP \text{MSG-club.ACC} \text{salt.ACC}}  
'The club was/will be used to beat the salt.'

In (8)–(9) above it occurs with a transitive stative stem. It may also occur on basic stative intransitive stems, as in (14). Note that despite the translation with 'people' in several of the following, the sentences cannot be interpreted as referring to a particular person or particular people.

(14)  
\textit{e-dó-i.}  
\textit{3-be.red-IMP}  
'People are red./'Being red happens.' (**'He/it is red.' **)They are red.**'

Examples (15–17) may be deponent Middle stems or involve otherwise defective roots (i.e. no synchronic non-Middle counterpart is found or the stem always carries some derivational suffix).

(15)  
\textit{e-gír'á á-l}  
\textit{3-be.quiet-IMP}  
'People are silent.'

(16)  
\textit{e-shíp'á-l}  
\textit{3-be.happy-IMP}  
'People are happy.'

(17)  
\textit{E-ta-wuas-átek-l.}  
\textit{3-PF-pride?-PF.PL.MOD-IMP}  
'People were/became proud.'

---

8. The non-Middle form \textit{a-wúás} means 'to shoot an arrow, fire a gun.' I do not know if this is historically related to the Non-Perfect(ive) Middle \textit{a-wuásá} 'to be proud.'
The Impersonal may occur with active intransitive stems:

(18) \( e\text{-}ba\text{-}翁\text{-}i. \)
    3-reach-VENT-IMP
    'Arriving will happen.'

(19) \( N\text{-}é\text{-}buak\text{-}i \)
    CN-3-shout/bark-IMP
    '(And) people make noise'/'(And) noise is made.'

It may also occur on Antipassive stems, which are usually semantically active:

(20) \( Ïre\ osi\ en\text{-}kátá naáíjó ená \\
     \quad \text{now always FSG\text{-}time.ACC like this.ACC} \\
     ná\text{-}shá n-é\text{-}un\text{-}ishô\text{-}i. \\
     \quad \text{f.rel-rel.3.ACC\text{-}rain CN-3\text{-}plant-APASS-IMP} \\
    'In a season like this of rain, planting is done.' (Pk)

(21) \( e\text{-}bel\text{-}ish\text{'}s\text{-}i \)
    3-break-APASS-IMP
    'The breaking is being done.'

Altogether, -\( i \) is clearly neutral for stative versus active semantics and it may occur with stems of any lexical transitivity value.

Though -\( i \) can occur on (di)transitive or intransitive stems, the Impersonal construction itself is detransitivized. If the stem is (di)transitive, one or more Accusative NPs can occur to express non-Agent(s) involved in the event or situation. Secondly, on the syntactic level, the verb carries a bound pronominal prefix as if it had a third person subject. But no Nominative NP is possible – not even a semantically unspecified NP like itwjanak 'people.' Thus, the "subject" does not display all the properties of a prototypical subject, as cross-referencing of the subject with a full Nominative NP or free pronoun is otherwise completely possible. Thus, the construction does not allow overt expression of the full complement of participants that the stem without -\( i \) would otherwise allow.

On a semantic level, the Impersonal construction displays a reduced transitivity relative to prototypical well-individuated participants (cf. Hopper & Thompson 1980). As a number of the preceding example translations suggest, this construction appears to semantically highlight the occurrence of an action or situation that "people" participate in. However, no particular people, much less a singular individuated entity nor an inanimate agent, can be referenced. This semantic restriction certainly reflects the origins of the -\( i \) morpheme in some type of 3rd person plural morpheme (cf. §5).

We now explore whether an Accusative NP in the Impersonal construction has acquired any subject properties (§3), and then turn to explore the semantic and functional side of the construction in detail (§4).
3. Subject properties and the accusative NP

We saw in §2 that the verb of the Impersonal construction is always marked as if there was a third person subject. In other Maa clausal constructions, a Nominative NP or free pronoun can optionally cross-reference the participant marked as a Subject in the verbal morphology. However, we have seen that no Nominative NP can occur in this construction (examples (2) and (7)). This raises the question of whether an Accusative-marked NP in an otherwise Impersonal clause might have acquired any Subject properties; that is, is there any indication of reanalysis towards a promote-structural in which the O takes on (some) subject properties (i.e. is the O becoming the S of a syntactic detransitive) – which appears to be happening with some Bantu languages in which a passive may involve what was historically 3rd person plural morphology (cf. Kawasha 2007; see also §5 of this paper). To explore this, we will compare properties of non-Impersonal and Impersonal constructions. The general conclusion from all tests is that the Accusative-marked argument has not acquired any subject properties in Maa. Thus, the only possible subject of the construction is indeed an “impersonal” third person. But since the 3rd person pronominal prefix cannot be doubled by an NP, it cannot be said to be like a normal subject.

3.1 Control of bound-pronominal marking on the verb

One property of a subject full NP or free pronoun in Maa is that it correlates with (co-indexes) the bound pronominal prefix on the verb. To illustrate this, I will start with first person singular forms. For first person singular, the bound pronominal subject prefix is á- (see Table 1).

(22) Non-Impersonal; intransitive subject (S) = 1st person singular
    Á-síŋ náŋó.
    1sg-sneeze 1sg.NOM
    ‘I will sneeze.’

(23) Non-Impersonal; transitive subject (A) = 1st person singular
    Á-dúŋ náŋó en-kíné.
    1sg-cut-inst 1sg.NOM FSG.goat.ACC
    ‘I will cut the goat.’

If a first person singular free pronoun is O and the A is a third person or plural, then áa- occurs on the verb.

(24) Non-Impersonal; object (O) = 1st person singular
    Áa-dúŋ nánó 3l-tónáñí.
    3>1sg-cut 1sg.ACC MSG-person.NOM
    ‘The man will cut me.’
Now we look at the Impersonal construction to see whether an Accusative free pronoun correlates with the normal subject bound pronominals. In (25a–b) we get áa- instead of á-. The áa- prefix is not the form that would correlate with an Accusative first person singular being the “subject”, but instead with it being the object.

(25) Impersonal
   a. Áa-duŋ-í nánó
      3>1SG-cut-IMP 1SG.ACC
      'I will be cut.'
   b. *Á-dúŋ-í nánó.
      1SG-cut-IMP 1SG.ACC
      ('I will be cut'; or any other meaning)

Thus, we conclude the construction is non-promotional relative to at least bound pronominal marking.

3.2 Control of relative clause prefix if fronted

Cross-linguistically word order and interclausal syntax are some of the first grammatical mechanisms to show the historical process of non-subjects taking on subject properties. Word order is not so telling for Maa, as VAO, VOA, and VS patterns occur (Payne, Hamaya & Jacobs 1994), and any argument can occur before the verb. Regarding interclausal syntax, however, we may consider relativization (§3.2) and control of number in infinitival clauses (§3.3).

When an S or A is fronted before the verb in the simplest type of non-Impersonal fronting construction, it triggers a relative clause prefix on the verb. Also, the normal bound pronominal prefixes from Table 1 do not occur. Any fronted S or A phrase will occur in the so-called Accusative tone pattern just like a fronted O, rather than in the Nominative tone pattern characteristic of postverbal subjects. That is, there is case-marking neutralization across S, A and O. For clarity, in each example I highlight the fronted NP. Example (26) illustrates a fronted S within a complement clause (bracketed), and (27)–(28) illustrate fronted A’s. All are non-Impersonal constructions.

(26) n-é-yioló-u apá áajo [enk-áí na-ishú]
    CN-3-know-INCHO long.ago that.PL FSG-God.ACC FSG.REL-live
    'and they realized that God lived' (enkai.016)

(27) Nme Múita 5-ta-ŋor-ó ol-púaa.
    NEG Muita MREL-PF-stab-PF MSG-antelope.ACC
    'It was not MUITA who shot the antelope.'

9. The Nominative form of the proper name would have the distinct tone pattern Múita, confirmed by audio recordings and native speaker intuition.
(28) a. Ōre akē peē ɛ-ɪtɛr-ónyɛ ol-tûrrûr
after 3-begin-VENT.MID MSG-group.ACC
l-ɔɔ ŋk-ayiɔɔ,
MPSD-PL.PSR F-boys.ACC
‘As the [new initiation] group of boys begins,

b. nàa il’ tûrrûr tátà tô-rik-û
and/FOC that.M.ACC group.ACC now MSG.REL.NOM-lead-VENT
‘that group is now [the one] who leads’ (bulunoto.031)

When an O or other non-Subject element is fronted, the verb takes its normal bound pronominal prefix and no relative clause prefix occurs on the verb.

(29)  amó mnyё ki-tm-à.
because 3SG.ACC 1PL-pass-PF
‘… because it is the one we have passed.’  (enamuke2.007b)
(lit: ‘… because it we have passed.’)

(30) Àmme al-púua ė-tâ-ŋôr-ô Máità.
NEG MSG-antelope.ACC 3-PF-stab-PF Muita.NOM
‘It was not an ANTELOPE that Muita shot.’  (W)

The following has the predicate complement ‘people who are dreadful’ fronted before the inflected verb ki-râ ‘we are.’ The verb in question does not take a relative clause prefix.

(31)  amó il-tônànà o-itu-rēisho ki-râ
because MPL-people.ACC MSG.REL.ACC-CAUS-afraid-APASS 1PL-be.PL
‘because we are dreadful people’  (Camus4.392)
(lit: ‘because people who are dreadful we are’)

Turning now to the Impersonal construction, it is possible to front what would normally be a post-verbal Accusative NP. The fronting does not trigger a relative clause prefix on the verb. Hence, the fronted Accusative phrase acts like a non-subject, rather than a subject.

(32) Transitive – fronting the only object
nàa āiny'ɛɛ e-yieún-i peē i-lô?
and what 3-want-IMP so.that SUBJN.2-go
‘and what is required (lit: wanted) for you to go?’  (Camus1.009)

(33) Ditransitive – fronting of the Theme object
nàa in-kishû ɛ-ished-i, …
and FPL-cattle.ACC 3-give-IMP
‘and they are given cows, …’  (aisinani.003b)
3.3 Control of Infinitive singular versus plural form

A further inter-clausal syntax property concerns Infinitive forms. A Maa Infinitive is not marked for person, but agrees in number with the subject of some preceding fully-inflected verb. For example, in the non-Impersonal construction in (34), the bound pronominal prefix on the Progressive auxiliary verb *gira* is first person plural. The subsequent verb ‘try’ (historically related to ‘say’) occurs in the plural infinitive form, which begins in *aa*- plus Low tone over the rest of the verb.

\[(34) \text{έκτ-}gir\text{á} \, \text{áa-}j\text{o} \, \text{má-ta-yioló} \]
\[1\text{PL-PROG INF.PL-try SUBJN-SUBJ-know} \]
\[i-rôrei \, l-\text{t} \, l-Máásäi \]
\[\text{PL-words.ACC M.PSD-PL.PSR.ACC M-Maasai.PL.ACC} \]

‘We are trying to learn about Maasai words.’

In contrast, the singular infinitive form has short *a*- plus a LH pattern over the verb:

\[(35) \epsilon-gira \, Kônené \, a-kêd \, ol-döinyó \]
\[3\text{-PROG K INFSG-climb MSG-mountain.ACC} \]

‘Konene is (currently trying) to climb the mountain.’

Now observe (36). Here the Progressive auxiliary has the Impersonal -i suffix. It requires the following Infinitive to be in the Plural form, as in (34) above, and does not allow the singular Infinitive form seen in (35). This is true even though the sole free NP, which is the accusative form *enk-ají* ‘house’, is singular in (36). That is, the singular Accusative phrase *enk-ají* does not control number in the infinitive. Example (37) is similar but with a different Impersonal matrix verb instead of the Progressive auxiliary.

\[(36) \epsilon-girâ-l \, áa-un \, enk-ají. \]
\[3\text{-PROG-IMP INF.PL-erect FSG-house.ACC} \]

‘The house is being erected.’ (Pk)

\[(37) Õre \, sîi \, ntnyé \, nê-m-ê-itôk\text{în-i} \, áa-ship-akin-o. \]
\[\text{DSCN also 3SG.ACC CN-NEG-3-repeat-IMP INF.PL-happy-DAT-MID.NPF} \]

‘And even him, he will not be happily received’ (elengon2.010)

Together these tests show no evidence that the Accusative-marked NP in the Impersonal construction has acquired any grammatical subject properties. The construction appears to be thoroughly non-promotional.

4. Functions of the Maa Impersonal construction

Functionally, the Maa Impersonal construction expresses three kinds of situations: when the communicative intent concerns the event or situation rather than particular
participants in the situation (§4.1); in passive situations such as when a non-Agent is topical or the Agent is unknown (§4.2); and for existentials formed on the root ata 'have' (§4.3).

4.1 Focus on the event or situation

The Impersonal construction is first used when the communicative focus is on the event or situation, as in a number of examples above and the following. As noted earlier, any sense of a possible Agent or (especially for intransitives) of an entity that might exist or participate in the situation can only be of a non-specific and non-referential human “people”; but even a sense of “people” as an Agent is often quite vacuous to non-existent. Nevertheless, to date we have found no evidence of the Impersonal occurring with meteorological verbs.

(38) E-gólˈí
3-be.strong-IMP
'(The act/state of) being strong exists.'

(39) E-bóitáˈí ó l-kísí
3-be.together-IMP assoc M-Kisi
'Being together with the Kisii happens./(People) are together with the Kisii.'

(40) enk-ánty oshí e-rik-íɛ-kí en-kóp.
fs-g-respect.acc always 3-rule-INST-IMP fs-g-land.acc
'By respect the world is ruled./(By respect ruling happens.'

Example (41) is a more extensive excerpt from a conversation which contains several Impersonal constructions. Though (41a) asks about “you”, lines (41b–c) are not really about a referential addressee, but rather about the general circumstances necessary for anyone to go to a particular more advanced training program.

(41) ⇒ a. Speaker 1: Náa áñynˈ35 e-yieún-i peë i-ló?
And what 3-want-IMP so.that subjn.2-go
'And what is required for you to go?' (Camus1.009)

b. Speaker 2: Óre dúóó en-tóki n-a-bá
ds-cn relevant fs-g-thing.acc rel.f-fsg.rel.acc-reach

⇒ ánnaa e-n-a-yieún-i,
like fs-g-rel.f-fsg.rel.acc-want-IMP
'The thing that is required,

⇒ c. k-é-yieún-i nabó tenáá i-ishˈu-á Bible School
ds-cn-3-want-IMP one.fem.acc if 2-finish-PF
'it is required one that you have completed Bible School’ (Camus1.010)

The following somewhat complex example is the first line of a text that describes the general way in which Maasai historically trained a subsequent warrior age-set, and how these activities have been curtailed over time by the national government. No
particular referential entities are referred to here – neither specific community elders that would do the supervising, nor any particular boys who have just been initiated. Rather, the communicative focus is on describing a general series of events that happens from time to time.\textsuperscript{10}

\begin{equation}
(42) \quad N\text{-}é\text{-álkó} \quad těnáa \quad n\text{-}á\text{-}a\text{-}j\text{-}í \quad k\text{-}é\text{-itúrrúrr\text{-}í}
\end{equation}

\text{CN-3-become} \quad \text{if} \quad \text{CN-FPL.REL.NOM-say-IMP} \quad \text{DSCN-3-gather-IMP}

\begin{equation}
n\text{-}á\text{-}a\text{-}j\text{-}í \quad \text{mk\text{-}ayôk} \quad áå \quad \text{tl\text{-}álbártak, ...}
\end{equation}

\text{CN-FPL.REL.NOM-say-IMP} \quad \text{FPL-boys.ACC} \quad \text{be} \quad \text{MPL-initiates.ACC}

‘Now if boys who are new initiates could be put together,...’ (aibartisho 001)

4.2 Functional passive situations

In (42) just above, one might argue that the functional situation is not much different from that of a passive if the text is primarily about young men rather than the work of elders, even though the young men are non-referential. Indeed, the Impersonal construction is used in a range of passive situations. A functionally passive situation is typically defined as one in which (a) a non-Agent is more individuated or topical in the discourse than a (known) Agent; or (b) the Agent is unknown or somehow not well individuated; or (c) for some social or communicative reason the speaker does not want to mention a (known) Agent (see Abraham 2006 for a recent summary of claims about passives). Typically more than one of these features is present in any particular instance. However, I will attempt to discuss them separately. In the process of discussing the passive function(s) of the Maa Impersonal construction, I will try to correct the record on one claim occasionally made in the literature about the Maasai “passive”, but which appears to be wrong (§4.2.4).

4.2.1 Topical non-Agent

The following simple elicited example is reflective of a situation where a non-Agent would be highly topical, and the Agent might be known but is comparatively irrelevant. This happens to be an (Impersonal) External Possession construction, in which the first person Possessor is marked as the grammatical object on the verb, reflecting its high topicality or high perceived “locus of affect” (Payne 1997). The ‘goat’ is overtly expressed in an Accusative NP, but it is not treated as the object insofar as the bound pronominal is concerned. (Likely the real-world Agent of such an event would be the speaker or speaker’s designee, but this cannot be expressed in the construction.)

\textsuperscript{10} The relativized ‘say’ verbs in this sentence give modal force; in isolation they could be translated roughly as ‘possible’.
The following extended authentic excerpt comes from a text about how a (generic) poor person who has died is handled. Translations of the two clauses immediately preceding (44a) are given in italics. The poor person taken outside the kraal in clauses (e) and (f) is the topic of the entire text, and these clauses are Impersonal constructions.11 Though technically all characters in the text are non-referential, we might say that participants can come to take on a certain degree of pragmatic topicality (if not referentiality) in the context, to the point that the dead person is referred to with demonstrative forms like ‘that previously-mentioned’ one.

They hire two [generic unspecified non-referential] people. Those two people pick him up and are given cows.

a. \( n\-\mathring{e}\-\mathring{y}a \quad \text{\'ala}\-\mathring{d}\-\mathring{u}\-\mathring{o} \quad \text{tonj\-\'ani}, \)
\( \text{CN}-3\text{-take} \quad \text{MSG}-\text{previously.mentioned} \quad \text{person.acc} \)
‘and they take that previously-mentioned [poor] person,

b. \( \text{\'aa}\-\text{ipan-\'ie} \quad \text{na\-\'a}\-\mathring{d}\-\mathring{u} \quad \text{ak\-\'e}\-\text{n\-\'inc}, \)
\( \text{INF.PL}-\text{exit-cause} \quad \text{FOCUS} \quad \text{just they.nom} \)
‘to remove him (outside),

c. \( \text{\'aa}\-\text{ipan-\'ie} \quad \text{te}\-\text{\'an}. \)
\( \text{INF.PL}-\text{exit-cause} \quad \text{obl} \quad \text{home.nom} \)
‘to remove him from the kraal.

d. \( \text{am\-\'o} \quad \text{\'al-\text{tonj\-\'ani}} \quad \text{naa} \quad \text{\'il\-\'o} \quad \text{\'le-\text{m-\'e-\'ata}} \)
\( \text{because} \quad \text{MSG-person.acc} \quad \text{FOCUS} \quad \text{that.m.nom} \quad \text{REL.M-NEG-3-have} \)
\( \text{\'il-\text{tonj\-\'anak}} \quad \text{l\-\'eny\-\'enak}. \)
\( \text{MPL-people.acc} \quad \text{MPSD-3PL.poss.acc} \)
‘because that is a man who doesn’t have his own people.’

e. \( \text{\'N\-\text{\'e-\text{ipan-\'e\-\'ek-\'i}}} \quad \text{\'ala}\-\mathring{d}\-\mathring{u}\-\mathring{o} \quad \text{tonj\-\'ani}, \)
\( \text{CN}-3\text{-exit-cause-imp} \quad \text{MSG}-\text{previously.mentioned} \quad \text{person.acc} \)
‘That previously-mentioned [poor] person is removed’

f. \( \text{\'n\-\text{\'e-\text{y\-\'a-\'i}}} \quad \text{na\-\'a} \quad \text{\'nj\-\'ata} \)
\( \text{CN}-3\text{-take-imp} \quad \text{FOCUS} \quad \text{outside.acc} \)
‘and he is taken outside’ (aisinani 003–005)

11. A truly poor person is one who doesn’t have any “people” who take care of him or her. Generosity is a traditional cultural value, and normally a person does not lack a personal network.
The understood agent in clauses in (e) and (f) is non-referential “people”, as this text describes the general way a poor dead person is handled. Though non-referential, the “people” perhaps come to take on a certain degree of discourse topicality due to the clauses immediately preceding (44a). Note that the Impersonal construction is not used here, even though “the poor person” is the primary topic of the discourse. This suggests the significance of particular focus on the non-Agent or on the event/situation for choosing the Impersonal construction, and thus it is not appropriate in (44a) due to the local topicality of the “two people”.

4.2.2 Unknown or non-existent Agent
The Maa Impersonal construction can also be used when any conceivable Agent is unknown, or even does not exist. There is no known or existing Agent in (45b), in the sense of a volitional or instigating entity. Rather, the use of the Impersonal ‘say’ verb is to indicate that the content of (45c) is culturally typical. (Also, the ‘you’ in the English translation of (45c) is based on the plural imperative form, but it is not a referential mention of an addressee in the context of the conversation.)

(45) a. Speaker 1:  
   K-átnyó'ó dói 
   dscn-what indeed 
   il-toñánák il-tóya? 
   mpl-peoples.nom rel.m-mpl.rels.nom-take 
   ‘What kind of people will take him?’

⇒ b. Speaker 2:  
   n-é-j-i, 
   cn-3-say-imp-pl.pf 
   ‘It is said,’

c.  
   Én-yíañ-ó il-tóñánák. 
   pl.subj-buy.from-ventive mpl-peoples.acc 
   ‘You should buy (hire) people’ [who are poor and have no social network] (aisalani 001)

The following example, also involving an Impersonal form of jo ‘say’, focuses on a particular term or word used for an academic subject; (47) is similar. In neither is there a known Agent.

(46) amò e-tíi subject nábo n-á-j-i anthropology... 
   because 3-exist subject one.fem.nom cn-fsg.rel.allsay-imp 
   ‘because there is one subject called anthropology...’ (Camus1.021)

(47) n-é-ar il5 sháni 
   cn-3-make.dysfunctional that.m.nom shrub.nom 
⇒ ó-j-i ol-túléléi 
   msg.rel.allsay-imp msg-sodom.apple.acc 
   ‘that shrub that is called oltulelei invades’ (Camus4.312b)
4.2.3 Diffuse, non-specific human Agent

As we have already seen in (41), (42), etc., the Maa Impersonal construction normally implies a sense of general “people” with unknown or non-specific reference. The Impersonal construction does not normally evoke the concept of an inanimate agent, even though inanimate agents are completely possible subjects in other Maa constructions.

Example (41) above is a rather prototypical example, where the “Agent” must be the rather diffuse referent “all Maasai people across a large area”, who would be involved in seeing that the next initiation group is properly established. The following additional example comes from a story about what a generic child experiences in growing up in Maasai culture. Clearly it is “initiated Maasai boys in general” (if not all Maasai people) who would “do” the various things.

(48) ‘So now that is where they [the newly-initiated boys] will be told different things’

⇒ a. fpl-rel.f-3-do-vent-imp FSG-life.acc
    ‘how life is done’;

⇒ b. rel.f-3-do-vent-imp FSG-respect.acc
    ‘how respect is done’;

⇒ c. rel.f-3fpl.rel.acc-be.fitting CN-3-do-imp
    ‘what is proper to be done’;

(49) ‘It is said let us hold (their) initial ceremony’

The root jo ‘say’ is used to denote the action of ‘naming’. Here again it is most common to find the Impersonal suffix, precisely because the focus is on the action of naming rather than who might have done it, even though it always must be “people” who do the naming.

(50) ‘When he was named Konyek, …’

As in (46–47) above, the Impersonal is also highly common in expressions of the closely-associated result ‘to be called something’.

In summary, it is important to underscore that the various dimensions I have tried to tease apart in §4.2.1 through 4.2.3 correlate quite highly with each other.
Perhaps most saliently, it is nearly impossible to find an example in elicitation or texts where common-place understanding would not agree that “people” must be the probable effector, Agent, or Actor in nearly all situations referenced by the Impersonal construction.

4.2.4 Specific Agents and the Impersonal

In the corpus data used for this study, there are no instances where the Impersonal is clearly chosen in order to avoid mentioning a known agent for social reasons. However, this construction may -though with very low frequency- be used where a fairly specific discourse-referential agent is known from context. The following excerpt is one of these rare exceptions, taken from a “David-and-Goliath” kind of story in which boastful Olarinkoi is defeated by a young neophyte warrior. The understood Agent of hitting in the Impersonal in line (g) must be either the young warrior (or possibly ‘the club’ as an extension of the young warrior), which is reasonably recoverable from the context.

(51) a. E-tu-búl-ú-á taá apá
   3-PF-grow-VENT-PF be.PF long.ago
   ‘He grew up

b. é-táá nól-móráí,
   SUBJN.3-be.SUBJN MSG-warrior.ACC
   to be a warrior,

c. n-é-y’úk ola-duóó rinká l-5 létókát,
   CN-3-swing MSG-previous club.ACC MPSD-MSG.PSR Oletukat.ACC
   ‘he swang the club from Oletukat,’

d. n-é-y’úk aké,
   CN-3-swing again
   ‘he swang,’

e. n-é-y’úk aké,
   CN-3-swing again
   ‘he swang,’

f. páá k-é-íte-jo ɔl-árínkói
   so.then DSCN-3-CAUS-say Olarinkoi.ACC
   ti-áborí enk-alóéna, pau13
   OBL-below FSG-ear.NOM whack!
   ‘so he did to Olarinkoi below the ear, pau!’

12. This does not necessarily mean that such could not occur – I just don’t know.
13. This clause contains an ideophone construction that requires the verb jo ‘say’. What is presumably “said” is the ideophone. I give a fairly free translation.
Though rare, use of the Impersonal construction in contexts like (51g) raises the question of whether specific Agents could not actually be expressed in this construction – i.e. is there the possibility of an agentive passive? Indeed, the suggestion that an overt Agent can be included in the Maa passive has been occasionally asserted in the literature. This idea apparently originated with Tucker and Mpaayei (1955: 79, also pp. 175–176), who stated with regard to the “passive” that “The noun agent, when used, has the tone pattern of the subject.” The claim is quoted in Greenberg (1959), but with no new supporting data. Because of the general quality of Tucker and Mpaayei’s work, and the stature of Greenberg as a linguist, the idea that the Maa passive can express an agent is wont to be taken as authoritative. However, all evidence I have gathered across more than a decade of work indicates this is not correct. Here I will elaborate on the data presented by Tucker & Mpaayei to explore how questionable their claim is, and also to help evaluate to what extent there may be development of an agentive passive construction out of the Impersonal construction.

First, certain (sparse) examples in Tucker and Mpaayei (TM) might be – wrongly – interpreted as evidence for an agentive passive. Unfortunately, TM mark only some of their examples for tone, though tone is crucial for determining Nominative versus Accusative syntactic status. The first offending example is E-rik-i nkishu aainei lm~rran, translated as ‘My cows will be led by the young men.’ (TM, p. 81).\(^{14}\) In all the data I have elicited, such a sentence with Nominative tone pattern (i.e. TM’s “the tone pattern of the subject”) on either the noun nkishu ‘cows’ or lm~rran ‘warriors’ is rejected as ungrammatical. Also, no comparable sentence has surfaced in over 10,000 lines of text.

\(^{14}\) Tucker and Mpaayei use bold for writing -ATR vowels. I have changed the vowel transcription to match the system used in the rest of this paper but have not supplied tone to this example. The morpheme divisions here are from Tucker and Mpaayei (1955: 81). I would gloss their non-tone marked example as e-rik-i (3-lead-IMP) nkishu (cattle) aainei (mine.pl.psd) lm~rran (warriors).
A second example on which TM do write tone is:

(52)  
| é-ísis-f | Sironka |
| 3-praise-IMP | Sironka |

‘He is praised by Sironka.’  
(TM p. 176; parsing and glossing mine)

In my data I have *Sirónka* as the Nominative form of this proper name, which matches what Tucker and Mpaayei present as the Accusative form (cf. TM p. 175). The Nominative and Accusative forms for certain nouns are homophonous (though case homophony is really quite rare), and tonal case patterns can vary from region to region; so it is difficult to know what to make of their example (52).

The most compelling examples in TM that might be taken as support for the claim are the following (reflecting the effects of cross-word tone sandhi on the gender prefix; parsing, morpheme glosses, and asterisks are mine). Though TM present these examples, they have been rejected by speakers I have worked with (the ungrammaticality in my data is indicated by “*DP”).

(53)  
| a. (*DP) é-ipót-i | én-kerái |
| 3-call-IMP | FSG-child.NOM |

‘He is called by the child.’  
(TM p. 176)

| b. (*DP) é-ísis-f | l-tuynáná |
| 3-praise-IMP | M-people.NOM |

‘He is praised by the people.’  
(TM p. 176)

A further non-tone-marked example from TM, which might possibly be claimed to be an Agentive passive, is line (b) of the following:

(54)  
| a. Astnycc n-a-t-ar-a? |
| what | REL.F–F.SG.REL-PF–strike-PF |

‘What (was it) that killed him?’

| dscn-3-PF–strike-PF-IMP | OBL spear |

‘He was struck by a spear.’  
(TM p. 82, parsing and glossing mine)

Here we find an Oblique phrase *te remet*. Interestingly, the general Oblique preposition *te* always requires that its governed NP occur in what TM called the “Nominative” tone pattern (Tucker & Mpaayei 1955:175). Since it is cross-linguistically common to express passive agents in oblique phrases, one might suspect that (54b) is the agentive passive construction we are looking for in Maa. However, there is no reason to assume that the Oblique *te remet* phrase in (54b) is an (inanimate) Agent phrase, even though the so-called Nominative tone pattern would indeed occur here on *remet*. As just noted, the “Nominative” is the case form always required after the preposition *te*; this is true regardless of the semantic role conveyed by any particular
oblique *te* phrase. A non-Passive/non-Impersonal counterpart to (54b) would be as follows; in (55), ‘spear’ is an Oblique Instrument of a canonical transitive construction, and is not the Agent.

(55) \[ K-\varepsilon-\ellá-\á-r-\á \; \varepsilon l-morrání \; \textit{te \; rémet}. \]
3-PF-kill-PF MSG-warrior.NOM OBL spear.NOM

‘The warrior struck him with a spear.’

The non-promotional Impersonal in (54b) simply changes the structure in (55) by suppressing the personal Agent ‘warrior’, but otherwise leaves everything intact including the Oblique phrase. Altogether, there is no reason to assume that in Tucker and Mpaayei’s example the *te remet* phrase is anything other than an Oblique – and the sentence could just as easily have been translated ‘Some unspecified person(s) struck him with a spear/He was struck with a spear.’

4.3 Existential

We now come to the third major function of our target structure. What looks like the same Impersonal construction is used with the stative possessive verb *ata* ‘have’ to render an existential function, as in ‘There is X’ or ‘X exists’. For example:

(56) \[ N-\varepsilon-atá-i \; \textit{aké \; ninyé \; o\-otunó}. \]
CN-3-have-IMP still emphatic MSG-leader.of.ageset.ACC

‘There is still the leader of the age-set.’

[lit: ‘There has still the-one-who-planted’]. (aibartisho.004)

Text distribution of this existential construction suggests that it is not a canonical presentational construction, as it may be used for unimportant participants; nor is it necessarily used for arrival of a participant onto a scene or with a change of scene. It is used both to express negative and positive existence. In the majority of cases, it is used for talking about the “existence” of a non-specific person or persons, as in descriptive texts that talk about customs and practices. This is precisely the concept involved in (56): no particular referential leader is being talked about, but rather the institution that involves such an office-holder.

The following example shows use of *ata* (allomorph *eta*) ‘have’ both in its possessive sense (lines c, e, and f), and in its existential sense when combined with the
Impersonal suffix -i (lines b and d). The mentions in the existential clauses are of non-referential “people” (in this particular example, non-referentiality in (57b) and (57d) is also ensured by the scope of negation).

(57)  
a. Òre apá te nk-árakí ená lenón,  
   DISCN formerly OBL FSG-reason.NOM this.FEM.ACC generosity.ACC  
   ‘Now because of this generosity,
   \[⇒\]  b. né-m-é-át'á-l apá il-tónjáni  
   CN-NEG-3-have-IMP formerly MPL-people.ACC  
   ò-yá il-ámeyu  
   MPL.REL.ACC-take MSG-famine.NOM  
   ‘there were no people that drought took (killed)’
   c. é-eta il-kólñáí en-dáa.  
   TEMP.3-have MPL-others.NOM.M FSG-food.ACC  
   ‘when others had food:’
   \[⇒\]  d. Né-m-é-át'á-l il-tónjáni ó-lò  
   CN-NEG-3-have-IMP MSG-person.ACC MSG.REL.NOM-go  
   ‘There wasn’t a person who went,’
   e. m-é-át'á en-klù  
   NEG-3-have FSG-cloth.ACC  
   ‘[who] he didn’t have clothes’
   f. é-ata il-kólñáí in-kláníja.  
   TEMP.3-have MPL-others.NOM.M FPL-cloths.ACC  
   ‘when others had clothes.’  

Mentions of entities or concepts via the ata-i existential construction need not be negated. Clauses (58a–b) are contiguous positive existentials from a descriptive text. What are posited as existing in these examples are “archetypes”, rather than particular referential entities.

(58)  
\[⇒\]  a. é-át'á-l i-múain aré te n-keráí, aráki áå  
   3-have-IMP FPL-types.ACC two.F.ACC OBL FSG-child.NOM or be  
   to l-maréti l-5 l-payán.  
   OBL MSG-family.NOM MPSD-MSG.PSR MSG-elder.ACC  
   ‘There are two types in a child, or in a family of a man.
\[⇒\]  b. é-át'á-l en-kéráí né-m-é-bárñ-í  
   3-have-IMP FSG-child.ACC CN-NEG-3-shave-IMP  
   ‘There is a child that cannot be shaved …. ’  

In addition to human animates, (58a) and (59) demonstrate that this existential construction can be used with inanimate non-referential concepts.
The Maa (Eastern Nilotic) Impersonal construction

(59) \( k-\text{ê-ata}\-t \) \( e-j\text{êún-otó} \)
DSCN-3-have-IMP FSG-save-NMLZ.ACC

‘there is salvation’  \( (\text{ilomon.0169}) \)

It may also be used for non-referential but still fairly discourse-individuated inanimate referents. In (60b) what is posited to (not) exist is ‘that problem’, marked with a demonstrative.

(60) a. \( n-\text{ê-akó} \) \( mmè \) \( en-\text{tòkì} \) \( ínà \)
CN-3-become NEG FSG-thing.ACC that.F.ACC

\( n-\text{à-tó-nîy-ò} \) \( \text{in-kòlòñì} \) \( kùmòk \) \( dùóò \)
REL.F-1SG-PF-hear-PF FPL-days.ACC many.ACC indeed

‘So that is not something I have heard (in) many days

\( \Rightarrow \) b. \( e-\text{áta-}t \) \( înà \) \( nyamâlì \) \( n-\text{ê-}j-\text{i} \) \( e-\text{ta-jèwu-}o \)
3-have-IMP that.F.ACC problem.ACC CN-3-say-IMP 3-PF-save-PF

\( \text{en-}títò \) \( n-\text{ê-ar-akin-}t \) \( \text{nà-r-mòruo} \).
FSG-girl.NOM CN-3-beat-DAT-IMP MSG-old.man.ACC

‘(that) there is said to be that problem of a girl being saved
(lit: has saved) and is forced to marry a man’  \( (\text{ilomon 0181}) \)

Finally, though it appears that this existential construction is most frequently used for talking about the “existence” of non-referential and non-specific concepts, some text examples show that it can be used for specific referential mentions too, as in (61)–(62) below, certainly including humans (62).\(^{16}\)

(61) \( \Rightarrow \) \( pèé \) \( e-\text{átá} \-t \) \( e-n-kýp \)
so.that SUBJN.3-have-IMP FSG-land.NOM

\( n-a-j-\text{i} \) \( \text{til-âkípiak} \)
REL.F-FSG.REL.ACC-say-IMP MPL- Laikipia.ACC

‘that is why there is a place called Laikipiak’  \( (\text{emutata 053c}) \)

(62) \( N-\text{ê-atá-t} \) \( ënà \) \( tít'ó \ ...
CN-3-have-IMP this.F.ACC girl.ACC

‘There was this girl…’  \( (\text{enkukuu 003}) \)

In sum, the Impersonal existential appears to depart semantically from the cases discussed in §4.2 in that in many examples there is no sense whatsoever that a human “people” Agent or Actor could be conceptualized as plausibly involved. What is posited to exist, is expressed in an Accusative NP; and this entity may be referential-specific

\(^{16}\) An alternative existential construction, using the root \( tìi \) ‘be at’, is more frequent for specific referential mentions (Payne 2009).
or not, and human or not. These differences suggest there is a distinct Impersonal Existential construction, separate from but clearly related to the simple Impersonal.

5. Origins of the Impersonal suffix and construction

The fact that any sense of a possible Agent in the Impersonal construction is limited to animate “people” is most certainly related to the fact that the suffix -i derives historically from an old plural subject affix (Greenberg 1959), though it is debatable whether this was a specifically pronominal element or not. The plural subject status of the affix is shown by at least three facts.

Greenberg presents a first argument based on infinitive behavior, as we saw in §3.3: only plural infinitive forms can follow the Impersonal construction, even when the sense may be singular or is unspecified for number. This demonstrates that something about the Impersonal construction must have a formal plural feature attached to it. Greenberg attributes it to the -i suffix, as nothing else in the clause need be specifically plural.

Second, Greenberg (1959) observes that certain verbs have contrasting singular versus plural stem forms. The verb ‘go’ is a case in point: lo(t) ‘go singular’ versus puo(n) ‘go plural’. The root ‘go’ can be used as a periphrastic future auxiliary, followed by a semantically main verb in the infinitive form. Consider first the following singular non-Impersonal instance:

(63) Á-lót-ú  a-nyá  ená  dáa  te  n-á-shuk-ún-ye.

1sg-go-vent  inf.sg-eat  this.acc  food.acc  obl  cn-1sg-return-vent-mid

‘I will eat this food when I return.’ (W)

Now observe that in a similar Impersonal construction, the plural root form puo(n) occurs, even if the only specifically-named entity is singular. In the following, whoever does the betrothing would be general “people”.

(64) K-é-púúó-i  taá  apá  áa-sai.

dscn-3-go.pl-imp  be.subjn  distal  inf.pl-betroth

‘She [a generic baby girl] will be betrothed.’ (enkiama 006)

17. Greenberg (1959:172) states that ‘die’, ‘come’, and ‘sit’ also have contrasting singular and plural stem forms. He does not suggest what the contrasting forms might be for ‘sit’ and in our data we have an invariant root form tən. ‘Come’ is simply the root for ‘go’ plus the Ventive. The form he gives for plural ‘be dead’ includes the Plural Perfect(ive) suffix. Once this is removed, there is no difference between plural and singular stem forms.
Third, and perhaps most significantly, there is a homophonous plural suffix -i, as seen in (65)–(66). Neither (65) nor (66) is impersonal in sense. In (65), kí- is the 1st person plural – not Inverse – bound pronominal prefix; and (66) has a specifically plural Nominative NP (the singular counterpart would be bolúñotò and would require the singular demonstrative form enà). Importantly, note that as a synchronic plural suffix, -i cross-references an explicitly Nominative NP; the Accusative counterpart in (66) would have been bolunót. In sum, unlike the data in preceding sections, the uses of -i in (65)–(66) are not Impersonal.

(65) Kí-tm-i.  
1pl-sit-pl  
'We sit/are sitting.' (W)

(66) K-é-gól-i òná bolunot olény.  
dscn-3-be.hard-pl these.nom chapters.nom very  
'These chapters are very hard.' (W)

We may compare (65)–(66) with the following (also non-Impersonal) singular situations, which lack Plural -i.

(67) é-tön nínye ená kíne ti ány.  
3-sit 3.nom this.f.nom goat.nom obl home.nom  
'This goat will stay (lit: sit) at home.' (Pk)

(68) E-tó-dú-á Kimeli ajó k-é-gól en-tim-atá.  
3-pf-see-pf Kimeli comp dscn-3-be.hard fsg-test-nmlz.nom  
'Kimeli thought that the test would be hard.' (W)

As a plural suffix in non-Impersonal constructions, -i is primarily documented on intransitive stems (either basic or derived). However, (69b) suggests that the Plural -i is also acceptable on transitive stems, as it occurs with tum ‘meet’ in the discourse context of a clear referential plural Agent i-léwá ‘men.nom’. There is no passive or impersonal sense here.

(69) a. n-é-puo i-léwá, n-é-puo, n-é-puo,  
CN-3-go.pl mpl-men.nom CN-3-go.pl CN-3-go.pl  
‘and the men went, they went, they went'  
⇒ b. n-é-túm-i il-ánjéni  
CN-3-meet-pl mpl-wisemen.acc  
and met wisemen,  
⇒ c. n-é-p’úó-i áa-rev in-kíshú  
CN-3-go.pl-pl inf.pl-drive fpl-cattle.acc  
‘And they brought the cattle’ (arinkoi 017–018)
The following is excerpted from a text about a warrior whose favorite bull was taken by a British district commissioner; as a result, the warrior killed the British official, he was then arrested and executed, and his family was made to pay a large fine. Clearly, the warrior is the discourse topic. Thus, the following could probably equally-well be rendered in English as ‘And the police grabbed him and he was taken and tied.’ Here the Agent of ‘taking’ is completely clear (‘the police’). The example nicely shows a situation where topicality and plural agency do not devolve on the same participant. Exactly such a situation is what could give rise to historical reinterpretation of the Plural -i as a passive device.

Interestingly, Greenberg (1959: 174) seems to assume that -i does not occur as a pluralizer synchronically in Maasai, but rather he describes an etymologically-related third-person pluralizer -i in other Eastern Nilotic languages such as Lotuho (one of the most closely related languages to Maa). He argues that it must be a third person pluralizer, but we have seen that in at least some dialects of modern Maasai the -i Plural is not restricted just to third persons (cf. (65) above).

In the literature it is often noted that impersonal constructions arise via reanalysis of a pronominal element which can optionally have an impersonal referent (cf. many uses of English they); often these are first or third person plurals. If such a subject pronominal is reanalyzed as a specifically impersonal element, then one may get a non-promotional Impersonal construction, as the now-impersonal pronoun retains at least some of its formal “subject” prerogatives.

This raises the question of whether there is evidence that the Maasai -i was once pronominal (beyond just being a plural morpheme). We may infer that Greenberg’s comparative answer might have been “no”, as he discusses only internal and comparative evidence for it being a “pluralizer” for disambiguating number, given that the bound pronominal prefix ε- simply indicates third person (cf. Table 1). However, if it were just a pluralizer and not pronominal, we have no particular way of explaining why the (now, Impersonal) construction as a whole dis-allows a co-occurring Nominative NP. In contrast, if -i had been pronominal, the -i itself would have satisfied the formal Subject requirement in any clause, and hence a second “Subject” could not have been added.

However, we must note that synchronically -i does clearly co-occur with bound pronominal prefixes like ε- ‘3’; hence synchronically the -i cannot be said to be (fully) satisfying any formal subject requirement itself. Also, -i does not denote person (which most pronominals do) in its synchronic plural function, as it occurs at least with both 1st and 3rd persons. At the present, we must leave the historical conundrum of a possible pronominal status for -i for further research.
Finally, I will just end on one more teasing note related to the historical development of new constructions. There is another existential construction which is closely-related to that discussed in §4.3. This existential also contains the etymological root ata ‘have’ and allows ONLY an Accusative NP. Notably, there is no -i.

(71) amò m-ê-átâ ñl-toîjâni ó-îìîéû
because NEG-3-have MSG-person.ACC MSG.REL.ACC-dare 'because there is no one who can face him,' (arinkoi.009)

(72) amò k-ê-átà tåtå nêk-atînî apâ najé
because DSCN-3-have NOW FSG-story.ACC formerly certain.ACC ‘because there was a particular story’ (Camus1.028)

I do not know whether this existential construction arose by dropping an -i from the construction discussed in Section 4.3 above, or whether it arose directly just via an impersonal interpretation of the third person bound pronominal prefix ε-. The latter is a widely-attested route for developing impersonal constructions, as we have alluded to above regarding English they.

Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACC</td>
<td>accusative</td>
</tr>
<tr>
<td>APASS</td>
<td>antipassive</td>
</tr>
<tr>
<td>ASSOC</td>
<td>associative</td>
</tr>
<tr>
<td>CAUSE</td>
<td>causative</td>
</tr>
<tr>
<td>CN</td>
<td>continuous discourse connective</td>
</tr>
<tr>
<td>COND</td>
<td>conditional</td>
</tr>
<tr>
<td>DAT</td>
<td>dative</td>
</tr>
<tr>
<td>DSCN</td>
<td>discontinuous discourse connective</td>
</tr>
<tr>
<td>EMPH</td>
<td>emphatic</td>
</tr>
<tr>
<td>EP</td>
<td>epenthetic</td>
</tr>
<tr>
<td>F</td>
<td>feminine</td>
</tr>
<tr>
<td>IMP</td>
<td>impersonal</td>
</tr>
<tr>
<td>INCHO</td>
<td>inchoative</td>
</tr>
<tr>
<td>INF</td>
<td>infinitive</td>
</tr>
<tr>
<td>INST</td>
<td>instrument</td>
</tr>
<tr>
<td>LOC</td>
<td>locative</td>
</tr>
<tr>
<td>M</td>
<td>masculine</td>
</tr>
<tr>
<td>MID</td>
<td>middle</td>
</tr>
<tr>
<td>NEG</td>
<td>negative</td>
</tr>
<tr>
<td>NMLZ</td>
<td>nominalization</td>
</tr>
<tr>
<td>NOM</td>
<td>nominative</td>
</tr>
<tr>
<td>NPF</td>
<td>non-perfect(ive)</td>
</tr>
<tr>
<td>OBL</td>
<td>oblique</td>
</tr>
<tr>
<td>PF</td>
<td>perfect(ive)</td>
</tr>
<tr>
<td>PL</td>
<td>plural</td>
</tr>
<tr>
<td>PROG</td>
<td>progressive</td>
</tr>
<tr>
<td>PSD</td>
<td>possessed</td>
</tr>
<tr>
<td>PSR</td>
<td>possessor</td>
</tr>
<tr>
<td>REL</td>
<td>relative clause</td>
</tr>
<tr>
<td>SG</td>
<td>singular</td>
</tr>
<tr>
<td>SUBJN</td>
<td>subjunctive</td>
</tr>
<tr>
<td>TEMP</td>
<td>temporal mood</td>
</tr>
<tr>
<td>VENT</td>
<td>ventive</td>
</tr>
</tbody>
</table>

High tone is indicated by an acute accent, and Low tone by no diacritic. The raised exclamation mark indicates a within-word Downstep High tone. +ATR vowels are
written as i, e, o, u; and –ATR vowels are written as i, ε, ɔ, ə. Codes after examples refer to particular texts, or to geographical “section” names if data is from elicitation: unmarked elicited examples are from il-Keekonyokie, Pk = Purko (of Kajiado district), and W = il-Wuasinkishu. Names of Maa grammatical forms are capitalized (e.g. Accusative, Progressive).

References


Impersonal constructions in Jóola-Banjal

Alain-Christian Bassène & Denis Creissels
Université Cheikh Anta Diop de Dakar/Université Lyon 2

Jóola-Banjal (an Atlantic language spoken in Senegal) has several constructions, not mentioned in previous works on Jóola languages, which include no canonical subject. This paper puts forward a formal classification of these constructions and a description of their functions, emphasizing the contribution of this West African language to the general typology of impersonality. Most of the types of impersonal constructions found in Jóola-Banjal have already been identified in other languages, but there also several points on which Jóola-Banjal shows interesting particularities.

Keywords: impersonal constructions; Atlantic languages; Jóola

1. Introduction

This paper describes the impersonal constructions of Jóola-Banjal, an Atlantic language of Casamance (Senegal).

Impersonal constructions constitute a major topic in traditional descriptions of European languages, and some types of impersonal constructions have been widely discussed by generativists. Similar constructions have been described in languages spoken in various areas of the world and belonging to different language families, and in several recent typologically oriented publications, discussions of phenomena which typically fall under impersonal constructions are extended to languages outside Europe – see in particular Aikhenvald et al. (eds.) (2001), Bhaskararao & Subbarao (eds.) (2004); Creissels (2007); Donohue & Wichmann (eds.) (2008).

Discussions about Bantu presentational focus constructions constitute so far the main contribution of African languages to the theoretical study and the typology of impersonal constructions. Apart from arbitrary readings of personal pronouns, grammars of West African languages rarely recognize the existence of impersonal constructions, even in languages in which they do exist. For example, no explicit mention

1. On Bantu presentational focus constructions, see among others Bresnan & Kanerva (1989); Marten (2006); van der Wal (2008); Creissels (2009).
of the existence of impersonal constructions can be found in Hopkins (1995), which constitutes the most detailed syntactic description of another Jóola language published so far, although some of the examples used by Hopkins to illustrate other phenomena include constructions similar to those analyzed in this paper. With this paper, we would like to convince our readers that some West African languages at least have a variety of constructions the analysis of which brings an interesting contribution to the typology of impersonality.

The paper is organized as follows. §2 provides some general information on Jóola-Banjal. In §3, we propose a definition concretizing the general notion of impersonal constructions within the frame of Jóola-Banjal morphosyntax. In §4, we examine a construction of intransitive verbs devoid of any manifestation of the S argument. In §5, we examine constructions in which the S argument of intransitive verbs does not appear in the canonical position for subjects, and no subject marker is prefixed to the verb. In §6, we describe the impersonal use of the transitive possession verb ebaj 'have' in a construction expressing existence. §7 presents an impersonal construction involving the negative copula. In §8, we analyze a construction in which a pronoun in subject position resumes a topicalized clause, but no subject marker is prefixed to the verb. §9 is devoted to a construction involving a frozen expletive subject marker. §10 presents arbitrary subject constructions in which 2nd person singular and 3rd person plural do not have their usual deictic or anaphoric interpretation. In §11, we briefly mention the existence of verb forms devoid of a subject marker that are not involved in impersonal constructions, and are better explained as the result of isolated lexicalization processes. §12 summarizes the conclusions.

2. General information on Jóola-Banjal

Jóola-Banjal belongs to the Jóola subgroup of the Bak language group included in the Northern branch of the Atlantic language family, one of the language families that constitute the Niger-Congo phylum. Jóola-Banjal is spoken by approximately 7,000 speakers in the villages of Badiatte, Bandial, Batignère, Batignère Essil, Elubalir, Enampore, Essil, Ettama, Kameubeul, and Séléki, which constitute the territory traditionally called Mof Avvi ('the land of the king'). The area is a peninsula bounded by the Casamance River on the north, the Kameubeul Bolong on the west, and the Ziguinchor-Oussouye road on the south and on the east.

Bassène 2007 provides an overall presentation of the phonology, morphology and syntax of Jóola-Banjal. The transcription of Jóola-Banjal follows the conventions already used in practical orthographies of other Jóola languages, which depart from other West African orthographies in the notation of vowels: the vowel system of Jóola-Banjal is characterized by ATR harmony at word level, and the acute accent
Signals + ATR words (which cannot lead to any confusion, since Jóola languages, like the other Atlantic languages of this area, do not have tone). Consequently, each of the five letters i, e, a, o, and u has two possible values depending on the presence vs. absence of the acute accent marking the word as ±ATR.

Jóola-Banjal morphosyntax is characterized by a system of noun classes similar to that found in other Atlantic languages, manifested by noun prefixes and obligatory agreement between the noun and several types of noun dependents, and between the subject NP and the verb. The numbering of noun classes we use in this paper, taken from Bassène 2007, must be viewed as arbitrary, although it aims at preserving consistency with previous descriptions of other Jóola languages whenever possible. The only coincidence with the system used in Bantu linguistics is that the classes typically including human nouns are labeled 1 (singular) and 2 (plural).

Jóola-Banjal has no case-marking of nouns, but the indexation of arguments by means of verbal prefixes and suffixes provides a firm basis for recognizing a syntactic function ‘subject’ grouping together the single core argument S of intransitive verbs and the agent A of prototypical transitive verbs, contrasting with a syntactic function ‘object’ including the patient of prototypical transitive verbs.

Of particular relevance to the topic of this paper is the regularity according to which, in finite predicative constructions (i.e. predicative constructions that can give rise to independent clauses), verbs normally include an obligatory prefix representing the single core argument S of intransitive verbs and the agent A of prototypical transitive verbs. If a co-referent NP is present, this subject marker expresses either class agreement (with non-human NPs) or person-number agreement (with human NPs and personal pronouns). In the absence of a co-referent NP, subject markers that do not belong to the 1st or 2nd person are interpreted anaphorically, triggering the identification of the argument they represent to a contextually salient referent compatible with the class or person-number value expressed by the subject marker – examples (1) to (3).

(1) a. Atejo na-tiñ-e si-nnaŋ s-a-s-u. Atéjo 3SG-eat-PF CL4-rice CL4-ANA-CL4-DEIX₂³
‘Atéjo ate the rice.’

2. The distinction between class agreement with non-human NPs and person-agreement with human NPs follows from the fact that, with non-human subjects, the subject marker always reflects the class prefix of the noun, whereas human subjects that exceptionally do not belong to classes 1/2 are represented by the same subject markers as human nouns belonging to classes 1/2.

3. The demonstrative determiners and pronouns of Jóola-Banjal include a fixed element -au-, glossed dem, and a variable deictic element with 3 possible values: -e (in the sphere of the speaker, glossed DEIX₁), -u (in the sphere of the addressee, glossed DEIX₂), and -ua (vague,
(2) a. *Na-tiñ-e si-nnaŋ s-a-s-u.
   3sg-eat-pf cl4-rice cl4-ana-cl4-deix₂
   ‘(S)he ate the rice.’

   b. *Atejo tiñ-e si-nnaŋ s-a-s-u

(3) a. *Atejo tiñ-e si-nnaŋ s-a-s-u

   b. *Atejo tey-e

In other words, the canonical verbal predication in Jóola-Banjal can be schematized as follows:\(^4\)

\[(A/S) a/s-V (P) (X…)](A/S)\]

A detail that greatly facilitates the recognition and analysis of impersonal constructions is that none of the subject markers of Jóola-Banjal has a zero realization or a zero allomorph.

The P argument of transitive verbs is not obligatorily indexed, but Jóola-Banjal has weak object pronouns, i.e. object pronouns that do not constitute autonomous words. They are realized as verb suffixes – example (4), and the ability to be represented by weak pronouns suffixed to the verb is not limited to objects – example (5).

(4) Ni-tiñ-e gu-mango g-a-g-u. \(\rightarrow\) Ni-tiñ-e-go.
   1sg-eat-pf cl8-mango cl8-ana-cl8-deix₂ 1sg-eat-pf-cl8
   ‘I have eaten the mangos’  ‘I have eaten them’

(5) Ni-jug-e su-joba s-a-s-u ni bi-it.
   1sg-see-pf cl4-dog cl4-ana-cl4-deix₂ in cl5-rice.field
   ‘I have seen the dogs in the rice fields.’
   \(\rightarrow\) Ni-juk-so-bo.
   1sg-eat-pf-cl4-cl5
   ‘I have seen them there.’

glossed \textit{deix}_3). Jóola-Banjal also has an anaphoric determiner combining a specific formative \textit{-a} (glossed \textit{ana}) and the second deictic marker.

4. \(A = \) agent of transitive verbs, \(S = \) single core argument of intransitive verbs, \(a/s = \) verbal prefix representing the A or S argument, \(P = \) patient of transitive verbs, \(X = \) oblique.
In addition to vowel harmony, the clearest evidence that weak object pronouns are affixes comes from the fact that, in the inflected form of the verb expressing verb focalization, characterized by the reduplication of the stem, they are inserted between the stem and its copy, as in \textit{ni-tif-fo-tiñ} ‘I ate it (the mango).’

On the limitations to the possible combinations of weak pronouns suffixed to the same verb, see Bassène (2007: 92–3).

3. **The notion of impersonal construction in Jóola-Banjal**

As discussed in Creissels (2007) and in the introduction of Siewierska (ed.) (2008), the notion of impersonality underlying traditional analyses of European languages is broad and disparate; it does not characterize a well-defined and homogeneous type of construction, but rather a family of constructions that in some way or other lack a canonical subject. The delimitation of impersonal constructions depends on the notion one has of what is (or should be) a canonical subject, and the extension of the notion of impersonal construction to other languages is therefore particularly problematic in the case of languages for which the notion of subject itself is controversial.

However, in a language like Jóola-Banjal, the agreement prefixes of verbs provide clear evidence of the relevance of a syntactic function ‘subject’ subsuming the single core argument of intransitive verbs and the agent of transitive verbs, and the obvious manifestation of a ‘canonical subject’ is the obligatory subject marker triggering an anaphoric interpretation in the absence of a co-referent NP. The subject marker constitutes in Jóola-Banjal the clearest manifestation of a canonical subject. Consequently, in this language, impersonal constructions in the sense of constructions that lack the usual manifestations of a canonical subject can be delimitated as follows: in Jóola-Banjal, an impersonal construction is a finite predicative construction in the sense of a construction underlying independent assertive clauses, which however differs from the prototypical predicative construction described in §2 in that the morphological slot normally occupied by a subject marker is left empty, or is occupied by a morpheme that does not show the canonical behavior of a subject marker with respect to agreement and/or reference.

4. **Intransitive verbs in a construction including no manifestation of the S argument**

A first type of impersonal construction involves intransitive verbs in a construction including no manifestation of the S argument obligatorily expressed in the canonical construction of the same verbs. Semantically, this construction expresses
predication about a situation which is not elaborated by selecting an element that would be treated as the S argument of a predication. As illustrated by examples (6c) and (7c), this construction concerns verbs describing states, and is in particular used to describe meteorological states.5

(6)  a. *Fu-toŋ f-a-f-u f u-mo-moc.*
    cl7-room cl7-ANA-cl7-DEIX cl7-be.dark-be.dark
    ‘The room is dark.’

    b. *Fu-mo-moc.*
    cl7-be.dark-be.dark
    ‘It is dark (the room, or other salient class 7 referent).’

    c. *Mo-moc.*
    be.dark-be.dark
    ‘It is dark (speaking of the atmospheric conditions).’

(7)  a. *M-al m-a-m-u m u-jébi-jébi.*
    cl10-water cl10-ANA-cl10-DEIX cl10-be.cold-be.cold
    ‘The water is cold.’

    b. *Mu-jébi-jébi.*
    cl10-be.cold-be.cold
    ‘It is cold (the water, or other salient class 10 referent).’

    c. *Jama jébi-e.*
    today be.cold-PF
    ‘It is cold today.’

Although particularly common with reference to meteorological states, this use of intransitive verbs describing states is more generally possible whenever the speaker has in mind a relatively vague referent for which the context does not readily provide a designation. For example, an impersonal formulation would not be acceptable in the answer to a question formulated as in example (8), because a referent previously named as *f a f u ‘the river’* is understood, whereas the impersonal formulation is normal in example (9), because the referent is only vaguely identified as ‘the place we are going to.’ As indicated below, we have observed an interesting parallelism with French, which uses the subject clitic *ça* encoding vagueness of reference much in the same conditions as the subjectless construction of Jóola-Banjal, contrasting with the subject

5. The reduplicated form of the verb found in these examples has the same TAM value as the form characterized by the suffix -e (glossed PF); it differs from it by implying emphasis on the verb, whereas the e-form is compatible with the focalization of other constituents of the clause. Sensitivity of verb morphology to focus phenomena is a feature Jóola languages share with other groups of languages included in the Atlantic family.
Impersonal constructions in Jóola-Banjal

clitics *il/elle* obligatorily used in contexts that would trigger the presence of a subject marker in Jóola-Banjal.

(8) –  
\[ \text{F-al f-a-f-u fú-ráli-ráli?} \]
\[ \text{CL7-river CL7-ANA-CL7-DEIX\textsubscript{2} CL7-be.far-be.far} \]
‘Is the river far?’ (French *Il est loin, le fleuve?*)

–  
\[ \text{Xani, fú-ráli-ut. / *Xani, ráli-ut.} \]
\[ \text{no CL7-be.far-NEG} \]
‘No, it is not far.’ (French *Non, il n’est pas loin (le fleuve).*

(9) –  
\[ \text{B-o nu-ja-ale-me ráli-ráli?} \]
\[ \text{CL5-REL 1PL-go-1PL-sbd be.far-be.far} \]
‘The place we are going to, is it far?’ (French *C’est loin, là où on va?*)

–  
\[ \text{Xani, ráli-ut.} \]
\[ \text{no be.far-NEG} \]
‘No, it is not far.’ (French *Non, ce n’est pas loin./*Non, il n’est pas loin)

5. Intransitive verbs in constructions in which the S argument shows object properties

5.1 The impersonal construction of *eŋañño* ‘remain’

*Eŋañño* ‘remain’ is an intransitive verb occurring in canonical intransitive predication – example (10), but also in a construction in which no subject marker is prefixed. In this construction, the absence of the subject marker correlates with the postverbal position of the NP representing the S argument, which consequently occupies the position normally occupied by the objects of transitive verbs – example (11).

(10)  
\[ \text{Si-rálam s-a-s-u si-ŋañño-e.} \]
\[ \text{CL4-money CL4-ANA-CL4-DEIX\textsubscript{2} CL4-remain-PF} \]
‘The money remained.’

(11)  
\[ \text{Nñaño-e si-rálam.} \]
\[ \text{remain-PF CL4-money} \]
‘There remained some money.’

As reflected in the translation, this construction belongs to a type of construction attested in many languages in which the canonical constituent order in transitive predication is AVP, and S in the canonical construction of intransitive verbs is morphosyntactically aligned with A. In such languages, a more or less important subclass

6. On the notion of ‘sujet indistinct’ in French, see Corblin (1991).
of intransitive verbs may have an alternative construction encoding a 'presentational' (or 'thetic', 'sentence-focus') organization of predication, in which the S argument is represented by an NP in postverbal position (i.e. in the canonical position of objects). Depending on language-specific rules, in such constructions, the S argument of intransitive verbs may lose other subject properties (in particular, the control of verb agreement) and acquire object properties, as discussed among others in Lambrecht (2000) and Creissels (2007).

As illustrated by example (11), contrary to what is observed in languages that have 'locative inversion' constructions, in Jóola-Banjal, the special treatment of the S argument of eŋañño in this construction is not bound to the presence of a fronted locative expression.

In some languages (for example, Tswana), such a construction is possible for any intransitive verb without any particular restriction. In other languages (for example, French), it is common for some intransitive verbs only, but it has been shown that it does not really involve a division of intransitive verbs into two subclasses, since various factors (in particular, the presence of a locative adjunct in frame function) may improve its acceptability for verbs that at first sight do not accept it. Jóola-Banjal illustrates the borderline case of a language in which this type of construction is possible with a very restricted set of intransitive verbs only: according to our observations, eŋañño 'remain' is the only verb of Jóola-Banjal that can be found in this construction.

Interestingly, this situation is not unique to Jóola-Banjal among West African languages: in Wolof too, 'remain' seems to be the only verb attested in a presentational construction of this type (Sylvie Nouguier, p.c.), and the same is observed in Manding. The case of Manding is all the more interesting because Manding and Jóola are typologically very different, and belong to language families (Mande and Atlantic, respectively) that have only a very remote genetic relationship, if any. This could perhaps be interpreted as evidence that 'remain' has semantic properties resulting in a particular predisposition to be used in a presentational construction.

5.2 The impersonal construction of intransitive verbs with a clausal argument

We have seen in §4 that intransitive verbs describing states, like 'be dark', 'be cold' or 'be far' have an impersonal construction devoid of any manifestation of their single core argument, found in particular (but not exclusively) in meteorological constructions. This section deals with intransitive verbs describing states and combinable with a single argument of clausal nature, such as eári 'be good' or essuneni 'be shameful'. When their core argument is of clausal nature, either finite, as in (12b), or non-finite, as in (13b) and (14b), such verbs have an impersonal construction in which no subject marker is prefixed to the verb, and the clausal argument occurs in postverbal construction, i.e. in
the position typical for objects in transitive predication. Note that, in (14), the \( w \)- added to the stem in the impersonal construction is not a subject marker, but an epenthetic consonant the presence of which results from a morphophonological rule. If it were a subject marker, it would not be repeated between the stem and its copy.

(12) a. Áine au-m-u na-suneni-e. (CL1)man (CL1)DEM-CL1-DEIX\(_2\) 3SG-be.shameful-PF
   ‘This man[’s attitude] is shameful.’

b. Suneni-e na-jow gu-ñen gu-rakkel. shameful-PF 3SG-go CL8-hand CL8-empty
   ‘It is shameful for him to go with empty hands.’

(13) a. Bu-nuk b-au-b-u bu-ssu-ssum. cl5-palm.wine cl5-DEM-CL5-DEIX\(_2\) CL5-be.sweet-be.sweet
   ‘This palm wine is sweet.’

b. Su-ssum e-rem bu-nuk ni bujom. be.sweet-be.sweet CL3-drink CL5-palm.wine in morning
   ‘It is pleasant to drink palm wine in the morning.’

(14) a. Fu-mangu f-a-f-u fu-ári-ári. cl7-mango cl7-ANA-CL7-DEIX\(_2\) CL7-be.good-be.good
   ‘The mango is good.’

b. Wári-wári e-baj a-ññil áine. be.good-be.good CL3-have CL1-child (CL1)man
   ‘It is good to have a son.’

The same construction is possible with passives. Jóola-Banjal has a passive suffix \(-i\), canonically used in an agentless passive construction in which the P argument of a transitive verb is encoded as the subject of an intransitive predication, whereas the A argument cannot be expressed. With transitive verbs that accept a second argument of clausal nature, such as effas ‘know’ in (15a), the canonical passive construction illustrated by (15b) is not possible with clausal arguments. Clausal arguments can nevertheless occur in the impersonal passive construction illustrated by example (15c), in which they are treated like clausal constituents that constitute the sole core argument of intransitive verbs in examples (12b), (13b) and (14b).

(15) a. U-añ-a-a-w gu-fas-e cl6-cultivate-AGNR-ANA-CL6\(^7\) 3PL-know-PF
   bi-it b-a-b-u tu bu-om. cl5-rice.field cl5-ANA-CL5-DEIX\(_2\) where cl5-be.found
   ‘The farmers know where the rice field is.’

7. \(-aw\) can be analyzed as a cliticized variant of \( wawu \) (anaphoric determiner, class 6).
b. Áine au-m-u na-fas-i bánoban.
   (CL1)man (CL1)DEM-CL1-DEIX₂ 3SG-know-PASS everywhere
   ‘This man is known everywhere.’

c. Fas-i gáabuok áine
   know-PASS that (CL1)man
   au-m-u na-ssái-ssái.
   (CL1)DEM-CL1-DEIX₂ 3SG-be.a.sorcerer-be.a.sorcerer
   ‘It is known that this man is a sorcerer.’

Note that, with infinitives, the impersonal construction of intransitive verbs combined with clausal arguments is not the only possibility. The canonical predicative construction illustrated by examples (16) and (17) is also possible, with the infinitive phrases e-jow gu-ñen gu-rakkel and e-baj a-ñnil áine in subject role. This can be viewed as evidence that the infinitive phrases of Jóola-Banjal have an intermediate status between fully canonical NPs and clausal constituents unable to occupy the most typical nominal positions in the matrix clause.

(16) E-jow gu-ñen gu-rakkel e-suneni-e.
   cl3-go cl8-hand cl8-empty cl3-shameful-PF
   lit. ‘To go with empty hands is shameful.’

(17) E-baj a-ñnil áine e-ári-ári.
   cl3-have cl1-child (cl1)man cl3-be.good-be.good
   lit. ‘To have a son is good.’

6. Impersonal use of ebaj ‘have’ in existential predication

The transitive verb ebaj ‘have’ has a canonical construction in which it assigns the role of possessor to the subject NP, and the role of possessee to the object NP, as in examples (18) and (19).

(18) Ni-baj-e ji-iba (ni e-poc-om).
   1sg-have-PF cl11-knife in cl3-bag-1sg
   ‘I have a knife (in my bag).’

(19) Atéjó na-baj-e gu-ñnil futox.
   Atéjó 3sg-have-PF cl2-child five
   ‘Atéjó has five children.’

8. In Jóola-Banjal, infinitives are not immediately recognizable as such (hence the absence of the abbreviation INF in the glosses), since they consist of a noun class prefix and a verbal stem, like other deverbal nouns. They however differ from other deverbal nouns by retaining verbal characteristics in their ‘internal’ syntax, i.e. in the internal structure of the phrases they head – Bassène (2007:181–185).
The same verb expresses an existential meaning in an impersonal construction characterized by the absence of any manifestation of a subject, but in which an object NP is present in the same way as in the canonical construction of the same verb – examples (20) to (23).

(20) Baj-e ji-iba ni e-poc y-a-y-u.
    have-pf CL11-knife in CL3-bag CL3-ANA-CL3-DEIX₂
    'There is a knife (in the bag).'

    1sg-see-see that have-have CL2-person HAB 3PL-steal
    'I saw that there were people stealing.'

(22) Filay pan baj bu-kut Mof Avví.
    this.year ipf have CL5-initiation Mof Avvi
    'This year there will be initiation in the Mof Avvi.'

(23) Baj-ut si-ñaru ni ga-llak g-a-g-u.
    have-NEG CL4-monkey in CL9-field CL9-ANA-CL9-DEIX₂
    'There are no monkeys in the field.'

This impersonal use of 'have' is in particular a common strategy for introducing new participants in the discourse – example (24).

(24) Baj-en-e aíne a-cce ga-ja-o1 Kajaka.
    have-pst-pf (CL1)man CL1-other CL9-name-3SG Kajaka
    'There was another man called Kajaka.'

The use of 'have' verbs in impersonal constructions with an existential meaning is cross-linguistically common – Creissels (1979: 494–504), Creissels (2007). In some languages (Greek, Wolof), the form used with an existential meaning is simply marked as 3rd person singular, and is therefore ambiguous between an existential reading in which the 3rd person mark is not referential, and a possessive reading in which the same mark receives an anaphoric interpretation. Other languages (French) avoid the ambiguity by adding a non-referential locative element to the form of 'have' conveying an existential meaning. In Jóola-Banjal, the absence of the otherwise obligatory subject marker unambiguously distinguishes the existential reading of ebaj from the possessive reading.

7. Impersonal use of the negative identification copula

Jóola-Banjal has verbless predicative constructions expressing identification and localization. In the positive, identification can be expressed by mere juxtaposition of constituents, as in (25), whereas localization is expressed by means of a copula agreeing
with the subject, as in (26). The structure of this copula is $u$-CL-$e$/$u$/$ua$. The first formative is a constant element $u$-glossed cop. The second formative is a class agreement marker, and the third formative $-e$/$u$/$ua$ is a deictic element found also as a formative of demonstrative pronouns and determiners (see Note 3 above).

   Atéjo $cl1$-cultivate-$AGNR$  
   'Atéjo is a farmer.'

(26) $Si$-$ñaru$ $u$-$s$-$ua$ $ni$ $ga$-$llak$ $g$-$a$-$g$-$u$.  
   $cl4$-monkey $cop$-$cl4$-$DEIX_3$ in $cl9$-field $cl9$-$ANA$-$cl9$-$DEIX_2$  
   'Some monkeys are in the field.'

Clauses in which the identification, categorization, or localization of an entity is negated make use of the negative copula -let, prefixed with a subject marker, as in (27) and (28).

(27) $Atejo$ $a$-$let$ $a$-$añ$-$a$.  
   Atéjo $3$sg-$cop$-$NEG$ $cl1$-cultivate-$AGNR$  
   'Atéjo is not a farmer.'

(28) $Si$-$ñaru$ $s$-$a$-$s$-$u$ $si$-$let$ $ni$  
   $cl4$-monkey $cl4$-$ANA$-$cl4$-$DEIX_2$ $cl4$-$cop$-$NEG$ in $ga$-$llak$ $g$-$a$-$g$-$u$.  
   $cl9$-field $cl9$-$ANA$-$cl9$-$DEIX_2$  
   'The monkeys are not in the field.'

The preceding examples are not impersonal. However, in clauses expressing negative identification, the negative copula also has an alternative construction in which no subject marker is prefixed to it, and no NP designating the entity in question is present, as in examples (29) and (30).

(29) $Let$ $inje$.  
   $cop$-$NEG$ $1$sg  
   'It’s not me.'

(30) $Let$ $Atejo$.  
   $cop$-$NEG$ Atéjo  
   'It’s not Atéjo.'

This impersonal use of the negative copula combines with the focalization construction to express negative focalization of a constituent in verbal predication, as in examples (31c), (32c) and (33c). In addition to intonation, focalization is marked, either by the use of different subject markers (if the focalized constituent is the subject), or by fronting the focalized constituent (in the other cases).
(31) a. Atejo na-ttep-e y-aŋ y-a-y-u.
Atéjo 3sg-build-pf CL3-house CL3-ANA-CL3-DEIX
'Théjo has built the house.'
b. Atejo a-ttep-e y-aŋ y-a-y-u.
Atéjo 3sg.foc-build-pf CL3-house CL3-ANA-CL3-DEIX
'It is Atéjo that has built the house.'
c. Let Atejo a-ttep-e y-aŋ y-a-y-u.
COP.NEG Atéjo 3sg.foc-build-pf CL3-house CL3-ANA-CL3-DEIX
'It is not Atéjo that has built the house.'

1sg-see-3sg Dakar
'I saw him/her Dakar.'
b. Dakkar ni-jug-ol.
Dakar 1sg-see-3sg
'It is in Dakar that I saw him/her.'
c. Let Dakkar ni-jug-ol.
COP.NEG Dakar 1sg-see-3sg
'It is not in Dakar that I saw him.'

(33) a. A-ññil a-k-u u-m-e fatia
CL1-child ANA-CL1-DEIX COP-CL1-DEIX on
bu-ra b-a-b-u.
CL5-bed CL5-ANA-CL5-DEIX
'The child is on the bed.'
b. Fatia bu-ra b-a-b-u a-ññil
on CL5-bed CL5-ANA-CL5-DEIX CL1-child
a-k-u a-am.
ANA-CL1-DEIX 3sg-cop.foc
'It is on the bed that the child is.'
c. Let fatia bu-ra b-a-b-u a-ññil
COP.NEG on CL5-bed CL5-ANA-CL5-DEIX CL1-child
a-k-u a-am.
ANA-CL1-DEIX 3sg-cop.foc
'It is not on the bed that the child is.'

8. Lack of agreement in a construction including a subject pronoun

The construction dealt with in this section concerns three verbs, ekkan, ecil, and etek. These verbs occur in a canonical transitive construction, but they also occur in a
construction expressing causation in which they take a clausal complement expressing the caused event.

Examples (34) to (36) illustrate the canonical construction of *ekkan*, *ecil*, and *etek* with their literal meanings of ‘hit’, ‘possess’, and ‘do’, respectively.

(34) \[Na\-teg\-e\ a\-ñïï\ a\-k\-u.\]
3sg-hit-pf cl1-child ANA-CL1-DEIX\(_2\)
‘He hit the child.’

(35) \[Atejö\ a\-cil\-e\ y\-a\-ñ\ y\-a\-y\-u.\]
Atejö 3SG.FOC-possess-pf cl3-house cl3-ANA-CL3-DEIX\(_2\)
‘The house belongs to Atejö.’

(36) \[A-cïl\ a\-kkan\-e\ b\-u\-rokk\ b\-a\-b\-e.\]
3SG-EMPH 3SG.FOC-do-pf cl5-work cl5-DEM-CL5-DEIX\(_1\)
‘He did this work himself.’

As verbs of causation, *ekkan*, *ecil*, and *etek* can still occur in a canonical predicative construction, if the cause is encoded by phrases that have the ability to function as canonical subjects: either a canonical NP, as *gasómulol* ‘his/her sickness’ in example (37) and *sibé sasu* ‘the cows’ in example (38), or an infinitive phrase, as *gabajut bunaa* ‘the lack of sun’ in example (39).

(37) \[Ga\-sómu\-ol\ gu\-teg\-e\ a\-kkay\-ut\ e-añ.\]
cl9-sickness-3SG cl9-hit-pf 3SG-go-NEG cl3-cultivate
‘His/her sickness made that (s)he did not go cultivating.’

(38) \[Si\-bé\ s\-a\-s\-u\ si\-cil\-e\ ni\ gu\-ssanumo.\]
cl4-cow cl4-ANA-CL4-DEIX\(_2\) cl4-possess-pf sbd 3PL-get.rich
‘It is owing to the cows that they got rich.’

(39) \[Ga\-baj\-ut\ bu\-naa\ gu\-kkan\-e\ b\-u\-rokk\ b\-a\-b\-u\ ni\ b\-u\-ssum.\]
cl9-have-NEG cl5-sun cl9-do-pf cl5-work cl5-ANA-CL5-DEIX\(_2\)
SBD cl5-be.pleasant
‘The lack of sun made work pleasant.’

The impersonal construction, characterized by the fact that no subject marker is prefixed to the verb, is used when the cause is encoded as a topicalized clause resumed by the pronoun mó, as in examples (40) to (42).

(40) \[Na\-sómu\-sómut, m\-o\ teg\-e\ a\-kkay\-ut\ e-añ.\]
3SG-be.sick-be.sick cl10-PRO hit-pf 3SG-go-NEG cl3-cultivate
‘He is sick, that’s the reason why he did not go cultivating.’

(41) \[A\-añ\-a\ ámak, m\-o\ cil\-e\ na\-mmeŋ\ e\-mmano.\]
cl1-cultivate-AGNR (cl1)great cl10-PRO possess-pf 3SG-be.full.of cl3-rice
‘He is a great farmer, that’s why he has plenty of rice.’
Impersonal constructions in Jóola-Banjal

(42) Mata e-bekkan-ol e-lú-lú, m-o
    since CL3-bicycle-3SG CL3-be.pierced-be.pierced CL10-PRO
    kan-e na-bbañ.
do-PF 3SG-return

‘His bicycle had a puncture, that’s why he returned.’

Mo belongs to a paradigm of anaphoric pronouns formed by prefixing a class marker to a constant element -o, glossed pro. Such pronouns occupy nominal positions in the clause, and in subject position, they normally trigger verb agreement like any NP in the same position. In the construction illustrated by examples (41) to (43), the pronoun of class 10 mo does not represent a noun belonging to class 10, but a topicalized clause. Comparison with examples (38) to (40) supports analyzing it as the subject of a verb expressing causation, but it does not behave like a subject NP, since it is not resumed by a subject marker prefixed to the verb, and it cannot be analyzed as occupying the subject marker slot either, since it does not undergo vowel harmony.

9. A construction involving a frozen subject marker

In Jóola-Banjal, a construction with an obligatory but invariable subject marker that cannot be explained as expressing agreement with any of the NPs involved in the construction is attested with the verb ejju.

This verb occurs in a canonical predicative construction with the meaning ‘begin’ or ‘do something first’; this construction is a raising construction in which the dependent verb may be in the infinitive (if the desired meaning is ‘begin doing something’) or in a finite form (if the desired meaning is ‘do something first’). In both cases, ejju ‘begin’ shows a subject prefix representing the first argument of the dependent verb – examples (43) and (44). Note that, in example (43), the class 7 prefix in fi-tiñ functions as an infinitive marker.9

(43) A-ññil a-k-u filay na-jju-e fi-tiñ si-nnañ.
    cl1-child ana-cl1-deix2 this.year 3SG-begin-PF CL7-eat CL4-rice
‘The child began eating rice this year.’

(44) Gu-jju-e gu-rósor bala gu-kkay lekkol.
3PL-begin-PF 3PL-play before 3PL-go school
‘They first played before going to school.’

9. Infinitives in Jóola-Banjal are marked by a lexically determined noun class prefix. The class 3 prefix is particularly productive in infinitive marker function, but other noun class prefixes fulfill the same function with limited subsets of verbs.
Examples (43) and (44) illustrate the use of *ejju* ‘begin’ in a canonical raising construction. We now examine the impersonal construction, in which this verb invariably takes a subject marker of class 3 which must therefore be analyzed as an expletive. In this impersonal construction *ejju* expresses, either ‘have just occurred’, as in examples (45) and (46), or ‘occur for the first time’, as in example (47). The two meanings are not differentiated in the construction, and the choice is entirely context-dependent. Semantically, the valency of the verb combined with *ejju* undergoes no change. Formally, the argument normally encoded as the subject of the verb in the infinitive is optionally represented by an NP preceding *ejju*, and obligatorily indexed on *ejju* by means of object suffixes that control the missing subject of the dependent verb. In other words, in this impersonal construction, the subject argument of the verb in the infinitive has the behavior that, in canonical predications, characterizes topicalized objects.10

(45)  Aare a-k-u ejju-ol e-ci g-ul.

(cl1)woman ana-cl1-deix2 cl3-have.just.occurred-3sg cl3-arrive-VEN

‘The woman has just arrived.’
lit. ‘The woman, it has just occurred to her to arrive.’

(46)  E-jju-óli e-púr-ul ni bi-it.

cl3-have.just.occurred-1pl cl3-go.out-VEN from cl5-rice.field

‘We have just gone out from the rice fields.’
lit. ‘It has just occurred to us to go out from the rice fields.’


cl3-have.just.occurred-1sg cl7-eat cl3-meat cl3-monkey

~occur.for.the.first.time

(a)  ‘I have just eaten monkey meat.’

(b)  ‘I am eating monkey meat for the first time.’
lit. ‘It has just occurred/occurs for the first time to me to eat monkey meat.’

Given that *ejju* ‘begin’ combines in this construction with a verb in the infinitive, and most infinitives belong to class 3, it may well be that, historically, the expletive subject marker of class 3 originally expressed agreement with an infinitive in subject function. However, this analysis would not account for the synchronic facts since, as illustrated by example (47), infinitives belonging to other classes (such as *fi-tiñ*) do not trigger any change in the subject marker prefixed to *ejju*.

---

10. As observed by Andrej Malchukov, this construction shows some analogy with the impersonal construction of *slučit’šja* ‘occur, happen’ in Russian (*On prišel ‘he came’ → Emu slučilos’ prijti* lit. ‘Him (DAT) happened to come (INF).*
An interesting particularity of this construction is that, if the dependent verb is transitive, object markers indexing its P argument may attach, either to the infinitive, or to ejju, after the object marker representing the A argument, as illustrated by example (48).

(48) \[
\begin{align*}
E\text{-}jju\text{-}om & \quad e\text{-}jug\text{-}ol. \\
\text{cl3-have.just.occurred-1sg} & \quad \text{cl3-see-3sg}
\end{align*}
\]

\[
= \begin{align*}
E\text{-}jju\text{-}om\text{-}ol & \quad e\text{-}juk. \\
\text{cl3-have.just.occurred-1sg-3sg} & \quad \text{cl3-see}
\end{align*}
\]

~occur.for.the.first.time

(a) ‘I have just seen him.’

(b) ‘I am seeing him for the first time.’

The possibility to attach a second object marker representing the P argument of the dependent verb is however limited by a person hierarchy (1 > 2 > 3): the object marker representing the P argument of the dependent verb can ‘climb’ only if the A argument (obligatorily indexed on ejju) stands higher in person hierarchy.

10. Arbitrary reading of the second person singular and third person plural subject markers

In addition to constructions that formally depart from the usual constraints on subject markers, Jóola-Banjal has two cross-linguistically very common constructions that show no deviance from a strictly formal point of view, but in which second person singular is not interpreted as referring specifically to the addressee, and third person plural is not interpreted as referring to a specific group of individuals known to the addressee. As observed in typical ‘pro-drop’ languages like Spanish, the arbitrary reading of 2nd person singular and 3rd person plural is limited to person markers attached to the verb, in constructions in which the morphosyntactic slot for subject NPs is left empty. The corresponding independent pronouns always have a specific reading.

The semantics of arbitrary 2nd person singular and 3rd person plural subject markers in Jóola-Banjal is in accordance with what has been observed for similar constructions in other languages (see Cabredo Hofherr 2003 and references therein).

Clauses including an arbitrary 2nd person singular subject marker express generalizations (‘In general, or when certain conditions are met, it occurs that …’), and the arbitrary 2nd person subject marker is inclusive, in the sense that the generalizations expressed by such clauses may apply to the speaker and the addressee too. This impersonal use of the 2nd person singular is particularly common in utterances that give instructions not bound to a specific situation, or which describe the usual way
an activity is performed. Examples (49), (50), and (51) are taken from texts describing palm-wine making, fishing, and marriage customs, respectively.

(49) ... nu-kok ga-ndapa-i ni ñi-it ñ-a-ñ-u
   2SG-tie CL9-climbing.belt-2SG on CL12-palm.tree CL12-ANA-CL12-DEIX₂
   '... you tie your climbing-belt on the palm tree
   min u-pirik u-bes-ño.
   and 2SG-cut CL6-branche-CL12
   and you cut its branches.'

(50) U-ban-me nu-tos bi ti-cce min u-bet
   2SG-finish-sbd 2SG-move up to CL13-other and 2SG-throw
   y-a-ñ-u e-mbal.
   CL3-ANA-CL3-DEIX₂ CL3-fishing.net
   'When you have finished, you move to another place and you throw the other fishing-net.'

(51) Nø ñaare u-ŋes-ol ükki u-re
   CL15.pro (CL1)woman 2SG-fetch-3SG until 2SG-reach
   netut nu-remp-ol
   middle 2SG-get.engaged.with-3SG
   'In that time, a woman, once you had courted her up to a certain point, you asked for her hand.'

By contrast, clauses including an arbitrary 3rd person plural subject marker refer to specific events, and express indetermination as to the identity of the subject argument: either the precise identity of the subject argument is not known to the speaker, or for some reason (s)he does not want to make it explicit. The arbitrary 3rd person plural subject marker is exclusive, in the sense that the speaker and the addressee cannot be included in the group of people referred to. In the following examples, the 3rd person plural subject marker presents the same ambiguity between an indeterminate and a specific reading as they in the English translation.

(52) Gu-jo-jok a-kkú a-k-u.
   3PL-catch-catch CL1-thief ANA-CL1-DEIX₂
   'They have caught the thief.'
   or 'The thief has been caught.'

(53) Gu-ppeg-e bu-lago b-a-b-u b-a-a súndo.
   3PL-shut-pf CL5-road CL5-ANA-CL5-DEIX₂ CL5-GEN home
   'They have blocked the road that leads to our village.'
   or 'The road that leads to our village has been blocked.'
With transitive verbs, the use or an arbitrary 3rd person plural is functionally similar to the use of a passive construction in which the P argument is encoded as the subject of a derived verb form marked by the passive suffix -i. Some precisions about Jóola-Banjal passive are in order at this point. In Jóola-Banjal, the only possible functional equivalent of passive constructions with an agent phrase in oblique function is an active construction combining object topicalization and subject focalization, as in example (55).

(55) A-ññil a-k-u, e-joba e-rum-ol
     cl1-child ana-cl1-deix2 cl3-dog cl3-bite-3sg
'The child has been bitten by a dog.'
lit. 'The child, A DOG has bitten him.'

The passive suffix -i is found in the construction illustrated by example (56), in which the agent cannot be expressed but is semantically implied. The passive character of this suffix (in the sense that it semantically implies the participation of an agent, which however cannot be expressed), follows from the contrast with anticausative forms that occur in superficially similar constructions with different semantic implications, since they imply removing the agent from argument structure – Bassène (2007: 165–168).

(56) A-ññil a-k-u na-rum-i.
     cl1-child ana-cl1-deix2 3sg-bite-pass
'The child has been bitten'

Semantically, the only difference between this passive construction and the arbitrary 3rd person plural construction is that the latter implies a human agent, whereas the agent implied by the passive construction is not necessarily human, as illustrated by example (56).

11. **Lexicalization of verb forms devoid of subject marker**

For the sake of completeness, it must also be mentioned that Jóola-Banjal has isolated cases of verb forms devoid of subject marker. They may be used by themselves or included in frozen idioms, with a more or less transparent but in any case non-compositional meaning. This can be illustrated by *kakkan* ‘apparently’ (< *ekkan* ‘do’) or *imbi púr to* ‘after that’ (< *imbi* obligative marker, *epúr* ‘come out’, to pronoun of class 13).

Such expressions are not amenable to regularities that would justify recognizing additional constructions involving verb forms devoid of subject marker. They are best
analyzed as adverbs that must be listed in the lexicon and that, historically, result from isolated lexicalization processes having affected individual verb forms, much in the same way as French peut-être or English maybe. The verbal origin of such expressions is clearly not relevant to a synchronic account of impersonal constructions in Jóola-Banjal morphosyntax.

12. Conclusion

In this paper, we have shown that Jóola-Banjal has a variety of impersonal constructions comparable to that observed in European languages, and that the functional domains in which these constructions are found are largely the same as in European languages. There are however some interesting contrasts, and on several points the data of Jóola-Banjal bring an interesting contribution to the general discussion of impersonality:

a. None of the subject markers of Jóola-Banjal has a zero realization. Consequently, the absence of the otherwise obligatory subject markers in several of the impersonal constructions of Jóola-Banjal facilitates the recognition of these constructions as impersonal. This situation contrasts with that of languages in which the forms found in the impersonal constructions are identical to forms triggering an anaphoric interpretation of the subject in canonical predications.

b. In many languages, the status of clauses describing meteorological events as impersonal is controversial. Jóola-Banjal has constructions describing meteorological events that are uncontroversially impersonal in the sense that they include no subject marker. The same impersonal construction of intransitive verbs is however also found in situations in which the argument structure clearly includes an S argument, but an S argument which is only vaguely identified. Consequently, the idea that the absence of the subject marker should always straightforwardly reflect an argument structure including no S argument must be abandoned. This aspect of the impersonal constructions of Jóola-Banjal supports the idea that, as already suggested for the meteorological constructions of other languages, verbs in meteorological expressions (whatever their formal make-up) are not characterized by the absence of an S argument, but rather by a low degree of referentiality of their S argument.

c. Possible variations in the extension of the subset of intransitive verbs that can occur in a presentational focus construction of the type found for example in French have been widely discussed in the unaccusativity literature. Jóola-Banjal contributes to the discussion by providing an example of a language in which such a construction has been observed with one intransitive verb only.

d. Among the cross-linguistically common functional types of impersonals, Jóola-Banjal entirely lacks affective impersonals, i.e. impersonal constructions motivated by the presence of an experiencer in argument structure.
Impersonal constructions in Jóola-Banjal

Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Meaning</th>
<th>Abbreviation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGNR</td>
<td>agent nominalizer</td>
<td>HAB</td>
<td>habitual</td>
</tr>
<tr>
<td>ANA</td>
<td>anaphoric determiner</td>
<td>IPF</td>
<td>imperfective</td>
</tr>
<tr>
<td>CL</td>
<td>noun class</td>
<td>NEG</td>
<td>negation</td>
</tr>
<tr>
<td>COP</td>
<td>copula</td>
<td>PASS</td>
<td>passive</td>
</tr>
<tr>
<td>DEIX₁</td>
<td>in the sphere of the speaker</td>
<td>PF</td>
<td>perfective</td>
</tr>
<tr>
<td>DEIX₂</td>
<td>in the sphere of the addressee</td>
<td>PL</td>
<td>plural</td>
</tr>
<tr>
<td>DEIX₃</td>
<td>‘vague’ deixis</td>
<td>PRO</td>
<td>pronoun</td>
</tr>
<tr>
<td>DEM</td>
<td>demonstrative</td>
<td>REL</td>
<td>relativizer</td>
</tr>
<tr>
<td>EMPH</td>
<td>emphatic pronoun or determinant</td>
<td>SBD</td>
<td>subordinator</td>
</tr>
<tr>
<td>FOC</td>
<td>focalization</td>
<td>SG</td>
<td>singular</td>
</tr>
<tr>
<td>GEN</td>
<td>genitive</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

References


Impersonal configurations and theticity

The case of meteorological predications in Afroasiatic

Amina Mettouchi & Mauro Tosco
EPHE, Paris/University of Turin

Impersonal predications are often viewed as structures expressing either agent defocusing, or lack of canonical subject properties. The study of one type of prototypical impersonal predication, meteorological predicates, in various Afroasiatic branches suggests that the subject or agent may not be centrally associated to the notion of impersonal. Rather, defocussing or backgrounding can concern either the entity or the event, resulting not only in subjectless structures and non-canonical subjects, but also in verbless structures and non-canonical predicates. What unifies those structures, rather than lack of canonical subjecthood or agent defocusing, is theticity, which may also be at play in other impersonal types than meteorological predicates.

Keywords: impersonal; Afroasiatic; theticity; meteorological predications; weather verbs

1. Preliminaries

Most studies on impersonals are based on European languages, and originally started with a class of structures whose characteristics are to have either unspeciﬁed agent pronouns (‘on’ in French, ‘man’ in German, etc.) or non-referential expletive third-person pronouns (‘il’ in French, ‘it’ in English, ‘es’ in German, etc.). Among the prototypical structures that have been studied, we find meteorological predications,¹ existential sentences, experiencer sentences, reflexive constructions, and sentences with an extraposed clausal argument. From this heterogeneous series of constructions, the notion that there could be a domain of “impersonal

predication”, around the notion of non-referential or non-specific status of the subject, has emerged.

This in turn has led linguists to find general criteria common to the various constructions listed above. Two directions have recently been taken (Siewierska): one is the analysis of impersonal predication as involving non-canonical subjecthood (Aikhenvald et al. 2001), the other one is the analysis of impersonalization as involving agent-defocusing. According to Creissels (Creissels 2006:325) for instance,2 ‘the constructions that are usually labelled impersonal constitute a heterogeneous set, the delimitation of which is the object of controversies. What is however constant in the use of the term impersonal construction, is that it is somehow or other difficult to apply to those constructions the notion of subject.’

Both trends in turn give rise to the possible inclusion under the label “impersonal” of various constructions that were not necessarily considered as such traditionally, such as action nominalizations, anticausatives, etc. Where should we draw the borders of the domain of impersonal predication, supposing that such a domain can indeed be delimited? At some point, is it not simpler to just define the domains we are studying as either that of non-canonical subjecthood, or that of agent-demotion, rather than as the domain of impersonal predication? Those questions are clearly beyond the scope of this paper, but they underline the fact that the underlying features found when investigating impersonal constructions always extend beyond the original boundaries of the domain. Such is also the case for the feature of theticity, which we would like to bring forward as a central feature of at least some impersonal predications.

Because we wished to have a precise starting point for the study of impersonality in Afroasiatic languages, we settled on the study of a subset of constructions that are considered as prototypical impersonals, namely meteorological predications describing raining events. We show that the formal strategies to encode raining predications are varied, and correspond to the backgrounding of either the entity or the situation. Further investigations concerning another type of meteorological constructions, external temperature predications, reinforce the claim we make that rain and temperature predications are impersonal constructions that centrally involve theticity. Theticity being a type of information packaging, several morphosyntactic encodings are possible. We argue that the fact that the subject or agent is involved

2. ‘[l]es constructions couramment désignées comme impersonnelles constituent un ensemble hétérogène, dont la délimitation est l’objet de controverses, mais ce qui est constant dans l’utilisation du terme de construction impersonnelle, c’est qu’il y a d’une manière ou d’une autre une difficulté à appliquer aux constructions ainsi désignées la notion de sujet’ (Creissels 2006:325).
in the coding of impersonals is language-specific, and depends on the properties of particular languages with respect to the referentiality of personal indexes for instance, or the existence of non-verbal predications.

Our preliminary investigation is based on various languages belonging to most branches of the Afroasiatic phylum: Cushitic, Berber, Semitic and Chadic. Afroasiatic languages are spoken in the northern and eastern parts of Africa and in the Near and Middle East. In the first part of this paper we show that raining predications involve partial or total backgrounding of either the entity or the process involved. In the second part we focus on two languages for which we have first-hand data, and show that raining and external temperature predications are expressed by constructions that are typically thetic, and that some of them are similar to predications belonging to other domains, such as epistemic modality, or attributive/equative predication. We also show that the grammatical elements involved in the construction of raining or external temperature predications are varied, and not limited to non-canonical subjecthood or agent-demotion strategies.

2. Atmospheric predications and impersonal constructions

Creissels (2006: 328) notes that ‘it is often, though not always, the case that meteorological predications simply conform to the subject + predicate format. Problems regularly arise however, and they are due to the the fact that it is difficult to recognize a ‘participant-event’ schema in all those situations.’ Creissels’ observations underline the fact that it is not specifically the subject, but rather the categorical (topic-comment) format which is problematic for the expression of meteorological phenomena.

The next section will briefly show the richness of these backgrounding processes in the atmospheric predications of different Afroasiatic languages. Backgrounding can

3. ‘[i]l est relativement courant dans les langues du monde que la description des phénomènes météorologiques se coule tout simplement dans le moule syntaxique sujet + verbe, mais la réduction des phénomènes météorologiques au schème sujet + verbe ne se fait pas toujours de la même façon, ce qui est déjà révélateur d’une difficulté à reconnaître dans ces situations une articulation événement – participant(s); et même dans des cas où une construction syntaxique canonique doit être reconnue, il n’est pas rare d’observer des particularités qui suggèrent la possibilité de dériver vers une construction où la reconnaissance d’un schème sujet + verbe serait problématique’ (Creissels 2006:328).
affect the entity or the process, and may be partial or total. This double articulation of backgrounding may be represented graphically as:

<table>
<thead>
<tr>
<th></th>
<th>partial</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>entity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>process</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1. Representation of the backgrounding possibilities

It seems that each backgrounding possibility may be actualized only once: in other words, partial backgrounding of an element and total backgrounding of another cannot cooccur. There are therefore four logically possible backgrounding types. Backgrounding of both the entity and the process is excluded – it would amount to nothing else than the absence of predication. Absence of backgrounding corresponds instead to the global apprehension of the predication, to which we turn in the next section.

Our analysis will concentrate on the different strategies used in the expression of “raining events”. One may object that raining is not a prototypical atmospheric predication. This is certainly true: for one thing, it conveys much less a state than predications like ‘to be sunny/cloudy’ or even of ‘to dawn/set (of sun)’; raining is much more something which happens against the backdrop of its absence, and this is probably true everywhere. It is an “event” much more than a state of affairs; it is dynamic rather than static.4

On the other hand, precisely this exceptionality of raining events (as well as of other less frequent atmospheric events like snowing, hailing, and the like) seems conducive to a wide array of possible linguistic realizations, where either the entity or the process are fore- and backgrounded. It is the ideal playground of backgrounding processes, as the following sections will show.

2.1 Global apprehension: The “the rain rains” strategy

When both the entity and the process are apprehended in toto, globally, there is no backgrounding. We have therefore constructions of the type “the rain rains”, which are well represented in the languages of the Horn of Africa, both in the Cushitic, represented here by Oromo, and the Semitic branch of Afroasiatic and represented here by Amharic:

4. As to Afroasiatic languages, they are mostly and traditionally spoken in areas where annual rainfall is well below the world average, and which are classified either as arid or semi-arid. A partial exception is provided by limited areas of the highlands of the Horn of Africa (where Semitic, Cushitic, and Omotic languages are spoken), and by certain tropical wet parts of West Africa where Chadic languages are found.
Impersonal configurations and theticity

(1) _bokee_ roob-ē³
rain  rain-PFV.3M
‘The rain rained’ > ‘it rained’  (Standard [Western] Oromo)

(2) _bokee_ hir-roow-a
rain  FOC-rain-PFV.3M
‘The rain will rain’ > ‘it will rain’  (Waata Oromo; Stroomer 1987: 381)

(3) _znab_ yɔ-zänbal
rain  PFV.3M-rain
‘The rain is raining’ > ‘it is raining’  (Amharic)

It will be noted that in 3. the entity and the process share the same root, contrary to 1. and 2. While such a construction is syntactically canonical in its entity-predication duality, it is semantically odd: one could easily speak here, following again Creissels (2006: 343), of the subject of such a configuration as a “prolongement du verbe”, an “internal subject” of the verb itself. Likewise, one can envisage that the predicate is here the “internal verb” of the subject.

One finds a similar strategy in Wandala (Central Chadic, Biu-Mandara; Frajzyngier in press), where a verb that means ‘to fall’, but is exclusively used with atmospheric predications, takes a noun meaning water or hail:

(4) á  là nálandzé
3sg fall hail
‘The hail is falling’ > ‘it is hailing’  (Wandala; Frajzyngier in press)

2.2 Partial backgrounding of the entity: The “the world rains” strategy

As soon as there is backgrounding, one of the two parts of the predication will be affected. The backgrounding of the entity corresponds to the well-known use of generic subjects (such as “sky”, “world”, “God”, or a semantically more abstract “state, situation”).

5. Transcription follows the sources or the standard orthography of the languages. Glosses have been modified when necessary. The following abbreviations are used:

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABS</td>
<td>absolute state</td>
</tr>
<tr>
<td>ANNX</td>
<td>annexion state</td>
</tr>
<tr>
<td>AOR</td>
<td>aorist</td>
</tr>
<tr>
<td>ART</td>
<td>article</td>
</tr>
<tr>
<td>ASSOC</td>
<td>associative</td>
</tr>
<tr>
<td>COMPL</td>
<td>complementizer</td>
</tr>
<tr>
<td>COP</td>
<td>copula</td>
</tr>
<tr>
<td>DECL</td>
<td>declarative marker</td>
</tr>
<tr>
<td>DEF</td>
<td>definite</td>
</tr>
<tr>
<td>F</td>
<td>feminine</td>
</tr>
<tr>
<td>FOC</td>
<td>focus marker</td>
</tr>
<tr>
<td>INGR</td>
<td>ingressive</td>
</tr>
<tr>
<td>IPFV</td>
<td>imperfective</td>
</tr>
<tr>
<td>IRR</td>
<td>irrealis</td>
</tr>
<tr>
<td>M</td>
<td>masculine</td>
</tr>
<tr>
<td>MID</td>
<td>middle</td>
</tr>
<tr>
<td>PFV</td>
<td>perfective</td>
</tr>
<tr>
<td>PROG</td>
<td>progressive</td>
</tr>
<tr>
<td>SBJ</td>
<td>subject</td>
</tr>
<tr>
<td>SG</td>
<td>singular</td>
</tr>
<tr>
<td>SPEC</td>
<td>specific</td>
</tr>
<tr>
<td>SUJN</td>
<td>subjunctive</td>
</tr>
<tr>
<td>UNM</td>
<td>unmarked</td>
</tr>
</tbody>
</table>

It is noteworthy that in 3. the entity and the process share the same root, contrary to 1. and 2. While such a construction is syntactically canonical in its entity-predication duality, it is semantically odd: one could easily speak here, following again Creissels (2006: 343), of the subject of such a configuration as a “prolongement du verbe”, an “internal subject” of the verb itself. Likewise, one can envisage that the predicate is here the “internal verb” of the subject.

One finds a similar strategy in Wandala (Central Chadic, Biu-Mandara; Frajzyngier in press), where a verb that means ‘to fall’, but is exclusively used with atmospheric predications, takes a noun meaning water or hail:

(4) á  là nálandzé
3sg fall hail
‘The hail is falling’ > ‘it is hailing’  (Wandala; Frajzyngier in press)

2.2 Partial backgrounding of the entity: The “the world rains” strategy

As soon as there is backgrounding, one of the two parts of the predication will be affected. The backgrounding of the entity corresponds to the well-known use of generic subjects (such as “sky”, “world”, “God”, or a semantically more abstract “state, situation”).

5. Transcription follows the sources or the standard orthography of the languages. Glosses have been modified when necessary. The following abbreviations are used:

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABS</td>
<td>absolute state</td>
</tr>
<tr>
<td>ANNX</td>
<td>annexion state</td>
</tr>
<tr>
<td>AOR</td>
<td>aorist</td>
</tr>
<tr>
<td>ART</td>
<td>article</td>
</tr>
<tr>
<td>ASSOC</td>
<td>associative</td>
</tr>
<tr>
<td>COMPL</td>
<td>complementizer</td>
</tr>
<tr>
<td>COP</td>
<td>copula</td>
</tr>
<tr>
<td>DECL</td>
<td>declarative marker</td>
</tr>
<tr>
<td>DEF</td>
<td>definite</td>
</tr>
<tr>
<td>F</td>
<td>feminine</td>
</tr>
<tr>
<td>FOC</td>
<td>focus marker</td>
</tr>
<tr>
<td>INGR</td>
<td>ingressive</td>
</tr>
<tr>
<td>IPFV</td>
<td>imperfective</td>
</tr>
<tr>
<td>IRR</td>
<td>irrealis</td>
</tr>
<tr>
<td>M</td>
<td>masculine</td>
</tr>
<tr>
<td>MID</td>
<td>middle</td>
</tr>
<tr>
<td>PFV</td>
<td>perfective</td>
</tr>
<tr>
<td>PROG</td>
<td>progressive</td>
</tr>
<tr>
<td>SBJ</td>
<td>subject</td>
</tr>
<tr>
<td>SG</td>
<td>singular</td>
</tr>
<tr>
<td>SPEC</td>
<td>specific</td>
</tr>
<tr>
<td>SUJN</td>
<td>subjunctive</td>
</tr>
<tr>
<td>UNM</td>
<td>unmarked</td>
</tr>
</tbody>
</table>
At the same time that the entity is backgrounded, the process comes by necessity to the foreground. This is a very common model in our languages. Among the many possible examples, it is the one found in Oromo (East Cushitic, Oromoid):

(5) waak’ii nu-u roob-e
   God us-to rain-PFV.3M
   ‘God/the sky has rained upon us’ > ‘it has rained for us’

(Borana Oromo; Stroomer 1987: 381)

In Chadic, we find similar examples for instance in Lele (East Chadic), where ‘[s]ome serial verb constructions have become fixed lexicalized expressions. The term for rain, clouds, is the same as the term for God, kumno. The expression ‘it rains’ is interesting in that it has a serial verb construction consisting of the verb bá (báá [W[eibegué] & P[alayer] 1982]) ‘fall’ and the verb ongi ‘to push as in childbirth, and ‘to rain’) in association with bá ‘all’” (Frajzyngier 2001: 122)

(6) kumno se bá ongi
    rain (= God) INCEPT fall rain (= push as in childbirth)
    ‘It started to rain’

(Lele; Frajzyngier 2001: 122)

In Kabyle (Berber), this strategy is at the root of meteorological expressions containing the subject lḥal ‘situation’, which is coindexed, as the Annexed state shows,6 with the third person masculine singular prefix on the verb:

(7) ye-ḥma lḥal
    subj3msg-be_hot.PFV situation.annx
    ‘It (the weather) is hot’

(Kabyle)

Partial backgrounding of the entity may acquire an intensive meaning, as in modern varieties of Arabic:

(8) id-dinya b-it-maṭṭar
    art-world prs-IPFV.3M-rain
    “The world rains” > ‘it is raining a lot’

(Egyptian Arabic)

while total entity backgrounding (the “it rains” model, cf. 2.4. below) has a more neutral meaning.

The partial backgrounding of the entity does not need its substitution with a generic entity: the natural entity may well be present, but lose at the same time its

6. The Annexed state is one of the two forms a noun can take in most Berber languages. It is obligatory in a number of contexts whose common denominator is dependence on a previous element for interpretation. For further details, see Mettouchi (2008).
Impersonal configurations and theticity

agentivity. Reduced to the role of syntactic object of a transitive verb, it may carry the semantics of the predication (the act of raining, for example), while both the syntactic subject and the predicate are external to it. This is the configuration “X (“the sky, God”) Y (“makes, hits”) the rain”. Although marginal in our domain, it is nevertheless found in Sidamo (East Cushitic, Highland):

(9) \( (Magan-u)/(gord-u) \) xeen-a \( gana-nno \)

\( \text{God-SBJ /sky-SBJ rain-ABS hit-IPFV.3M} \)

“(God/the sky) hits the rain” > ‘it is raining’  
(Sidamo; Gasparini 1983: 112)

(10) \( (Magan-u)/(gord-u) \) xeen-a \( \text{birr-i} \)

\( \text{God-SBJ /sky-SBJ rain-ABS rain-PFV.3M} \)

“(God/the sky) has rained the rain” > ‘it rained’  
(Sidamo; Gasparini 1983: 325)

\( Magan-u \) and \( gord-u \) are the Subject-case forms of \( magan-o \) ‘God’ and \( gord-o \) ‘sky’, respectively. They can be omitted, yielding, e.g. \( xeeana \) \( \text{birr-i} \). But in neither case the “natural” entity involved in the state of affairs (\( xeen-a \) ‘rain’) is the grammatical subject of the sentence. In this configuration the rain is rather the object, as shown by the presence of a transitive verb and, most of all, by the caseform of the noun itself: this appears in the Absolutive (citation) case, not in the Subject case. We find, e.g. \( xeen-a \), ‘rain’, not \( xeen-i \): *\( xeen-i \) \( \text{birr-i} \) is not acceptable.

2.3 Partial backgrounding of the process: The “the rain falls/hits” strategy

When backgrounding affects the process, it results in the use of a generic predicate, like “to fall”, or “to happen”, or, alternatively, in the use of a deictic particle. The consequence of the backgrounding of the process is of course, conversely, the foregrounding of the entity. This is the common model “the rain falls”, “the wind blows”, etc. found in many languages of the Horn of Africa, such as Somali (East Cushitic, Omo-Tana branch):

(11) \( \) roob baa da’-ay-a

\( \text{rain FOC fall-PROG-PRS.3M.FOC} \)

“The rain is falling” > ‘it is raining’  
(Somali)

as well as in Ts’ämakko (East Cushitic, Dullay):

(12) \( \) ?err-o-se \( \text{dib-i} \)

\( \text{rain-M-DEF fall-3SGM.UNM} \)

“The rain fell” > ‘it rained’  
(Ts’ämakko; Savà 2005: 84)

7. In the Sidamo orthography \( \langle x \rangle \) stands for ejective \( /t’/ \).
Interestingly, neither Somali *da’* [da?] nor Ts’amakko *dîb* are the usual verbs for ‘to fall’ (which are rather expressed by *dhac* [daʃ] in Somali and *pûl* in Ts’amakko). We have seen above (cf. (4)) that, likewise, in Central Chadic Wandala a special verb for “atmospheric fall” is used. Somali *da’* may actually be used for any falling liquid; its use for dripping water is fully acceptable:

(13)  
\[\text{biyo baa da’-ay-a}\]  
water FOC fall-prog-prs.3m.foc  
‘Water is falling’  
(Somali)

Other atmospheric entities will use other semantically more pertinent generic verbs:

(14)  
\[\text{danab baa dhacay}\]  
thunder FOC fall-pst.3m  
‘A thunder fell’ > ‘it thundered’  
(Somali)

(15)  
\[\text{dabaysha baa socota}\]  
wind-art.f FOC walk-mid-prs.3f  
‘The wind walks/comes’ > ‘the wind blows’  
(Somali)

(16)  
\[\text{cadceedda baa soo baxatay/dhacaday}\]  
sun-art.f FOC here go_out-mid-pst.3f/fall-mid-pst.3f  
‘The sun came out/fell’ > ‘the sun dawned/set’  
(Somali)

For less dynamic atmospheric predications other than ‘to rain’ other strategies will be used, as will be seen in §3.1. below with data from Gawwada, another East Cushitic language.

The “Rain falls” strategy is very widespread; instead of ‘to fall,’ the generic verb is often ‘to hit’ (which has been encountered upon above (9) in Sidamo “God hits the rain”. In East Cushitic, this use is attested, e.g. in Dhaasanac (Omo-Tana branch):

(18)  
\[\text{?îr kà tutuna}\]  
rain here hit-red-ipfv.a  
‘Rain is hitting’ > ‘it is raining’  
(Dhaasanac; Tosco 2001: 530)

The use of a ‘to hit’ verb is further found in Berber and Chadic:

(19)  
\[\text{te-kkat lehwa}\]  
subj3sg-hit.ipfv rain.annx  
‘Rain is hitting’ > ‘it is raining’  
(Kabyle)

(19’)  
\[\text{ye-kkat wedfel/wâdu}\]  
subj3msg-hit.ipfv snow.annx/wind.annx  
‘Snow/wind is hitting’ > ‘it is snowing/the wind is blowing’  
(Kabyle)
Note that in Berber the verb must be in the imperfective aspect, and the subject must always be postverbal (this will be elaborated on below, in §3). In Gidar (Central Chadic), we also find the verb ‘to fall’ used in conjunction with the noun ‘rain’:

\[
\text{(20) } \text{sômbò ná-n ã-gàp-şk éngîlì bûn dô-r’dì á wrà }
\]
\[
\text{Sombo GEN-3M 3M-arrive-PRF PREP.home rain 3M-fall PREP bush}
\]
\[
\text{à ká-n bà }
\]
\[
\text{PREP ON.3M NEG}
\]

‘As for Sombo, he arrived home, the rain did not fall on him in the bush’  
(Frajzyngier 2008: 478)

2.4 Total backgrounding of the entity: The “(it) rains” strategy

So far, backgrounding has been partial – either the entity of the process have been lexically and semantically backgrounded through the use of a more general, less specific lexical entry. But backgrounding can be pushed to the complete obliteration of either the entity or the process. Total backgrounding of the entity results of course in a subjectless predication, of the type “rains”. This model may be represented in the Horn of Africa by North Cushitic Beja:

\[
\text{(21) } \text{bireet-iya }
\]
\[
\text{rain-PST.3M}
\]

‘It rained’  
(Beja; Wedekind, Wedekind & Musa 2007: 164)

\[
\text{(22) } \text{bireet-iini }
\]
\[
\text{rain-PRS.3M}
\]

‘It is raining’  
(Beja; Wedekind, Wedekind & Musa 2007: 164)

2.5 Total backgrounding of the process: The “it is rain” strategy

Conversely, total backgrounding of the process entails the obliteration of the verbal character of the predication, yielding a structure of the “copula rain” strategy. Within the Afroasiatic languages, one finds the model in Kabyle Berber:

\[
\text{(23) } \text{d ageffur }
\]
\[
\text{COP rain.ABS}
\]

‘It is/was raining’  
(Kabyle)

This strategy is also at play for various other meteorological predicates in Kabyle (\textit{d azyal}, \text{COP heat.ABS, ‘it is hot’}, \textit{d asemmiḍ}, \text{COP cold.ABS ‘it is cold’}, \textit{d agu}, \text{COP fog.ABS ‘it is foggy’}, \textit{d ṭṭlam}, \text{COP darkness.ABS ‘it is night’, etc.). The copula is of deictic origin and is traditionally labelled ‘predicative particle’.
2.6 Modeling the backgrounding

As far as the general presence of backgrounding is concerned, we have therefore three levels:

<table>
<thead>
<tr>
<th>Zero Degree</th>
<th>(global apprehension, absence of backgrounding): “rain rains”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Backgrounding</td>
<td>total</td>
</tr>
<tr>
<td>entity</td>
<td>partial</td>
</tr>
<tr>
<td>“the sky rains”</td>
<td>“it rains”</td>
</tr>
<tr>
<td>process</td>
<td>partial</td>
</tr>
<tr>
<td>“the rain falls”</td>
<td>“it is rain”</td>
</tr>
</tbody>
</table>

| Full Degree | (total backgrounding of both entity and process): — (absence of speech) |

Figure 2. Backgrounding levels

3. Competing strategies and the role of theticity

The very fact that backgrounding (either partial or total) is central in meteorological predications of the ‘rain’ type poses the question of their status in terms of pragmatic organization: what are the consequences of the backgrounding of the entity or the process for the information structure of the predication?

The answer seems indeed to be that the topic-comment structure does not apply here. In other words, meteorological predications are instances of thetic sentences – sentences in which the bipartite organisation of the sentence into a presupposed and a non-presupposed portion is by definition absent: ‘The thetic statement forms a unit with respect to what it contributes to the discourse at a given point. It expresses a pragmatically unanalyzed state of affairs and presents it as a piece of complex information’ (Sasse 1987: 558); synthetically and more forcefully, Sasse (1995: 4) calls a thetic predication monomial. In Lambrecht’s (1994) theory of focus, in thetic sentences (in his terminology ‘sentence-focus’) no element is predictable or presupposed; rather, ‘both the subject and the predicate are in focus. The focus domain is the sentence, minus any topical non-subject arguments’ (Lambrecht 2000: 617).

We are now going to show the centrality of theticity in the coding of meteorological predications in Gawwada and Kabyle.

3.1 Gawwada

In Gawwada (another East Cushitic language of the Dullay group, and very close to Ts’amakko, exemplified above in 2.3), we find different strategies being used, on the basis of their pragmatic implications.
The “the rain falls” strategy is possible:

(24) IRR-IW-0  DIP-I
      RAIN-M  FALL-PFV.3M
   "The rain fell" > ‘it rained’  (Gawwada)

But the most frequent construction is rather of “the world rained” type, in which a noun (here IRR-IW-0 ‘rain’) is verbalized:

(25) PIY-E  I=IRR-IW-TI
      WORLD-F  SPEC=RAIN-PFV.3F
   “The world rained” > ‘it rained’  (Gawwada)

This construction is the only one in use for less dynamic atmospheric predications; in the following example, the noun KUYA-S-KO ‘day’ is verbalized:

(26) PIY-E  I=KUYAS-KI
      WORLD-F  SPEC=DAY-PFV.3F
   “The world became day” > ‘it dawned’  (Gawwada)

Denominal verbalization provides another simply strategy; in the following example the noun AWN-E ‘night’ is verbalized by a productive Ingressive extension (-uy-); syntactically the subject noun PIY-E ‘the world’ is here elided (and thereby the structure comes to resemble total backgrounding of the entity; cf. 2.4. above), but it still triggers agreement on the verb (which appears in the 3rd Feminine form):

(27) I=AWN-UY-TI
      SPEC=NIGHT-TEGR-PFV.3F
   “It nighted” > ‘it became night, night fell’  (Gawwada)

What are the structural features of those two constructions? While in (25) and (26) we find the by-now usual backgrounding of the entity though a generic noun, the structure exemplified in (24) is the one used to introduce new, and therefore non-topical subjects. The same applies to the Somali structure “the rain falls” shown above in (12): it is a structure usually called “subject focus” in Cushitic studies, characterized by the absence of subject case-marking on the subject noun and by the suspension of subject-verb agreement: the verb appears then in an invariable form (usually the third singular masculine, but in Somali a reduced agreement pattern applies, rather than the total suspension of agreement). Finally, the focus marker (baa in Somali) is found after the subject in its bare form without the subject clitics which normally appear before the verb (cf. Saeed 1999; cf. also Tosco 2003 for a text-based analysis of pragmatic marking in Somali). Paralleling ‘it rained’ above we’ll find therefore a sentence like the following, where a singular verbal form agrees with a subject plural noun:

(28) NIMAN  BAA  YIMI
      MEN  FOC  3M-COME.PFV
   ‘(Some) men came’  (Somali)
The same applies in Gawwada: here, no focus-marking particle exists, but the verb still is in an invariable third singular masculine form and no Subject Clitic is found. In the case of ꜰירrückw ’rain’, itself a Masculine noun, absence of the Subject Clitic only will mark the theticity of the sentence. Suspension of subject-verb agreement will be visible when a subject noun is feminine or plural, as in the following sentence, where the subject is a feminine noun; still the verb appears in the default third singular masculine form; there is no Specific subject-clitic in front of the verb, and a common (but not obligatory) left-dislocation of the adverbial of place ꜰ⇡ StringBuffer ‘in the calabash’ occurs:

(29) ꜰ⇡ StringBuffer ꜰราว-Francis Ꜳࡓ-Francis
     calabash-ASSOC.M beer-F be_there-IPFV.3M
     ‘There is beer in the calabash’

vs. the corresponding categorical, topic-comment sentence:

(29’) Ꜳࡓ-Francis ꜰ⇡ StringBuffer Ꜳࡓ=Francis Ꜳࡓ-Francis
     beer-F calabash-ASSOC.M spec=be_there-IPFV.3F
     ‘The beer is in the calabash’

3.2 Kabyle

Partially similar is the situation in Kabyle (Berber), where different strategies are used according to pragmatic implications.

In order to underline the proximity of meteorological predicates with other thetic predications, let us first have a look at presentative structures and sentence focus in general. Those constructions are characterized by a VS order (Mettouchi 2008) whenever a verb is present (30), and by the use of the accusative set of pronouns to refer to the main participant in the situation, with presentative non-verbal predicates (31).

(30) ye-FindObject=d wayzen
    sbj3msg-arrive.pfv=prox ogre.annx
    ‘The ogre arrived’

(31) ha-t wergaz-nni
    loc-acc3msg man.annx-anaph
    ‘Here comes the man’

The VS order is strictly respected for all meteorological predicates: the SV order, which characterizes in Kabyle the topic-comment format (Mettouchi 2008), is ungrammatical.

Example (32) involves a dynamic entity or process: the verb ‘to hit’ and the “the rain falls” strategy is used.

(32) te-kkat lehw
    sbj3fsg-hit.pfv rain.annx
    ‘It is/was raining’
The word order is VS, and the formulation is not acceptable if the word order is SV (unless we want to characterize rain: ‘rain falls, it is in the nature of rain to fall’):

(32') *lehma  te-kkat
    rain.ABS  SBJ3SG-hit.IPFV
    ‘It is/was raining’  (Kabyle)

Example (33) is of “the world rains” type. It is found with the expression of atmospheric predicates characterized by a strong experiential component. It also implies a VS word order, SV being ungrammatical and uninterpretable.

(33) ye-hma  lhal
    3MSG.SBJ-be_hot.IPFV situation.ANNX
    ‘It is hot’ (“the weather is hot”)  (Kabyle)

This structure is composed of a verb and its obligatory subject marker (here ye- for a Masculine Singular subject), and followed by a noun coreferential with that subject marker, lhal (itself a loan from Arabic al-ḥāl) with the meaning ‘state of affairs, situation, condition’. The verbal form ye-hma cannot be used by itself as an atmospheric predication: lhal must be coreferenced to the subject marker on the verb, and is obligatorily postverbal. Those verbs can nevertheless be used without a coreferential noun (as in 34) or with a noun referring to the entity whose temperature is evaluated (as in 34'):

(34) ye-hma
    3MSG.SBJ-be_hot.IPFV
    ‘It is hot’ (i.e. “Something (an object) is hot”)  (Kabyle)

(34') ye-hma  udajin
    3MSG.SBJ-be_hot.IPFV  tajine.ANNX
    ‘The tajine (cooking dish) is hot’  (Kabyle)

Those examples are important in that they underline the fact that person affixes are necessarily specific in Berber (cf. Mettouchi 2005), and that for a generic reading, such as the one needed for atmospheric predications (where the entity is difficult to delimit), a coreferential noun with vague reference is needed. We therefore have three elements here: the nature of the lexical noun (generic reference), its obligatoriness, and the VS word order. The first two components are needed to induce a non-specific interpretation of the person affix, the third one characterizes thetic predications (cf. Mettouchi 2006 and 2008).

This coreferential generic noun also appears in epistemic predications (for more details on modal predicates in Kabyle, see Mettouchi 2009), that is, when the speaker assesses a situation:

(35) y-uy  lhal  t-ruh
    SB13MSG-take.IPFV  situation.ANNX  SBJ3SG-go.IPFV
    ‘It happens that she left’  (Kabyle)
Here again, this coreferential noun always appears in postverbal position, which is the position for thetic predications in Kabyle. The absolute ungrammaticality of the SV order (topic-comment interpretation) is evidence in favor of the thetic interpretation of the predication.

We therefore have a configuration that is common to meteorological predicates, and some modal predicates, where the generic noun lḥal represents the entity, which is coindexed to a third person masculine singular that cannot in itself have a ‘dummy’ interpretation.

We have mentioned above (in 2.5.) another format for meteorological predicates in Kabyle: the non-verbal copular structure, in which the invariable copula d is followed by a noun in the Absolute state:

(37)  
\[ d \text{ azyal} \]  
\[ \text{COP heatwave.ABS} \]  
'It is/was very hot'  
(Kabyle)

This structure is also the one used for equational clauses (and clefts):

(38)  
\[ d \text{ argaz} \]  
\[ \text{COP man.ABS} \]  
'It/he is/was a man'  
(Kabyle)

(39)  
\[ d \text{ amellal} \]  
\[ \text{COP white.ABS} \]  
'It/he is/was white'  
(Kabyle)

(40)  
\[ d \text{ argaz i ye-wwet} \]  
\[ \text{COP man.ABS REL SBJ3MSG-hit.PFV} \]  
'It is a/the man that he hit'  
(Kabyle)

We find here another grouping, parallel to that between the ‘situation’ atmospheric predicates and the epistemic modal predicates: a grouping between the ‘deictic’ atmospheric predicates and the attributive predicates. The translations of (38) and (39) must not hide the fact that those predicates are strictly monomial, in that they consist of the qualifying noun and the copula. No clausal topic is expressed here. ‘COP + noun’ non-verbal clauses therefore are thetic predications.

We find similar situations in Somali, where the declarative marker waa behaves in a way reminiscent of the copula in Kabyle:

(41)  
\[ waa \text{ run} \]  
\[ \text{DECL truth} \]  
“It is truth” > ‘it is true’  
(Somali)
All those configurations are used not only for meteorological predications, but also for other structures that in other languages are called impersonal, and involve the use of empty pronouns for instance. In Afroasiatic languages, or at least in the languages described in this paper, impersonal strategies do not involve empty pronouns, because third person pronouns are generally specific and referential. Other strategies are used, which involve generic coreferential nouns, or non-verbal structures, both strategies being characterized by their thetic (in the sense of monomial) dimension, which we propose to consider as an essential component of at least some impersonal predications.

4. Conclusions

Atmospheric predications are the prototypical topos of impersonal predications: the difficulty to clearly separate the entity and the process strongly encourages backgrounding processes.

We have proposed to link the problematic separation of entity and process (at the cognitive level), to the thetic format as a monomial predication (at the linguistic level).

The partial survey we have conducted on Kabyle (Berber) and Gawwada (Cushitic), based mainly on raining and temperature predications, shows that when languages have special constructions for thetic predications, atmospheric predicates belong to that type. Moreover, in Kabyle atmospheric predicates share some features with epistemic modal predications, and with attributive ones (as in many languages).

The fact that attributive and atmospheric predicates do not involve dummy pronouns (as in English), but rather non-verbal predications, shows that the subject function is not centrally involved in impersonal constructions: rather, it is the backgrounding process that gives rise to constructions that may differ among languages. The ‘dummy pronoun’ strategy is only one among many other options.

We therefore propose to consider that the differences observed in impersonal structures are due to typological constraints (syntactic status of person markers, word order flexibility, presence of grammaticalized non-verbal predications, etc.): it seems to us that the category of the Impersonal only finds its unity at the semantic/cognitive level. This notwithstanding, some constant features can be found cross-linguistically.
References


Revisiting impersonal constructions in Modern Hebrew

Discourse-based perspectives*

Ruth A. Berman
Tel Aviv University

The study focuses on three impersonal constructions in Modern Hebrew: subjectless sentences with 3rd person plural main verbs, subjectless sentences with modal operators that take a complement clause, and sentences with generic pronoun subjects. Structural and semantic analyses elaborate on earlier studies in a discourse-embedded functional perspective based on authentic adult-child conversational interchanges and extended texts elicited from schoolchildren, adolescents, and adults in Hebrew and other languages. These serve to demonstrate the effects of such usage-based factors as genre, age-schooling development, as well as target language typology. The study concludes by arguing for a confluence of structural devices that combine to form a cline of impersonalization in the expression of a more or less depersonalized discourse stance.

Keywords: discourse; genre; Hebrew; impersonals; language development; subjectless

1. Introduction

The paper elaborates on prior studies of impersonal and related constructions in Israeli Hebrew in structural, functional, and discourse-based perspectives. Structuralist analysis of two subjectless predicate-initial impersonal constructions (with 3rd person plural verbs and with modal operators taking complement clauses), led to

* The author is grateful to Talmy Givón and Bracha Nir for helpful discussion of issues, to Ora Schwarzwald for aid with references, and to the editors of this volume and two anonymous reviewers for their perceptive comments on an earlier draft. Responsibility for remaining inadequacies rests with the author.
the characterization of Modern Hebrew as “an (S)VO language” (Berman 1980).¹ In another, functionally-oriented study, subjectless plural-verb impersonals were defined as higher in agentivity as compared with two other “agentless” constructions in Modern Hebrew – passive and middle-voice – in which verb-pattern morphology expresses alternations of transitivity and voice (Berman 1979). Developmental research on young children’s use of Hebrew-specific devices in conversational and oral narrative discourse showed that they tend to adopt an agent- rather than a patient-oriented perspective on events (Berman 1993a, b; Berman & Neeman 1994). More recently, generic use of a 2nd person pronoun subject in extended texts was analyzed for Hebrew and other languages as a means for expressing a relatively depersonalized discourse stance (Berman 2005).

The present study evolves out of and expands on this earlier work in the following directions. Structurally, concern is with three types of impersonal constructions in Modern Hebrew: subjectless sentences with plural verbs or taking modal operators plus complement clauses and sentences with a generic subject. Functionally, these constructions are analyzed as expressing a depersonalized “discourse stance”, defined as a pragmatic frame of reference for characterizing how people use language to position themselves with respect to a piece of discourse in a given set of circumstances (Berman, Ragnarsdóttir & Strömqvist 2002). In keeping with a usage-based view of linguistic analysis and language development (Bybee 2006; Tomasello 2003), occurrence of the target constructions is examined here in authentic, unedited language materials. The data-base is a large sample of extended written texts, both narrative and expository, elicited from schoolchildren, adolescents, and adults in Hebrew and other languages (Berman 2008; Berman & Verhoeven 2002) – hence concerned with “later language development” (Berman 2007; Tolchinsky 2004) – supplemented by longitudinal data from the naturalistic speech output of young pre-school children and their caretakers (CHILDES archive).

Following a brief outline of relevant features of Modern Hebrew (§2), the bulk of the study describes the structural and semantic properties and the discourse functions of the three target constructions – subjectless main clauses with plural verbs (§3.1), subjectless modal operators with complement clauses (3.2), and clauses with generic 2nd person subjects (3.3). Occurrence of these constructions in the data-base is then related to the variables of discourse genre (§4.1), developmental level (4.2), and target language (4.3), concluding with discussion of the “confluence of devices” involved in

---

¹ The label “Hebrew as an (S)VO language” was subsequently extended to subject-pronoun elision in Hebrew, analyzed as a partially pro-drop language, where verbs inflected for person serve the discourse function of topic maintenance in conversational interactions and oral narratives of young children compared with adults (Berman 1990).
expressing an impersonal discourse stance in Hebrew and suggestions for a “cline of depersonalization” (§5).

2. Relevant properties of Modern Hebrew

As background to the (re-)analysis of impersonal constructions considered in this chapter, this section briefly reviews relevant features of Modern Hebrew: Word order, Predicate-initial constructions, Verb-pattern morphology in voice and valence-changing operations, and grammatical Tense and Mood.

The basic word order of current Hebrew is SVO, hence subject-initial, with lexical or pronominal subject NPs and with lexical or copular predicates (Givón 1979, 1994; Schwarzwald 2001). The language also has a range of Predicate-initial constructions, some of which can be assigned a referential subject, even if not necessarily realized on the surface. Thus (i) VS order may alternate with SV(O) constructions, typically with change-of-state or so-called “unaccusative” verbs (Friedmann & Costa 2011; Levin & Rappaport Hovav 1995), and in presentative contexts in discourse (Berman & Neeman 1994). (ii) Existential and possessive constructions are basically predicate-initial, expressed in Hebrew as a non-habere language by a special existential operator yeš in present tense and the copular verb haya ‘be’ elsewhere; in possessives, this is followed by a dative-marked possessor and a subject-like possessee argument, e.g. yeš lanu harbe sfarim ‘be to-us many books = we have lots of books’ (Ravid 1977; Ziv 1976). And (iii) verbs inflected for person typically occur with a surface subject – a property that in Hebrew, unlike more typically “pro-drop” languages, is confined to 1st and 2nd person in Past and Future tense (Berman 1990).

Directly related to the theme of this volume is the existence of (iv) predicate-initial subjectless constructions, traditionally characterized as “incomplete” or “indefinite” (Gesenius 1910). From a “subject-centered” point of view, with “ impersonalization associated with the lack of a canonical subject” (Siewierska 2008), Hebrew impersonal, non-referential propositions include weather and other circumstantial or evaluative propositions, where a subject-requiring language like English or French would use an expletive or dummy subject. As in other languages, in Hebrew these are typically modified by temporal or locative expressions, e.g. xam (po) ha-yom ‘hot (here) today = it’s hot here/it’s hot today’, haya ná’im eclexem ‘was pleasant by-you = it was nice at your place’. And they include what Barðdal (2004) terms impersonal predicates that “select for oblique subjects” in Scandinavian languages. In Hebrew, where these involve an optional oblique argument, it is invariably confined to dative marking of

2. Unless otherwise noted, Hebrew data are presented in broad phonemic transcription, representing current pronunciation, with non-final word-stress marked by an accent aigu.
the experiencer role (cf. *xam li ha-yom* 'hot to-me today = I'm hot today'). This is true, too, of (v) evaluative-type predicates, e.g. *xaval li alav* 'pity to-me for-him = I'm sorry for him', *nimás lánú me-hitnahaguto* 'sick to-them from-behavior-his = they're fed-up with his behavior', *meša‘āmem la-yeladim ba-babáyit* 'boring to-the-kids at-home' (see, further, Berman 1981).

This last type of construction shares numerous structural and functional properties with predicate-initial modal operators plus complement clause analyzed in §3.2 below. Speakers also sometimes use the pronominal *ze* 'it, this' as a “dummy” or “expletive” subject. This usage reflects the fact that Modern Hebrew is susceptible to rapid processes of change, partly due to the special circumstances of its relatively recent revival as a medium of everyday spoken interchange (Berman 1997; Harshav 1993; Ravid 1995). Use of *ze* occurs even with the most canonic type of impersonals – (iv) above – as in the examples in (1) overheard on a bus in a conversation between two native Hebrew-speaking women.

(1) a. *ze meod gašum axšav ba-xuc*
   it very rainy now outside
   = 'It's raining very hard now outside'

   b. *ze haya nora cašuf šam*
   it was horrible crowded there
   = 'It was horribly crowded there'

This expletive pronoun also occurs with evaluative predicates of the kind noted here, as in the examples in (2) and (3) – of a woman addressing her two-year-old daughter from the longitudinal child language sample and from the written essays of two 4th-grade 9-year-olds respectively.

(2) a. *ze naxon še-ha-bad hu pepita*
   it right that the-cloth gingham
   = 'It's true that the cloth is gingham'

3. This pronoun has a special status in the language, marking it off from its 3rd person singular counterparts masculine *hu*, feminine *hi* 'he/she = it'. For example, (i) *ze* is not suffixed inflectionally to prepositions in non-nominative position, compare *hu* 'he' ~lo 'to-him, to-it', *alav* 'on-him, on-it' but *le-ze* 'to-it', *al ze* 'on it'; and (ii) it can function as a pro-copular with sentential subjects, e.g. *le’ašen ze asur* 'to-smoke it=is forbidden', *še-tavo ze ya’azor li* 'that-you'll come will-help me = it will help me if you come'. On the other hand, *ze* does not mark neuter as against masculine or feminine gender.

4. For example, while more typically verb-framed than verb-satellite (Slobin 2004; Talmy 1985, 2000). Hebrew manifests some features of what have been termed "equipollent" languages. For example, it allows the equivalents of "he ran into the house", but not "he swam across the lake", or "the bottle floated along the river" but not "the bottle floated down the river".
b. ze lo tov lehaxnis stam kol davar la-pe
   it not good to-insert just all thing to-the-mouth
   = ‘It’s not good to put just anything into your mouth’

(3) a. ani rō’e līf’amim yeladim še-megalim sodot leyad
   I see sometimes children that-reveal secrets next-to
   ha-xaverim šelahem
   the-friends of-them
   = ‘I sometimes see kids that tell secrets when (they are) near their friends’

ve-ze mē’dōd lo yafe legalot sodot ba-xevra
and-it very not nice to tell secrets in company
   = ‘and it is very rude to tell secrets in front of other people’

b. ze lo tov lā’asot dvarim kā’yle
   it not good to-do things like those
   = ‘It is not good to-do things like that’

Two other facets of the structure of Modern Hebrew that impinge on the topic of ‘impersonalization’ concern the domains of valency, voice, and tense. First, the binyan system of verb-morphology functions in valence-changing relations such as with intransitive reflexive or change-of-state predicates or transitive causative constructions, also serving to express a more or less agent-oriented perspective on events. Compare, for example, ha-xalon nīšbar ~ ha-yēled šavār et ha-xalon ‘the-window broke ~ the-boy acc broke the-window’ with the shared verb-root š-b-r in two different patterns; hu gīgel et ha-agala ba-midron ~ ha-agala hitgalgela ba-midron ‘he rolled the-cart down-the-slope ~ the cart rolled down-the-slope’ from the shared root g-l-g-l; ha-acic nafal ~ ha-xatul hipīl et ha-acic ‘the vase fell ~ the-cat caused-fall the-vase’ from the shared root n-p-l (Berman 1993a). The same system of seven morphological patterns also serves for alternations of voice between active, middle, and passive: Compare ha-mehandesim pītxu šīta xadaša ‘the-engineers developed (a) new method’ ~ šīta xadaša hitpatxa be-mēšex ha-šanim (a) new method developed over the years’ ~ šīta xadaša putxa (al ydey ha-mehandesim) ‘(a) new method was-developed (by the-engineers)’ – with all three constructions based on verbs with the shared root p-t-x in three different binyan patterns (Berman 1979).

Another relevant factor is that a relatively more or less impersonal point of view also interacts with predicates in use of different TAM (Tense-Aspect-Mood) categories (Bybee & Fleischman 1995; Hopper 1982; Timberlake 2005). In her analysis of discourse-based temporality in English, Hebrew, and Spanish, Kupersmitt (2006) points out that different degrees of generality/specificity of nominal reference interact with the predicate-oriented domain of TAM and Voice, such that past perfective tense/aspect is associated with more specific, hence immediately involved episodic information, whereas timeless or habitual present and/or use of hypothetical irrealis
mood typically reflects a generalized, detached impersonal stance. Relevant features of Modern Hebrew structure in this respect are, first, that it lacks grammaticized marking of Aspect, so that the benoni ‘intermediate’ form of verbs refers to both generic or extended present and to immediate or ongoing events; second, it lacks grammatical marking of subjunctive and conditional moods, so that future tense marking is extended to a range of irrealis categories, supplemented by the modal operators analyzed in §3.2 below; and, third, use of binyan verb morphology, as noted, makes it possible to express a less personally involved, non-agentive point of view by means of middle voice constructions as well as by syntactic passives.

3. Hebrew impersonal constructions

This section details three types of impersonal constructions, illustrated from adult-child conversational interchanges and extended texts written by schoolchildren, adolescents, and adults: Subjectless main clauses with 3rd person plural predications, subjectless clauses with modal operators and complement clauses, and clauses with generic 2nd person pronominals. Each section first outlines structural, morphosyntactic features of the relevant construction, followed by semantic analyses of type or degree of referentiality and discourse-based functional considerations of agency-defocusing and expressing a more or less depersonalized “discourse stance”.

3.1 Subjectless clauses with plural predicates

The first construction analyzed here is termed mišpat stami ‘(an) indefinite’ sentence’ in traditional school grammars listed in Schwarzwald (1978), who characterizes such sentences as “syntactically incomplete and semantically impersonal”. In the invented examples in (4), the verb (in bold type) is invariably in masculine plural form, as shown by the underlined suffixes, present tense -im, past tense -u.

(4) 3rd Masculine Plural Impersonals

a. šotim horbe mic be-Yisrael ba-kāyic
   drink+PLUR much juice in-Israel in-the-summer
   = ‘People drink lots ~ Lots of juice is/gets drunk in Israel …’

b. šatu lanu et kol ha-konyak ba-msiba
   drank+PLUR to-us ACC all the-cognac at-the-party
   = ‘They drank all our cognac ~ All the cognac got drunk at the party’

These sentences, as roughly translated into English by a generic subject like ‘they’, ‘people’ or by passive voice, were defined in my earlier study as “strictly subjectless” and non-referring, inter alia because they cannot have a pronominal referent as antecedent
Discourse-based perspectives on Hebrew impersonals

(Berman 1980). For example, (5a) below is ungrammatical, unless hem ‘they’ can be interpreted as referring to some specific people mentioned earlier, a sub-set of the universe implied by the impersonal predicate šatu ‘drank+++plur’. In contrast, in (5b), the (implicit) subject of the verb axlu ‘ate’ is necessarily co-referential with the overt pronominal subject hem ‘they’ of the verb šatu by equi-NP ellipsis.

(5) a. šatu lanu et kol ha-konyak ve hem kaasu drank+plur to-us ACC all the-cognac and they were-angry
   = ‘All the cognac got drunk, and they were angry’

   b. hem šatu lanu et kol ha-konyak aval axlu they drank+plur to-us ACC all the-cognac but ate+plur rak me’at only little
   = ‘They drank all the cognac but ate very little’

Third-person plural impersonals are common at different levels of Israeli Hebrew style, including journalese and prose fiction (Taube 2007). Their usage in colloquial Hebrew is attested to by caregiver input to young children in the examples in (6), from different Hebrew-speaking adults addressing their two-year-old children (from the Berman corpus on the CHILDES archive). Labels in brackets indicate the name and age of the child.

(6) a. Aunt: eyx kor’im la-báyit šel ha-kélev? [Leor 1;9]
   How call+plur the-house of the-dog
   = what’s it called?

   b. Fath: ma omrim? [Leor 1;9]
   What say+plur
   = What does one say? Child: Please
   Child: (be)vakaša
   ‘Please’.
   Fath: lo, omrim toda.
   No, say+plur thanks
   = No, you/we/people say thank you

   c. Moth: naxon, nitracim ha-yom. roxacim right, we-wash+fut today, wash+plur
   gam rosh also hair
   = ‘Yes, we’ll wash today. Washing hair, too’

   d. Fat: axšav lo ro’im televizya, axšav now not see+plur television, now
   holxim lišon go+plur to-sleep
   Now we ~ you’re not watching TV, now we ~ you’re going to sleep
Exchanges like these occurred in nearly every transcript of our extensive data-base of conversational interactions between adults and young children, together with predicates that involve personal reference – with an incorporated pronoun subject in (6c) *nitraxec ‘we-will-wash’, and with the pronominal subject *hem ‘they’ pronominalizing ‘dogs’ in (6e). They were also common in children's speech, as in the examples in (7).

(7) a. *kaxa olim al ha-géšer*  
   *so go+PLUR on the-bridge*  
   “That’s how we climb onto the bridge”

b. *ze lo taim ha-xol ... bol’im et ze,*  
   *it not tasty the-sand ... swallow+PLUR ACC it,*  
   *ve-moridim et ha-xol me-ha-roš*  
   *and take-down+PLUR ACC the-sand from-the-head*  
   “The sand doesn't taste good, (you ~ we) swallow it and take it off your head” [said when having her hair washed]

c. *im marbicim li ani marbic xazara*  
   *if hit+PLUR to-me I hit back*  
   = ‘If someone hits me, I hit back’

Semantically, these Hebrew constructions, while “strictly subjectless”, share many properties with 3rd plural constructions that take a third person plural pronominal like English ‘they’. That is, “from the semantic perspective [they] are constructions with a non-referential human subject which excludes the speaker and the addressee” (Siewierska & Papastathi 2008) and, as such, they are functionally akin to the class of what Malchukov and Ogawa (2008) define as “R-impersonals” that are triggered by lack of referentiality. However, such constructions, whether lacking a surface subject as in Hebrew (and also Greek, Hungarian, Italian, Polish, and Spanish in the sample analyzed by Siewierska and Papastathi) or requiring a pronominal subject (English, Dutch, French, German), do not, strictly speaking, lack any implication of agentivity. Rather, they invariably imply agency, to the class not merely of animates, but of human beings. For example, in the dialogic excerpt in (6e), the mother is telling her daughter
that she cannot take her dog along because they, kids and their parents, people in general, do not take a dog with them when they go to the store.\(^5\)

The scope of reference in 3rd person subjectless constructions is further restricted to people within the specific domain of discourse, either implied or explicit. Thus, in the invented examples in (4), drinking juice or cognac is constrained by location as attributable to any but only people residing in Israel in (4a) and by the event – to people present at the party hosted by the speaker in (4b). The same is true of the example from a little boy in (7d), whose comment on what they were told that day at nursery-school implies that he heard this from his teachers or other people associated with that context. Similarly restricted scope of reference of subjectless impersonals is illustrated by the excerpts in (8) and (9) from texts produced by schoolchildren and adolescents in the framework of a cross-linguistic project on developing text construction abilities, (as detailed in Berman 2008; Berman & Verhoeven 2002).\(^6\) Thus, the sphere of activity referred to in (8), from an essay on violence written by a 9-year-old girl, is clearly what takes place in school and reference is to her fellow-students – and while the subjectless impersonals in bold alternate with the overt subject pronoun \textit{anáxnu} ‘we’.

\begin{verbatim}
(8) lefi da’áti alimut ze ha-davar haxi nora še-yeš, še-ravim kol ha-zman,
    ve-še-marbicim ve-‘olakim ve-xuli. kol yom ro’im be-beyt ha-séfer béhayot še-
    marbicim ve-‘mekalelim ve-kol miney dvarim še-lefi da’áti xayavim lehipasek.
    im anáxnu rocim xayim tovim yoter, anáxnu xayavim lehafsik et ha-alimut
    hazot bimhera … [G13]

‘In my opinion violence is the worst thing there is, that 0 fight+plur all the time, and 0 hit+plur and 0 yell+plur and so on. Every day 0 see+plur at school problems that 0 hit+plur and 0 curse+plur and all kinds of things that in my opinion must stop+plur+pass. If we want a better life, we have to stop this violence right away …’
\end{verbatim}

These constructions also occur, less commonly, in the narrative texts elicited in the same project, where the same participants recounted an incident where they had experienced interpersonal conflict, as the excerpts in (9).

\(^{5}\) Thus, a sentence like \textit{yešenim} harbe ba-xóref ‘sleep+plur much in-winter’ could not refer to bears or other such creatures, but means something like \textit{people sleep a lot in the winter}; and the statement \textit{xayavim lehagiš et ha-bexinot bi-zman} ‘must+ masc-plur submit acc the-tests in-time’ is acceptable in the context of an all-girls’ school, in the sense of ‘You = all girls must hand in their tests in time ~ tests must be submitted on schedule’.

\(^{6}\) Participants are identified by age-group and serial number out of 20 per age in square brackets: \textit{G} stands for grade school children aged 9 to 10 years, \textit{J} for junior high students aged 12 to 13 years, \textit{H} for high school students aged 16 to 17 years, \textit{A} for adult graduate university students in their 20s and 30s. Impersonal predicates are bolded, with the plural suffixes underlined present tense \textit{-im} and past tense \textit{-u}.
The excerpts in (9), from stories written by three boys at different levels of age-schooling, occur in the typically episodic and specific rather than generalized style of discourse of a personal-experience narrative. Like those from the more obviously impersonal and more distanced context of expository discourse in (8), they bear all the hallmarks of the impersonal constructions at issue here: lacking an overt grammatical subject, taking 3rd person masculine plural predicates, and non-personal in reference. Yet their scope of reference is restricted, even if implicitly: in (9b) to the other kids playing soccer with the narrator, specifically those on the other team, and in (9c), the ill-treatment the narrator suffered at the hands of others in his new school, specifically his classmates.

The same also holds for the examples in (10) from children's naturalistic speech output: (10a) a little boy telling his aunt what happened at his nursery school that day, (10b) a little girl telling her parents why she and her friends got yelled at; and (10c) from peer-interaction of a girl at kindergarten.
Thus, while in Hebrew, as in many typologically unrelated languages, 3rd plural subjectless impersonal constructions do not explicitly specify an agent, the range of their agentive reference is often defined by the specific locative and/or temporal context under discussion – here, at nursery school that same day or at a picnic whose whereabouts and participants are (assumed) known to the addressees.

From the point of view of discourse function, we take as a point of departure Sierwierska’s (2008) introductory comment to the effect that “The notion of impersonality is a broad and disparate one” to argue that an impersonal “discourse stances” serves the two related purposes of agency downgrading and of generalizing about a habitual state of affairs. Sierwierska further points out that “From the structural point of view, impersonalization is associated with the lack of a canonical subject, from the functional perspective with agent defocusing … in the sense of diminishing the prominence or salience from what is assumed to be the norm or, in the terminology of Langacker (1991), archetype.” This view is consistent with what Givón (1994) refers to as “suppression of agency” in discussing voice and inversion, and it relates to general processes of agency alternation, defined as “patterns in the use of grammatical constructions that express differing amounts of involvement of the agents (or causes) of the states, activities, or events referred to … in the course of ongoing text production” (Tolchinsky & Rosado 2005).

In her comparison of 3rd plural impersonals and passives in Modern Hebrew, Taube (2007) characterizes the former as “actional in nature”: It “expresses agentivity and marks the actualization of the event”, in contrast to passive voice, which she describes as being “unmarked as concerns actionality and thus allows for focusing on the state of the undergoer” (2007: 282). My earlier analysis of Hebrew described suppression of agency as ranging on a descending cline from (3rd person plural) active to passive and middle voice respectively (Berman 1979). These contrasts are elaborated on below (§3.3) in comparing means used by speaker-writers of different languages for expressing a more or less depersonalized discourse stance.

The scope of reference of 3rd plural impersonal constructions also interacts with the domain of Tense/Aspect. Thus, use of an imperfective or extended present form of the verb – as in examples (6) through (8), with the present-tense plural

---

7. This example is from the Blum-Kulka (2009) corpus of conversational peer-interaction.
suffix -im – lends the proposition as a whole a more habitually generic reading. In contrast, use of a perfective or past tense verb – as in the invented sentences in (4b) and the narrative excerpts in (9b), (9c), and (10) – constrains the scope of the proposition to a more episodic time-frame, hence to a more specific, less generalized frame of reference. The preference for present-tense predicates with these constructions is also noted by Taube (2007) who points out that “Many of the 3rd masculine impersonals in our corpus [of newspaper and prose fiction – RAB] are in the present participle form, and they usually describe a rule, a custom, a routine” (2007: 289). I propose that these represent the “default” or least marked instance of the construction in question. For example, when speakers are asked to provide a so-called mişpat stami ‘indefinite sentence’, they invariably use the (timeless or generic) present tense, although past/perfective and future/irrealis constructions would be grammatically well-formed and semantically plausible. This indicates that a key function of subjectless plural-verb constructions is to express generalized propositions about a regular or habitual state of affairs, combining subjectlessness for non-specified agency and extended present tense for non-specified temporality.

3.2 Subjectless constructions with modal operators

A second group of subjectless impersonal constructions take the form of a modal operator followed by an infinitival or tensed ‘that’ clause as complement, as illustrated in (11) from the longitudinal child-language sample and Blum-Kulka (11d), and in (12) from the written texts of schoolchildren and adolescents.

(11) a. oy, ze meluxlax nora. carix lizrok et ze la-kvisa it to-the-laundry
   Oh, it dirty terrible, must to-throw ACC
   ‘It’s terribly dirty, it has to go into/we have to throw it into the wash’

b. at yodā'at im sear kacar i efšar
   You+fem know+fem with hair short not possible
   la’asot cannot
   to-make braids
   ‘You know you can’t make braids with short hair’

c. rak ba-madregot efšar laredet
   only in-the-stairs possible to-descend
   ‘It’s only possible/you can only go down by the stairway’

d. kodem tikre’u la-xaverim. carix
   first call+IMP, PLUR to-the-friends, must
   likro la-xaverim
   to-call to the friends
   ‘First call the other kids, we/you have to call the other kids’
The examples in (11) illustrate the two modal operators that were commonest in our data-base, carix ‘must, have to’ and (iy)efšar ‘(im)possible = can(t)’, followed by a verb in the infinitive. Those in (12) are rather more sophisticated and less everyday, taken from expository essays written on the topic of interpersonal conflict by students from grade- to high-school age, and they include a finite tensed ‘that’ clause in (12d).

(12)  

a. ma carix la‘asot kdey lo lehagia le-makot ve-le-klalot? [G-12]  
what must to-do for not to-reach to-blows and-to-curses  
‘What must we do/what must be done so as not to get to blows and curses?’

b. ze kmo maxala še-nidbeket la-or ve-i-efšar  
it like disease that sticks to-the-skin and not possible  
lehištaxrer mimena to-get free from it  
‘It’s like a disease that sticks to the skin and that you can’t get rid of’

c. et ha-bé’ayot efšar liftor be-štay draxim [J-07]  
acc the-problems possible to-solve in two ways  
‘It is possible to solve problems in two ways, problems can be solved in two ways’

d. yasol liheyot še-meuxar miday lexanex yeladim [H-01]  
can to-be that late too to-educated children  
‘It could/might be too late for educating kids’

In morphological form, these modal operators in subjectless impersonal constructions are invariant (3rd person masculine singular), often in the shape of homonymous verbs, adjectives, or adverbs. Thus carix ‘must, have to’ in (11a) is a verbal form – that also has (suppletive) past and future tense forms. Unlike most of these modal operators, carix can also occur with what Heine (1995) describes as “agent-oriented subjects with specific reference” – in which case it agrees with a surface subject in number and gender, and takes an infinitival complement. For example, Assaf, aged 3;3 says ani carix liknot li sus ‘I must to-buy to-me horse = I must buy myself a horse; his sister; Sivan [4;4] says ha-géšem carix lehaškot et ha-ecim ‘The rain must to-water acc the-trees’, and [age 4;7] bney-adam lo crixim kol ha-zman laāvod ‘people not must+PLUR all the-time to-work = people shouldn’t have to work all the time’). In contrast, (i) efšar ‘(im)possible’ is an adverbial that alternates with the adjective efšari – which inflects for number and gender. Two other terms commonly used where English might have an agent-oriented modal auxiliary are also suppletive, adverb-like elements: yitaxen ‘likely’ (e.g. yitaxen še-hu codek ‘likely that-he right = he may be right’) and kday ‘be worthwhile’ (e.g. kday še-nemaher ~ lanu lemaher ‘worthwhile that-we’ll-hurry ~ to-us to-hurry = we should hurry’).

In syntactic structure, these invariant subjectless modal constructions, as noted, can take infinitival and tensed ‘that’ clause complements, with the two sometimes in competition (for example, kday ‘(be) worthwhile’ in the sense of ‘should’). This
alternation is lexically restricted: Some modal operators take only infinitival complements (e.g. high register use of the existential marker yeš ‘be’ in the sense of ‘(it is) necessary’ or the preposition al + N ‘on + N’ in the sense of ‘(it is) incumbent upon’) while others take only tensed še- ‘that’ clauses (including yaxol liheyot ‘can to-be, maybe, is possible’ in (12c) or yitaxen ‘(it is) likely’). The infinitival complement appears more basic, binding the modal operator to the following proposition more tightly than its tensed counterpart, in a construction traditionally termed nasu murxav ‘extended predicate’, rather than constituting a separate clause. Sharvit’s (1997) analysis of modal expressions in Mishnaic Hebrew is confined to infinitival constructions, but further analysis is required based on contemporary discourse usage.

As subjectless predicate-initial operators with clausal complements, modals share the syntactic structure of evaluative predicates like boring, annoying noted in §1 above. This also applies to impersonal passives – for example, besofo šel davar huskam še-tov laxtof anašim kmo X ‘eventually (it was) agreed that (it is) good to-kidnap people like X’ (from an article in Ha-arets, 3.10.2009) – where the complement of passive huskam ‘was agreed’ is a subjectless evaluative clause with tov ‘good’. Taube (2007) characterizes the Hebrew impersonal passive as “an invariable verbal form … [that] is, however, rather infrequent, occurring in fixed formulas” (p. 278). The fact that they are common, as Taube notes, in newspaper writing and legalese yet failed to occur in the data-base of the present study, suggests that (non-formulaic) passive impersonals represent high-register, more formal alternatives to their active-voice plural-verb counterparts. That is, in everyday colloquial usage, Taube’s example of huxlat lidxot et ha-mišpat ‘(It) was-decided to-postpone the-trial’ could be replaced by hixlitu lidxot et ha-mišpat ‘decided+plur to-postpone the-trial = They decided …’

In case-marking assignment of role-properties (Malchukov & Ogawa 2008), Hebrew modal operators lie somewhere between evaluative and impersonal passive constructions. Unlike evaluatives (and like impersonal passives), modal operators typically fail to assign an oblique case role in the form of a dative experiencer. The adverbial kday ‘worth(while), pays, beneficial’ is an exception, since it can occur both with and without a dative argument (e.g. kday lexa ~ lanu la’asot zot ‘(it is) worthwhile for-you ~ for-us to-do that’ in the sense of ‘You ~ We should do it’) – suggesting that it is semantically closer to an evaluative than to a modal operator, hence less strictly impersonal. That is, lack of even oblique argument assignment underscores the impersonal nature of modal operators.

Not only do the basic modal operators – carix ‘must, have to’ and efšar ‘possible’ – not allow a dative experiencer, they rarely if ever take a non-normative expletive ze ‘it’, illustrated with evaluative predicates in (1) to (3) above. These two operators, which occur relatively early in child speech, reflect the basic semantics of irrealis modality: deontic necessity and epistemic possibility respectively (Bybee & Fleishman 1995; Reilly et al. 2002). On the other hand, these semantic dimensions also reflect
the “mixed” nature of contemporary Hebrew, as follows: Compared to their Spanish counterparts *deber, poder* (Silva-Corvalán 1995), only *carix* ‘must, have to’ occurs in both subjectless and “agent-oriented” constructions. Basic ability as well as possibility are expressed either by the sentence-initial operator *efšar* ‘be possible’ or by the verb *yaxol* ‘can, be-able’: In subjectless constructions, *yaxol* is invariant and requires a copular complement *liheyot* ‘to-be’ as in (12d) above. More typically, it occurs in agentive constructions, agreeing morphologically with the surface subject (e.g. *hu yaxol la’azor* ‘he can to-help’, *ani yaxóltí la’azor* ‘I could+1st to-help = I was able to help’, *hem yaxlu la’azor* ‘they were-able to-help’). And it can also occur in plural-verb subjectless impersonals: for example, Shachar, aged 3;9, says of somebody on the telephone *hu haya medaber, yexolím lišmoa oto* ‘he was talking = used to talk, can+plur to-hear him = we are able to hear him’.

Modal operators in Hebrew are thus a mixed group of items, syntactically, morphologically, and lexically. As such, they differ markedly from the grammatically distinct set of “agent-oriented” modal auxiliaries *can, must, should* etc. or their semi-modal alternatives *be able to, have to, ought to* respectively of English. In terms of discourse function, constructions with modal operators were earlier analyzed as encoding irrealis “propositional attitudes” (Reilly et al. 2002), that range from subjectively personal affective, via more generalized deontic to quite distanced and detached epistemic attitudes. This interacts with target language morphology, such that sentence-initial operators may express a relatively depersonalized, less involved or self-committing discourse stance compared with their agent-oriented alternatives in Hebrew or the modal auxiliaries of English. This is shown by comparing the same kind of judgmentally prescriptive attitudes as expressed in different languages by 9-year-old children asked to discuss the issue of “problems between people”: English – *When you fight, you can hurt the person’s feelings … so you should always be nice and respectful …*; French – *Quand on se fait racketer, c’est une exception. Il faut s’éloigner le plus possible de la personne*; Spanish – *Hay muchos niños que se fijan en los demás y tienen que mejorar su actitud*; Hebrew – *asur la-morim leharbic la-talmidim, carix la’asot hakol kdey lo leha-gia le-makot* ‘forbidden to-teachers to-hit to-students, must to-do everything in-order not to-reach to-blows = teachers mustn’t ~ aren’t allowed to hit students, people ~ we must do everything not to get to the point of blows’.

As an interim summary, the examples in (13) – from the expository essays written by schoolchildren and high-school students on the topic of interpersonal conflict – illustrate both types of subjectless impersonal constructions, with 3rd plural verbs and with modal operators followed by a complement clause occurring in the same segment of discourse.

(13) a. *im ró’m xaver bi-šéát cara carix la’azor lo* [G-18]  
   if *see+plur friend in-time trouble must to-help him*  
   ‘If you see a friend in trouble, (you) should help him’
b. mi-beayot londim hamon gam bizman ha-beayot, from problems learn+PLUR both at-time+GEN the-problems, ve gam eyx še-potrim otam. efsar lilmod and also how that solve+PLUR them. Possible to-learn la-xayim ve-la-pa’am ha-ba’à for-life and for-next time ‘You learn a lot from problems, both at the time and also how people solve them. You can learn for life and for next time’

c. ani xoševet še-lifney še-doršim šivyon bi-zxuyot, I think that-before demand+PLUR equality in-rights, carix še-yiheye kodem kol shivyon xovot must that-will-be first-of all equality-of duties ‘I think that before demanding ~ we demand ~ one demands equal rights, (it is) necessary that ~ there must be ~ we need to have equal obligations’

d. yaxol liheyot še-lo meyaxasim le-xax maspik xašivut can to-be that-not relate+PLUR to-it enough importance = ‘It could be that not enough importance is attributed to it ~ People might not relate enough importance to it’

It is no chance that these excerpts, combining the two structurally and semantically distinct subjectless constructions, are from non-narrative, expository-type texts. As discussed below (§3.1), this discourse genre is particularly suited to the generalized, non-specific, non-agentive, less involved discourse stance embodied by subjectless impersonals.

3.3 (Pro)nominal expression of generic reference

Hebrew speakers can also express a depersonalized discourse stance by using generic categorial terms such as people, a person, like those mentioned by Shemesh (2009) as functioning as impersonal subjects in Mishnaic Hebrew: adam ‘person’, iš ‘man, iš exad ‘man one = a man’, iš plóni ‘some man’. As an example from our data-base, the excerpt in (14) is from a talk given by a woman [A-01] on the topic of interpersonal violence, using the category label ben-adam ‘son+GEN Adam = (a) person’, pronominalized by nominative hu ‘he’ or by suffixal -o in subsequent mentions.

(14) ha-ben-adam carix laxšov ba-roš šelo ma ha-plúsim ve-ma ha-minusim, hu carix liheyot meyuman ba-maxšava ma haxi tov bišvilo, im ma hu yiheye haxi šalem ‘The man [= a person] must think in his mind what are the pluses and minuses, he must be skilled in thinking what is best for-him, what he can feel most at one with’
Such usages occurred mainly in more formal essay-type contexts (see §3.1), mainly by older speaker-writers in our study. Far commoner was reliance on the 2nd person masculine singular pronoun – nominative ata or suffixal -xa – analogously to use of Dutch je, English you, or Spanish tú in comparable samples – as analyzed for these languages in parallel data bases by van Hell, Verhoeven, Tak & Oosterhout (2005); Reilly, Zamora, & McGivern (2005); and Tolchinsky & Rosado (2005), respectively. In Hebrew, generic ata is invariable, whereas when used as a personal pronoun it contrasts with feminine singular at and with plural atem. Generic use of 2nd person pronouns in Hebrew is illustrated by the bolded forms in excerpts from oral expository texts of a 9-year-old girl in (15a) and a 12-year-old boy in (15b): The 2nd person pronoun takes the free nominative form ata ‘you’ in subject position, alternating elsewhere with the oblique form -xa suffixed to a preposition.

(15) a. nagid mišehu ba laševet leyad-xa, ve-hu mag’il ot-xa, az lehagid lo, še-ata lo roce liyot xaver šelo
   ‘(let’s) say somebody comes to-sit next-to-you, and he disgusts acc+you, so tell him that you do not want to-be friends with-him’

b. lamrot še-ani xošev še-ze lo mašehu še-carix la’asot,
   im ata lo yodea šešela axat …
   ‘Even though I think it [= cribbing] is not something that should to-do [=that should be done, that people should do], if you don’t know one question ….’

c. ze taluy be-eyx še-ata mityaxes la-inyan, ani lo yodea, tni
   li beaya še-ani agdir lax
   ‘It depends how you relate to-the-matter, I don’t know, give +imper, 2nd fem me (a) problem that I’ll-define for-you+2nd fem’

These examples underscore the generic sense of 2nd person ata, since in (15a), it co-occurs with the indefinite pronoun mišehu ‘someone, somebody’ (literally ‘who-that-hi’, elided to mišu in current speech), pronominalized by hu ‘he’ in the same sentence, as a colloquial variant of the more formal generic noun ben-adam ‘(a) person’ in the adult’s text in (14). In (15b) the 2nd person pronoun occurs in the same context as the indefinite pronoun mašehu ‘something’ and, moreover, contrasts overtly when used generically versus the same pronoun used in a deictic, personal sense in the feminine forms of the imperative verb tni ‘give’ and suffixal -ax ‘for you’ (cf. masculine lex). When the boy addresses his teacher, requesting her guidance on how to discuss the topic of “problems between people”.

While clearly used in a generic sense, contrasting with specific personal reference in these examples, Hebrew ata is typical of more colloquial, everyday use in contrast to higher-register 3rd person generic subjects as in (14). 2nd person pronouns evidently express a relatively interactive orientation, even when used generically. And, indeed,
across the Hebrew-language sample, these forms occurred almost exclusively in the **spoken texts** (see §3.1), reflecting a relatively informal, personalized type of discourse stance compared with the more distanced and detached written language (Berman & Ravid 2009).

4. **Comparative trends**

A usage-based perspective implies that how different constructions are deployed for expressing an impersonal stance on events will be affected by factors such as communicative context, level of literacy, and target language typology. Below, observations emerging from the Hebrew-language data-base used for this study are reviewed and integrated with findings from studies on comparable samples in other languages – to compare occurrence of the three constructions described above in relation to the variables of genre (4.1), development (4.2), and language (4.3).

4.1 **Type of discourse**

Research shows children to be genre-sensitive from an early age. For example, young preschoolers can distinguish between scripts and personal-experience narratives (Hudson & Shapiro 1991) or between fictional narrative and description (Sandbank 2002), while 9-year-old speaker-writers distinguish clearly between the linguistic means they use in their expository essays compared with personal-experience narratives (Berman & Nir-Sagiv 2007).

Discourse-based comparisons indicate that use of the three target constructions described above for making non-agentive, impersonal, or generic reference depends critically on communicative context. **Subjectless Impersonals**, as illustrated earlier, are common in adult-child interactive input and output from a very early age, typically to refer to generic states of affairs in extended present. Subsequently, they are extended to reporting on past events in restricted contexts such as school or the playground. Moreover, in a large sample of written texts produced by grade-school, middle-school, and high-school students (N=36 in each group), they were used significantly more, and by more respondents, in the expository essays compared with the personal experience narratives elicited from the same participants on the shared topics of violence in schools or interpersonal conflicts (Berman 2003; Berman & Nir-Sagiv 2004). Analogous findings emerged for use of **Subjectless Modal** constructions: These are documented in Hebrew adult-child conversational interactions from pre-school age, they are typical of expository essays rather than personal-experience written narratives in different languages (Reilly et al. 2002), and they occur at least once in half the expository essays of school-age students, but in fewer than 10% of the stories that they wrote (Berman 2003). Use of a **2nd Person Generic Pronoun** interacts with the factors of modality and register as well as with discourse genre: Across the data-base
of 320 Hebrew extended texts, *ata* as a generic subject was used at least once, often several times, by one-fifth [16 out of 80] of the respondents in the four age-groups – but exclusively in spoken and not written texts, typically in the expository genre, although it can be found occasionally in oral peer interactions from late pre-school age (in the data-base of Blum-Kulka, 2009). Further, while Hebrew speakers can make use of lexical nouns for generic reference – among older speaker-writers in the form of a high-register term like *ben-adam* '(a) person' (as in (14) above) and among children, by more colloquial, plural generics like *anašim* 'people', *yeladim* 'children', these are less widely used than subjectless constructions (Ravid et al. 2002). The latter emerge as a favored means for expressing an impersonal, non-agentive stance in Hebrew from early on in development.

Favoring of impersonal constructions in expository prose compared with personal-experience narratives can quite obviously be explained as a general feature of this type of discourse. And it confirms findings for use of diverse linguistic forms – including personal versus generic and expletive pronouns, middle and passive voice, modal expressions – in a comparable sample of texts in different languages revealing clear and consistent differences between the direct, immediate, subjectively personalized perspective of personal-experience narratives and the more detached, distanced, abstract, and impersonal style of expository discourse (Berman 2005) Expository discourse is a natural site for the expression of a depersonalized discourse stance, hence for use of impersonal construction, since it is typically topic-oriented, with a focus on concepts and ideas, whereas narrative discourse is concerned mainly with people, their actions and motivations, hence more personalized and subjective (Britton 1994; Longacre 1996). Comparisons of expository discussions with personal-experience narratives in both speech and writing in our data-base revealed an intersection of genre and modality ranging along a clear language-independent hierarchy – with oral narratives at one end and expository essays at the other – in domains of linguistic register, agent downgrading, and overall discourse stance (Nir-Sagiv, Bar-Ilan & Berman 2008 – for English; Ravid & Berman 2009 – for Hebrew; Tolchinsky & Rosado 2005 – for Spanish). These convergent findings point to a broad, *genre-related continuum of impersonalization*, extending out to interactive conversation at one end, via personal experience and fictive narratives to informative texts, expository discussions, and research papers, at the other.

### 4.2 Age-schooling related factors

Age-related developments in the acquisition and use of the three target constructions indicate that all three appear before school-age. As noted, Subjectless Impersonals are common in adult input even to 2-year-olds, and in child speech from as young as age 3 years; Modals occur early in Child Directed Speech, but children use them with clausal complements only from age 4 to 5 years – as part of the development of
complex syntax; while generic *ata* as a depersonalized mode of reference appears at late pre-school age. The extended texts in our data-base show two major developments in expression of a detached, depersonalized stance, as in other areas of language acquisition and use (Berman & Slobin 1994): (1) a broadening of the range of expressive options and (2) a change in the content of these different devices.

With respect to the broader range of devices deployed with age and schooling, in a recent study comparing development of a “depersonalized stance” in French and Spanish, Jisa and Tolchinsky (2009) point to an “increase in the diversity [emphasis mine, RAB] of linguistic means” appropriate to expository discourse. In general, with age and increased literacy, related research on different languages (in the Berman 2005 special issue) reveals a general shift to more formal, often more typically written styles of expression. These include greater reliance on passive voice and a move from *get* to *be* passives in English; use of *se* middles and passive voice rather than generic subjects like *on*, *nous*, *tout le monde* in French; infrequent use of generic *maður* in Icelandic by mature participants compared with their younger counterparts; decrease in reliance on 2nd person pronouns used generically in favor of other, more formal devices such as the pronoun *men* in Dutch; and greater reliance on *se* impersonals, middles, and passives in Spanish. A key development in this respect in Hebrew not previously noted is increased use of intransitive middle-voice *binyan* verb-pattern morphology to express a non-agentive perspective on events. The *bolded* verbs in the examples in (16) are in the *hitpael* verb-pattern typical of change-of-state achievement predicates, those in (17) are in the *nif'al* pattern, commonly used for adjectival passives or so-called unaccusative predicates (Berman 1993a; Berman & Neeman 1994).

(16) a. *alimut ze davar še-mitgaber im ha zaman* [J-03 expos]

violence it thing that *gets-bigger* with the-time

= ‘Violence is something that increases with time’

b. *ve-az hitpateax mi-ze riv* [H-03 narr]

and then *developed* from-it quarrel

= ‘And then there developed a quarrel out of that’

c. *ze taluy bi-yexolto šel ha-oved le-hištalev ba-cévet* [A-19 expos]

it depends in-ability-his of the-worker *to-integrate* in-the-team

= ‘It depends on the worker’s ability to become part of the group’

(17) a. *beayot ben anašim ze kmo maxala še-nidbéket la’or* [G-07 expos]

problems between people it like disease *that-sticks* to-the-skin

= ‘Problems between people are like a disease that infects the skin’
Use of the two typically intransitive verb-patterns hitpael and nif'al out of the 7 binyan forms in the Hebrew sample of 320 written and spoken expository and narrative texts increased consistently with age, as follows: from 6.8% of all (lexical) verbs used by 4th-grade 9-to-10-year-olds to 14% in 7th grade, to nearly 20% in high school; among adults, 17% intransitive pattern usage was supplemented by nearly 5% use of the two passive voice patterns – rare until high school age, when they account for only 2.0% of all verbs. This age-related increase in intransitive verb-forms reflects developing ability to adopt a patient–rather than an agent-oriented perspective on events, as shown earlier for children’s oral picture-book narratives in Hebrew and other languages (Berman 1993b; Berman & Slobin 1994).

Increased age-related diversity in expression of an impersonal discourse stance is reflected not only between alternating systems (say active versus middle or passive voice), but also by expansion within a given linguistic system. This was noted above in the extension of generic present tense use of subjectless plural-verb impersonals to more temporally specified past (and also future) tense, and it is observed in a shift from use of colloquial-style generic lexical subjects like yeladim ‘children’, anašim ‘people’ to higher-register terms like ben-adam ‘(a) person’ (Berman & Ravid 2009). A particularly dramatic development in this respect, documented for English and French as well as Hebrew, was a significant age-related shift in the type of propositional attitudes expressed by modal terms (Reilly et al. 2002). Nine- to 12-year-olds use modal terms mainly to express prescriptively normative, socially dictated attitudes to the topic of “problems between people” – in Hebrew, by means of operators such as carix ‘must, have to’, lo carix ‘shouldn’t, ought not’, often by (early acquired) modals of prohibition or allowance – asur ~ mutar ‘forbidden = you can’t, mustn’t’ ~ ‘allowed ‘you can, may’ respectively. In contrast, older speaker-writers resort to epistemic, cognitively motivated modal expressions, referring to as possible or probable contingencies as eventu-

alities that can, may, or will arise out of interpersonal conflict – in terms such as yaxol liheyot ‘can to-be = it’s possible’, yitaxen ‘likely’. This development in use of subjectless modals as an impersonal construction is consistent with other studies that show age-related changes not only in linguistic structure but also in the thematic content and pragmatic perspectives expressed by older, more literate, even if non-expert speaker-writers in discussing their ideas and relating their experiences (see, for example, for Hebrew, Ravid 2006; Ravid & Cahana-Amitay 2005).
Research into developing discourse abilities from preschool age to adolescence and beyond suggests two related explanations for these developments. The first is the universal progression where, starting out as “child speakers”, language learners early on move to becoming “native speakers”, but only much later, with increased social-cognitive and literacy-related development, advance to the level of “proficient speaker/writers” of a given native language (Slobin 1997, 2001). The second, related factor is that of “later language development”: It takes until adolescence or beyond for speaker-writers to evolve the cognitive skills and communicative sophistication necessary for deploying a fully developed repertoire of linguistic forms flexibly and appropriately in varied discursive contexts (Berman 2007; in press; Ravid & Tolchinsky 2002; Tolchinsky 2004). This kind of alternation is illustrated by the bolded impersonal, passive, and modal forms in the short excerpt in (18), from the of a high-school boy [H-10].

(18) aval ba-zman ha-axaron menasim+PLUR lehaxdir la-yeladim but in-the-time last try+PLUR to-introduce to-the-kids
(gila‘ey arba-esre šloš-esere) et ha-nose šel soylanut (ages+GEN thirteen fourteen) ACC the topic of tolerance
ha-nose muxdar be-emcaut ha-tiksőret ...
the-topic introduced+PASS in-medium+GEN the-media
aval be-gila‘im yoter gvohim crixim ladun be-ze but in-ages more high+PLUR must+PLUR to-discuss in-it
be-kvucot lefi bxirat ha-ne‘arim in-groups by choice+GEN the lads

= ‘But lately (they’ve been) trying to introduce the topic of tolerance to kids aged 13 or 14. The topic is introduced through the media, but in older age-groups, (people) should discuss it (it should be discussed) in groups of the young people’s choosing’

As I have argued elsewhere for other facets of Hebrew-language development, children are from early on familiar with most of the forms available in their language – in the case in point, plural impersonals, modals plus complements, generic pronouns, passive voice, etc. (Berman 2004). But it takes many years, often to high school age, before speaker-writers can extend and integrate use of diverse forms to meet the function of expressing an impersonal stance in a stylistically felicitous and communicatively appropriate fashion.

4.3 Cross-linguistic factors in selecting means for depersonalization

The free translations into English of the Hebrew examples provided in this chapter, including for the excerpt in (18), where the English version supplies surface subjects and extends use of passive voice, are indicative of critical cross-linguistic differences
in expression of an impersonal stance. This is highlighted by the different options selected for expressing the same content in English, French, and Hebrew in (19), with subject (pro)nouns in italics and verbs underlined.

(19)  

<table>
<thead>
<tr>
<th>Passive Patient Subject:</th>
<th>English spoken here</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generic Pronoun Subject:</td>
<td>İci on parle français</td>
</tr>
<tr>
<td>Subjectless 3rd Plural:</td>
<td>0 kan medabrim ivrit</td>
</tr>
</tbody>
</table>

These alternative formulations show how, if a language does not have readily available subjectless constructions for meeting the function of "impersonalization" then, quite obviously, other options will be sought.

The cross-linguistic comparisons proposed below are based on a large and uniquely comparable sample of the same text types derived by the same elicitation methods from the same age-groups of native speaker-writers in different languages. A clear divide emerged between the means favored for expressing depersonalization and agency downgrading in languages tolerant of subjectless constructions like Hebrew and Spanish compared with languages that require a surface subject like English and French. Hebrew relies on the three options detailed in §2, supplemented by morphological options for agent downgrading by means of intransitive middle-voice verb-morphology (§3.2), the latter system also used, but far less frequently, for passive voice. Spanish likewise has a rich range of options for expressing a depersonalized discourse stance: These include, along with a generic 2nd person pronoun subject, common reliance on *se* marked subjectless impersonals and, less commonly, *se* marked middle and passive voice constructions as well as periphrastic or syntactic passives with a non-agentive subject (Jisa & Tolchinsky 2009; Tolchinsky & Rosado 2005). In contrast, Reilly et al. (2005) describe English as using passive voice constructions – where younger children favor *get* passives compared with more mature reliance on *be* passives (Berman & Slobin 1994) – or else expletive pronouns *it, there* and generic subjects such as *people* alternating with generic pronouns like *you, they, one* – with these alternating in what appear to be rather haphazard ways. For example, in the course of a single short essay, a Californian high-school student [eH02] alternates generic reference as follows: “Problems are only as big as *one* makes them out to be. If *you* ignore a problem, or try not to make it into a big deal, *the other person* is forced to do the same thing. As *one* grows up, *they* learn to get along with people ...”

Other subject-requiring languages in the sample differ from English in having what we term a “dedicated” generic pronoun. This is the case with French *on*, which gives way with increased age and literacy to greater reliance on passive constructions for expressing a depersonalized stance (Jisa 2004; Jisa & Viguié 2005), while in Swedish, the analogous item *man* remains favored across the board as against its counterpart *maður* in the typologically closely related Icelandic, which prefers use of passive
constructions (Ragnarsdóttir & Strömqvist 2005). And in fact, reliance on passive voice, where “downgrading the topicality of agents raises the topicality a non-agent participant” (Givón 1994) in the form of a surface subject nominal, reveals clear cross-linguistic distinctions. Thus, comparison of passive voice usage in the written texts in five different languages differed markedly in this respect, with Dutch and English making the greatest use compared with Hebrew and Spanish, and French somewhere in the middle (Jisa et al. 2002).

These language-particular differences in use of passive voice for agency down-grading are of interest since, as noted, the discourse data-base was closely parallel in all seven languages (elicited on the same topic, by the same procedures, and at similar levels of age and schooling). Besides, all these languages have structurally productive passive voice constructions. The relatively low reliance on passives for agency down-grading in Hebrew, then, as in Spanish, is due not to structural productivity per se, but rather to the availability of alternative rhetorical options for expressing this same discourse function in the form of subjectless impersonals along with middle-voice constructions in which intransitive morphology combines with typically inanimate, non-agentive subjects. Basic expressive options for expressing a depersonalized discourse stance are accessible to even quite young children, in the form of a dedicated generic pronoun in subject-requiring languages (French on, Swedish man), and subjectless impersonals in Spanish (with se) and Hebrew (with 3rd person plural verbs). These constructions are typical of relatively informal, everyday discourse rather than the more mature, high-register style where passive constructions tend to occur in Hebrew (Ravid & Berman 2009; Taube 2007).

Analogous findings emerged in analysis of complex syntactic constructions in the same data-base. Although English, Hebrew, and Spanish share much the same repertoire of structural devices for linking clauses in extended discourse (by coordinate, complement, adverbial, and relative clauses), speaker-writers differed in their use of such constructions in parallel corpora: English relied relatively more on nonfinite subordination, Hebrew was largely paratactic, while Spanish texts favored complex embedding and interdependency of clauses (Berman & Nir-Sagiv 2009; Berman, Nir & 2009). Taken together, these patterns suggest that target language typology interacts with “rhetorical style” in determining the means favored by speaker-writers for meeting discourse functions such as event-construal (Berman & Slobin 1994: 622–639),

8. Assouline (2010) describes an interesting case of change in the Yiddish spoken by members of the ultra-orthodox speakers of the Jerusalem dialect, who use the 1st person plural mir as a “dedicated” generic pronoun, alternating with the impersonal pronoun me(n) and contrasting with occasional use of the pronoun undz to refer to a specific social group, generally expressing the opposition between “us” versus “them”.
clause-combining connectivity or, in the domain at issue here, expression of a deper-
sonalized discourse stance by means of passive voice, generic pronouns, or subjectless
impersonal constructions.

Another theme, one that echoes Slobin’s (1982, 2001) cross-linguistic insights
into child language development, is the effect of grammaticization of a given set of
options in a particular language. Reilly et al.’s (2002) comparison of modal expressions
occurring in written texts elicited on the same topic in different languages shows that
the younger English-speaking children use modal expressions significantly more than
their French-, Hebrew-, and Spanish-speaking peers – a finding we attributed to the
relative salience and accessibility of modals as a closed class of grammaticized terms in
English.9 Moreover, across the texts of children, adolescents, and adults, modal expres-
sions were typically subject/agent-oriented in English, less so in French, and even less
so in Hebrew and Spanish, two languages that do not require a surface subject in such
constructions. English can use expletive subjects to express impersonal, hence non-
referential, non agent-oriented modality (e.g. it is possible, necessary, likely), but these
were far more restricted in the texts we examined than their subjectless counterparts
in corresponding samples in Hebrew and Spanish. On the other hand, even advanced
level second language speakers of English use such rather heavy and stiff-sounding
constructions extensively to express an impersonal stance, where a subject/agent-or-
iented construction might sound more natural.

5. Concluding discussion

Two general conclusions emerge from a discourse-anchored consideration of Hebrew
impersonal constructions: First, that speaker-writers typically rely on a “confluence
of devices” from the structural options available to them in a given language for
expressing a discourse function such as a depersonalized stance; and, second, that
the available devices can be ranged along a continuum of differing degrees from
most to least impersonal.

The two texts reproduced in (20) and (21) illustrate the weaving together of
numerous alternative means for expressing an impersonal stance in Hebrew. Consider,
first, the oral text of 12-year-old girl [J-11] asked to give a talk discussing the topic of
“problems between people”. Numbered brackets indicate clause boundaries, with dep-
ersonalizing forms in bold.

9. Degree of grammaticization might also explain relatively wider distributions of the
generic 3rd-person pronoun subjects on and man in French and Swedish compared with the
partial occurrence of 2nd person ata in the texts of Hebrew 9-year-olds (4 out of 20 children).
Nearly every one of the 15 clauses in the original Hebrew version of this oral expository text produced by a young teenager is bolded, indicating that it is entirely impersonal and generic in overall stance – the only exception being the formulaic *ani xošëvet* ‘I think’ [cf. also Fr. je trouve, Sp. creo] in the opening line, a genre-typical marker of generalized discussion of an abstract topic. Even the translated version in (20’) illustrates a removed, distant, and totally generalized impersonal discourse stance. It demonstrates that by early adolescence, Hebrew speakers have recourse to numerous alternative constructions to express a depersonalized perspective on situations: by means of subjectless 3rd personal plural and modal constructions as well as by use of generic 2nd person pronoun (in Clause #6) – coupled with generic we (in Clauses #6 and #13) and lexical reference to generic entities, typically in the plural (e.g. people, problems, solutions), The text contains no personal pronouns or specific reference, while its impersonal, generalized, non-specific, and non-episodic tenor is underscored predicatively by use of timeless present tense plus irrealis mood and infinitival and other non-finite predicates.

Rather different, although no less varied linguistic devices for expressing an impersonal stance are illustrated by the expository essay written by a woman [A06] on the same topic in (21).
Di scourse-based perspectives on Hebrew impersonals 349

lefatéax et ha tikšóret ke-kišur xayim kax ] 7 še-kol adam yiheye kašuv ve-me’unyan lezahot et ha-téder ] 8 še-alav mešader xavero, ] 9 ve-kax timanana béáyot rabot. ] 10

(21’) Translation of Woman’s Essay into English

To my regret in our society in general and in Israel in particular (there) are very many problems between people and in different domains.1 The crux of the problems derives from difficulties in communication, both verbal and non-verbal 2 and frequently I find 3 that two people communicate on completely different wave-lengths4 leading to a breakdown.5 (It is) very important in my opinion 6 to promote interpersonal communication as a life skill 7 so that every person will be attentive and ready to identify the frequency 8 on which his-fellow communicates 9 and in that way many problems may be avoided. ] 10

Quite typically, and unlike schoolchildren, adults alternate perspectives between a distanced, objectively impersonal perspective and their own personal point of view. The adult essay in (21) (quite typically) reveals a variegated rather than homogeneous discourse stance, ranging from less to more personalized and back again – for example, in the discourse-marking comments “to my regret”, “I find”, “in my opinion” – reflecting a skillful alternation in perspective and point of view beyond the abilities of even young teenagers (Berman & Nir-Sagiv 2007). Like the girl’s talk in (20), here, too, the expository mode of discourse elicits atemporal predications rather than reference to specific times, through use of the extended habitual present and irrealis modality. But the adult’s text is in a consistently higher register, including use of 3rd person singular generic nominals like kol adam ‘each person’, xavero ‘his fellow(man)’ and passive-voice verbs.

In contrast, the example in (22), culled from an e-mail message sent by a highly literate, native Hebrew-speaking colleague, demonstrates how impersonal reference can function in a less formal style of discourse.

(22) od me’at holxim la-kalfi[1, ve-kcat mad’ig oti] 2 še-yeš soon go+PLUR to-the-polls, ] 1 and rather worries me 2 that be od anašim] 3 še-lo hištaxne’u bixlal] 4 še-ze still people [3 that-not convinced at-all]4 that it xašuv ] 5 le-hacbia ] 6 important[5 to-vote ] 6

‘We will soon be going to the polls, and it worries me a bit that there are still people that haven’t been altogether convinced that it’s important to vote.’

This short excerpt illustrates different depersonalizing constructions discussed in this chapter – 3rd person plural subjectless impersonals in [1], a subjectless evaluative
predicate in [2], an intransitive middle-voice construction in [3], combined with non-normative use of an expletive ze ‘it’ as subject in [4]. The casually unbuttoned medium of e-mail communication elicits from the educated speaker-writer of contemporary Hebrew a rich variety of discourse-stance encodings within a single proposition. Packaged together here are (1) a totally impersonal non-agentive going to the polls (here confined to the universe of Israelis with the right to vote), followed by (2) a subjectless evaluative clause personalized by use of accusative case in the causative construction mad’ig oti še … ‘(it) worries me that’ (cf. the equally well-formed passive ani mud’ag ‘I (am) worried’); then follows (3) a complement clause with middle-voice morphology used with the verb for ‘convince’ (again, cf. the passive alternant šuxne’u ‘were convinced (by X)’); ending (4) with a complement clause taking an overt expletive pronoun ze where the subjectless equivalent še 0 xašuv lehacbia ‘that (is) important to vote’ would be normatively more acceptable (cf. examples (1) to (3) in §1 above). Insertion of an overt expletive pronoun here is still confined to more colloquial spoken (and written on electronic mail) usage of educated native-speakers of Hebrew. But these are precisely the contexts in which language change will first be realized, as has been shown for other aspects of Modern Hebrew, too (Ravid 1995).

The excerpt in (22) thus demonstrates how a confluence of devices for nominal reference in expressing a depersonalized discourse stance interacts with different degrees of generality and specificity in types of predicating elements as well. This anecdotal instance is confirmed by Kupersmitt’s (2006) indepth analysis of the same data-base of extended texts as considered here, demonstrating that in English, Spanish, and Hebrew – three languages that differ markedly in grammatical TAM – the predicate-oriented domains of tense, aspect, mood, and voice interact clearly with overall discourse context, with past perfective tense/aspect and active voice preferred for conveying more specific, hence immediately involved episodic information as against reliance on timeless or habitual present, use of irrealis mood, and middle or passive voice predicates for expression of a generalized, detached impersonal stance.

In sum, a functionally discourse-based analysis of Hebrew impersonals combines with the idea of a confluence of structural devices – subjectless constructions, pronominals, case-marking, and tense/mood/voice – as conspiring together to express varying degrees of “impersonalization” (Siewierska 2008). Cross-linguistic comparisons of how these different devices are deployed in contrasting discourse genres suggest a continuum of depersonalization combined with agency downgrading, ranging from totally impersonal, non-referential via generic to specific reference: At one end are strictly subjectless constructions like Hebrew 3rd person plurals or Spanish se- constructions or their counterparts with plural pronouns like impersonal they in
English; these align with other totally impersonal subjectless constructions such as weather expressions and modal or evaluative propositions where subject-requiring languages resort to expletive subjects like English *it* and where in Hebrew, an experiencer role may be expressed by dative-marked (pro)nominals. Less impersonal, expressing a more inclusive rather than a fully detached discourse stance, although also non-specific in reference, are generic pronominals, either dedicated terms like French *on*, Swedish *man*, or else 2nd person singular pronouns used non-personally, alternating with 3rd person generic nouns like *people*, *a person* or a variety of pronouns used generically such as *we*, *one*, etc. At the personalized end of the continuum are means for making specific reference to individuals or classes of individuals and entities, by fully specific, personal forms of reference, both deictic 1st and 2nd person pronouns, as well as lexical noun phrases and anaphoric pronouns. These different means of nominal reference for conveying varying degrees of generality/specificity converge in a usage-based perspective with the predicate-oriented domains of tense, mood, and voice in the expression of a more or less impersonal discourse stance.

**Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACC</td>
<td>Accusative Direct Object marking prepositions</td>
</tr>
<tr>
<td>GEN</td>
<td>Genitive Suffix on Initial Head Noun in Bound Compounds</td>
</tr>
<tr>
<td>IMP</td>
<td>Imperative, marked for 2nd person number and gender</td>
</tr>
<tr>
<td>PASS</td>
<td>Passive voice forms of verbs</td>
</tr>
<tr>
<td>PLUR</td>
<td>Plural ending on verbs, nouns, adjectives</td>
</tr>
<tr>
<td>TAM</td>
<td>Tense, Aspect, Mood</td>
</tr>
</tbody>
</table>

**References**


CHILDES archive. (http://childes.psy.cmu.edu/data/Other/Hebrew/BermanLong.zip).


The elephant in the room

The impersonal -ne/-te construction in Polish

Anna Kibort
University of Cambridge & Surrey Morphology Group

This chapter offers an overview of the little studied Polish impersonal construction which is made up of an auxiliary or copular verb in 3sg.n and a passive or resultative -n/-t- participle bearing the singular neuter inflection -e. I describe the morphological form and syntactic behaviour of this construction, demonstrate that it is not only functionally impersonal but also syntactically subjectless, and discuss its relation to the impersonal passive of the intransitive as well as to the impersonal predicative adverbial construction. I conclude that the form of the impersonal -ne/-te construction is underspecified with regard to which of the two syntactic constructions (passive verbal or non-passive adverbial) it instantiates. Hence, syntactic frameworks need to provide an appropriate model for this type of underspecification.

Keywords: Polish; impersonal; resultative participle; passive of intransitive; predicative adverbial construction

1. Introduction

Polish has a wide range of impersonal constructions identified on the basis of their lack of a referential subject. The aim of this chapter is to give an overview of one that has been surprisingly little studied: the impersonal construction which uses the past/passive -n/-t- participle bearing the singular neuter inflection -e. The construction is exemplified in (1) and (2), under the assumption that no neuter singular referent can be found in the context of these sentences which could be interpreted as the antecedent of their ‘dropped’ subject:

1. I gratefully acknowledge the support of a Postdoctoral Fellowship from the British Academy, which has made this research possible. I also wish to thank Anna Siewierska and Andrej Malchukov for their helpful comments.
Sentences in (1) and (2) lack an overt subject – either lexical or a ‘dummy’ expletive one, since Polish does not have expletives. The sg.n inflection of the predicate is recognisable as a morphological strategy for situations when subject-predicate agreement breaks down either due to the lack of a subject or due to inadequate morphosyntactic properties of the subject. In Polish, as in many other languages, breakdown of subject-predicate agreement leads to the use of the (3)sg.n marking as the ‘exceptional case default’ inflection on the verb.² Importantly, as I show in §5, the location participant in (1b) and (2b) cannot be regarded a subject. Therefore, all four sentences in (1) and (2) are truly subjectless, in addition to being functionally impersonal by virtue of their non-elaboration of the causal participant of the event (Siewierska 2008: 121).

The impersonal -ne/-te construction poses an interesting problem of analysis, as it is not clear whether it is passive or non-passive. It can be argued to be an instance of the impersonal passive of the intransitive, since we can establish the following active-passive alternation:

(3) a. **Firma** codziennie sprzątała w pokojach.
*company*(f).nom every-day tidied.3sg.f in rooms
‘The (professional) company did the cleaning in the rooms every day.’

b. **W pokojach było codziennie sprzątane** (przez firmę).
in rooms was.3sg.n every-day tidy.part.sg.n (by company)
‘[It] was cleaned every day in the rooms (by the company).’
(4) a. *Piotr* *ładnie posprzątał w pokoju.*
    Peter(M).NOM nicely tied-up.3SG.M in room
    ‘Peter tided up/cleaned nicely in the room.’

    in room was.3SG.N nicely tidy-up.PART.SG.N (by Peter)
    ‘[It] was nicely cleaned/tidied up (by Peter) in the room.’

On this analysis, the -ne/-te participle in (3b) and (4b) performs a predicative function: it is the main verb of the passive predicate, accompanied by an auxiliary ‘be’.

On the other hand, the impersonal -ne/-te construction can be argued to be an instance of a predicative adverbial construction, a typical example of which is given below in (5a). This analysis seems particularly appropriate for sentences such as those in (2), since we observe the following analogy:

(5) a. *W pokoju było czysto.*
    in room was.3SG.N clean
    ‘[It] was clean in the room.’

    b. *W pokoju było posprzątane.*
    in room was.3SG.N tidy-up.PART.SG.N
    ‘[It] was cleaned/tidied up in the room.’

    c. *W pokoju było czysto i posprzątane.*
    in room was.3SG.N clean and tidy-up.PART.SG.N
    ‘[It] was clean and tidied up in the room.’

On this analysis, the -ne/-te participle in (5b–c) also performs a predicative function; however, it is not the main verb but a predicate adverb, accompanied by a copula ‘be’ functioning as the main verb.

While all the above examples with the -ne/-te participle can be considered instances of the so-called ‘objective resultative’, examples (6) and (7) illustrate the analogous use of the -ne/-te participle in the ‘possessive resultative’:

(6) a. *Miał codziennie sprzątane.*
    had.3SG.M every-day tidy.PART.SG.N
    ‘He had the cleaning done every day.’

    b. *Miał codziennie sprzątane w pokoju.*
    had.3SG.M every-day tidy.PART.SG.N in room(M).LOC
    ‘He had the cleaning done in his room every day.’

(7) a. *Miał ładnie posprzątane.*
    had.3SG.M nicely tidy-up.PART.SG.N
    ‘He had [it] nicely cleaned/tidied up.’

    b. *Miał ładnie posprzątane w pokoju.*
    had.3SG.M nicely tidy-up.PART.SG.N in room(M).LOC
    ‘He had it nicely cleaned/tidied up in his room.’
In (6) and (7) the -ne/-te participle is part of a secondary predicate which in turn is part of the clausal object of a personal verb (‘have’). The -ne/-te participle in these sentences shows default non-agreement inflection (sg.n), because the clausal complement of which it is a part has no head – that is, the small clause has no subject.

Regardless of whether the impersonal -ne/-te construction is analysed as passive verbal or non-passive adverbial, it is clear that in both cases the syntax makes use of one and the same -n-/-t- participial form. It is the same form which is also commonly found in personal sentences such as (8a–b), in which the -ne/-te participle shows agreement with the subject noun phrase:

\[(8)\]
\[\begin{align*}
\text{a. } & \text{Pomieszczenie było codziennie sprzątane.} \\
& \text{chamber(n).nom was.3sg.n every-day tidy.part.sg.n.nom} \\
& \text{‘The room was cleaned every day.’}
\end{align*}\]
\[\begin{align*}
\text{b. } & \text{Pomieszczenie było ładnie posprzątane.} \\
& \text{chamber(n).nom was.3sg.n nicely tidy-up.part.sg.n.nom} \\
& \text{‘The room was nicely cleaned/tidied up.’}
\end{align*}\]

Like (1) and (2), sentences in (8a–b) pose a similar analytical problem with regard to their status as passive or non-passive. Thus, in (8a–b) the -ne/-te participle can be analysed either as the main verb of a complex passive predicate, or as the predicative complement of the subject, and it matches the subject noun’s gender (here: neuter), number (here: singular), and case (here: nominative). If the subject has different inflectional properties, the participle matches them accordingly, as in the following examples with a feminine singular subject:

\[(9)\]
\[\begin{align*}
\text{a. } & \text{Sala była codziennie sprzątana.} \\
& \text{hall/ward(f).nom was.3sg.f every-day tidy.part.sg.f.nom} \\
& \text{‘The hall/ward was cleaned every day.’}
\end{align*}\]
\[\begin{align*}
\text{b. } & \text{Sala była ładnie posprzątana.} \\
& \text{hall/ward(f).nom was.3sg.f nicely tidy-up.part.sg.f.nom} \\
& \text{‘The hall/ward was nicely cleaned/tidied up.’}
\end{align*}\]

3. The reason for my suggestion that even in undisputably passive sentences the participle can be argued to inflect for case are examples such as the following, where the participle bears genitive case agreeing with the case of the quantified noun of the subject noun phrase:

\[(i)\]
\[\begin{align*}
\text{Pięciu żołnierzy zostało zastrzelenych przez snajperów.} \\
& \text{five.gen soldiers(mhum).gen became.3sg.n shoot.part.pl.gen by snipers} \\
& \text{‘Five soldiers were shot by snipers.’}
\end{align*}\]

Note that the case value of the numeral in such subject phrases is disputed; a discussion of this issue is, however, beyond the scope of this chapter.
Furthermore, while (8) and (9) are personal varieties of the ‘objective resultative’, examples (10) and (11) illustrate personal varieties of the ‘possessive resultative’:

(10) a. Miał to pomieszczenie codziennie sprzątane.
    had.3SG.M this.N.ACC chamber(N).ACC every-day tidy.PART.SG.N.ACC
    ‘He had the room cleaned every day.’

    b. Miał to pomieszczenie ładnie posprzątane.
    had.3SG.M this.N.ACC chamber(N).ACC nicely tidy-up.PART.SG.N.ACC
    ‘He had the room nicely cleaned/tidied up.’

(11) a. Mam już wszystkie egzaminy pozdawane.
    have.1SG already all.NONMHUM.ACC exams(NONMHUM).ACC take.PART.PL.NONMHUM.ACC
    ‘I already have all the exams taken.’ (meaning I have already taken all
    of the exams’) (Rothstein 1993:715)

    b. Miała już projekt prawie wykończony.
    had.3SG.F already design(M).ACC almost finish.PART.SG.M.ACC
    ‘She already had the design almost finished.’ (meaning She had almost
    finished the design’)

As in (8), in (10)–(11) the -n/-t- participle is an agreeing element: its inflection matches
that of its head noun which is also the subject of the small clause. The main verb of the
matrix clause does not have to be ‘have’ – there are many more verbs that would accept
a small clause of this type as an object (not only ‘get’, but also ‘see’, ‘give’, ‘seem’, etc.).

In (8)–(11) I have shown personal uses of the -n/-t- participle analogous to the
impersonal uses of this participle – that is, they were all predicative uses. However,
personal (but not impersonal) constructions additionally have a common variant
where the -n/-t- participle can be used attributively, as in the following examples; compare (12a–b) with (8a–b):

(12) a. codziennie sprzątane pomieszczenie
    every-day tidy.PART.SG.N NOM chamber(N).NOM
    ‘a/the daily-cleaned room’

    b. ładnie posprzątane pomieszczenie
    nicely tidy-up.PART.SG.N NOM chamber(N).NOM
    ‘a/the nicely cleaned/tidied up room’

4. I have taken this sentence and (13b–c) from Dombrowski (2006:28), but as they all illust-
trate common phenomena, many more examples could easily be provided.
c. *złamana* noża
   break.PART.SG.F.NOM leg(F).NOM
   ‘a/the broken leg’

d. *zgubione* banknoty
   lose.PART.PL.NONMHUM.NOM banknotes(NONMHUM).NOM
   ‘(the) lost banknotes’

These noun phrases with the -n/-t- participle used attributively also commonly occur in the ‘possessive resultative’, as in the following examples:

(13) b. Wojtek ma złamaną nogę.

   Wojtek(M).NOM has break.PART.SG.F.ACC leg(F).ACC
   ‘Wojtek has a broken leg.’ (meaning ‘Wojtek has broken his leg’)
   (Lempp 1986: 126)

c. Mam zgubione banknoty.

   have.1SG lose.PART.PL.NONMHUM.ACC banknotes(NONMHUM).ACC
   ‘I’ve got some banknotes lost.’ (meaning ‘I have lost some banknotes’)
   (Lempp 1986: 126)

In the sections below I first discuss the extent of the impersonal -ne/-te construction in Polish, then the morphology of the -ne/-te participle and the syntax of the impersonal construction in which it can be found. In particular, I clarify the overlap of the -ne/-te impersonal construction with the impersonal passive of the intransitive and with the predicative adverbial construction. I also briefly discuss the status of the locative argument in this construction, arguing that it cannot be considered its subject. Apart from providing illustrative constructed examples, I cite a large number of examples from naturally occurring written and spoken discourse extracted from the web and from two large corpora of Polish language: the IPI PAN Corpus (of over 250 million segments) developed by the Linguistic Engineering Group at the Institute of Computer Science, Polish Academy of Sciences, and the PELCRA Reference Corpus of Polish (of around 93 million words) developed by the Department of English at the University of Łódź, Poland. Both corpora are available for searching online.

2. The extent of the -ne/-te construction in Polish

Searches for the -ne/-te participle in Polish corpora and on the web bring up many impersonal uses of the -ne/-te participle formed from semantically transitive predicates implying an external, mostly human, agent. The following is a range of verbs that can be found in the impersonal -ne/-te construction, matched with their -ne/-te participles showing various aspects, prefixes, and sometimes negative polarity. Both lists, of verbs and of their participles, are representative though certainly not exhaustive:
The impersonal -ne/-te construction is commonly associated with verbs denoting household activities such as sprzątać ‘tidy’, gotować ‘cook’, piec ‘bake’, prać ‘launder’,
prasować ‘iron’, ścierać ‘wipe’, wytrzeć ‘wipe’, zmyć ‘wash-up’, etc. Examples abound, and the following is a small selection from the Polish corpora and texts found on the web. Note that example (15d) contains a headless -ne/-te participle formed from the verb złożyć ‘fold’, which can be taken as evidence that the construction is indeed productive:

(15) a. A moja mama mi zawsze na to: ‘a ja z czworgiem and my mother me always to this and I with four dzieci magisterkę pisałam i zawsze było posprzątane, children master’s-thesis wrote and always was cleaned obiad z dwóch dań i upieczone, pozmywane – a dinner of two courses and baked washed-up and zmywarek nie było”.
dishwashers not were

‘And my mother always replies to me: “and I was writing my master’s thesis with four children and [it] was always cleaned, dinner of two courses, and [it was] baked, washed up – and there were no dishwashers”.

b. Nie było ugotowane, posprzątane, a dziecko 3 godziny not was cooked cleaned and child three hours siedziało w huśtawce bo nie miał jej kto z sat in swing because not was her who from niej wyciągnąć.
it take-out

‘[It] wasn’t cooked, cleaned, and the child had been sitting in a swing for three hours because there was no one to take her out of it.’

c. Wtedy zobaczyłam swoje życie jako bieganina za obowiązkami, then saw my life as running after duties aby było czysto, aby było poprasowane i poprane… so-that was clean so-that was ironed and laundered tylko, po co?
only for what

‘Then I saw my life as a race against duties, making sure [it] was clean, ironed and laundered… but what for?’

d. Lista życzeń: 15. Deska do prasowania – najlepiej taka, żeby list of-wishes 15. board for ironing best such that wystarczyło położyć i żeby już było wyprasowane. be-enough lay and that already was ironed

‘A wish list: 15. An ironing board – the best one would be such that it would be enough to put [the ironing on it] and [it] would already be ironed.’
e. *Ma obowiązek szeroko pojętego dbania o dom, czyli*  
that be cooked laundered and folded into cube  
that he/she has the duty of a broadly understood care for the household,  
that is, [to make sure] that [it] would be cooked, laundered and folded  
into a cube.'  

f. *jeszcze nie pościerane?! (…) szybciej!*  
‘[It] has not yet [been] wiped?! (…) hurry up!’  

In all these and similar examples, verbs denoting household activities are used intran-  
sitively, and their objects are understood (or, generic). It would not be surprising to  
find that a corresponding class of verbs behaves in a similar way in other languages,  
both in the active – dropping the understood object, and in the impersonal – dropping  
the understood subject or head noun (as in the examples above). An example is Hun-  
garian, where a large class of verbs with a similar denotation to the Polish ones just dis-  
cussed can be used both transitively and intransitively: *kitakarít* ‘clean up’, *kimos* ‘wash’,  
*kivasal* ‘iron’, *elmosogat* ‘do the washing’, *bepakol* ‘pack’, *tálat* ‘serve (food)’,  
*felmos* ‘wash the floor’, *befűt* ‘heat up’, *kiszellőztet* ‘air’, *bevásárol* ‘shop’, etc., both in the active and in  
the passive or resultative impersonal (Tóth 2000: 251–252). The following example,  
from Tóth (2000: 252, example 24), compares the personal variant of the sentence with  
the resultative participle in -*va* with the subjectless variant (where the location argu-  
ment is not a subject, but an optional oblique):

In three seconds [it] was washed-up, dried, and [they went quickly]  
round the rooms.'  

(16)  
a. A *szobák* ki *van* takarít-*va*.
  the rooms-NOM PV_{out}^{\text{5}} are clean-*VA
  ‘The rooms are cleaned.’

b. (A *szobák-ban*) ki *van* takarít-*va*.
  the rooms-INE PV_{out} is clean-*VA
  ‘It is cleaned (in the rooms).’

Among some other functions, this element marks the perfectivity of the predicate.’
The impersonal -ne/-te construction in Polish may also include an optional location argument which is not a subject (see §5). A location argument such as w domu 'at home' or another appropriate one could easily be added to most impersonal -ne/-te sentences with participles denoting household activities. Other impersonal -ne/-te sentences can be formed with an oblique location argument (or adjunct) wszędzie 'everywhere', or with other appropriate expressions of location:

(17) a. Wszędzie było ładnie udekorowane.
   everywhere was.3sg.n nicely decorate.part.sg.n
   'Everywhere was nicely decorated.'

b. Wszędzie było dobrze oznakowane.
   everywhere was.3sg.n well signpost.part.sg.n
   'Everywhere was well signposted.'

c. Wszędzie było strzeżone.
   everywhere was.3sg.n guard.part.sg.n
   'Everywhere was guarded.'

d. Wszędzie było zajęte/pozajmowane.
   everywhere was.3sg.n occupy.part.sg.n
   'Everywhere was occupied.' (about parking spaces, toilets, etc.)

e. Wiem, że wczoraj podłoga była mokra na dolnej kondygnacji, know that yesterday floor was wet on bottom floor
   ale na wyższych nie było zmyte.
   but on higher-ones not was washed
   'I know that yesterday the floor was wet on the lower floor, but on the higher ones [it] had not been washed.'

f. Patrzę, że na rurze sterowej jest dookoła porysowane (taka look that on tube steering is around scratched this
   jedna linia), zaglądam do główki ramy i w tym samym one line peep into head frame and in this same
   miejscu jest też porysowane ale nie dookoła tylko taki place is also scratched but not around only this
   półkrąg. Ale w miejscu gdzie na amorze było porysowane half-circle but in place where on fork was scratched
   to nic nie było.
   then nothing not was
   'I see that on the steering tube [it] is scratched [all] around (just one line). I peep into the head of the frame and in the same place [it] is also scratched but not around, just a half-circle. However, in the place where [it] was scratched on the fork – there was nothing.' (about bike maintenance)
g. Akurat tak było, że w jednym pomieszczeniu to były deski, just so was that in one room that were boards 
a w drugim było nie wykończone, to glina była. and in second was not finished that clay was 
‘It was just like that, in one room there were boards, and in the other [it] was not finished, there was clay.’

h. Dreads się doczepia w ten sposób najlepiej, że się rozwala dreads refl attach in this way best that refl break-up końcówkę i się je doszydelkowuje i nawet nie widać ending and refl them crochet-hook and even not see w którym miejscu było przedłużane. in which place was extended
‘Dreads are best attached in this way that [you] break up the ending and [you] crochet-hook them and [then you] cannot even see in which place [it] was extended.’

The following sentences exemplify common uses of the impersonal -ne/-te construction formed from other verbs which were given in the list in (14):

(18) a. Nie pozostaje ci nic innego jak wziąć biling do ręki not remains you nothing else than take bill into hand i sprawdzić kiedy było dzwonione, o której godzinie, i and check when was phoned at which hour and zobaczyć kto wtedy był w domu itp. see who then was at home etc.
‘Nothing else remains for you [to do] but to take the bill in your hand, check when [it] was called, at what time, and check who was at home at that time, etc.’

b. Już skończone – odezwała się spokojnie, próbując already finished pronounced refl calmly trying doprowadzić mój ubiór do ładu. bring my garment to order
‘[It’s] already finished – she said calmly, trying to get my garment in order.’

c. Jеще же сть нідокінчоне. wiesz. jeszcze niedokończone. still is unfinished you-know still unfinished bo nie ma wiesz. dywanik. zasłonki. tylko spłukani because not has you-know carpet curtains only flat-broke jesteśmy zupełnie. we-are completely
‘[It] is still unfinished, you know, still unfinished, because there is not, you know, a carpet, curtains, but we are completely broke.’
d. Jeśli ktoś miał niezależnie pobrane świadczenia ale było potrącone z bieżących świadczeń to też wpisuje to deducted from current benefits then also fills this in first row

'If one had taken their benefits independently but [it] was deducted from their current benefits, then they also fill this in the first row.' (about filling out a tax form)

e. Samemu można zrezygnować (aby nie było przedłużane) w oneself can opt-out so-that not was extended at desk cash

'One can opt out oneself (so that [it] is not extended) at the cash desk.' (about licence fees)

f. No i co z tego, że było otwarte? Czy zamknięcie now and what of this that was open(ed) INTERR closing sklepu to jakiś hold?

shop this-is some homage

'And so what that [it] was open? Does the closing of a shop [constitute] some homage?'

g. A placone było nieźle, bo pośpiech był.

and paid was not-badly because urgency was 'And [it] was paid rather well, because there was urgency.'

h. Zatem – finis. Skończone z Bellonem i skończone ze thus 'finis' finished with Bellon and finished with wszystkimi. Żegnam panów.

everyone bid-farewell gentlemen

That is it, then – the end. [It is] finished with Bellon and finished with everyone. I bid you farewell, gentlemen.'

i. Nie, nie, w biosie było sprawdzane i włączane.

no no in bios was checked and switched-on 'No, no, in bios [it] was checked and enabled.'

j. Mówię mu że jest zajęte, a on na to: “zamknij mordę”.

tell him that is occupied and he on this shut gob 'I tell him that [it] is taken, and he [replies] to this: “shut your trap”.

k. Byłem tam w poniedziałek i było zamknięte.

was there on Monday and was closed ‘I was there on Monday and [it] was closed.’
l. Poprzez swą ogólndostępność kusi tych co
through own-self general-accessibility tempts those that
prawdopodobnie by nie spróbowali gdyby było zabronione.
probably would not tried if was forbidden

‘Through its general accessibility [it] tempts those who would probably
not try [it/them] if [it] was forbidden.’ (about drugs including ecstasy)

Many of the -ne/-te participles listed in (14) can be followed by the complementiser że
‘that’ and a finite clause, as in (19a). In those cases, I treat the clausal complement as
the grammatical subject of the sentence with the participial predicate. Clausal subjects
do not have the inflectional properties of controllers of agreement and are therefore
non-canonical. Sentences with the -ne/-te participle may also have other non-canonical
subjects which lack the inflectional properties of controllers of agreement, as in (19b).
Examples such as (19a–b) therefore qualify as impersonal, but not as subjectless (in
contrast with the impersonal and subjectless clauses in (15), (17), and (18)):

(19) a. Jest napisane, że biuro jest czynne.
   is written that office is active(=open)
   ‘[It] is written that the office is open.’

b. Było napisane “20”.
   was written “20”
   ‘[It] was written “20”’.

However, some of those participles may equally easily occur without such subjects,
but instead with adverbial elements: either the particle-adverb jak ‘as’ which relates
its clause to another clause, as illustrated in (20), or complex adverbial complements
which themselves involve finite clauses, as illustrated in (21).

(20) Jak już bylo/zostało powiedziane, ....
    as already was/became said
    ‘As [it] was already said, ....’

Other -ne/-te participles which can be found in this pattern include: odnotowane
‘noted’, podkreślone ‘emphasised’, przypomniane ‘reminded’, skomentowane ‘com-
mented on’, wspomniane ‘mentioned’, wyeksponowane ‘highlighted’, zasugerowane ‘sug-
gested’, zauważone ‘noticed’, and so on.

(21) Tak bylo/zostało powiedziane, że wyszło, że przepłaciliśmy.
    so was/became said that turned-out that overpaid
    ‘[It] was said in such a way that it turned out that we overpaid.’

And again, other -ne/-te participles which can be found in this pattern include: eks-
ponowane ‘highlighted’, napisane/wypisane ‘written,’ obliczane/obliczone ‘calculated’,
planowane ‘planned’, płacone ‘paid’, pomyślane/przemyślane ‘thought through,’ potrącane

Many of the *-ne/-te* participles listed above are found in the ‘possessive resultative’ construction which I first illustrated in examples (6) and (7), repeated here as (22) and (23):

(22) a. *Miał codziennie sprzątane.*

\[
\text{had.3sg.m every-day tidy.part.sg.n}
\]

‘He had the cleaning done every day.’

b. *Miał codziennie sprzątane w pokoju.*

\[
\text{had.3sg.m every-day tidy.part.sg.n in room(m).loc}
\]

‘He had the cleaning done in his room every day.’

(23) a. *Miał ładnie osprzątane.*

\[
\text{had.3sg.m nicely tidy-up.part.sg.n}
\]

‘He had [it] nicely cleaned/tidied up.’

b. *Miał ładnie osprzątane w pokoju.*

\[
\text{had.3sg.m nicely tidy-up.part.sg.n in room(m).loc}
\]

‘He had it nicely cleaned/tidied up in his room.’

In these sentences, the clausal complement of which the *-ne/-te* participle is a part has no head – that is, if we analyse the clausal complement as a small clause, it has no subject. Examples of this variant of the impersonal *-ne/-te* construction are also numerous both in Polish corpora and on the web, and here is a small handful of examples:

(24) a. *Szpitale mają płacone od pacjenta.*

\[
\text{hospitals have paid from patient}
\]

‘Hospitals have [it] paid per patient.’ (meaning ‘Hospitals are paid per patient’)

b. *Ja mam tak wypisane na ulotce.*

\[
\text{I have so written on leaflet}
\]

‘I have [it] written in this way on the leaflet.’

c. *A w święta majowe będziesz miał otwarte?* and in holiday May you-will have open ‘And during May holiday are you going to have [it] open(ed)?’

d. *W głowie miał zawsze przewrócone.*

\[
\text{in head had always tumbled}
\]

‘He always had [it] tumbled in the head.’ (meaning ‘He was a bit of a nutcase’).
The last example, (24d), which is close to being an idiom (although it does have an active personal variant), brings me finally to a small but open and productive semantic class of verbs which do not seem to be used as main verbs in transitive personal clauses, but are found only as -ne/-te participles with the prefix prze- (roughly) ‘through’ in the impersonal construction with ‘have’, and only in the colloquial language:

(25)  
Piotr ma przechłapane/przerąbane/prześwistane.  
Peter has splashed/hacked/whistled  
‘Peter has [it] splashed/hacked/whistled.’ (meaning: ‘Peter is in trouble’)

3. The morphology of the -ne/-te construction

In the following subsections I discuss the key properties and the status of the elements making up the complex verb of the -ne/-te construction: the -n/-t- participle, and the finite verb.

3.1 The -n/-t- participle: form and meaning

The -n/-t- participle underlies a set of forms which make up an inflectional paradigm identical to that of the adjective (see e.g. Laskowski 1998a: 268–269). For example, the number and gender forms of the participle sprzątan- ‘tidy.part’ in the nominative, as compared with the forms of the adjective czerwon- ‘red’, are:

(26)  

<table>
<thead>
<tr>
<th>Form Type</th>
<th>Pronoun</th>
<th>Participle Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>[SG].M</td>
<td>czerwon-y</td>
<td>sprzątan-y</td>
</tr>
<tr>
<td>[SG].F</td>
<td>czerwon-a</td>
<td>sprzątan-a</td>
</tr>
<tr>
<td>[SG].N</td>
<td>czerwon-e</td>
<td>sprzątan-e</td>
</tr>
<tr>
<td>[PL].MHUM</td>
<td>czerwon-i</td>
<td>sprzątan-i</td>
</tr>
<tr>
<td>[PL].NONMHUM</td>
<td>czerwon-e</td>
<td>sprzątan-e</td>
</tr>
</tbody>
</table>

Furthermore, the -ne/-te participle falls in the same declensional paradigm as the largest inflectional class of adjectives.7

---

6. The choice of the -n- vs -t- variant of the suffix in the formation of the participial stem depends on the phonological makeup of the verb stem to which it is attached. In an academic textbook of morphology, Laskowski (1998a: 268–269) breaks up the participial suffix into four variants: -n/-t/-on/-ęt-, and specifies in detail the conditions for their choice as well as the additional sound alterations in the verb stem that may be induced by the formation of the participle.

7. Additionally, the the -n/-t- participle can occur with the ending -o which falls outside the number-gender-case adjectival paradigm and realises a unique impersonal ending as well as turns the participle into a finite verb form with past tense meaning, e.g. sprzątan-o ‘tidied.impers’.
The -n/-t- participle is often conventionally labelled a ‘passive participle’ (Pol. imiesłów bierny, e.g. Bartnicka 1970:13, 50; Laskowski 1998a: 268, 1998b: 202; Kallas 1998:473; Nagórko 1998:92, 125; Saloni & Świdziński 1998:192–193; also the Polish IPI PAN Corpus). However, at the same time it is widely acknowledged that the ‘passive’ label does not fit all participles in this class, and authors of academic grammars always clarify that the label ‘passive participle’ merely groups all participles sharing the same morphological form:

We will note here the class of adjectival participles including ucieszony ['one that has become joyous; overjoyed'], zmartwiony ['worried, concerned'], zadomowiony ['ensconced'], rozpędzony ['one that has gathered speed']. These are productive formations made in a regular fashion from perfective verbs with się ['refl'] which have inceptive meaning i.e. which denote the beginning of a certain state. … Participles of this type, whose morphological form does not differ from that of passive participles, are nevertheless fundamentally different from the latter as regards their function. Participles such as ucieszony ['one that has become joyous; overjoyed'], as well as personal forms of verbs they are formed from, represent unmarked diathesis. (Laskowski 1998b: 202–203; my translation and emphasis)

We will treat as adjectives also those lexemes which have the form of passive participles, but which are formed from intransitive verbs: wyspany ['one that has slept enough'], uśmiechnięty ['one that has smiled/been smiling']. (Kallas 1998:473; my translation and emphasis)

In order to distinguish passive participles from the other type of identically formed participles, some authors resort to calling the latter type ‘morphologically passive participles with active meaning’ (Bartnicka 1970:52), or ‘irregular’ (as opposed to ‘regular’) passive participles (Saloni & Świdziński 1998:193).

Many Polish grammars also acknowledge – even though this statement might not be prominent in any particular textbook – that the distinction between ‘true’ passive participles and morphologically passive participles with active meaning is, in fact, not clear cut. On the one hand, participles such as ukradziony ‘stolen’ or recytowany ‘recited’ are clearly passive in the sense that list jest/był ukradziony ‘the letter is/was stolen’ corresponds to ktoś ukradł list ‘someone stole/has stolen the letter’, and liturgia jest recytowana ‘liturgy is recited’ corresponds to ktoś recytuje liturgię ‘someone recites/is reciting the liturgy’ – and this type of correspondence between constructions which use related verb forms is normally treated as a canonical passive alternation. On the other hand, participles such as wyspany ‘one that has slept (well/enough)’ or popękany ‘cracked’ are clearly non-passive in the sense that on jest wyspany ‘he is well-slept’ cannot correspond to the ill-formed ‘He has been slept by someone (e.g. his mother)’, and wazon jest popękany ‘vase is cracked’ cannot correspond to ‘The vase has been cracked by someone’. Instead, on jest wyspany ‘he is well-slept’ corresponds to on się wyspał ‘he has slept well/enough’, and wazon jest popękany ‘vase is cracked’
corresponds to *wazon popękał* ‘the vase has cracked’ – hence, in these cases, no passive alternation can be established.

However, apart from these two clear classes, there is a host of -n/-t- participles whose status as passive or active cannot be resolved unambiguously. For example, *on jest zmartwiony* ‘he is worried/concerned’ could be considered active on the basis of its correspondence to the inchoative *on zmartwił się* ‘he has (become) worried’, or passive on the basis of its correspondence to the causative *ten problem go zmartwił* ‘this problem has worried him/has got him worried’; likewise, *silnik jest zepsuty* ‘the engine is broken’ could be considered active on the basis of its correspondence to the inchoative *silnik zepsuł się* ‘engine has broken’, or passive on the basis of its correspondence to the causative *ktos zepsuł silnik* ‘someone has broken the engine’. Bartnicka (1970: 50–85) discusses in detail a very large number of actual examples of written and spoken Polish which illustrate this ambiguity, and informally refers to the ambiguous participles as ‘incompletely passive’ *niezupełnie bierne*. This class of ambiguous participles is by no means a marginal phenomenon – on the contrary, it appears to be representative of the class of -n/-t- participles.

Contrary to the Polish descriptive tradition, but following typological work, I argue that all Polish -n/-t- participles (not just the ‘active’ ones), which obviously result from the same morphological derivation, correspond to what in typological work is known as the *resultative participle*. Nedjalkov & Jaxontov (1988: 6), who undertook a cross-linguistic study of resultative constructions, define the term ‘resultative’ as indicating ‘those verb forms that express a state implying a previous event’. A resultative participle, therefore, characterises its head ‘by expressing a state that results from a previous event’. A resultative participle, therefore, characterises its head ‘by expressing a state that results from a previous event’ (Haspelmath 1994: 159). Although in the Polish descriptive tradition the term ‘resultative’ is normally used only with reference to ‘active’ -n/-t- participles (see the first quote above from Laskowski 1998b) and to a smaller class of participles with a similar function formed from the -l- stem (e.g. *zbiegły* ‘escaped’, *zwiędły* ‘withered’; see e.g. Laskowski 1998b: 203, and Cetnarowska 2000), I argue that the term should be properly extended to all -n/-t- participles, whether ‘passive’, ‘active’, or ambiguous, since in fact all of them are one and the same type of participle. This suggestion is not merely a terminological correction, but – as I will demonstrate below – it helps clarify the important distinction between passive and non-passive constructions, and identify the syntactic status of the impersonal -ne/-te construction in Polish.

3.2 The uses of the -n/-t- participle as a resultative participle

It is widely known that adjectival ‘past’/‘perfect’/‘resultative’ participles (e.g. English *eaten*, *sung*, *fallen*), like many deverbal adjectives (e.g. English *dreadful*, *fearful*), have a semantic ‘orientation’ – a notion that seems to originate from
Lehmann (1984: 152). Syntactically, they modify their head noun or complement their subject. Semantically, ‘past’/‘perfect’/‘resultative’ participles are oriented towards the affected participant.

In semantically transitive events the affected participant is usually the patient/theme, hence resultative participles formed from transitive verbs are typically patient/theme-oriented, and therefore the eaten dog is understood as ‘the dog that has been eaten’. In semantically intransitive events where the only participant is a patient/undergoer/experiencer, that participant is also typically recognised as affected, hence resultative participles are naturally formed from intransitive unaccusative verbs, and the fallen leaf is understood as ‘the leaf that has fallen’. However, if a transitive or an intransitive event can be construed as affecting the agent, agent-oriented resultative participles can also be formed from both transitive and intransitive unergative verbs. Although this construal may perhaps occur less frequently, it is nevertheless widely attested particularly with certain classes of verbs (in particular, verbs of obtaining, wearing, ingestion, and ‘mental ingestion’, see Nedjalkov & Jaxontov 1988: 9, cf. Haspelmath 1994: 174, footnote 10). English examples include drunk, which can be used to refer to either the liquid or the person; and the informatively enriched agent-oriented participles in an unbelievably learned architect, a confessed killer, a run-away slave, an over-exercised athlete, a well-read person, etc. (examples from Bresnan 2001: 34–36, though used by her to support different argumentation).

Since the operation deriving the resultative participle from the verb involves semantics – the operation takes as its input a semantically unoriented lexical form and outputs a semantically oriented lexical form – it is no surprise that there may be semantic restrictions on the formation of the resultative participle. These will not be discussed here in detail, but need to be mentioned briefly since they apply to Polish -n/-t- participles as well. First of all, there is the well-established semantic restriction involving telicity: while telic unaccusatives do form resultative participles (a fallen leaf, a recently appeared book), atelic unaccusatives typically do not (*the remained boy). In addition to these, there are also pragmatic considerations, for example involving informativeness (e.g. a prepared teacher vs a well-prepared teacher).

All types of Polish resultative participles in -n/-t- can be used as adjectives, and therefore are found in both the attributive position, modifying their head noun, and in the predicative position, complementing the subject of the clause (Bartnicka; see also Szupryczyńska 1980: 38–45). Since Polish syntax does, in principle, allow subjectless constructions, it follows that if the speaker wishes to use an adjectival/adverbial predicate with reference to the ambient quality or location, but without specifying or identifying the referent of the predication, the syntax should allow it. And indeed it does, through the use of subjectless sentences with impersonal adjectival predicates such as the following, where (b) and (c) are repeated from (2):
Furthermore, some resultative participles in 

\[-n/-t-\] can be used as main verbs in

the passive construction. This applies only to a subset of all resultative participles – specifically, to those which are formed from syntactically unergative verbs – since only unergative predicates can be syntactically passive, as is proposed in the Unaccusative Hypothesis first formulated by Perlmutter (1978). It follows, then, that if the speaker wishes to foreground an intransitively construed event at the cost of not elaborating the agent of this event (and having the option of including it in the clause only as an oblique argument), the syntax also allows this. The following examples contrast a personal variant of the passive in (a) with impersonal passive variants in (b) and (c), all of which are adapted from earlier examples in (8a) and (3b):

\[
\begin{align*}
\text{(27) a. } & \text{ Było } \text{} \text{ladnie/czysto}. \\
& \text{was.3SG.N } \text{nicely/cleanly} \\
& \text{‘[It] was nice/clean [there].’} \\
\text{b. } & \text{Było } \text{(ladnie) posprzątane.} \\
& \text{was.3SG.N (nicely) tidy-up.PART.SG.N} \\
& \text{‘[It] was (nicely) cleaned/tidied up.’} \\
\text{c. } & \text{W pokój był (ladnie) posprzątane.} \\
& \text{in room(M).LOC was.3SG.N (nicely) tidy-up.PART.SG.N} \\
& \text{‘It was (nicely) cleaned/tidied up in the room.’}
\end{align*}
\]

I take a closer look at both types of impersonal \[-n/-t-\] construction in §4 and exemplify them with sentences found in Polish corpora and the web. For a more detailed discussion of the form and uses of the resultative participle illustrated mostly from English, see Kibort (2005).
3.3 The finite verb

In its impersonal uses, the -ne/-te participle co-occurs with several different types of finite verb which may be analysed as a copula or as an auxiliary depending on the analysis of the structure they are in.

Following Szupryczyńska (1980: 35–38), I recognise the following verbs as copulas which can co-occur with predicate adjectives including resultative participles: być/bywać ‘be’, and stać się/stawać się ‘become, come to be’. Additionally, one more verb can be found with adjectival (i.e. resultative) participles derived from perfective verbs, but not with non-derived adjectives (Szupryczyńska 1980: 37): zostać/zostawać ‘become, get’. This suggests that, in sentences with adjectival (as opposed to nominal) predicates the verb zostać/zostawać ‘become, get’ is not a copula, but an auxiliary in a passive construction. In a recent corpus study of passive constructions in Polish, Górski (2008: 61–63, 67–69) confirms this hypothesis and supports it with examples which demonstrate that zostać/zostawać ‘become, get’ plus a resultative participle can only be interpreted as a passive construction alternating with a corresponding active.

The following verbs are unarguably considered auxiliaries of the passive construction (Siewierska 1984: 126–130; Laskowski 1998b: 194–197; Saloni & Świdziński 1998: 320; Górski 2008: 48–50): być/bywać ‘be’, and zostać/zostawać ‘become, get’. Furthermore, some scholars (e.g. Korytkowska 1993) have suggested that mieć ‘have’, as in examples (10)–(11) above, can also be analysed as a passive auxiliary. This hypothesis is supported by the following correspondence (examples from Górski 2008: 44 who cites the example in (29a) from Korytkowska 1993: 172):

(29) a. Pacjentka ma/miała zrobioną operację
patient(f).nom has/had.3sg.f do.part.sg.f.acc operation(f).acc

przez znanego chirurga.
by known surgeon

‘The patient had [her] operation done by a well-known surgeon.’

b. Znany chirurg zrobił
known.m.nom surgeon(m).nom did.3sg.m

pacjentce operację.
patient(f).dat operation(f).acc

‘A/The well-known surgeon performed an operation on the patient.’

It appears that the sentence in (29a) fulfils several characteristics attributed to the passive. However, Górski (2008: 44) remarks that the verbal form ma zrobioną ‘has do.part.sg.f.acc’ cannot be regarded as related to zrobić/zrobić ‘did.3sg.m/do.inf’ by any diathetical morphological derivation. I agree with Górski’s statement that the construction which uses ‘have’ and the resultative participle is grammaticalised (perhaps as an emerging new tense, the ‘perfect’), but in my view the alternation in the assignment
of grammatical functions to the arguments of these two predicates is too different from the alternation holding between an active and a passive predicate. I therefore consider (29a) as distinct from the passive, and the -n-/t- participle in this example as a member of a secondary predicate which is part of the clausal object of the personal and active verb ‘have’.

4. The syntax of constructions with the -ne/-te participle

In the following subsections I discuss the possible syntactic analyses of the impersonal -ne/-te construction. I regard the construction made up of the finite verb ‘be’ and the resultative participle as syntactically ambiguous, and demonstrate that it overlaps with both the impersonal passive of the intransitive (§4.2) and the predicative adverbial construction (§4.3).

4.1 Ambiguity of the ‘be’ + resultative participle construction

It is evident from the long-standing discussion in the literature (most recently, Górski 2008:61–64ff) that personal sentences with być ‘be’ and the -n-/t- participle do not have a universally accepted analysis. The problem lies, of course, with those sentences where the -n-/t- participle is derived from a transitive verb which denotes a situation that can be construed either causatively or inchoatively. I already mentioned this issue in §3.1 while discussing the meaning of the -n-/t- participle. I now reiterate and relate this observation to the analysis of the syntactic structure in which the participle is found. Compare sentence (30a) with its inchoative (or, anticausative) counterpart in (b), the passive in (c), and the active (causative) counterpart of the passive in (d):

\[(30)\]
\[
\begin{array}{ll}
\text{a. } & \text{Silnik } \text{był } \text{zepsuty.} \\
& \text{engine(M).nom } \text{was.3sg.m } \text{break.part.sg.m.nom} \\
& \text{‘The engine was broken.’} \\
\text{b. } & \text{Silnik } \text{zepsuł się.} \\
& \text{engine(M).nom } \text{broke.3sg.m } \text{refl} \\
& \text{‘The engine had/has broken.’} \\
\text{c. } & \text{Silnik } \text{był } \text{zepsuty } \text{przez } \text{mechanika.} \\
& \text{engine(M).nom } \text{was.3sg.m } \text{break.part.sg.m.nom } \text{by } \text{mechanic} \\
& \text{‘The engine was/got broken by the mechanic.’} \\
\text{d. } & \text{Mechanik } \text{zepsuł silnik.} \\
& \text{mechanic(M).nom } \text{broke.3sg.m } \text{engine(M).acc} \\
& \text{‘The mechanic broke the engine.’} \\
\end{array}
\]

Since (30a) cannot unambiguously be interpreted as corresponding to either the inchoative or the causative variant of zepsuć ‘break’, its morphological form must be considered underdetermined between an active and a passive syntactic structure.
I assume that, in principle, an impersonal -ne/-te construction may be similarly ambiguous. However, unfortunately, the same ambiguity test is not as easily applicable to the impersonal variant of the construction with być ‘be’ and the -ne/-te participle, because the impersonal -ne/-te construction overwhelmingly tends to be formed from semantically transitive predicates which imply an agent and therefore are unlikely to be construed inchoatively. The following attempt to construct sentences corresponding to the ones in (30) is marginally successful, though the test might not be similarly felicitous with verbs other than ‘tidy up’:

(31)  a.  
Było posprzątane.
    was.3SG.N tidy-up.PART.SG.N
    ‘[It] was cleaned/tidied up.’

b.  
Posprzątało się (samo).
tidied-up.3SG.N REFL (own-self.N)
    ‘[It] tidied up (by itself)/[The place] got tidied up (by itself).’

c.  
Było posprzątane przez Piotra.
    was.3SG.N tidy-up.PART.SG.N by Peter
    ‘[It] was tidied up by Peter.’

d.  
To Piotr posprzątał.
    this Peter(m).nom tidied-up.3SG.M
    ‘It was Peter who tidied up.’

4.2 Overlap with impersonal passives of intransitives

I consider passivisation to be a morphosyntactic alternation relating two predicates, an active one and its corresponding passive variant. More specifically, following syntactic analyses deriving from the Relational Grammar tradition, I analyse the passive as an operation on the argument structure of a predicate which downgrades the unergative argument, typically expressing an agent, to the grammatical function of an oblique. In transitive predicates, the argument expressing the patient or theme is granted the grammatical function of the subject of the passive clause. In intransitive predicates, which have no argument expressing a patient or theme, the resulting passive clause is subjectless. Although passivisation is not restricted to transitive predicates, it is applicable only to unergative predicates (as first argued by Perlmutter 1978). At first glance, it appears therefore that many, perhaps most, instances of the impersonal -ne/-te construction could be analysed as impersonal passive.

Although it seems obvious that the impersonal -ne/-te construction should feature in descriptions of the Polish passive construction, very few publications actually mention and illustrate Polish impersonal passives of intransitives at all. The notable exceptions include Lewicki (1964), Brajerski (1979), and more recently Słoń (2008).
Lewicki offers a list of Polish constructions expressing ‘activities and human states impersonally’, and one of the constructions included in the ‘impersonal predicates’ category is the impersonal passive. Lewicki remarks that this construction is found mostly in dialects, as well as in standard spoken Polish; it is rare in the written language (particularly in the academic prose) because it is ‘imprecise’. He cites two examples of impersonal passives from literary Polish (both are taken from Pisma wybrane, vol. 1, by Maria Dąbrowska, published in Warsaw in 1956), and emphasises that both instances are stylistic devices (1964: 316):

(32) a. Namawiał dalej, ażeby stąd wyjechała. (…) encouraged further so-that from-here departed

 Że nie wymówione? Ale wymówić można, kiedy chcąc.
that not given-notice but give-notice can when want
‘He was encouraging [her] further to leave this place. (…) That [it has] not [been] given notice? But one can give notice when one wants to.’

b. Mieli też niemało swoich domowych frasunków, na które w had also not-little own household problems to which in

 odezwach i namowach docierających do wsi nie było appeals and prompts arriving at village not was

 wskazane żadnej dobrej ani prędkiej rady.
indicated/offered any good or quick advice
‘They also had quite a few household problems of their own for
[the solution of] which [it] was not offered any good or quick advice
in the appeals and prompts arriving at the village.’

Without further diachronic study it is impossible to tell whether Lewicki was right in stating that the impersonal passive in Polish in the mid-20th century was indeed so severely restricted to dialects and stylisations. One possibility is that – rather like nowadays – the construction may have been more frequent than assumed, but somehow did not enter into the academic descriptions of the language; another is that other impersonal sentences with the -ne/-te participle (if there were indeed any) were classified as non-passive predicative adverbial constructions (see §4.3). It is also important to note that while example (32a) corresponds to many contemporary examples of the impersonal -ne/-te construction, such as (33), the same cannot be said of example (32b).

(33) Bo nie było powiedziane konkretnie. Pan minister wypowiedział because not was said specifically mister minister expressed

 się, że jeden z wariantów…
refl that one of options
‘Because [it] was not said specifically. The minister expressed himself
that one of the options…’
In contemporary Polish, example (32b) would instead have the form of:

(34) \[ \text{Nie było wskazanej żadnej dobrej rady.} \]
\[ \text{NEG was.3SG.N offer.PART.SG.F.GEN any.SG.F.GEN} \]
\[ \text{good SG.F.GEN advice(F).GEN} \]
\[ '\text{There was no good advice offered.'} \]

Example (34) is a personal clause with a ‘non-agreeing’ genitive subject. It illustrates the contemporary way of expressing existential negation, as opposed to standard or verbal negation exemplified in (35) which has a canonical nominative subject:

(35) \[ \text{Nie była wskazana żadna dobra rada.} \]
\[ \text{NEG was3SG.F offer.PART.SG.F.NOM any.SG.F.NOM} \]
\[ \text{good SG.F.NOM advice(F).NOM} \]
\[ '\text{No good advice was offered.'} \]

I hypothesise that the structure in (32b) is an instance of an earlier form of existential negation in the passive, where both the auxiliary and the participle were in the ‘default’ non-agreeing form. Although this form seems to have been replaced now by the participle agreeing with the noun in the genitive case, rare modern examples such as (36) found in the PELCRA corpus might perhaps be considered fossils of the earlier structure:

(36) \[ \text{Nie ma na tym napisane.} \]
\[ \text{NEG has on this write.PART.SG.N} \]
\[ '\text{There isn’t [anything] written on this.'} \]

I leave this interesting question for further study.

Brajerski (1979), in his article on the impersonal predicates ending in -no/-to, mentions the impersonal passive of the intransitive in Footnote 30 (p. 96). He argues that impersonal passives formed from perfective verbs, such as the ones below, have meanings very close to predicative adverbials (see §4.3):

(37) a. \[ \text{W piecach już napolone.} \]
\[ \text{in stoves already lit} \]
\[ '\text{[It has] already [been] lit in stoves.'} \]

b. \[ \text{Tam pod lasem już zaorane.} \]
\[ \text{there by forest already ploughed} \]
\[ '\text{There by the forest [it is] already ploughed.'} \]

c. \[ \text{W pokoju było posprzątane.} \]
\[ \text{in room was cleaned} \]
\[ '\text{In the room [it] was cleaned/tidied up.'} \]
On the other hand, he argues that impersonal passives formed from imperfective verbs have meanings which are clearly verbal:

(38) a. *W piecach jest teraz właśnie palone.*
   in stoves is now just lit
   ‘[It] is [being] lit in stoves just now.’

b. *Pod lasem jeszcze nie oranego.*
   by forest yet not ploughed
   ‘[It has] not yet [been] ploughed by the forest.’

c. *Proszę nie wchodzić, bo tu jest teraz sprzątane.*
   please not enter because here is now cleaned
   ‘Please do not come in, as [it] is [being] cleaned here now.’

Other authors and a recent corpus study of the Polish passive by Górski (2008) do not confirm Brajerski’s hypothesis of the simple division into a non-passive adverbial construction and a passive verbal construction on the basis of the aspect of the verb. But at least Brajerski explicitly mentions and illustrates the impersonal passive in his paper. In a large number of other studies of the impersonal passive, whether by Polish or foreign authors, the Polish impersonal passive is misrepresented by the -no/-to construction which, though impersonal, is not structurally or syntactically passive (see Kibort 2004, 2006, 2008 for detailed argumentation). The distinction between the impersonal -no/-to construction and the impersonal -ne/-te construction is rightfully recognised and approached from the cognitive grammar perspective in the recent work by Słoń (for example 2008).

The reason why impersonal passives are, or appear to be, less common in Polish than morphological impersonals is that both types of construction largely compete for the same communicative space. However, since morphological impersonals entail an agent which is either generic or indefinite, impersonal passives are found naturally in situations where the speaker wants to foreground an intransitively construed event. Consider the following selection of examples from this point of view:

(39) a. A: *Było już sprawdzane u Państwa?*
   was.3sg.N already check.part.sg.n at ladies-and-gentlemen
   B: *Tak, tak, było sprawdzane.*
   yes yes was.3sg.n check.part.sg.n
   ‘Has there already been checking at yours? Yes, yes, there has been checking.’ (an exchange between the ticket conductor and passengers on the train)

b. *Było codziennie sprzątane.*
   was.3sg.n daily clean.part.sg.n
   ‘There was cleaning every day.’ (from a customer review of a hotel)
It can be demonstrated that, unlike morphological -no/-to impersonals, impersonal passives do not have a subject which could participate in syntactic control or binding:

(40) a. *Było sprawdzane przejeżdżając przez Poznań.  
   was.3SG.N check.PART.SG.N pass.PART_CONTEMP  
   ‘There was checking [the tickets were checked] while passing through Poznań.’

cf. b. Było sprawdzane gdy przejeżdżaliśmy przez Poznań.  
   was.3SG.N check.PART.SG.N when passed.1PL.MHUM  
   ‘There was checking [the tickets were checked] while we were passing through Poznań.’

c. *Tutaj było palone czekając na egzamin.  
   here was.3SG.N smoke.PART.SG.N wait.PART_CONTEMP  
   ‘It was smoked here while waiting for the exam.’

(41) a. *Było codziennie sprzątane we wszystkich swoich pokojach.  
   was.3SG.N daily clean.PART.SG.N in all.PL.LOC own[REFL].PL.LOC rooms(NONMHUM).LOC  
   ‘There was cleaning every day in all of one’s own rooms.’
cf. b. Było codziennie sprzątane we wszystkich naszych/pokojach.  
was.3SG.N daily clean.PART.SG.N in all.PL.LOC our.LOC/their.LOC rooms(NONMHUM).LOC  
‘There was cleaning every day in all of our/their rooms.’

However, despite having no subject, impersonal passives have an agent which may be expressed overtly in a prepositional phrase, like in the personal passive:

(42) a. Było sprzątane, przez firmę.  
was.3SG.N clean.PART.SG.N by company  
‘The cleaning has already been done, by a [professional] company.’

b. Nie widać, żeby tutaj było sprzątane przez firmę.  
NEG see.[NONPERS] compl.[3SG] here was.3SG.N clean.PART.SG.N by company  
‘It does not look as if this place was cleaned by a [professional] company.’

c. Było sprawdzane, przez innego konduktora.  
was.3SG.N check.PART.SG.N by different conductor  
‘There has [already] been checking [of tickets], by a different conductor.’

d. Czy na tej ulicy już było sypane przez kogokolwiek?  
interr on this street already was.3SG.N grit.PART.SG.N by anyone  
‘Has there already been spreading [of grit] on this street by anyone?’

The overt expression of the oblique agent in the impersonal passive does not seem to be as easily acceptable as in the personal passive, but this is more likely to be due to pragmatic and information structure considerations rather than syntax. It has also been observed that impersonal passives ‘often have an implicitly human interpretation, which suggests that this interpretation is associated with subjectless forms of personal verbs, irrespective of the syntactic source of that subjectlessness’ (Blevins 2003: 489). However, unlike in the Polish -no/-to impersonal, but similar to the reflexive impersonal, this default human agent interpretation can be overridden in the impersonal passive, as is demonstrated in example (43):

(43) Ptak sprawdza każdą próbówkę. W tej było już sprawdzane.  
bird checks every test-tube in this-one was.3SG.N already check.PART.SG.N  
‘The bird checks every test tube [for food]. In this one the checking has already been done.’ (from a description of an experiment)
This is an important point which shows that the (frequently tacit) assumption about impersonal passives being restricted to the human agent interpretation is incorrect. This restriction may perhaps be a tendency, but it is certainly not universal.

Regarding the syntactic restriction on passives being limited to unergative predicates, but being inapplicable to unaccusative predicates, there is every evidence that this is upheld in impersonal passives. Questionable (i.e. apparently unaccusative) instances of impersonal passives could be analysed as non-passive predicative adverbial constructions. However, it appears that in practice -ne/-te participles used impersonally tend to be formed only from unergative verbs, so the question of potentially compromising the analysis of the passive does not seem to arise.8

Apart from the syntactic features of the passive construction discussed above, other features ascribed to the passive do not apply distinctively to this construction, therefore I do not elaborate on them here. One such feature is the apparent verbal as opposed to adjectival (or, adverbial, in the impersonal variant) character of the participle. On the basis of an extensive corpus study of the Polish personal passive, Górski (2008: 61–64) concludes that the same lexeme may have either a verbal or an adjectival interpretation in different contexts. Furthermore, he draws the same conclusion for the well-known distinction between stative and actional passive by demonstrating that sometimes it simply cannot be established (2008: 73–78). Similarly, aspectual differences do not seem to reflect any definitive structural differences between constructions. I have not researched in detail any aspectual issues pertaining to the Polish impersonal passive, but see Abraham and Leiss (2006a,b) for an extensive discussion of aspectual restrictions on impersonal passives in German, which may serve as a guide in the study of this area of Polish passives.

For these reasons, numerous Polish authors treat the passive as a purely syntactic phenomenon pertaining to the clause, rather than a morphological phenomenon pertaining to the verb (see e.g. Szupryczyńska 1973: 73–91; Laskowski 1998b: 194–197; Saloni & Świdziński 1998: 100, 320; Górski 2008: 49–50). My own analysis of the formation of resultative participles from verbs as a morphological derivation, and of the passive as a morphosyntactic operation on the argument structure of predicates, is compatible with the conclusions of these authors and locates the overlap between the ‘be’-passive and the adverbial predicative construction at the level of the formal expression of the two constructions.

8. Note that Abraham (current volume) offers an interesting discussion of the semantic and syntactic correlates of unaccusativeness in relation to impersonal constructions, both passive and non-passive (such as the Polish morphological impersonal ending in -no/-to).
4.3 Overlap with the predicative adverbial construction

Since the passive construction in Polish – as well as in English – uses the resultative participle as its main verb (with an auxiliary ‘be’ or ‘become’), the verbal complex in ‘be’-passives has an identical form to the copular construction with an adjectival/adverbial predicate. The following are garden variety examples of Polish sentences with a copula and a predicate adjective or adverb, respectively:

(44) a. Był zimny/ciemny/straszny/przyjemny.
    was.3sg.m cold/dark/horrible/pleasant.sgm
    ‘He/It was cold/dark/horrible/pleasant.’

    was.3sg.n coldly/darkly/horribly/pleasantly
    ‘[It] was cold/dark/horrible/pleasant.’

Standard descriptions of Polish, such as Laskowski (1998b: 187–197), state explicitly that jest czytany ‘is read.part.sg.m’ has the same form as jest młody ‘is young.sg.m’.

Frajzyngier (1978: 149–150) regards ‘be’-passives as formally a subclass of subject-complement clauses which he calls ‘nominal sentences’. He defines ‘nominal sentences’ as sentences with nominal predicates, or ‘copular constructions’, whose logical structure can be either X = Y (as in: Elizabeth II is the present Queen of England) or X ∈ Y (as in: Salt is white) (Frajzyngier cites both examples from Suppes 1957: 101). Frajzyngier argues that ‘be’-passives differ from other nominal sentences only in the fact that the predicate in ‘be’-passives (i.e. the participle following the copula/auxiliary) is morphologically derived from the lexical class of verbs, while in other nominal sentences it does not have to be so derived: ‘we might have languages, such as Semitic, in which not only verbal adjectives are derived from verbs but such nominal categories as agent, instrument, name of action and place of action’ (1978: 150). Therefore, formally there is no distinction between ‘be’-passives and other nominal sentences. Based on the analysis of a sample of over thirty languages chosen at random from several language families, Frajzyngier further points out that there are no languages that have ‘be’-passives but do not have nominal sentences formed with a copula. Moreover, the passive form in a language will contain the equivalent of ‘be’ only if the nominal sentence contains ‘be’. Finally, diachronic analysis shows that ‘be’-passives are, generally, more recent forms than other passives or statives. The most natural explanation of the similarity between ‘be’-passives and static nominal sentences is, therefore, that the former developed from the latter, and this happened because nominal sentences with a copula presented a suitable structure for the realisation of the passive. ‘In languages for which the ‘be’-passives are attested in the oldest available texts, one can claim that actually there is no distinction between be-passives and nominal sentences’ (Frajzyngier 1978: 154).
Returning to example (44), while (44a) is personal, it is uncontested that (44b) is genuinely impersonal, with no possibility of reconstructing any hypothetical omitted subject (e.g. infinitival). Although in any particular context it may be possible to interpret the adverb as a modifier relating to some more or less abstract referent which is present in the context of the utterance – for example, the ‘air’, ‘situation’, etc. – there is no trace of a subject or instigator in the syntactic or semantic structure of this clause.

Adverbs in modern Polish, such as those in (44b), have the ending -o/-ie distinct from the 3SG.N adjectival ending in -e. In contrast, -n/-t- participles used impersonally have the ‘default non-agreeing’ 3SG.N ending in -e. The -n/-t- participial form ending in -o does exist, but in modern Polish it has a different function: as mentioned earlier, the -no/-to participle is used (currently without an auxiliary) as the main verb of the morphological impersonal which does not accept an overt expression of a subject but has a syntactically active covert subject which participates in control, raising, and reflexive binding (see Kibort 2004, 2006, 2008). Bartnicka (1970:158) remarks that if the -n/-t- participles had the ending -o when used as predicate adverbs, they would be indistinguishable from the special impersonal forms. She hypothesises that this must have been the reason why -n/-t- participles used as predicate adverbs almost exclusively have the ‘default non-agreeing’ 3SG.N inflection -e.

The fact that -ne/-te participles in subjectless sentences can function as adverbs seems to be confirmed by the possibility of conjoining the two types of element. The following examples, found on the web, are unquestionably grammatical:

(45) a. Żeby było sprawiedliwie to napisać, że generalnie było czysto i posprzątane.
   so that was fair then write that generally was cleanly and tidied-up
   ‘In order that [it] is fair, I will write that generally [it] was clean and tidied up.’

b. W domu wszędzie było czysto i posprzątane…
   in home everywhere was cleanly and tidied-up
   wydawało się, że żyliśmy dostatnio.
   seemed refl that lived affluenty
   ‘Everywhere at home [it] was clean and tidied up… it looked like we were living in affluence.’

c. Raz tylko po sztormie były na plaży głiny, ale po południu już było czysto i posprzątane.
   once only after storm were on beach algae but after noon already was cleanly and tidied-up
   ‘Only once after a storm there were algae on the beach, but already in the afternoon [it] was clean and tidied up.’
d. Smutne... bo naprawdę się staram żeby było czysto i sad because really strive that was cleany and

posprzątane, ugotowane, ale dziecko jest tak absorbujące, że... tidied-up cooked but child is so absorbing that

‘[It's] sad... because I really make an effort so that [it] is clean and tidied up, cooked, but the child is so absorbing that...’

e. Mam wrażenie, że w całym domu mam brudno have impression that in whole house have dirty-ly

i nieposprzątane. and not-tidied-up

‘I have an impression that in the whole house I have [it] dirty and not tidied up.’

f. Jakiś gość zwrócił kelnerowi uwagę, że jest brudno i some fellow told waiter remark that is dirty-ly and

nieposprzątane ze stołu po poprzednich gośćach. not-tidied-up from table after previous guests

‘Some guy made a remark to the waiter that [it] is dirty and not tidied up from the table after the previous guests.’

g. Jak w domu masz brudno i nieposprzątane to kogo to when in home have dirty-ly and not-tidied-up then whose it

jest wina? is fault

‘When you have [it] dirty and not tidied up at home, then whose fault is it?’

h. Kiedyś tam byłem to pusto, cicho i pozamykane. sometime there was then empty-ly quiet-ly and closed

‘I was there some time ago and [it was] empty, quiet and [all] closed.’

i. Zbadał mnie na fotelu no i jak “dziewica” jestem, examined me on chair so and like virgin am

ciasno, cicho i pozamykane. tightly quiet-ly and closed

‘He examined me on the chair, and so I am like a “virgin”, [it is] tight, quiet and closed.’

The fact that the impersonal passive of the intransitive and the impersonal adverbial construction may both occur with ‘be’ makes it impossible to determine whether some sentences are passive or non-passive. I suggest that the best solution is to treat sentences with ‘be’ and resultative participles as underspecified with regard to which construction they instantiate, and for syntactic frameworks to provide an appropriate model of such an underspecification.
5. Location arguments in -ne/-te and predicative adverbial impersonals

In this section I briefly address the question of the grammatical status of the location arguments in the -ne/-te impersonal construction and in the predicative adverbial impersonal. I assume that prepositional phrases expressing locations of the foregrounded event or ambient quality in -ne/-te and predicative adverbial impersonals are oblique arguments rather than adjuncts. A detailed discussion of their syntactic status is beyond the scope of this chapter, however I want to emphasise that they should not be considered syntactic subjects of those sentences.

Impersonal sentences such as (46a–b), with oblique location arguments, clearly contrast with personal sentences in (47a–b) in which the location arguments are expressed as subjects:

(46) a. W pokoju było codziennie sprzątane.
in room(m).loc was.3sg.n every-day tidy,part.sgl.n
‘[It] has been cleaned every day in the room./There has been cleaning in the room every day.’
b. W pokoju było ładnie posprzątane.
in room(m).loc was.3sg.n nicely tidy-up,part.sgl.n
‘It was nicely cleaned/tidied up in the room.’

(47) a. Pokój był codziennie sprzątany.
room(m).nom was.3sg.m every-day tidy,part.sgl.m.nom
‘The room was cleaned every day.’
b. Pokój był ładnie posprzątany.
room(m).nom was.3sg.m nicely tidy-up,part.sgl.m.nom
‘The room was nicely cleaned/tidied up.’

In (47), both ‘be’ and the -n/-t- participle show agreement with the subject noun phrase, while in (46) such agreement is not established.

I demonstrated in §4.2 that impersonal passives do not have a subject which could participate in syntactic control and binding. The existence of a prepositional phrase expressing the location argument does not change this, since the prepositional phrase does not function as the grammatical subject of these sentences. Example (49) is repeated from (41):

(48) a. *W pokoju było sprzątane
in room(m).loc was.3sg.n tidy,part.sgl.n
ogładyając telewizję.
watch,part,contemp television(f).acc
‘There was cleaning in the room while watching television.’
It seems clear that the ‘cleaning’ event can be conceptualised in several different ways, and examples (46) and (47) demonstrate that the ‘room’ can be conceptualised as either a location or as a patient/theme of the ‘cleaning’. Thus, we observe an argument alternation which yields two different constructions. Nevertheless, when the ‘room’ is expressed through a prepositional phrase, it is not a subject.

6. Summary and conclusions

My goal for this chapter was to give an overview of the little studied Polish impersonal construction which uses the participle -n/-te bearing the singular neuter inflection -e.

I began by showing the general distribution of the -ne/-te participle in both personal and impersonal sentences and focused on the impersonal sentences such as:

(50) a. *Było codziennie sprawdzane we wszystkich swoich pokojach.

‘It has been checked./The checking has been done.’

b. Było codziennie sprawdzane we wszystkich naszych pokojach.

‘There was cleaning every day in all of one’s own rooms.’

c. The cleaning lady was cleaning in the room while watching television.’
b. _Jest sprawdzone._

is check.PART.SG.N

‘[It] is checked./[It] has been checked.’

On a communicative-functional view, the construction exemplified in (50) involves agent or instigator defocusing, while on a structural view it lacks a subject. More specifically, the impersonal _-ne/-te_ construction has neither an overt, nor an omitted or covert syntactic subject which could participate in syntactic operations such as control or binding. In other words, this construction is not only functionally impersonal, but also subjectless.

In terms of syntax, the impersonal _-ne/-te_ construction overlaps with both the impersonal passive of the intransitive, and with the impersonal predicative adverbial construction such as:

(51) _Było zimno._

was.3SG.N coldly

‘[It] was cold.’

In terms of form, the impersonal _-ne/-te_ construction is made up of a finite auxiliary verb or copula and a resultative participle. The resultative participle can be derived from both unergative and unaccusative verbs (both transitive and intransitive). If we came across an impersonal _-ne/-te_ construction with the resultative participle derived from an unaccusative verb, it would be analysed as a non-passive copular adverbial clause. However, there does not seem to be any functional motivation to produce such sentences. On the other hand, if we came across an impersonal _-ne/-te_ construction with the auxiliary _zostać/zostawać_ ‘become, get’ plus a resultative participle, it would be analysed as an impersonal passive clause, since it could be interpreted only as a passive construction alternating with a corresponding active.

Unfortunately, the situation is not so simple. The impersonal _-ne/-te_ construction is very common, but its instances which are most commonly found are ambiguous between the two interpretations and analyses. I therefore argue that, just like the form of the personal passive with ‘be’, the form of the impersonal _-ne/-te_ construction with ‘be’ and a resultative participle is underspecified with regard to which of the two syntactic constructions (passive verbal or non-passive adverbial) it instantiates. Hence, syntactic frameworks need to provide an appropriate model for this type of underspecification.

Finally, it is important to note that the impersonal _-ne/-te_ construction provides one more piece of clear evidence against constraints, proposed in most theoretical syntactic frameworks, that require all clauses to have subjects (including null or shared subjects).
Abbreviations

1 first person  M masculine
2 second person  MHUM masculine human
3 third person  N neuter
ACC accusative  NEG negative
COMPL complementiser  NOM nominative
DAT dative  NONMHAM non-masculine human
F feminine  NONPERS non-personal
FUT future  PART participle
GEN genitive  PART\textsubscript{CONTEMP} contemporaneous participle
INE inessive  PL plural
INF infinitive  PV preverbal affix
INTERR interrogative  REFL reflexive marker
IMPERSON impersonal  SG singular
LOC locative  VA (Hungarian) -\textit{va} suffix

References


Meteorological verbs in Uralic languages – are there any impersonal structures to be found*

Merja Salo
University of Helsinki

Finno-Ugric and Samoyed meteorological verbs form an interesting and insufficiently studied group with much internal diversity. Some verbs have zero valence, others display a more or less semantically faded subject, yet others feature an object. There are also expressions which have simply fossilized. With causative transitive verbs, the prevailing restriction seems to be that either a subject or object is possible, but not both. Purely nominal utterances and those consisting of a noun and the verb ‘to be’ have generally been disregarded. The study draws on data from 14 Uralic languages and as such constitutes the most comprehensive account to date of the expression meteorological phenomena in Uralic.

Keywords: weather verbs; Uralic; aivalent verbs; varying valencies; cognate constructions; impersonal (?)

1. Introduction

The Uralic peoples form a medium-sized community of around thirty million people. For the most part, they live scattered across a vast area stretching from Scandinavia to the Taimyr Peninsula in Central Siberia and from the Carpathian basin to the Arctic Ocean. The strongest link between Uralic peoples is linguistic; they do not share a unified culture. Finns, Estonians and the speakers of minor Balto-Finnic languages live in Balto-Scandia, while the Mordvins, Mari, Udmurts and southern Komi belong to the Central Russian agricultural area. The Saami, Mansi, Khanty, northern Komi

* I am very grateful to the staff of the Department of Finno-Ugrian and Nordic Studies at the University of Helsinki and several native speakers of various Uralic languages for their many useful comments and additions. Some of the comments have been added to the text itself. I would also like to thank Raija Bartens, Pål Eriksen, Ulla-Maija Forsberg (née Kulonen), Leena Kolehmainen, and the editors of this volume for reading and improving my paper and Leena Huima for technical assistance and Kimberli Mäkäräinen for revising it.
and Samoyeds share many common arctic or subarctic features in their traditional cultures, while the Hungarian civilization is typically East Central European. Since these groups have lived apart from the rest of their linguistic relatives for so long, they have continuously borrowed and adopted various linguistic features from the non-Uralic groups around them. On the European side, the Indo-European language family dominates, since the Finno-Permic branch and Hungarian are surrounded by people speaking Germanic, Slavic or Turkic languages. The indigenous people in Siberia, however, belong to ten language families of which Turkic, Tungusic, Yeniseic and nowadays Slavic languages are widely spoken in and around the places that the Ob-Ugric and Samoyed peoples live in.

Unlike in the Indo-European languages to the west, impersonal meteorological or physiological structures consisting of a verb and some kind of pronominal “dummy” subject have hardly ever been attested in any of the Uralic literary languages. Furthermore, impersonal passives with pronoun subjects are almost unknown. This is not to say, however, that the Uralic languages lack meteorological constructions. On the contrary, such constructions are very well represented in these languages. In fact, Uralic meteorological verbs form an interesting group that exhibits great internal diversity. Some verbs have zero valence, others display a more or less semantically faded subject, yet others feature an object. In addition, there are expressions that have simply fossilized over time. With causative transitive verbs, the prevailing restriction seems to be that either a subject or object is possible but not both.

This paper is organized as follows: In §2, I will provide a brief overview of the nature and origins of the basic meteorological vocabulary found in these languages. The following sections will discuss the different argument structure patterns found when expressing meteorological phenomena in Uralic. §3 will be devoted to what is often considered to be the primary meteorological construction used in Uralic, avalent constructions featuring avalent predicates. Then in §4, I will discuss a range of constructions that do feature an argument, albeit one that is often optional. §5 will focus on the subject in meteorological constructions, its origins in mythological concepts and its actual expression. The discussion will close with some concluding remarks in §6.

Given that the Uralic languages are less well known to the general linguistic public than Indo-European languages are, I will first give a short overview of the classification and genetic relationships of the 14 languages of the family to be discussed in this study.

The traditional genetic tree of Uralic places the origins of the family at a time depth of 5000 years before the present (Janhunen 2009:68). The deepest dividing

1. Updated information on the members of the Uralic language family: http://www.helsinki.fi/~tasalmin/fu.html
line within the family is assumed to exist between western Finno-Ugric and eastern Samoyed, of which Tundra Nenets, Selkup, Nganasan and Kamas – all spoken in Siberia – are relatively well documented. Very little material is available on the other members of the Samoyed branch (Enets, Yurats and Mator). The easternmost branch of the Finno-Ugric languages is made up of the Ugric languages Khanty, Mansi and Hungarian; the first two, spoken in Siberia, form the Ob-Ugric sub-branch. It is, however, generally assumed that Hungarian maintained a connection with Mansi longer than it did with Khanty. In reality, though, the Ob-Ugric languages form a continuum of several dialects that vary greatly grammatically and lexically. For example, the Khanty language nowadays has four more or less standardized literary languages: two Eastern and two Northern variants. It is not impossible that there has even existed a common Ugric protolanguage for all these entities. The remaining Finno-Ugric languages to be presented in this study are spoken in Europe. The Permic languages, Komi in the north and Udmurt in Central Russia, are linguistically relatively unified and show no major dialectal differences. The next languages geographically west of Udmurt are the Volgaic languages Mari and Mordvin. Mari has two literary variants Hill and Meadow Mari (or Western and Eastern Mari according to their own terminology) as well as Mordvin: Moksha in the west and Erzya in the east. The existence of a Volgaic protolanguage has been rejected and nowadays it is generally assumed that Mari broke away from the Finno-Volgaic branch on its own, followed by Mordvin and then by Saami. Janhunen (2009: 65) even suggests that Mordvin might be more closely related to the Finnic languages than Saamic.

At present, the Saami languages form a dialect continuum of nine variants spanning four countries: Norway, Sweden, Finland and Russia. Although neighbouring dialects are relatively close to each other, the main Saami dialects are so different that they warrant being called separate languages. There are six literary Saami languages, of which five use a Latin-based alphabet and one a variant of Cyrillic. The largest language in terms of number of speakers is North Saami. Lule, Kildin, South, Inari, and Skolt Saami have fewer speakers than North Saami does, while Ume, Pite and Ter Saami are spoken essentially only by the older generations. Several Saami languages are already extinct (Sammallahti 1998: 1–2). The main Saami language referred to in this study is North Saami, although some comments will also be made on the other Saami languages.

The Finnic languages along the shores of the Baltic Sea have seen the development of many written standards. Purely from the linguistic point of view, the Finnic branch is a dialect continuum comparable to Khanty, Saami or even possibly the Selkup dialects. It is usually divided into at least seven languages: Livonian, Estonian, Votic, Ingrian, Finnish, Karelian and Veps. The Southern Finnic group comprises Livonian, Votic, and Estonian, including South Estonian. This group, which evolved by the end of the first millennium, shares a number of common innovations, of which the
change of their morphology into the fusional type may be regarded as central (Pajusalu 2009:95–96). The Northern Finnic group comprises Finnish, Karelian and four of its dialects, Ingrian and Veps. For this study, Finnish and Estonian have been chosen to represent the Finnic language branch.

2. **Uralic meteorological verbs**

2.1 **Previous studies**

So far, there has been only one relatively large comparative study of meteorological verbs in Uralic, namely that of Bartens (1995). It covers material from nine Finno-Ugric languages and constitutes a rich source of important information on the topic. A much smaller comparative study contrasting Finnish and Estonian meteorological utterances with those found in some neighbouring Indo-European languages has been carried out by Hakanen (2001). Saarinen (1997) has explored the single expression ‘it rains’ in various European languages; this study, however, is restricted geographically in that it only includes languages from Saami to Komi. The most recent study of meteorological verbs is that of Kolehmainen (2010a),2 which is confined to Finnish. It documents that one and the same participant can be expressed with a single meteorological verb in many ways: “grammar encodes different types of conceptualization, not different types of situation, and the same situation can be conceptualized in different ways” (Wierzbicka 1996:423). One cross-linguistic investigation of meteorological expressions in different language families has also just been published (Eriksen-Kittilä-Kolehmainen 2010).

2.2 **Basic Uralic meteorological words and their origins**

While some parts of the basic meteorological vocabulary are usually viewed as being old and having originated with earlier protolanguages, other lexical items have been borrowed relatively recently. There are also words whose origin we know nothing about. Quite a few words, however, do not have any information pertaining to their origin. Since there has always been reason to talk and mostly complain about the weather in the northern part of Eurasia, meteorological expressions constitute an important feature of the vocabulary of these gatherers, hunters, fishermen, reindeer-breeders, and farmers. At the same time, meteorological vocabulary is inclined to change, especially when the language community stops being monolingual. Bilingual individuals easily import words and structures from other languages.

---

2. Another study by the same author (Kolehmainen 2010b) compares Finnish constructions with German ones and concludes that they are quite similar.
In view of the above, relatively few meteorological words enjoy a wide distribution among the various branches of Uralic languages. Very often their occurrences are dispersed and gaps are filled with words belonging to newer layers of the vocabulary. For example, the Finno-Ugric word *pilwe ‘cloud’ is missing from the Samoyed languages, but on the Finno-Ugric side it is well established (UEW I 381, SSA II 367). The word *päjwä ‘fire’, which at least dates back to Finno-Ugric and possibly all the way back to Uralic, has its counterparts in the Finnish päivä and the North Saami beaivi meaning ‘day, sun’, and possibly in the Komi bi ‘fire’ and the Khanty paj ‘thunder, stormy weather’. Nenets and Selkup have a verb meaning ‘to warm, to warm up’, which originates with the Proto-Samoyed word *pejwä ‘warm, warmth’ (UEW I 360, SSA II 456), that might be connected to these words, too. The meaning ‘day, sun’ is currently represented by many different stems: for the Finnish novelty aurinko ‘sun’ it is difficult to find other counterparts, one uncertain comparison has been made to the Saami languages and so it has remained without a good etymology (SSA I 90). The Volgaic nouns Erzya Mordvin či, Moksha Mordvin ši and Mari keče are connected to the Finno-Ugric word *kečä meaning ‘circle, round, ring’, which in other languages is widely known as a loop in a snare (UEW I 141–142, SSA I 337). The Udmurt nunal ‘day’ and Komi lun ‘id.’ have been connected to the Finnish words louna ‘southwest, mida’ and lounas ‘lunch’ (UEW II 693, SSA II 97). In the Samoyed languages there are two common etyma for this heavenly body: *jälä ‘light, day, sun’ now meaning in the four languages treated here ‘sun’ and *kâjâ ‘sun’ having the same meaning in Nenets xajer” and in Kamas kuja (Janhunen 1977: 40, 58).

The Finnic word sää ‘weather, air; storm’ has only one solid counterpart, the Hungarian ég ‘sky’ (obsol.) air’ which entitles it to be regarded as the Finno-Ugric *sâñe ‘air’ (UEW I 435, SSA III 244). The Finnic ilma is has an extremely broad semantic range, meaning ‘air, weather, storm; world; cardinal point’; the same range also occurs in the Saami albmi. (There is also another Saami word, belonging to this domain, ilmi ‘weather’, although this is a Finnish loanword.) On the other hand, its range in the Permic languages has started to yeer off in a different direction: Udmurt in (: inm-) ‘sky’ and Komi jen (: Jenner) ‘god’. Komi also has the compound word jen-ež ‘sky’, where the second part has the meaning of ‘skin, covering’. In Ob-Ugric, the original meaning has been better preserved: Mansi ēlәm ‘weather’ and Khanty jelәm ‘world; cardinal point’ have no cognates in their Eastern dialects (UEW I 81–82, SSA I 224–225). Here, though, a case of internal borrowing within the language family has been attested: the Northern Mansi word jelәm is borrowed from its neighbour, Northern Khanty. Some Saami languages have borrowed the Finnic word taivas ‘sky’. The origin of this word is somewhat controversial. It is quite likely that it originates with the Baltic languages, from the proto-Baltic *deivas ‘god’ (SSA III 256). The Finnic word jumala ‘god; wise man, sorcerer’ has a Finno-Volgaic distribution. In Mordvin, it exists as a part of the name of one god in the pantheon and in a compound word jondol ‘lightning’.
(act. ‘god’s fire’). In Mari, it exists in jumo ‘sky; uppergod, picture of god, icon’ (UEW II 638, SSA I 247 which contradicts its proposed Aryan origin).

The Finno-Ugric *jäŋe ‘ice’ has kept its meaning well in the other daughter languages except in Khanty, where in most of the dialects, it has taken on the new meaning of ‘water’, which is still close to the Finno-Ugric meaning. Similar names for ice also exist in neighbouring languages belonging to other language families (UEW I 93, SSA I 260). The counterparts for the Uralic word *lume ‘snow’ are missing in the Ob-Ugric languages and on the Samoyed side they mean ‘snowfall’ (UEW I 253–254, SSA I 102). Interestingly enough, the only weather verb for which it is possible to reconstruct an Indo-European etymon is ‘snowing’, even though five different roots represent the words in the contemporary daughter languages (Bauer 2000:99). Of the several nouns meaning ‘snowstorm’, one seems to be quite old, *purku, which is rich in nominal and verbal counterparts in most of the Uralic languages. This word has been borrowed by Russian from the Finnic languages and by Komi via Russian. There are still doubts that some of these Uralic words might have separate descriptive origins (UEW I 406, SSA II 436). Some etymological sources consider the Finno-Permic noun-verb *tule- ‘wind; blow’ to be even older, possibly Uralic *tuxlj, in which case it might originate with the proto-Indo-European *dhuH-li ‘blowing, dusty’. It is missing in Saami, Mordvin, Mansi, Khanty and Hungarian (UEW II 800, SSA III 340–341).

One of the oldest words in the Uralic vocabulary is definitely the Finnish verb sataa ‘to rain’, which also used to mean ‘to fall down; to yield a crop’. It has two possible counterparts in Saami. Its next counterparts are on the other side of the language family, in the Samoyed branch, where it enjoys a wide distribution among all nine of the known Samoyed languages (SSA III 159–160, missing in UEW). On the Indo-European side, the impersonal verb for ‘to rain’ displays six different roots across the daughter languages (Bauer 2000:98). Some meteorological nouns along with their areal distribution are presented at the end of this study in Table 3.

3. Basically avalent constructions

As mentioned in the introduction, the Uralic languages are rich in meteorological constructions that feature avalent predicates. As we shall see below, some of these may co-occur with various locational expressions and others with facultative adjuncts and even without any arguments.

3.1 Purely verbal structure, avalent verbs

3.1.1 Finnish

Since the early 20th century, the Finnish grammatical tradition has favoured the term ‘unipersonality’ over ‘impersonality’ (Helasvuo & Vilkuna 2008:219). The concept
of unipersonality is more strictly oriented to grammatical form than impersonality. Unipersonal verbs appear in the 3rd person singular, and even some multipersonal verbs are sometimes used unipersonally. According to Airila (1928:3), these verbs are less common in Estonian or Hungarian. The number of these extremely reduced utterances is always very limited in the language. There are two types of unipersonal verbs, of which only a part of the second, smaller group is discussed here. These unipersonal verbs or predicates exclude any subject argument, thus instantiating ‘deep’ impersonality. These include such denominal verbs or old noun-verbs related to the weather verbs as tuulee '[the wind] is blowing' ~ tuuli ‘wind’, myrskyää ‘[the wind] is blowing hard’ ~ myrsky ‘storm’, ukkostaa ‘it’s thundering’ ~ ukkonen ‘thunder’, salamoi ‘it’s lightning’ < salama ‘lightning’, tyyntyy ‘[the wind/sea] calms down’ < tynni ‘calm’, pakastaa or pakastuu ‘it’s getting colder’ ~ pakkonen ‘frost’, kylmenee ‘it’s getting colder’ < kylmä ‘cold’, viileenee or viilentyy ‘it’s cooling down’ < viileä ‘cool’, lämpeenee ‘it’s getting warmer’ < lämmin ‘warm’, pyryttää or tuiskattaa ‘it’s snowing heavily’ < pyry or tuisku ‘snowstorm’, nietostaa or kinostaa ‘snowdrifts are forming’ < nietos or kinos ‘snowdrift’, pilvistyy ‘it’s getting cloudy’ < pilvi ‘cloud’, pouantaantu ‘the rain stops, the weather turns dry’ < pouta ‘rainless, dry weather’, kirkastuu, selkenee or seestyy ‘it’s clearing up’ < kirkas, selkeä, sees ‘clear’, as well as certain verbs or predicates denoting emotions or physical sensations.

3.1.2 Estonian

The grammar of Estonian resembles the grammar of Finnish in many respects; the greatest differences are in the vocabulary. Estonian meteorological expressions have rarely been the focus of a separate study and even grammatical sources are very laconic about the subject, considering that only two examples were provided in the new comprehensive handbook dedicated to the language. They were presented under the title Aluseta laused ‘Subjectless sentences’ (Erelt-Erelt-Ross 2007:470–471). Purely verbal sentences are: tormab ‘it’s storming’, müristab ‘it’s thundering’, külmetab ‘it’s getting colder’, hämärdub ‘it’s getting dark’ (Hakanen 2001:64). Estonian has no verb corresponding to the Finnish tuulla ‘blow’, instead it has an expression ‘to be windy’, e.g. välas on tuuline ‘outdoors, it’s windy’ (SES 651).

3.1.3 Saami languages

In North Saami, as a rule, formal subjects are not used with verbs describing natural phenomena: arvá ‘it’s raining’ ~ arvi ‘rain’, muohttá ‘it’s snowing’ ~ muolta ‘snow’, bieggá ‘the wind is blowing’ ~ biegga ‘wind’, beaivvádahtii ‘it got sunny’ < beaivi ‘day, sun’, sevnnjodii ‘it was getting dark’ < seavdnjat ‘dark, darkness’ (Nickel 1994:397; Sammallahti 1998:96). Inari Saami has also plenty of denominal examples: polviiškiδij ‘it had started to get cloudy’ < polva ‘cloud’, te taalvaj ‘it’s hat gewintert/winter has
arrived’ < tälvi ‘winter’, lijim vyelgimin, mut ko purgij ‘I wanted to leave, but it began to snow heavily’ < purga ‘heavy snowfall’ (IW entries 3405, 4901 and 3538).

3.1.4 Mordvin languages

There are very few cases where Mordvin verbs can form a clause by themselves. The following cases can be found in Moksha šopәd'i/Erzya čopoťi or čopod'i ‘it’s growing dark’ ~ E ěpod'a ‘dark’, E sundergali ‘id.’, E kelmevt'i/M kelmafti ‘it’s cooling down’ < E kelme/M kel'ma ‘cold’, E ašolgadi/M valdaškadi ‘dawn is breaking’ < E ašo ‘white’, M valda ‘light, clear’, M vaźmәdi ‘id.’. Larger grammatical descriptions occasionally save a few pages for this matter as Koljadënkov did (1954: 80–83; 1959: 199), when he calls these sentences impersonals (bezličnye) following the Russian linguistic tradition. Alëškina has also listed several types of impersonal sentences, including meteorological verbs, which she classifies into three groups: (1) change(s) in weather condition; (2) atmospheric-meteorological phenomena; (3) natural phenomena and processes of the environment as manifestations of natural forces (Alëškina 1973: 28).

3.1.5 Mari

Impersonal structures in Mari have been little studied. It is clear, though, that meteorological expressions can consist of a single verb such as jüštemeš ‘it’s getting cold’, jükšemeš ‘id.’, jükšemda ‘id.’ < jüstö ‘cold’, jualya⁴ ‘it is cooling down’ < ju, jualye ‘cool’, kõlmәkta ‘it is freezing’ < kõlme ‘frozen, frostbitten’, kasešteš ‘the night is falling’ < jas ‘evening’, mardežla ‘the wind is picking up’ < mardež ‘wind’, pәčkemәšteš ‘dawn is breaking’, pәlyoŋdo ‘light’, pәckemәšalteš ‘it’s getting late’ < pәckemәš ‘dark, dusky’, rïmbalya ‘it’s getting late’ < rïmbalye ‘dark, dusky’, ojara ‘it’s clearing up’ ojarlana ‘id.’, ojarešteš ‘id.’ < ojar ‘clear weather’ [loanword from Chuvash], jonɡә dәmes ‘it’s clearing up’ < jonɡәdo ‘wide, cloudless, clear, fresh’, lebešta ‘it’s warming up’, lәβәrta ‘id.’ < lebe ‘warm’, umәrta ‘it’s warming up’ < umәr ‘warm, clear, calm (weather)’, (Učaev 1956: 75–77; Bartens 1995: 44).

3.1.6 Udmurt

The oldest studies dealing with impersonal sentences in Udmurt date back approximately 70 years, one of which could be found in storage at the University of Helsinki

3. This verb has an interesting derivational suffix t. More information in M. Salo 2010.

4. In translitterating Mari – as well as the other languages studied – the Finno-Ugric scientific phonological system has been followed. In the case of Mari, this means that the Cyrillic alphabet has not been slavishly transliterated, as the voiced Cyrillic stops b, d, g are actually voiced fricatives β, δ, γ respectively, and Cyrillic back-i is actually the reduced vowel ă.
Meteorological verbs in Uralic languages

Library. The author studied quite a few older grammars in German and gave examples of impersonal sentences in different Indo-European languages. He reduces the number of impersonal sentence types in Udmurt to four, of which natural phenomena constitute one minor group (Alatyrev 1941:13). The same division also occurs in his grammar later on (Alatyrev 1970:199). It has been calculated that Udmurt has approximately 30 avalent meteorological verbs, of which many can take an informal subject anyway, which makes their classification rather troublesome (Pozdeev 1973:263). Subjectless verbs include: lişim ja ‘it’s snowing’ < lişim ‘snow’, dżarde ‘dawn is breaking’ ~ dżardon ‘dawn’, sakte ‘dawn is breaking’ ~ sakton ‘dawn’, těla ‘the wind is blowing’, tělaške ‘id.’ ~ těla ‘wind’, vořekja ‘the stars are twinkling’, džonekte ‘it’s getting dimmer’ ~ džomit ‘dark, dim’, pejmița ‘it’s getting dark’ ~ pejmiț ‘dark, dim’ (Bartens 1995:46–47).

3.1.7 Komi

At least two studies focusing on Komi impersonal verbs and their argument structure have been written (Ievleva 1984; Fedina 1999). The author of the latter study enumerates four groups of impersonal verbs, of which two meteorological verbs and verbs describing the physiological or psychological state of a human being or his organs are more archaic (Fedina 1999:163). The author of the former study discusses a few meteorological verbs under the subheading avalentnye bezličnye glagoly ‘avalent impersonal verbs’ and declares them to be semantically narrow, even though they can also be used figuratively (Ievleva 1984:70–74). She has also noticed that many Komi impersonal denominal verbs have similar denominal counterparts in Russian, e.g. čaremalis ~ zanastilo ‘it was frozen [of the surface of snow]’, liismaalis ~ zarosilo ‘the dew was falling’. Such cases are not rare, especially when the verb has incorporated a qualifier, e.g. vojtevtis ‘the northern wind had risen’ < voj ‘northern; night’ + těvtni ‘blow [of a wind]’ ~ zasevernovetrilo (Ievleva 1984:71–72). In many respects, Komi and Udmurt are very similar. More than 50 avalent meteorological verbs have been documented in the Komi literary language, not including the rarer verbs found in various dialects (Ievleva 1984:71), e.g.; jugde ‘it’s dawning’, pendë ‘it’s getting late’, remde ‘id., zerë ‘it’s raining’, limjalë ‘it’s snowing’, těvžise ‘the wind has started to blow’, kędzedede ‘it’s getting cold’, irknițë ‘[the air] is starting to be crisp’, pužjale ‘[the ground] is frosting over’, purkedë ‘[the wind] is picking up [snow]’, sile ‘it’s melting’, soneđed ‘it’s getting warm’, pęže ‘it’s hot’, (Saharova 1967:150; Ludykova 1983:69; Bartens 1995:49).

5. The Komi chapter has been revised and commented on by researcher Paula Kokkonen, an expert in Komi, additions are made by Galina Misharina, doctoral student and specialist in Komi folklore at the University of Helsinki.
3.1.8 Northern Khanty

Eleven avalent verbs have been documented in Northern Khanty (Čeremisina-Solovar 1990: 22–23). Of these, nine are meteorological, e.g. rimxәmtәl ‘it’s getting dark’ < rimax ‘darkness’, pătlә or pătlәtә ‘id. (mom.)’ ~ pătlam ‘darkness; dark’, nuwemәt ‘dawn is breaking’ < nuwi ‘white, light’, iškәtәl ‘it’s getting cold’ < iški ‘cold’, jılәtәl ‘it’s cooling down’ < jıl ‘cool’, melkәtәl ‘it’s getting warmer’ < melәk ‘warm’, tewәntәl ‘it’s calming down [a wind]’, < tewәn ‘calm, windless’, etәrmәtәl ‘it’s clearing up (and also getting colder)’ < etәr ‘clear’, jәrtәl ‘it’s raining’ ~ jәrt ‘rain’. The remaining two verbs are used to describe the effects of the weather or human activities in his dwelling: jәtxәmtәl ‘it’s getting cooler [in a dwelling]’ < jәtx ‘cool’ and xoxәtәl ‘it’s clattering [because of the heating]’. Of these three – xoxәtti, nuwemәti and rimxәmtti do not take any subject argument (Čeremisina-Solovar 1990: 23). Nearly all of them, except the last one and the verb jәrtti ‘to rain’ which is a typical old noun-verb case having no derivational suffix in either of the forms, have a denominal origin.

3.1.9 Northern Mansi

The leading figure in Mansi studies, Rombandeeva, states that verbs describing natural phenomena are often old denominal verbs or old noun-verbs, e.g.: īțimli ‘night is coming’ < īți ‘night’, wōti ‘it’s blowing’, wōtiślәnti ‘it’s blowing gently’ < wōt ‘wind’ (a loanword from the Iranian languages), rakwi ‘it’s raining’ < rakw ‘rain’, tūji ‘it’s snowing’ < tūjt ‘snow’, satumli ‘it’s getting dark’ < satum ‘dark [of a day]’, turmanli ‘it’s getting dark’ < turman ‘dark’, tultili ‘to get cloudy’ < tul ‘cloud’, solи ‘it’s covered with frost’ < sol ‘hoarfrost, [white] frost’, polи ‘it has crusted over’ < pol ‘the frozen surface of snow’, xәtәli ‘dawn is breaking’ < xәtol ‘sun, day’, postиli ‘dawn is breaking’ < pos ‘light’ (Rombandeeva 1979: 35, 50, 117; Rombandeeva-Vahruševa 1989: 101–102; Skribnik-Afanas’eva 2004: 53). These sentences have previously been considered impersonals and can still be considered such, even though Rombandeeva (1979: 50) believes them

6. All the Northern Khanty words from different sources have been standardized to the phonemic transcription of the Kazym dialect. Vowel shortness is marked with a breve above the letter, ә represents the reduced vowel, ә an open front vowel (which distinguishes the Kazym dialect from the Shuryshkair dialect), ɪ the voiceless lateral, and x the voiceless velar fricative. All of the examples have been approved by two native speakers, researcher Andrej Kaksin, a recognized specialist of Kazym Khanty, from the Ob-Ugrian Research Institute in Khanty-Mansiysk and Professor Evdokiya Nëmysova, a well known pedagog from the Yugra State University in Khanty-Mansiysk.

7. The examples from different sources have been slightly unified according to the commonly accepted phonemic system of Northern Mansi (see e.g. Skribnik–Afanas’eva 2004: 9–13). x represents the voiceless velar fricative, ɣ the voiced velar fricative not occurring word-initially,ŋ the velar nasal occurring only medially and in the word-final position.
to be personals and completes. In her opinion, the subject has been omitted in short sentences in order to avoid tautology, as all of the necessary information is already embedded in the predicate verb.

3.1.10  Hungarian

Hungarian does not have many purely avalent meteorological verbs; some examples include fagy ‘it’s getting cold’ ~ fagy ‘cold, frost’, havazik ‘it’s snowing’ < hó [ː hava-] ‘snow’, villámlik ‘it’s lightning’ < villám ‘light’, viharzik ‘a gale is blowing’ < vihar ‘storm’, hajnalodik ‘dawn is breaking’ < hajnal ‘dawn’, alkonyodik ‘it’s getting dark’ < alkony ‘dusk, evening’, [be]esteledik ‘it’s getting late’ < este ‘evening’. Many of these denominal verbs end in -ik, which is the marker of intransitivity. Most of these verbs can take an argument as a subject. When used metaphorically, the first group can also take a subject: villámlott a szeme ‘eyes (sg in dual meaning) flashed fire’ or dörgött a hangja ‘his voice thundered’ (Kugler 2000: 409).

3.1.11  Tundra Nenets8

So far, Samoyedologists have concentrated on questions of the complex morphology and morphophonology of the language, while overlooking the syntax, in spite of it also being a very relevant topic. All Nenets verbs can be divided into two groups according to their temporal relations. For momentaneous or perfective verbs, the indicative aorist expresses the immediate past, and the indicative preterite the more remote past. For continuous or imperfective verbs, the indicative aorist expresses the present and the indicative preterite the simple past. The narrative mood essentially expresses the opposition between the perfect and the pluperfect (Salminen 1998: 531). Nenets impersonals can be considered verbs, e.g. jalumź9 ‘to dawn’ (> jalumdz10 ‘dawn’), pewśumź ‘to get dark’ (> pewśumbi ‘evening’) < pewă(ś), ‘to be dark’, tīdělěć ‘to dusk’, xadumź ‘to begin to snow heavily’ < xad ‘snowstorm’ and their various derivatives such as jalumdźană(ś),

8. Information on Nenets and the other Samoyed languages has been revised and improved by the prominent Samoyedologist Tapani Salminen, Academy Research Fellow and university lecturer at the University of Helsinki.

9. The examples are transliterated by the author from Cyrillic sources; their morphological analyses have been revised by two specialists in Nenets: doctoral student Lotta Jalava, and visiting researcher Roza Laptander, both from the Finno-Ugrian language Studies at the University of Helsinki.

10. The Nenets glottal stop has a dual marking. In the normative orthography, it is written with separate letters for the different sandhi variants: <”> or in some sources q means non-nasalizable glottal stop, <’> or in some sources h means nasalizable glottal stop (more closely in Salminen 1998: 522–526).
'to be dawning', *jalumdānalč* ‘to begin to dawn’ and many others. The usual tempus is aorist, as in *jalumdāna* ‘dawn was breaking’ (Tereščenko 1973: 108).

### 3.1.12 Selkup

The Selkup language area covers much of the vast West Siberian taiga between the Ob and Yenisei rivers. Although the number of Selkup speakers is relatively small, they are divided into three distinct groups. Selkup-Russian bilingualism is widespread, as well as the knowledge of other neighbouring Siberian indigenous languages. Most of the ethnic Selkups in the younger generations are Russian monolinguals. Up to the beginning of the last century, Selkup was used as a lingua franca, and thus also spoken by non-Selkups. According to Helimski (1998b: 549), this may have contributed to the relative simplicity of its grammatical system in comparison with the other Samoyed languages. Structurally, Selkup seems closer to the Ugric languages in many respects than it does to its northern Samoyed relatives; this may be at least partly due to secondary contacts between Selkup and Ob-Ugric (Helimski 1998b: 550). Impersonal verbs are used in Selkup: *lipkimšt̪qo*\(^{11}\) ‘to get dark’, *čēlimpiqo* ‘to dawn’, *tajamejgo* ‘to turn into summer’, *tūrimpīqo* ‘to come (the time)’ etc. (Kuznecova & al. 2002: 318–319). The usual tempus is aorist, as in *sōra* ‘it’s raining’ (Kuznecova & al. 2002: 23) and *merka* ‘the wind is blowing’ (Helimski 1998b: 553).

### 3.1.13 Nganasan

The Nganasan language is spoken by the northernmost ethnic group in Eurasia. Due to the intense Russification of younger generations, the total number of Nganasan native speakers only totalled a few hundred in 1990. Many Nganasans are or were able to speak Dolgan, Yakut, Enets or Evenki, the languages of their neighbours (Helimski 1998a: 480–481). In Nganasan, impersonal denominal verbs are used, e.g. *d'irbamsa* ‘to become frosty’ < *d'irba* ‘frost’, *xojmәgimsa* ‘to darken’ < *xojmәga* ‘dark’, *kotumsa* ‘to begin to snow heavily’, < *kōdu* ‘snowstorm’, *katgimša* ‘to dawn’ < *katgә* ‘light (adj.)’, *česәgimši* ‘to get colder’ < *česәga* ‘cold’ (Tereščenko 1973: 108, 245–246). Although avalent verbs are not common in Nganasan textbooks, some examples do exist; for instance, *katgimšt̪tu* ‘dawn is breaking’ (Tereščenko 1973: 109) is used to present the present tense.

### 3.1.14 Kamas

Information about Kamas is provided in §5.1.

### 3.2 Avalent verbs with expanded valencies: Temporal and local adverbs

It is generally known that Finnish avoids verb-initial sentences and that almost any nominal constituent can assume this crucial clause initial position, which is why

---

11. The Selkup examples are revised by Jarmo Alatalo, a specialist in Selkup.
weather verbs often take an adverbial in front of them. The adverbial can be temporal or locative. Syntactically, these structures are similar to existential sentences (Hakanen 2001:65). The same word order seems to apply to most of the Uralic languages. It is possible that they have been influenced to some degree by Russian absolute-impersonal verbs that never appear in any form but the 3.sg (cf. the examples in Leinonen 1985:80–81). The most common sentence pattern in Leinonen's material is an active verb with an optional locative constituent. According to her observations, even if the locative expression is not explicit, we have to assume that it is semantically there. The Uralic material often has a local constituent present, as in (1) through (3); occasionally, it has other constituents such as temporal expressions.

(1) Fin Tänän/ ulkona sata-a ja tuule-e.
   Today/ outdoors rain-PRES.3SG and blow-PRES.3SG
   ‘Today it is raining and blowing/outside.’

(2) Est Väljas tuiska-b.
   outdoors snow-heavily-PRES.3SG
   ‘Outdoors, it is snowing heavily.’ (Erelt-Erelt-Ross 2007:471)

(3) Saa Dá-sa muht-ii isorasat.
   this-LAT snow-PRET.3SG terribly
   ‘It snowed terribly here.’ (Sammallahti 1998:100)

Similar examples of single predicate verbs with adverbs can be found in Inari Saami: iho myettäjä ‘at night, it has been snowing a lot’ (Bartens 1995:38, IW entry 2574). In Mordvin, adverbial constituents can take definite declension forms; at least they do so in Erzya (4). The Erzya word ušo is more of an adverbial ‘outside, outdoors’, although in Moksha, ušә also means ‘weather’ (cf. the Komi in example 73).

(4) E Ušo-so-ńń jakšavt-i.
    M Ušә-sә jakšәft-i.
    out-INESS-DEF freeze-PRES.3SG
    ‘It’s freezing outside.’ (Koljaděnkov 1959:199)

The example (5) shows a typical feature of Mari grammar, i.e. converbs, which have developed under the influence of the Turkic languages.

    in.the.evening and in.the.morning freeze-GER set.up-PRES.3SG
    ‘In the evening(s) and morning(s), it is freezing.’ (Učaev 1956:77)

12. If the adverb follows the verb, it reveals the length or manner of the situation or action: Fin Satoi koko yön kaatamalla. ‘It was pouring all night.’ Est Sadas kogu öö nagu oavarrest. ‘id.’
The same pattern is followed also in the Ugric and Samoyed languages (8)–(10); often an argument indicating the time or place is present. One of the peculiarities of Nenets syntax is shown in example (11), were no subject is actually present and the sentence only has a temporal adverbial in the genitive, in this case, tėj jala’ ‘yesterday’.

(8) Kha Tamxătәl wεra iškat-әs.
    today very be.cold-preterit
    ‘Today, it has been getting much colder.’ (Čeremisina–Solovar 2000: 23)

(9) Man Kon xötәl-әs, män in j năj-әl unl-әw.
    outside dawn-preterit we still fire-sit-present.1pl
    ‘Outdoors, the dawn was [already] breaking, [but] we were still sitting by
    the light [lamp, fire, chip].’ (Rombandeeva 1979: 35)

(10) Hun Egész nap havaz-ott.
    whole day snow-preterit
    ‘It has been snowing the whole day.’ (Kovács)\(^\text{13}\)

    yesterday’s day-gen be.hot-preterit.3sg
    ‘Yesterday, it was hot.’ (NRS 828)

(12) Sel Wraqiq čēlimpa.
    quite dawn.aor.3sg
    ‘It was quite light [at dawn].’ (Kuznecova & al. 2002: 319)

(13) Nga Tîminia česa”ľikumi”-ә.
    now get.cold-aor.3sg
    ‘Now it got colder.’ (Tereščenko 1979: 128)

3.3 Avalent verbs with adverbials, translatival + ‘become’-construction

According to Lyskova (1999: 49), the predicate forms the core of the Ob-Ugric sentence, although it does not have to necessarily be verbal; it can also be nominal. Both languages have a special meteorological construction consisting of a verb and its
nominal argument in the translative case. This seems to be an Ob-Ugric innovation (Bartens 1995: 53, 56). In Northern Khanty (the Kazym and Shuryshkar literary languages), the model of a noun in the lative case -a + jiti, as in (14) and (75), is used only with weather and time expressions (Čeremisina–Solovar 1990: 17). Its Eastern Khanty counterpart has the noun in the translative case (-γ) and the northernmost dialect, Obdorsk also has the noun in the same case, but with a different ending (-i), which geminates the preceding word final consonant. Both of these represent the pre-Khanty translative ending (*-yi). The subject is facultative. Northern Mansi has a corresponding construction with a noun in the translative case (-jyg) + jēmtunjakwe in (15). Particular mention should be made of the fact that this is the way the passing of time (with a px)\textsuperscript{14} is described in folklore texts.

\begin{align*}
(14) & \quad \textit{Sora pătlam-a jìti } \textit{pit-әs.} \\
& \text{ soon dark-LAT become-INF begin-PRET.3SG} \\
& \text{‘Soon it was going to be dark.’} \quad \text{(Lyskova 1988: 170)}
\end{align*}

\begin{align*}
(15) & \quad \textit{Tēl-jy} \ jēmt-i, \ ašәrm \ saw \ tot-әw, \\
& \text{winter-TRANS become-PRES.3SG cold suffer bring-PRES.1PL} \\
& \text{tuj-jy} \ jēmt-i \ os \ ĭomwoj \ saw \ tot-әw.} \text{summer-TRANS become-PRES.3SG again mosquito s. b.} \\
& \text{‘When winter comes, we suffer from the cold, when summer comes, we suffer from mosquitoes.’} \quad \text{(Rombandeeva 1979: 62–63)}
\end{align*}

The next Tundra Nenets example bears considerable resemblance to the Ob-Ugrian structure of a noun in the lative/translative + ‘become’ as described above. The special suffix -ŋe has been added to the nominal stem, the verb is xes ‘become; turn into’. Salminen and Jalava name this essive and consider it to be a derivational category rather than a grammatical case, since it does not inflect for number. Following a different grammatical tradition, Laptander refers to it as the translative. Historically, it is an old agglutinated modal gerund form of the verb ‘to be’.\textsuperscript{15} The subject is voluntary.

\begin{align*}
(16) & \quad \textit{Num-da } \textit{ŋeʃojye } \textit{xaja.} \\
& \text{weather-PX.3SG autumnal become.AOR.3SG} \\
& \text{‘Autumn weather has arrived./The weather became autumnal.’} \quad \text{(Tereščenko 1973: 109)}
\end{align*}

---

\textsuperscript{14} In my material, pxs were never present.

\textsuperscript{15} It has also other functions depending on the semantics of the predicate verb, e.g. padnananje manzara ‘[s]he works as a secretary’ (NRS 432), sawawna xaʃenaje tara ‘he is a good hunter (lit. he acts well as a hunter)’ (NRS 632) or ſerçaxa, saʃoxoʃa ſaŋona” jeʃerm’ xengdɔr ye mečetiwa”’ ‘when the wind is blowing and it is raining, we use the sail from our boat as shelter’ (NRS 802).
4. Verbs having object, subject facultative

Meteorological constructions occurring with an overt argument are not always easy to analyze in Uralic since the argument in question in many of the languages may be interpretable as a subject or object. This is most obviously the case with verbs that are essentially transitive when they occur in the objective conjugation. Problems of interpretation may, however, still arise even when the intransitive verbs occur in the subjective conjugation.

4.1 Verb + Object, subjective conjugation

There are only a few exceptional cases when intransitive Finnish meteorological verbs can take on a transitive meaning and have objects as in (17) and (22), which presents a resultative construction without a subject. Even a structure with two objects, one in the partitive and the other as the total object is possible. This kind of variation cannot be found outside the Finnic languages, due to the absence of a partitive.16 Estonian has the same structures in (18) and (20) as can be found in the Finnish in (19) and (21).

(17) Milloin-ko-han sata-a ensi lume-n/ emätyks-en?
    when-INTER-CLC rain-PRES.3SG first snow-A/G record-A/G
    ‘When it is snowing the first snow/raining the [new] record?’

(18) Sada-s paks-u lum-e ma-ha.
    rain-PRET.3SG thick-A/G snow-A/G ground-ILL

(19) Sato-i paksu-n lume-n maa-han.
    rain-PRET.3SG thick-A/G snow-A/G ground-ILL
    ‘The thick snow has been falling down.’
    (Pajusalu)17

(20) Nüüd saja-b hein-a märja-ks.
    now rain-PRES.3SG hay-A/G wet-TRANSL

(21) Nyt sata-a heinä-n märä-ksi.
    now rain-PRES.3SG hay-A/G wet-TRANSL
    ‘It’s raining now the hay wet.’
    (Pajusalu)

(22) Yöllä o-n pyryttä-nyt polu-n ump-een.
    night-ADE be-PRES.3SG snow.heavily-PAST.PTC trail-A/G up-ILL
    ‘Last night, a heavy snowfall covered the trail.’

16. Although some of the Saami languages have the partitive, it is not used to express the object.
17. Renate Pajusalu, visiting professor of Estonian language and culture at the University of Helsinki.
The following North Saami examples consist of a transitive predicate verb and an object. Nowadays most nouns have identical accusative and genitive forms. Subjectless sentences like these are quite common in the Saami languages. The verb ‘blow’ can also be used transitively with an object, as seen in (23).

(23)  
\[
\begin{array}{ll}
\text{Dāl } & \text{biegga} \quad \text{maddel-iid.} \\
\text{now } & \text{blow-PRES.3SG south.wind-PL.A/G} \\
\end{array}
\]
'It is now blowing (for a long time) from south.'

(LD\textsuperscript{18} 162–163)

(24)  
\[
\begin{array}{ll}
\text{Bij-ai } & \text{muohtta-ga.} \\
\text{put-PRET.3SG snow-A/G} \\
\end{array}
\]
'It was snowing.'

(Nickel 1994: 397)

Inari Saami has also parallel examples: \textit{ráálu} (ACC) \textit{kiško} ‘it’s breaking the ice [in spring]’, \textit{tää sättä jaayrijd} (PL.ACC) \textit{kalmeð} ‘now, it can let the lakes freeze’, \textit{tää lii purgam raðe} (ACC) ‘it has blocked the road with snow’ (IW entries 1523, 1178 and 3538). Actually, Inari Saami has a quite number of meteorological metaphors referring to a living being, e.g. \textit{kaggad/pajedið oovtijd} (PL.ACC) ‘pry/raise showers (rain or snow)’ (Idström 2010:149). It has been reported that Kildin Saami has a different attitude about such causative constructions as \textit{raaji} (ACC) \textit{cuonjjo} ‘delaet nast/the snow is being covered with a thin crust of ice’, or \textit{muśt} (LOC) \textit{tautidt} (PL.ACC) \textit{mordta} ‘my bones are aching’. The younger generations consider them to be normal impersonal constructions, while the older generations believe that a supernatural being with magical power is behind these meteorological and physiological constructions (Senkevič-Gudkova 1975:185).

The same causative structure as in Finnish (25), Saami (e.g. Inari Saami above) and Mordvin (with objective conjugation, see (34) and (36) in the §4.2) also exists in Mari (26), where it is not limited to meteorology though. It is equally possible to say \textit{βürәm} (ACC) \textit{kõlмâkt-a} ‘the blood is freezing’ or \textit{moγә́rm} (ACC) \textit{kõlмâkt-a} ‘I am [act. the body is] freezing’. In this context, the subject is never present. In Udmurt causative constructions, the object can be in the nominative (27) or in the accusative. These structures also occur in Komi, as seen in (28), and in Mansi (29), although the morphology does not show it.

(25)  
\[
\begin{array}{ll}
\text{jäätä-ä } & \text{järve-n.} \\
\text{freeze-PRES.3SG lake-ACC} \\
\end{array}
\]

(26)  
\[
\begin{array}{ll}
\text{Jer-әm } & \text{kõlмâkt-a.} \\
\text{lake-ACC freeze-PRES.3SG} \\
\end{array}
\]
'It freezes up the lake.'

(Bartens 1995:45)

---

18. This dictionary uses the scientific Nielsen orthography, other Saami sources the new orthography.
In Khanty sentences (30) and (31), it is difficult to decide if the verbs are transitive and as such have only taken an object and omitted the subject, or if the verbs are intransitive and have a subject. The basic word order SOV does not help much, due to the fact that both the subject and the indefinite object are unmarked. For the Eastern Khanty dialects, it has definitely been attested, that the word meaning ‘God, sky’ cannot be used with the objective conjugation forms (Bartens 1995: 54). The basic verb werti ‘to make’ has in some older minor sources had only a transitive meaning. The situation has changed now, however, and a meteorological intransitive meaning ‘to form, dawn, to be getting’ has been added to the description of the lexical item (HRS 62).

The normal transitivity of the Mansi verb wāruŋkwe with its usual meaning of ‘machen, schaffen, verursachen/to make, to create, to cause’ changes from transitive to intransitive ‘entstehen, werden (Zeit, Wetter)/to become (time, weather)’ only in meteorological connections (WW 718, see also example (59)), and Bartens assumes that these Mansi structures were originally transitive, too. Quite a bit of evidence from related languages backs this up, as seen in North Saami dahkat borgga/avvi (A/G object) ‘to begin to snow heavily/to begin to rain (SSS 97), act. to make the snowfall/the rain’ which has a good parallel in Finnish tehdä sadetta (part). The object can also be in plural (33). The strongest fact, however, is that the same verb in Mansi folklore texts can have two arguments, both a subject and an object. In the archaic language of the older sources, there is no doubt about the transitivity of these verbs in a meteorological context (Bartens 1995: 58–59).
4.2 Objective conjugation, subjects sometimes forbidden

Occasionally, some meteorological verbs can conjugate according to the objective conjugation forms and take objects. In Mordvin, the objective conjugation is more accurate than it is in Mansi, Khanty or Nenets, as it can express both person and number. Naturally, there are many incorporated forms in the verbal paradigm, and more so for the plural forms. Examples (34) and (36) represent the situation where the subject ‘weather’ is not expressed, although it can nevertheless be inferred from the context.

The predicate verb in the objective conjugation can also form a sentence without the object, although this is rare (cf. (36)). If the object is in the plural, the predicate verb often has a frequentative suffix -ś as in (35), where the complete SVO is visible.

(34) E Ėči pižeme, valske kev-ks
    today rain tomorrow stone-TRANS
    mastor-oňt kelmevt-si.
    earth-DEF.GEN freeze-OBJ.CONJ(3SG).PRES.3SG

‘Today rain, tomorrow it freezes the earth hard as a rock.’

(Alēškina 1973: 28, 35)

(35) M Varma-ś ufś-ắźon
    wind-DEF.NOM blow(freq.)-OBJ.CONJ(3PL)PRET.3SG
    vorgaz-ôń ki-tie-ń.
    wolf-GEN track-DEF.PL-GEN

‘The wind has covered the tracks of the wolf.’

(MRS 792)

Almost the same situation as in (34) is expressed in (36), but without the object, which is nonetheless present in the verb inflection. Occasionally similar cases with the verb ‘freeze’ also occur in other Uralic languages, e.g. in Selkup (37), where nom ‘weather’ is the subject.

(36) E Ve-nze lańks, ulema,
    night-GEN.PX.3SG pop(over) probably
    tago kelmevt-si
    again freeze-OBJ.CONJ(3SG).PRES.3SG

‘At night, it’ll probably be freezing again.’

(ERS 250)

(37) Nom kandeptomba-t.
    weather freeze(durat.)aor-OBJ.CONJ(SG).3SG

‘It got colder/It’s freezing’

(SōW 291)
In Khanty, the definite object could be deleted, in which case the objective conjugation is used instead; however, the objective conjugation forms are very rare in the meteorological sentences. Two such examples have been found, both of which have taken the argument in the lative (-a): *wotasa nertәs* (subj. conjug.) or: *neratsәlli* '(suddenly) a heavy snowstorm rose' (OW 595, Bartens 1995:55). The newer dictionary knows only the verb *nerijti* 'to behave (badly)' (HRS 191), to which according to Kaksin the word *atma* 'badly' can optionally be added. The second example describes the same semantic field: *wota poukәmsәlli* '(suddenly) a wind rose' (OW 661). In contemporary sources the verb does not occur, only its nominal counterpart *pәkәtni* 'bad, rainy weather' (HRS 220). Obviously these two meteorological verbs do not belong to the active vocabulary any more as Kaksin confirms. One week later he was able to produce more good examples, but he was quite cautious in saying that anyhow he would prefer passive structure instead of the (40) with the subject and the object. He also said that in (38) any object is out of the question, even though the objective conjugation can be used.

(38) **Tәrm-ew atәm, isa ši pu-l-al-le.**
    weather-PX1PL bad always so blow-PRES-OBJ.CONJ(SG).3SG
    'Pogoda u nas ploxaja: vsë vremja duet.' (Kaksin)
    'We are experiencing bad weather, it's blowing the whole time.'

(39) **Mosәŋ, pa ješa jert-l-al-le turn-lal.**
    Maybe more a.little rain-PRES-OBJ.CONJ(SG).3SG grass-PL.PX3SG
    'Možet byt', eščë nemnogo [nebo] podoždit svoi travy.' (Kaksin)
    'Maybe it's watering the grass a bit more.'

(40) **Uwәs wot kәtra lupsә laŋal pu-l-al-le.**
    northern wind old shed roof blow-PRET-OBJ.CONJ(SG).3SG
    'Severnyj veter kryšu starogo labaza sdul.' (Kaksin)
    'The north wind has blown away the roof of the old shed.'

A remarkable feature of the Samoyed meteorological expressions is the abundance of objective conjugation forms, even if no object is present. All the following Tundra Nenets verbs are transitive and if the object is missing, the use of objective conjugation is obligatory. If the object is present, the subjective conjugation would be the most likely choice. A focused object that contributes new information triggers subjective conjugation, as can be seen in (43). All Nenets verbs are either transitive or intransitive, e.g. the verb *pәś* 'begin' is always transitive. The predicate verb 'blow' for 'wind' seen in (42) and (43) is\(^{19}\) durative and transitive; therefore, it always take

---

\(^{19}\) It has been suggested that "the higher the transitivity of the situation in which the verb occurs in, the greater the probability that it will be conjugated in the determinative
an object, even if it is no more than in the form of a suffix for the objective conjugation. Strangely enough, meteorological expressions in Nenets seem, to some degree, to favour transitive verbs, whereas the same verbs would be intransitive in other languages. Hungarian meteorological expressions having objective conjugation forms are very unusual, but SVO is possible: *a szél fujja a zászlót* (ACC) ‘the wind blows the flag’ (Kovács) and it shows the opposite strategy in comparison with Nenets in (43), which is a normal SOV case.

(41) ˍ甯ermint xad-[da] păngribta-da.

north snowstorm-[PX.3SG] rise.AOR-OBJ.CONJ(SG).3SG

‘A northern snowstorm broke out.’

(Tereščenko 1973: 109, NRS 445 without the px)

(42) ˍSi"jw jamb jala-’ merča lăbibi-da.

seven long day-GEN wind blow(durat.).AOR-OBJ.CONJ(SG).3SG

‘The wind was blowing seven long days.’

(NRS 169)

(43) ˍMercia jeja-’ pan-m’ lăbibi.

wind tarpaulin-GEN edge-ACC blow(durat.).AOR.3SG

‘The wind is making the edge of the tent’s tarpaulin flap.’

(NRS 169)

Interestingly enough, in Selkup, the objective conjugation predicate, as in the next example *pōtimpat* ‘it is warm’, can also form alone a complete sentence (see e.g. Tereščenko 1973: 117); however, it is more common in SV expressions as in (37) *nom kandeptәmbat* ‘it’s freezing (durative)’ or as can be seen in (44):

(44) ˍČēlit čēli-ti somak pōtimpa-t.

in.daytime sun-PX.3SG quite warm.AOR-OBJ.CONJ(SG).3SG

‘In daytime, the sun is quite warm.’

(Tereščenko 1973: 118)

4.3 The participating member in the partitive and other different ‘rain/snow’ structures in Finnish

All the different Finnish ‘rain/snow’ verbs (*pyryttää, ripeksiä, sadella, sataa, tihkuttaa, tihuttaa, tuiskuta, tuiskuttaa, vihmoa*) can have a noun in the partitive as their only argument (Kolehmainen 2010a: 8). But whether this partitive member is a subject or an object has been a matter of dispute for decades. The problem is that it is

[or: objective] conjugation” (Körtvély 2005: 152). This is not, however, really true for these meteorological examples. Obviously the only what is for the time being possible to say – is that the principles according to which the forms of different conjugation (probably) in all of the Samoyed languages are used are not yet well understood.

20. Jalava and Laptander suggest that the north is the subject and the snowstorm the object, L. explains that it is erroneously lacking the accusative ending.
rather difficult to decide what the subject properties are. Partly the partitive member behaves like an object, even though it does not agree in number and it mainly occupies the post-verbal location. The partitive member as the only supplement to the verb has been considered the subject of the intransitive existential sentence as an alternative category to the nominative case. These partitive subjects have been described as neutralized members being neither proper subjects nor proper objects. Vilkuna, however, accepts that exceptionally intransitive verbs can have an object as their only argument (Vilkuna 2000: 156). Both the nominative and the partitive NPs are referring to divisible states of water (Kolehmainen 2010a: 20–21). The partitive forms also occur in Estonian:

(45) Est Saja-b vihm-a/ rahe-t/ lõrts-i/ lun-d.
    Fin Sata-a vet-tä/ rake-i-tä/ räntä-ä/ lun-ta.
    rain-pres.3sg water-part hail-pl-part sleet-part snow-part
    ‘It’s raining/hailing/sleeting/snowing.’

(Hakanen 2001: 64)

The situation as in Khanty where the subject in one example (46) and the object in the other example (47) describe the same situation, is very peculiar. There are also other cases when one and the same participant can be manifested in different ways, as subject: pilvi satoi vettä ‘the cloud was raining’ or as adverb: pilvestä satoi vettä ‘it was raining from the cloud’ (Kolehmainen 2010a: 10).

(46) Lumi sat-o-i-kin/ tul-i-kin vete-nä.
    snow rain-past.3sg-clc come-past.3sg-clc water-ess
    ‘The snow fall/came down as water.’

(47) Sato-i-kin lume-n vete-nä/ räntä-nä.
    rain-past.3sg-clc snow-a/g water-ess sleet-ess
    ‘The snow fell as water/sleet.’

5. Constructions with a subject

In the following sections, the constructions have been categorized according to the nature of the prevailing subject. In the main, each language features only under one type.

5.1 Plain subject +V

Often the words taivas ‘sky’, sää ‘weather’ or ilma ‘air, weather’ function as a subject, even though they are facultative in nature: (ilma/sää) lämpenee ‘it is getting warmer’, (taivas/sää/ilma) kirkastuu ‘it is clearing up’. Temporal verbs such as sarastaa ‘dawn is breaking’, or its synonym valkenee ‘it dawns’ and its antonyms hämärtää, pimenee
or *synkkenee* ‘it grows dark’ are more inclined to take a subject such as *aamu* ‘the morning’ for the first one, *ilta* ‘evening’ for the second one; the word *päivä* ‘day’, however, is suitable for use in both contexts. These words occur quite frequently as independent verbs, too. The latter two are not just restricted to being used when evening or night is falling; they can also be used to refer to changes in the sky during the day, too.

In meteorological expressions, Finnish verbs that are usually transitive act intrinsively, and thus do not take an object as seen in (48). Finnish has two words for ‘sun’. The older, Uralic *väi*, which also means ‘day’, is more fixed in such idioms as *väi paistaa*, so that just *paistaa* or *ulkona paistaa* ‘outdoors [the sun] is shining’ without the subject are also quite normal. The newer word *aurinko* can be substituted for *väi* in this example. Estonian has almost parallel structures, in (49) and (50).

(48) (Päivä) paista-a jo lämpimästi.
(sun) scorch-PRES.3SG already warmly
‘(The sun, or:) it’s already warmly shining.’

(49) Hommik-ul paisti-s veel päike, ...
morning-adess scorch-PRET.3SG only sun
‘The sun only shone in the morning, …’ (Hakanen 2000: 68)

(50) Kõu21 mürista-b.
thunder thunder-PRES.3SG
‘It’s thundering.’ (Airila 1928: 7)

The etymological counterparts for the Finnish word *ilma* ‘weather, air’ in North Saami *albmi* ‘sky; snowstorm; heaven (in a religious sense); air (in dialects)’ (SSS 3), in Inari Saami *alme* ‘sky, bad weather’ (IW entry 62), in Kildin Saami *alm* ‘sky, snowstorm’ (SRS 24) seem to have kept their meaning, at least in dictionaries. Inari Saami speakers also have a living being named *Tálvášáś* “Jack Frost”, who plays a very active role in the rotation of the seasons. Their thunder is also animated: *äijih čiärgu* “the old man is roaring” (Idström 2010: 133). This is familiar to Finns and Mordvins, too: *ukkonen*22 *jyrisee*, M

21. This word has its only counterpart in Finnish *kouko* ‘ghost; bear, wolf, beast; louse; tall person’ and the assumed origins are a Baltic word meaning ‘brownie, gnome’ or an Indo-European word *kouko* (SSA I 414).

22. In Estonian *äike(ne)* ‘thunder’ is derived from the word *äi* ‘father-in-law’. In North Saami, *addjá* means ‘grandfather’ and *addyán!* is a mild expletive (SSS 10). In the central Saami languages, its counterpart also means ‘thunder’ (SSA III: 494). The North Saami word for ‘thunder’, *baján*, is derived from *bajá- ‘upper’ (SSS 23).
ativat toraj, E purgine šzerši, where ukkonen and atam are equally derived from ukko23 'old man' and aţa 'id.' and purgine is an old loanword from Baltic languages meaning 'devil' and is where the Finnish swearword perkele originates from (SSA II 340).

The Maris are considered to be among the last people still believing in their ancient spirits and deities and especially the easternmost Maris living in the diaspora near the Ural Mountains have kept their old beliefs. According to Chesnakova24 the thunder tends to have the name paβaj 'uncle (father’s brother)' and when it is thundering the Maris say jum-paβaj kīđörtä. When the children do something nasty, they are scolded by saying jum-paβaj pāra ... 'Uncle thunder hits [you, if you don’t…]'. In Mari meteorological sentences, the subjects can be present, and very often jumo,25 which originally meant ‘God’ or ‘sky; icon’ (TW 194) is expressly used. The examples in (51) and (52) are archaic nowadays (Bartens 1995:44–46). These days, it is hard to find examples with jumo as speakers seem to have become alienated from using it. It is very rarely used in the locative case, as it has been in (53) where it definitely means ‘sky’. A suitable alternative for sentences (51–53) could also be kaβa ‘sky’. Another word suitable for the subject position is keče ‘day; sun; weather’ as seen in (54); this word, however, seems to be totally optional.

(51) Jumo kūδört-a.
God thunder-pres.3sg
‘It’s thundering.’ (Airila 1928: 7; Bartens 1995: 45, TW 194)

(52) Jumo/ keče βolyalt-ešt.
God day dawn-pres.3sg
‘Der Tag (“Gott”) graut/The day (“God”) is dawning.’ (TW 48)

(53) Jumә-što sūðr peš-ak šuko.
sky-LOC star very-clc much
‘In the sky, there are a lot of stars.’ (SMJa 10: 174)

23. Finnish folk poems and other sources describe Ukko as also being the God of Rain and Clouds. It is not impossible that Ukko may have had some traits of Father God from the very beginning, even if there is evidence of an obvious Christian influence as in the phrase Herran ilma ‘the Lord’s weather’. The standard Finnish word ukkonen has a dialectal counterpart Pitkä(i) nen ‘the good, tall one’ derived from pitkä ‘long, tall’ appering in Estonian in the forms pitkne, pikne and pikker ‘thunder, lightning’. In addition, other euphemistic names for the God of Thunder have been used (U. Salo 2006:5–12). According to Castrén (1852:137–140), semantic equivalents for Ukko, the highest God, are found widely throughout northern Eurasia.

24. Sonya Chesnakova, Ph.D., specialist in literature, born in Bashkortostan.

25. Although it has etymological counterparts in the Finnish (jumal[a]) and the Saami languages (ipmel), they are not used in meteorological expressions. These words have been questionably connected to the Mordvin god’s name Jumi (SSA I 247).
(54) Keče leβešt-a / jūkšemd-š/ kasešt-eš.
weather warm-PRES.3SG cool-PRET.3SG get.late-PRES.3SG
'It's warming up/it cooled down/it's getting late.'

(Učaev 1956:75, 77; Bartens 1995:45)

According to Alatyrev (1970:198), Udmurt expressions representing natural phenomena can equally take the subject kuaž\(^{26}\) 'nature, weather; God' or not, which is actually quite normal for Uralic verbs having varying valencies. When the subject is used, it can also take a px as a definitive marker. Pozdeev (1973:253) is inclined to consider the word kuaž to be obsolete and states incorrectly that the verbs are avalent. The meaning 'God' is mainly present in folklore texts (Alatyrev 1941:8). Bartens was assisted by an Udmurt, who provided her with a large number of sentences where kuaž and inmar, the mythological God of the Udmurts, were interchangeable. She thus draws the conclusion that both words have lost their original meaning for contemporary speakers and that they are on the brink of becoming formal subjects, especially in the case of kuaž (Bartens 1995:48); this is, in some respects, erroneous, as the use of these subjects is possible only in meteorological contexts, where the subjects are optional depending on the speaker's choice with predicate verbs having varying valencies. The new Udmurt-Finnish dictionary (USS) is based on cooperation with native speakers of both languages, which makes its information reliable. Quite often, the authors give alternative expressions, as in (55). Determinating pxs occur randomly, px.2sg and px.3sg are possible. The noun in (56) jugit can be a substantive or an adjective meaning 'light, clear'.

(55) (Kuaž) zor-e / sakt-e.
weather rain-PRES.3SG dawn-PRES.3SG
'Sataa (vettä)./It is raining/Dawn is breaking.' (Alatyrev 1970:198, USS 152)\(^{27}\)

(56) Inmar\(^{28}\) gudirja/ jugit lu-e.
God thunder-PRES.3SG light be-PRES.3SG
'It is thundering/Dawn is breaking.' (Airila 1928:7)

---

26. Meaning 'priroda, pogoda, totem/nature, weather, totem' (Alatyrev 1941:7, 9); 'priroda (pogoda)/nature (weather)' (Alatyrev 1970:198); 'pogoda; gospod', bog/weather; lord, God' (URS 220); 'sää, ilma; jumala/weather, air; God' (USS 224).

27. Another dictionary gives this example without the parantheses (URS 155, 220).

28. The Udmurt god's name Inmar is derived from the word in [: inm-] meaning 'sky'. The derivational relationship is equal to that of the Finnish Ilmari or in the folkloric Ilmarinen (originally meaning 'the god of the air, weather'), which can be seen in the form it derived from: ilma 'air; weather, storm; world; cardinal point'. This word also has counterparts in other Finno-Ugric languages, as already seen in §2.2 (also Ajhenval’d–Petruhin–Helimskij 1982:189).
Khanty meteorological verbs often have the word tŏrәm ‘God; icon; sky; weather’ (OW 1015–1018) as their first argument. The same word could also have the additional meanings of ‘century, time; world, space’ as in the sentence tŏrәm pa janas wŏs ‘and times [sg + px.3sg] were different [then]’ (HRS 270), (§5.3 has more examples with pxs in (74) and (77)). This basic figure of Khanty mythology occurs in lexical items describing the ancient Siberian conception of the tripartite world in num29 tŏrәm ‘God, sky’, the highest celestial God in their traditional beliefs as compared to tăm tŏrәm ‘this air, world’ and tŏm tŏrәm ‘that air; the underworld’. Today it is recognized that Russians and Komi have significantly influenced the meaning of tŏrәm as the uppermost god in the celestial hierarchy. According to the theory supported by many scientists, the Ob-Ugric belief about an all-powerful celestial god is a rather late phenomenon, although its views about a bearer of the weather are quite ancient (Kulemzin 2006: 141). To modern-day Khanty speakers, nŭm tŏrәm only means ‘space’ (HRS 189). Some Khanty meteorological subjects are less inclined to take pxs, as seen in (57) and (58).

(57) Ajkeman xošәm wot pul-as.
    gently warm wind blow-pret.3sg
    ‘The warm wind was gently blowing.’
    (Čeremisina–Solovar 1990: 72)

(58) Aľәŋ30 xŏńl (śi) tiw-әs.
    morning dawn so form(intr.)-pret.3sg
    ‘The red sky of dawn was breaking up.’
    (HRS 16, 266)

According to Rombandeeva, a native speaker of Mansi, the subject tŏrәm ‘sky’ can be omitted as a matter of course, but it can also be present as witnessed in (59)–(60). The subject can also sometimes be tŏrәm-šiškwe ‘sky-mother’, at least in the Eastern Mansi dialects (WW 492), even though God is mainly considered to be masculine. As with its counterparts in other related languages, the word tŏrәm has many different meanings nowadays: ‘Gott, Himmel; Zeit, Jahreszeit; Welt, Wetter/God, sky; time, season; world, weather’ (WW 657). In the folklore texts, it has its parallel in word kworeә or kors ‘sky, time; age’, actually a mythologic person, numi-tŏrәm’s father. Another god, šaxәl-tŏrәm

29. Some researchers even consider this Ob-Ugrian-Samoyed word num or numi to be an older name for God. To be more precise, the original meaning of the Ob-Ugrian word is ‘high, celestial’ and the Samoyed word is connected with the Sogdian num and Proto-Turkic nom meaning ‘religion, religious worship, religious rules’ (Ajhenval’d–Petruhin–Helimskij 1982: 189). The Ob-Ugrian words can not be borrowed from the Samoyed languages due to the fact that their meanings are wider on the Ob-Ugrian side (UEW I 308–309).
30. It can be ‘morning’ or ‘in the morning’.
'thunder god’ is either numi-törәm’s younger brother or his son (Lyutsidarskaya 2008). There are examples, where the törәm is present and even very active: törәm wōtį y ti pokapas ‘the sky got irritated and turned windy’ (Balandin 1967: 304, WW 448). For Mansi speakers, the clouds can also be active participants when it is thundering as in (61) (cf. also the Finnish example in §4.3.).

(59) Törәm turap wār-s.
    sky cloudy make-PRET.3SG
    ‘The weather turned cloudy.’ (WW 657, 684)

(60) Xosa tōrәm ēl-ēyən, săw suns-ēyən.
    long time live-PRES.2SG much see-PRES.2SG
    ‘[If] you have a long life, you are going to see a lot.’
    (Rombandeeva–Vahruševa 1989: 116)

(61) Śaxәl kuryaltaxt-əs, tuwał rakwaltaxt-əs.
    thundercloud thunder(inch.)-PRET.3SG then rain(inch.)-PRET.3SG
    ‘It began to thunder and rain.’ (RMS 99)

Details concerning the syntax of the Kamas language are very limited, although chronologically (ca. 250 years) it is perhaps one of the best documented of the Samoyed languages. The texts and grammatical paradigms have been collected at a stage when the community of speakers was heavily reduced and the last native speakers were bilingual. The entire attested lexis of Kamas consists of only a few thousand words and about one-third of the lexicon has been borrowed from the Turkic languages (Künnap 1999: 5). Predictably, the verbal morphology is complex as it is in the other Samoyed languages. In addition to encoding person and number, Kamas verb stems also encode tense, mood and the category of the definite object. Aspect, however, is expressed by syntactic means, i.e. with the help of auxiliaries; this contrasts sharply with its closest related languages (Simoncsics 1998: 590). Such aspectual syntagms consist of the gerund of the lexical verb + auxiliary verbs such as those in (64) and (65). Based on the very scanty material that has been collected on Kamas, it seems that the language has no avalent verbs and that all of the basic meteorological expressions are SV. Furthermore, the subjects do not take pxs.

(62) Sәre uzziuf-e.
    snow fall-GER
    ‘It is snowing.’ (KW 82)

(63) Num kuurum-ňe.
    sky thunder-GER
    ‘It is thundering.’ (KW 35, 46)
(64)  *Num d’žál>mna-mna*[^31]/* dår>mna-mna*
   sky   dawn-AOR.AUX.3SG  lighten-AOR.AUX.3SG
'Dawn is breaking.'  (KW 14)

(65)  *Kuja d’omda-laा-’bә.*[^32]
   sun   shine-GER-AUX.3SG
'The sun is shining.'  (Simoncsics 1998: 590)

5.2 Sdef + V

The definite declension for nouns is a feature common to only Mordvin and Hungarian, although their stem-dependent location varies. In Mordvin, nouns can have the definite ending (of pronominal origin)/ś in the singular nominative and -ńte, -ńte in the plural nominative; in the oblique cases, the forms differ.[^33] It is possible that this is an areal feature, as it is also found in the northern dialects of Russian, Veps, and the Saami languages of the Kola Peninsula (Itkonen 1966: 257; Bartens 1999: 83–84; Stadnik-Holzer 2006). This way of expressing determinacy can be considered to be a western feature, parallel to the eastern (Turkic) way of expressing the determinacy with pxs (Luutonen 2010: 108). Both of these patterns exist side by side in the Mordvin languages. The syntactic subject in meteorological expressions is always determinate, having either the definite post article -ś or the px 3.sg -ze, -zo in Erzya or -č in Moksha. However, determinate marking with a px seems to no longer be productive in Mordvin, although random examples can still be found, as seen in (66) and earlier in (36).

(66)  E *Mењә́l-еś[^34]*/ či-ś ašolgard-ś.
M *Mењә́l-ś/  śi-ś valдаšкәд-ś.
   sky-DEF.NOM  day-DEF.NOM  dawn-PRET.3SG
'The sky/the day was dawning.'  (Koljaděnkov 1959: 199)

The subject can be replaced by a definite spatial marker such as *ušoś* 'out, outdoors', which can also be in the definite inessive *ušosoń* 'id.' as seen in (4). It has been

[^31]: Actually, the meaning is ‘Morgenrot/zarja/dawn’ explained as “der Himmel dämmernd liegt”, consisting of a gerundial main verb and a stative auxiliary verb amna- ‘to sit, to lie’ (Klumpp 2002: 151, 168, 178).

[^32]: This is actually a double verb with a gerundial initial component like the Turkic languages have. The auxiliary ending is originally an independent verb form *i’bә* ‘lies’ (Künnap 1999: 34; Klumpp 2002: 141–144).

[^33]: In Erzya, the oblique ending -ńt is probably a combination of genitive ending -ń and a demonstrative *đe*-morpheme; Moksha uses analytic forms (Bartens 1999: 83–87).

[^34]: This word’s only etymological counterpart is the Hungarian *menny* ‘sky, heaven’ (UEW I 276).
proposed that these words and many others could be considered formal subjects (Bartens 1995:41–42; Molnár 2001). I, however, do not concur. The meteorological nouns are, indeed, semantically more or less empty subjects that are predictable from the context. But by no means can they be regarded as formal subjects, as their use is restricted to a single, semantically narrow field: meteorology. The idea that they could be a type of formal subjects has definitely arisen from translations into western Indo-European languages. The Mordvin sentence ‘it’s raining’ can be expressed with a single verb E piżi, M piši, or with a structure piżemęs moli, which can be considered to be similar “dummy (auxiliary) verb” construction as in Russian or Japanese (cf. Malchukov & Ogawa, in this volume). When it is snowing, the word ‘snow’ lov has to be added:

(67) E lov-oś piż-i/ pé-i/ moli.

snow-def.nom rain-pres.3sg/ fall-pres.3sg/ go-pres.3sg

‘The snow is falling/it’s snowing.’

Obviously the neighbouring Indo-European languages have influenced the Hungarian grammar. Currently, Hungarian has two articles: the indefinite egy (of numeral origin) and the definite a or az (of pronominal origin). Most of the Hungarian meteorological verbs can take an argument as a subject. The common word order is VS, but SV is also possible, as in (71). When used with a metaphorical meaning, the first group, avalent verbs, can also take a subject: villámlott a szeme ‘eyes (sg in dual meaning) flashed fire’ or dörgött a hangja ‘his voice thundered’ (Kugler 2000:409). Unlike some of the other Finno-Ugric languages, the Hungarian literary language does not have any attested meteorological structures that use the word meaning ‘God’ as their subject. For Hungarians, the sky is thundering dörög (az ég). Airila (1928:7), however, pointed out that the compound word isten[-]nyila ‘lightning (lit. God's arrow)’ does include the word ‘God’.

(68) Es-ik (az eső)/ a hó.

fall-pres.3sg def.art rain35 def.art snow

‘It’s raining/snowing.’ (MFSz 238, 351)

(69) Fúj a szél.

blow-pres.3sg def.art wind

‘It’s blowing.’ (MFSz 720)

(70) Hűvösöd-ik már az idő.

get.cold-pres.3sg already def.art weather

‘It’s already getting cold.’ (Bartens 1995:60)

35. The Hungarian word for ‘rain’ eső is actually a present participle form of the verb esik ‘fall’.
The sun is scorching hot

The use of Northern Khanty pxs depends on the situation, i.e. the persons participating in the discussion. The three examples in (74) through (76) describe common situations, even though all of the other pxs are possible, too. According to my informants, the px in (76) indicates that I should go out.

Sporadically, old sources show traces of ancient Siberian mythology, but these probably show the influence of Russian and Zyryan Christianity in the 17th to 18th centuries.

36. According to many sources, this is a determinating use of 2sg and 3sg. Although this should also be known from Mari, too, I have been unable to find these types of meteorological utterances in my material.
As stated previously, it is widely held that the Ob-Ugric belief about an all-powerful celestial god is a rather late phenomenon, although views about the bearer of the weather are quite ancient (Kulemzin 2006). The sky was a masculine God, as seen in the Mansi example (77).

(77) Tōrәm jäy-em rakw-i/ tūj-i.
sky father-px.1sg rain-pres.3sg snow-pres.3sg
‘It’s raining/snowing. (lit. My celestial father rains/snows.)’ (WW 492, 657, 674)
In the Samoyed languages, the formal subject is expressed by a noun that means ‘sky, weather’. In the case of Nenets, this noun is num’, which also refers to ‘god’ (Tereščenko 1973: 109, NRS 320). In most instances, this noun in Nenets takes a determining px.
Both the subject and the object can have pxs, as seen in (79), where contrary to many expectations the objective conjugation has been used. The determining px is also used in the Selkup example in (80).

(78) Num-da jaluma/ pewšuña.
sky-px.3sg dawng.aor.3sg/get.dark.aor.3sg

(79) Nara-’ nakuna xajer-ťa ņobpiña’
spring-gen pop('to') sun-px.3sg continuously
jal-a-m-da jambumamb’i-da.
day-acc-px.3sg lengthen(durat.).aor-.obj.conj(sg).3sg
‘Closer to the spring, the days (sg) are continuously getting longer.’
(Tereščenko 1973: 109)

(80) Siri-ťi aiča.
snow-px.3sg fall.aor.3sg
‘The snow is falling.’
(Tereščenko 1973: 109)
In the following Nganasan sentences, the 3rd person pxs -tu, -di, -du are mainly used for determining nouns such as ņuo ‘sky; god; weather’, mәu ‘earth’ and kou ‘sun’ (Tereščenko 1979: 95). However, the px is not obligatory.

(81) Dāļ-đi ċesgimta-tį.
day-px.3sg get.cold-pres.3sg
‘It’s getting colder.’
(Tereščenko 1979: 239)

(82) Soru-lbīu” ņuo-du xojmągim-se.
rain-subord sky-px.3sg darken-pret.3sg
‘When it was raining, it started to get dark.’
(Tereščenko 1973: 109)

(83) ņuo-du norumù-’o/ kotumù-’o.
weather-px.3sg begin.spring-aor.3sg begin.snow.heavily-aor.3sg
‘Spring weather has started to arrive/It began to snow heavily.’
(Tereščenko 1973: 109, 246)
5.4 Cognate constructions

In the Uralic languages, impersonal (or unipersonal) sentences consisting of only the predicate verb are considered the basic means of expressing weather phenomena (Itkonen 1966:302). It is therefore not surprising that Itkonen also views cognate constructions (CC) or figura etymologica as representing a later level of development. This strategy has been occasionally found in Indo-European, but it seems to be more widespread in other languages (Malchukov & Ogawa, this volume). These grammatical structures take a subject, which has a function similar to that of the corresponding “dummy” pronominal subject in the Indo-European languages. The dummy subject allows the impersonals to be changed into the normal sentence type for these languages, i.e. taking two members.

As a general rule Finnic languages do not favour CCs, where subjects and verbs originate from the same root. An exception to the rule has to be mentioned. There is a famous Finnish song composed by Jean Sibelius with lyrics by A.V. Forsman (Koskimies) where the verb *tuulla* ‘to blow’ has a tautological subject and it is even in the imperative mood, both of which are extremely marginal: *Tuule, tuuli, leppeämmin*... ‘Blow, the wind, more gently...’ With the noun *tuuli*, other verbs such as *puhaltaa, tuiverta* etc. usually occur. Estonian does not like using CCs. Pajusalu seems to remember that it might be possible to say *torm tormab* ‘the storm is raging’ in dialectal Estonian. According to Seurujärvi-Kari,37 CC cases are possible in Saami: *biegga biekkasti* ‘the wind was picking up’.

There are some cases with CC in Mordvin. The sentence ‘it is raining’ can be expressed with a single verb or the subject can be added: E *piźemeś piźi, M piźәm piśi* lit. ‘the rain is raining’. Based on the existing causative suffix -vt/-ft- Koljadënkov – a native speaker of Erzya – presumes that there used to be a subject as in E *keľmeś keľmevti/M keľmoš keľmofti* ‘moroz morozit/lit. the cold is cooling down’ or E *jakšamoś jakšavti/M jakšamś jakšafti* ‘xolod xolodit/lit. the freeze freezes’ (Koljadënkov 1959:199). Nowadays, only the latter case uses a tautological subject in Moksha.

In Mari, CC expressions are fairly common: *lum lumeś* ‘it’s snowing’ *jür jüreš* ‘it’s raining’, *küδә rčö küδә* ‘it’s thundering’, *βolγenče βolγalteš* ‘it’s lightning’. CC expressions are quite common in Udmurt, as well: *tel tela* ‘the wind is blowing’, *lijy ljija* ‘it’s snowing’, *zor zore* ‘it’s raining’. The single verb *zore* ‘it’s raining’ can take an argument in the instrumental case: *zoren zore* ‘it’s hailing’, but it cannot be used without the word *je* ‘ice’, as *zoren zore* ‘it’s raining’ is not proper Udmurt (Bartens 1995:48). CC expressions are less common in Komi, although they can still be found: *tev teлal*

37. Irja Seurujärvi-Kari, lecturer in Saami at the University of Helsinki.
é 'the wind is blowing', *gim gimčė*³⁸ ‘it’s thundering’, *lim limčalė* ‘it’s snowing’. Instead there are verbs having arguments in the instrumental case as (84), with the addition of the word *ji* ‘ice’ as in *jizerén zere* ‘it’s hailing’ (Bartens 1995: 49–50).

(84) *Ritjadorjš zerišt-i-s sonid zer-ên.*

at.nightfall rain(dim.)-PRET-3SG warm rain-INSTR

‘At nightfall, it rained a warm rain.’ (Saharova 1967: 152; Bartens 1995: 49)

In Northern Khanty, the tautological fossilized subject argument as seen in *jεrt jεrtәl* lit. ‘the rain is raining’ or *xo ˇs xo ˇńľ*³⁹ ‘the star is twinkling’ is possible with eight of eleven avalent verbs (Čeremisina–Solovar 1990: 23, Kaksin). Different types of CC are quite common especially in the folklore texts, e.g. *xătәŋ xăt* ‘sunny day’, *xujәŋ xuś xu* ‘masculine Russian man’, *wet lŭjpә lŭjәŋ kŭr* ‘five-toed foot with toes’, *iš potәrәn potәrәl* ‘(s)he speaks with these words’ (Steinitz 1976: 41–46). CC expressions are also possible in Northern Mansi, as evident from (85), or in an example from the Konda dialect: *wōt kėi niś wōti* ‘der Wind weht sanft/the wind is blowing gently’ (WW 741).

(85) *Tēli tūjt tūj-i, tūji rkw rkw-i.*

in.w inter snow rain-PRES.3SG in.summer rain rain-PRES.3SG

‘In winter, it snows; in summer, it rains.’

(Rombandeeva 1979: 35, Skribnik–Afanas’eva 2004: 54)

In Hungarian, on the other hand, the word order for some meteorological sentences that have subjects is direct: e.g. in *jégeső esik* ‘it’s hailing’ or *záporeső esik* ‘it’s pouring’ (MFSz 238). The focus is on the noun, which is why it comes before the verb, the opposite of the usual *esik az eső* ‘it’s raining’ (Bartens 1995: 60). Nenets also has CC expressions, e.g. *munota muno”ŋa*⁴⁰ ‘it’s thundering (NRS 262), derived from the word *mu* ‘sound’ (NRS 264–265) and *xarpŕi” xarpărŋa”* ‘the Northern lights are just shining’ (NRS 751). Nevertheless, the dictionary has chosen a different verb *măilid”* ‘[the Northern lights] are beginning to dance (or move quickly)’ in a longer example on the same page.

5.5 Formal subject

As a rule, the Uralic languages do not form meteorological expressions by means of “dummy” pronominal subjects. A pronominal element is, though, well known in Finnish dialects and the informal spoken language, even though it is foreign to the

---

³⁸. According to Misharina, it’s possible to say *Ilja prorokur gimčė* ‘Prophet Elia is thundering’ when it’s Elia’s day or the subject can also be the word *jen* ‘God’.

³⁹. The tempus marker -(ә)l has been assimilated to the stem.

⁴⁰. According to Salminen, this is a rare expression from the western dialects, the first part of which is a present participle of the verb ‘utter sounds’.
literary language and therefore overlooked in most grammatical descriptions. It usually expresses emotional attitude towards the surroundings or weather conditions. This kind of use of “light” pronoun as seen in (86)–(88) is clearly a feature affected by the Swedish det pronoun meaning ‘it’. Many other grammatical innovations in Finnish have also been influenced by the Swedish, which has led to Finnish grammar having switched course towards the Indo-European languages. The Finns lived for 700 years under Swedish rule and even when it was over, the Swedish language maintained its leading position for several decades as the language of culture, education and government.

(86) Pyryttel-i se teä-v viiko.
snow(freq.)-past.3sg it/ptl this-gen week-gen
‘It was snowing this week.’

(87) Tuule-e-pa-s se navakasti.
blow-pres.3sg-clc-clc it/ptl hard
‘It is blowing hard.’

(88) Kainka se jaks-a-kin sata-a koko päivä-n!
how it/ptl be.able-pres.3sg-clc rain-inf whole day-gen
‘How on earth can it have been raining all day long!’

Finnish sentences often lack a subject and some other phrase will willingly take its place if it can. In everyday language about weather, it is fairly common to see the “dummy” element se ‘it’, which also has a partitive counterpart ‘sitä’ in passive generic zero sentences; these two elements are conjugated forms of the same pronoun, and are, as such, both considered particles41 (Vilkuna 2000: 150). Similar place fillers used include pronominal and time expressions.

In Estonian linguistic works, there is no mention of the impersonal meteorological expressions of the type found in Finnish. But as a result of an educated guess these were discovered in spoken Estonian, too. Pajusalu says that only the short form of the pronoun ta or tema ‘it’ is possible in this case. A search on the Internet reveals that instead of the usual sajab, sajab vihm or vihm sajab ‘it’s raining’ the pronominal variant ta sajab is quite common in Estonian informal texts.

41. This impersonal pronominal structure is found in even broader contexts for other unipersonal connections, e.g. sitä sattuu kaikenlaista mukavaa, ‘all sorts of nice things can happen’, on se kumma, että… ‘it is strange that…’, sitä vääsy helposti ‘one gets tired easily’. The standard analysis of sitä can be found in (Hakulinen 1975). According to her, sitä is not a generic or deictic pronoun, and instead has become a particle. Actually, spoken Finnish is currently experiencing grammaticalization, with the demonstratives se (sg.) and ne (pl.) emerging as definite articles (Itkonen 1966: 257; Laury 1997).
Due to the fact that the basic word order for Saami is SVO, formal subjects sometimes also occur nowadays as a result of the strong Scandinavian influence on the language as in (91) (Nickel 1994: 397; Sammallahti 1998: 96). One of the older Saami grammars still categorically states that these sentences never have subjects and the author also points out that dat as in (92) is not a subject but an emphasizing particle (Nielsen 1926: 300, 360). For non-natives, it is not easy to distinguish these emphasizing particles from informal subjects.

6. Conclusions

A summary of the range of subjects found in the meteorological expressions of the fourteen languages covered by this study is presented in Table 1 and the corresponding range of sentence types in Table 2. There are some basic grammatical differences in Uralic languages that contribute to the dissimilarity of the columns. The only feature that all the fourteen languages have in common is that basic sentences are SV. Definite articles and possessive suffixes do not occur simultaneously in the meteorological expressions of a single language; only one or the other is used. It is interesting to

---

42. Seurujärvi-Kari, a native speaker of North Saami, says that natives do not create these types of sentences, and it is only non-natives (Scandinavians) who use these when they are trying to speak Saami.
note that the bare predicate structures considered to be a basic element of Uralic by Itkonen (1966: 302), are, in fact, missing in Kamas and Forest Enets. The presence of an objective conjugation is an eastern feature that unexpectedly occurs in Mordvin, as well. On the Siberian side, it has not been easy to find meteorological expressions that use forms of the objective conjugation, although this could be due to the limited sources available. The results of new fieldwork can, in some cases, shed light on this question. CCs also form an interesting group of their own; the sources typically tend to say that they are more or less common and only give a couple of examples. Unfortunately, our knowledge of the various languages is not equal, and thus the gaps in the tables. Nonetheless, these two tables provide a good starting point for future studies. Table 3 classifies four types of subjects in basic SV sentences showing how the older and newer vocabulary already discussed in §2.2 is dispersed throughout the contemporary languages. These semiformal subjects or mythological subjects are very typical of the Uralic languages.

Table 1. Subject types in meteorological expressions

<table>
<thead>
<tr>
<th></th>
<th>‘it’</th>
<th>N+def.art.</th>
<th>N+PX</th>
<th>CC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finnish</td>
<td>x</td>
<td>–</td>
<td>no (x)</td>
<td></td>
</tr>
<tr>
<td>Estonian</td>
<td>x</td>
<td>–</td>
<td>no (x)</td>
<td></td>
</tr>
<tr>
<td>Saami</td>
<td>x</td>
<td>–</td>
<td>no x</td>
<td></td>
</tr>
<tr>
<td>Mordvin</td>
<td>–</td>
<td>x</td>
<td>no x</td>
<td></td>
</tr>
<tr>
<td>Mari</td>
<td>–</td>
<td>–</td>
<td>–43 x (common)</td>
<td></td>
</tr>
<tr>
<td>Udmurt</td>
<td>–</td>
<td>–</td>
<td>x x</td>
<td></td>
</tr>
<tr>
<td>Komi</td>
<td>–</td>
<td>–</td>
<td>x x</td>
<td></td>
</tr>
<tr>
<td>Khanty</td>
<td>–</td>
<td>–</td>
<td>x x (common)</td>
<td></td>
</tr>
<tr>
<td>Mansi</td>
<td>–</td>
<td>–</td>
<td>x x (common)</td>
<td></td>
</tr>
<tr>
<td>Hungarian</td>
<td>–</td>
<td>x</td>
<td>no (x)</td>
<td></td>
</tr>
<tr>
<td>Nenets</td>
<td>–</td>
<td>–</td>
<td>x (x)</td>
<td></td>
</tr>
<tr>
<td>Selkup</td>
<td>–</td>
<td>–</td>
<td>x no?</td>
<td></td>
</tr>
<tr>
<td>Nganasan</td>
<td>–</td>
<td>–</td>
<td>x no?</td>
<td></td>
</tr>
<tr>
<td>Kamas</td>
<td>–</td>
<td>–</td>
<td>no? no?</td>
<td></td>
</tr>
</tbody>
</table>

43. Oral information from Florian Siegl, who was recently living among the last speakers of Forest Enets collecting data for his dissertation.

44. In Mari, the possessive suffixes in meteorological connections are used as emphasizing particles only.
Table 2. Sentence types in meteorological expressions

<table>
<thead>
<tr>
<th>Language</th>
<th>V</th>
<th>Vobj.conj.</th>
<th>SV or VS</th>
<th>VO or OV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finnish</td>
<td>x</td>
<td>–</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Estonian</td>
<td>x</td>
<td>–</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Saami</td>
<td>x</td>
<td>–</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Mordvin</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Mari</td>
<td>x</td>
<td>–</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Udmurt</td>
<td>x</td>
<td>–</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Komi</td>
<td>x</td>
<td>–</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Khanty</td>
<td>x</td>
<td>(x)</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Mansi</td>
<td>x</td>
<td>no</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Hungarian</td>
<td>x</td>
<td>(x)</td>
<td>x</td>
<td>–</td>
</tr>
<tr>
<td>Nenets</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Selkup</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>no?</td>
</tr>
<tr>
<td>Nganasan</td>
<td>x</td>
<td>no?</td>
<td>x</td>
<td>no?</td>
</tr>
<tr>
<td>Kamas</td>
<td>–</td>
<td>no?</td>
<td>x</td>
<td>no?</td>
</tr>
</tbody>
</table>

Explanations for the symbols in use:

- **x** present
- **(x)** occurs occasionally
- **–** not present
- **no** the feature in question does not occur in these expressions
- **no?** no data

Of the fourteen languages discussed above one has recently died out. Of the remaining thirteen, only three are used in multiple contexts in various ways according to the needs of a modern society; for example, texts that contain everyday language most commonly spoken are missing for most of the Uralic languages and beyond that, their literary cultivation functions at a bare minimum. The area that they are used in has considerably shrunk in the 20th century, mostly being replaced by Russian. The unvarnished truth for many minor or even middle-sized Uralic languages is that they are only used amongst family members and friends in the countryside. In towns and cities, minority languages suffer and often disappear within a few generations.

The neighbouring Turkic languages seem to have had meteorological structures similar to those in the Uralic languages: e.g. Bashkir *tundi ra* ‘it’s getting colder’, Tatar *jakti rá bashlad* ‘dawn was breaking’ (Gadžieva–Serebrennikov 1986: 89). Turkic dictionaries always have some cases of CC to be found, such as
the Turkish yağmur yağıyor ‘it’s raining’, Tatar jaŋğör java ‘id.’, or kük kükri45 ‘it’s thundering’. In addition, the Tungusic languages are characterized by the use of these types of meteorological structures, and examples even occur in Baltic and Slavic languages. Grammatical structures with a subject are very frequent, and this has probably had a great impact on the abundance of SV expressions in Mari and Udmurt (Bartens 1995: 62).

Table 3. Meteorological subjects according to their meaning in basic SV sentences.

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>‘day, sun’</td>
<td>‘weather45, (air)’</td>
<td>‘sky, (God)’</td>
<td>‘thunder, (God of thunder)’</td>
</tr>
<tr>
<td>Finnish</td>
<td>päivä, aurinko</td>
<td>sää, ilma</td>
<td>taivas</td>
</tr>
<tr>
<td>Estonian</td>
<td>päev, päike, pälkene</td>
<td>ilm</td>
<td>taevas</td>
</tr>
<tr>
<td>NSaami</td>
<td>beaivi</td>
<td>dálki, ilbmi</td>
<td>albmi</td>
</tr>
<tr>
<td>ISaami</td>
<td>pejrik</td>
<td>šonŋå</td>
<td>alme</td>
</tr>
<tr>
<td>Mordvin</td>
<td>či, ši</td>
<td>pogoda, veņelks, ušo[łks]</td>
<td>meñeľ</td>
</tr>
<tr>
<td>Mari</td>
<td>keče</td>
<td>iyeče</td>
<td>jumo, kaľa</td>
</tr>
<tr>
<td>Udmurt</td>
<td>nunal, šündj</td>
<td>kuaz, inkuaž</td>
<td>inmar</td>
</tr>
<tr>
<td>Komi</td>
<td>lun, šondi</td>
<td>povodi’a, pogoda</td>
<td>jen, jenež, ſebesa</td>
</tr>
<tr>
<td>Khanty</td>
<td>xatt, naj</td>
<td>jet đậmam, ilam, tŏräm</td>
<td>tŏräm, kwores</td>
</tr>
<tr>
<td>Mansi</td>
<td>xōtal, nāj</td>
<td>ľăm, jēlām, tŏräm, kwores</td>
<td>jīlīn, šaxal</td>
</tr>
<tr>
<td>Hungarian</td>
<td>nap</td>
<td>idő, időjárás</td>
<td>ľeg, menny</td>
</tr>
<tr>
<td>Nenets</td>
<td>jaľa, xajer*,</td>
<td>num*</td>
<td>num’</td>
</tr>
<tr>
<td>Selkup</td>
<td>čělį, tėl</td>
<td>num, nom, nop</td>
<td>num, nom, nop</td>
</tr>
<tr>
<td>Nganasan</td>
<td>d’alj, kou</td>
<td>ńuo</td>
<td>ńuo</td>
</tr>
<tr>
<td>Kamas</td>
<td>kuja, d’žatå, tš-</td>
<td>num, nom</td>
<td>num, nom, šud’zę?, tši?</td>
</tr>
</tbody>
</table>

* This Samoyed word is an early loanword form Ugric or per-Ugric (Janhunen 1998: 477).

45. This has been compared to the corresponding words in Mari, Udmurt and Kamas seen in Table 3 (EWT 306).

46. Some of the words have the meaning ‘bad weather, storm’ first and foremost.
Grammatical structures with an object as the only argument occur frequently in Russian, e.g. *dorogu* (acc.) *zamelo* ‘a heavy snowfall covered the road’ or with a facultative agent in the instructive case *dorogu zaneslo* (*snegom*) ‘id.’ (Leinonen 1985:83; Bartens 1995:63). In Russian, there is also a grammatical structure that uses the instrumental oblique as Udmurt, Komi and Mansi do: *duet tėplym vetrom* ‘a warm wind is blowing’ (Bartens 1995:63). The OV structures in several Uralic languages form an interesting group challenging that idea that subject and verb are the basic elements of a sentence.

In the Germanic languages, the informal subject pronoun is usually neuter, but there are also examples in Norwegian dialects in which the formal subject is the masculine pronoun *han* ‘he’, and in one Norwegian dialect even the feminine pronoun *ho* ‘she’ is found as the subject of meteorological expressions (Saarinen 1997:11). This is attested also on some Swedish dialects, Icelandic and Faroese (Eriksen-Kittilä-Kolehmainen 2010:574). Another interesting linguistic feature is also known in the Baltic Sea region: formal subjects are not obligatory. Swedish and German can form meteorological sentences with a reversed word order where no formal pronominal subject is needed (Hakanen 2001:26 with references).

Both the fact that different linguistic traditions and viewpoints have played an important role in the background of the material quoted here and that the great variability in regard to their size and age lead to the situation where meteorological expressions are by no means commensurably presented. The whole question of their impersonality in the Uralic languages is based on the comparison of equivalent expressions in Indo-European languages. Some researchers have been more eager to suggest that the subjects have sometimes faded from these expressions, others are more doubtful. In fact, the issue is that Uralic verbs can have a different number of valencies depending on the chosen point of view. In weather expressions, the subject is usually so obvious that it does not need to be expressed. I hope that I have illustrated that even though many of the constructions are impersonal functionally (i.e. they lack an agentic subject), they are not impersonal formally since they have an overt subject, albeit not a human one.

**Abbreviations**

<table>
<thead>
<tr>
<th>A/G</th>
<th>accusative-genitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACC</td>
<td>accusative</td>
</tr>
<tr>
<td>ADE</td>
<td>adessive</td>
</tr>
<tr>
<td>AOR</td>
<td>aorist</td>
</tr>
<tr>
<td>AUX</td>
<td>auxiliary</td>
</tr>
<tr>
<td>causat.</td>
<td>causative</td>
</tr>
<tr>
<td>CLC</td>
<td>clitic</td>
</tr>
<tr>
<td>DAT</td>
<td>dative</td>
</tr>
<tr>
<td>DEF</td>
<td>definite</td>
</tr>
<tr>
<td>DEF.ART</td>
<td>definite article</td>
</tr>
<tr>
<td>dim.</td>
<td>diminutive</td>
</tr>
<tr>
<td>durat.</td>
<td>durative</td>
</tr>
<tr>
<td>ELAT</td>
<td>elative</td>
</tr>
<tr>
<td>ESS</td>
<td>essive</td>
</tr>
</tbody>
</table>
Dictionary sources


References


Leinonen, M. 1985. *Impersonal Sentences in Finnish and Russian* [Slavica Helsingiensia 3]. Helsinki: Department of Slavic Languages of the University of Helsinki.

Leinonen, M. 1998. The postpositive particle -to of Northern Russian dialects, compared with Permic languages (Komi Zyryan). *Studia Slavica Finlandica XV*: 74–90.


Impersonal constructions in Ket

Edward Vajda, Andrey Nefedov & Andrej Malchukov
MPI EVA (Leipzig) and Western Washington University (Bellingham)

Agreement in Ket involves multiple techniques of verb-internal subject-object marking with intransitive subject (S) markers aligning in various ways with the transitive subject (A) or transitive object (P) markers. This article examines verbs where the S appears in slots normally reserved for P marking, while the slot normally reserved for A marking contains the invariant prefix da-. Many impersonal verbs with this da- prefix turn out to be versions of regular transitives that allow for substitution of da- by other personal subject markers. Interestingly, some intransitive verbs previously recorded as lacking da- in fact allow this prefix optionally. This suggests that the range of da- intransitives was originally broader, and that more, if not all, verbs in this class derive historically from regular transitives. Accordingly, this Ket pattern is analyzable as a ‘transimpersonal construction’, whereby a transitive verb takes an indefinite subject and experiencer object, although the transimpersonal in Ket has not spread consistently across the lexicon.

Keywords: subject-object marking; transimpersonal construction

1. Introduction

This paper examines impersonal constructions in Ket. In these verbs, the intransitive subject appears in slots normally reserved for object marking, while the slot normally reserved for transitive subject marking contains the invariant prefix da-. Many impersonal verbs with this da- prefix turn out to be versions of regular transitives that allow for substitution of da- by other personal subject markers. New fieldwork has revealed that some intransitive verbs previously recorded as lacking da- in fact allow this prefix optionally. This suggests that the range of da- intransitives was originally broader, and that more, if not all, verbs in this class derive historically from regular transitives. Far from being typologically exceptional, this Ket pattern matches what have recently been called ‘transimpersonal constructions’, whereby a transitive verb takes an indefinite subject and experiencer object (see Malchukov 2008; cf. Mithun 2008). Ket has simply reanalyzed a formerly transitive construction with a non-specific subject as a patient-subject intransitive clause. However, in keeping with the strong lexical
conditioning of agreement marking strategies in Modern Ket, the resultant pattern is not consistent across the lexicon. Some verbs display a ‘frozen’ subject marker, where the da- prefix cannot be replaced by actual person agreement markers, while others do not allow this prefix at all. Therefore, far from being typological anomalous, the subset of intransitive Ket verbs examined here follows a path of reanalysis well attested elsewhere, though the resultant pattern is nevertheless morphologically highly idiosyncratic and lexicalized.

The article is organized in the following way. §2 introduces key aspects of Ket finite verb morphology. §3 gives a concise overview of verb agreement configurations in Ket. §4 provides a general description of da-intransitives, while §5 discusses their typology and evolution in the light of new fieldwork data. §6 concludes the article with a summary.

2. Overview of Ket finite verb morphology

Following Vajda (2000, 2001, 2003, 2004, 2005, 2007, 2008), the Ket finite verb can be analyzed as consisting of the ten position classes, as shown in Figure 1. The caption ‘valence’ indicates potential subject or object positions, which are used variously depending on the lexical item in question:

<table>
<thead>
<tr>
<th>P8</th>
<th>P7</th>
<th>P6</th>
<th>P5</th>
<th>P4</th>
<th>P3</th>
<th>P2</th>
<th>P1</th>
<th>P0</th>
<th>P-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>valence</td>
<td>left</td>
<td>valence</td>
<td>thematic</td>
<td>tense/mood</td>
<td>valence</td>
<td>tense/ mood</td>
<td>valence</td>
<td>(right)</td>
<td>valence</td>
</tr>
<tr>
<td>base</td>
<td>(infinitive or incorporate)</td>
<td>consonant(s)</td>
<td>or valence</td>
<td>consonant</td>
<td>base</td>
<td>(verb root or aspect/transitivity marker)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1. Position classes in Ket finite verb form creation

Most verb stems are discontinuous, with lexical morphemes separated by intervening agreement (valence) or tense/mood affixes. All productive patterns of verb stem formation require that both P7 and P0 be filled, and many contain a morpheme in P5 as well, so that these three positions form the basic lexical stem. P5 contains a non-syllabic consonantal morpheme (in a few cases two such morphemes are concatenated in this position). These lexical elements will be best glossed as TH for “thematic consonant”, since their meaning in Modern Ket is generally opaque. Because most stems represent compounds of positionally disjunct morphemes, the tense/mood and subject/object affixes interspersed among them are logically interpretable as prefixes, suffixes or infixes depending upon the location of the stem’s semantic head. In most productive stem patterns, the semantic head occupies P7, while in unproductive stem patterns the root morpheme is generally found in P0.
Which valence position(s) a stem fills is a lexical idiosyncrasy, not predictable by any overarching rule of grammar, though the shape of the markers themselves predictably follows syntactic rules of agreement, as shown in Figure 2.

<table>
<thead>
<tr>
<th>Position</th>
<th>P8 (P7 left base)</th>
<th>P6 (P5 thematic cons.)</th>
<th>P4</th>
<th>P3 (P4/2 tense/mood)</th>
<th>P1 (right base)</th>
<th>P-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>agreement functions</td>
<td>(person/class)</td>
<td>(person/class)</td>
<td>(person/class)</td>
<td>(person/class)</td>
<td>(AN-class)</td>
<td>(AN-class)</td>
</tr>
<tr>
<td></td>
<td>all A</td>
<td>some Sa and Sb</td>
<td>some A, Sa and Sb</td>
<td>some Sa, Sb and Sp</td>
<td>some Sa, Sb and Sp</td>
<td>some Sa, Sb and Sp</td>
</tr>
<tr>
<td>1SG</td>
<td>di (d, t, t)</td>
<td>ba–bo</td>
<td>–</td>
<td>–</td>
<td>di (d, t, t)</td>
<td>–</td>
</tr>
<tr>
<td>2SG</td>
<td>ku (k, g; h)</td>
<td>ku (g, u)</td>
<td>–</td>
<td>–</td>
<td>ku (k, g; h)</td>
<td>ku (g, u)</td>
</tr>
<tr>
<td>3SG</td>
<td>da (d, t, t)</td>
<td>a–o–bu</td>
<td>a–o</td>
<td>a–o</td>
<td>a–o</td>
<td>a–o</td>
</tr>
<tr>
<td>3N (SG, PL)</td>
<td>da (da, dvo)</td>
<td>i–u–bu</td>
<td>(dit, dir, it)</td>
<td>–</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>1PL</td>
<td>di (d, t, t)</td>
<td>daŋ (taŋ, ᵇaŋ)</td>
<td>–</td>
<td>–</td>
<td>dan (taŋ, ᵇaŋ)</td>
<td>n</td>
</tr>
<tr>
<td>2PL</td>
<td>ku (k, g; h)</td>
<td>kaŋ (gaŋ, ᵇaŋ)</td>
<td>–</td>
<td>–</td>
<td>kaŋ (gaŋ, ᵇaŋ)</td>
<td>n</td>
</tr>
</tbody>
</table>

Figure 2. Ket subject/object agreement markers (allomorphs given in parentheses)
A = transitive subject, P = object, SA = active intransitive subject, Sp = inactive intransitive subject

The different agreement positions obviously lack a one-to-one correspondence with individual semantic roles or syntactic functions, and no single rule can be given to predict agreement strategy. Also, some stems require the subject be marked simultaneously in two different positions, once in P8 as well as again in either P6 or P1: e.g. dabu tsaq ‘she makes a quick round trip’ [da⁸-⁰bu⁶-⁰t⁵-⁰saq⁰ 3F.SBJ⁷-⁰3BJ⁶-⁰TH⁵-⁰NPST⁴-⁰go. once⁰], datisabo ‘she loads a gun’ (originally, ‘puts an arrow to a bowstring’) [da⁸-⁰ti⁷-⁰s⁴-⁰qo⁰ 3F.SBJ⁷-⁰bowstring⁵-⁰NPST⁴-⁰3SG.SBJ⁴-⁰stretch⁰]. Multi-site subject marking is likewise a lexical idiosyncrasy of the stem that cannot be varied to convey benefactivity, version, or similar grammatical categories (Vajda 2005).

Modern Ket uses these agreement markers to build seven productive patterns, which can be called ‘Agreement Configurations’. Analogous to lexically conditioned conjugations or declensions, each Agreement Configuration employs the various valence positions somewhat differently. Tense and mood marking is accomplished through a similar combination of semantically opaque morpheme shapes in positions 4 to 2 to build six productive tense/mood classes (cf. Vajda 2003, 2005).

3. Ket agreement configurations

Verb-internal subject/object referencing is greatly complicated by the fact that the choice of agreement positions is lexically conditioned and thus an idiosyncrasy of each individual stem. Modern Ket contains two productive transitive configurations and five productive intransitive configurations. The membership of any given verb in
one or another agreement configuration cannot be predicted by any general semantic or formal rule. In addition to the plethora of productive patterns for marking subject/object agreement, a residue of unproductive agreement position configurations also exists. These include two additional transitive configurations requiring multi-slot subject agreement. Despite its highly lexicalized nature, agreement configuration membership in productive stem types is often associated with certain combinations of semantic and morphological features. In unproductive stem types (e.g. stems that lack a morpheme in P7), agreement marking is less predictable. Often, the morphological or semantic trigger characteristic of the same pattern in productive stems is lacking in the unproductive stem types, which reflect an earlier state of verb agreement in Yeniseian (Vajda 2008).

3.1 Subject/object marker combinations in transitive verbs

Let us first examine Agreement Configurations found among productive transitive stem types. The marked productive transitive pattern can be found in morphological causatives made with P5 that mark their objects in P4/3/1, depending on the object's person and grammatical class. These verbs invariably contain an infinitive in P7, which serves as the verb's semantic head.

| SBJ-ref | laug-th-caus-th-npst-th-iter-tr-th-an-pl-sbj-ref
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>d^8dáq^7q^5a^4y^1ra^0 (I make you.S laugh)</td>
</tr>
<tr>
<td>2SG</td>
<td>k^8dáq^7q^5a^4d^1da^0 (you.S make me laugh)</td>
</tr>
<tr>
<td>3M</td>
<td>d^8dáq^7q^5a^4r^1da^0 (he makes him laugh)</td>
</tr>
<tr>
<td>3F</td>
<td>da^8daq^7q^5a^4v^ry^1a^0 (she makes it laugh)</td>
</tr>
</tbody>
</table>

Figure 3. Sample paradigm fragment of Ket Transitive Configuration I

Causatives-of-state made with P0 sin and a descriptive modifier in P7 also follow this pattern: dùttapsin ‘I fill it’ [d^8d^6-ut^7-t^5-a^4-b^3-sin^0 1SB^1-full^7-TH^5-npst^4-3N.OBJ^3-cause.to.become^0].

Among unproductive stem-building patterns, this agreement pattern represents the basic, unmarked type and is therefore not associated with any particular morphological or semantic trigger: dúptêt ‘he hits it’ [d^8-b^3-tet^0 1SB^1-3N.OBJ^3-hit^0], kúris ‘you.s dress me’ [kú^8-di^1-s^0 2SB^1-1SG.OBJ^1-dress^0].

The unmarked transitive agreement pattern in productive stem types (Ket Transitive II) uses the P8 + P-1 circumfix to mark the subject, and P6 to mark the object. This pattern is characteristic of all productive types of transitive stems except those that contain P5 causative q.
Impersonal constructions in Ket

Note that in transitive forms of any type, inanimate-class subjects take the same morpheme as feminine-class transitive subjects.

The basic nature of this pattern in Modern Ket is underscored by the fact that transitive stems deriving from Russian invariably conform to it, with the borrowed material (normally an infinitive form) appearing in P7. Examples include: 

- 

Figure 4. Sample paradigm fragment of Ket Transitive Configuration II

- 

Figure 5. Sample paradigm fragment of Ket Transitive Configuration III

A handful of transitive stems display anomalous agreement configurations, which include multi-site subject agreement in P8 and P1: 

- 

Figure 6. Sample paradigm fragment of Ket Transitive Configuration III
‘I extend my hand with it’) \([di^{§}\cdot u^{6}\cdot k^{5}\cdot di^{1}\cdot qa^{0}}\text{ SBJ}^{8}\cdot 3\text{N.OBJ}^{6}\)-with^{5}\cdot 1\text{Sg.SBJ}^{1}\)-extend.\) This rare pattern could be called Transitive Configuration IV and probably is a vestige of Common Yeniseian, where the P1 slot was the original subject marking position (Vajda 2008). One verb uses this pattern for plural subjects only: \(db\text{\textasciitilde}k\text{tajga}^{\prime}\text{he leads me around}^{\prime} (< ‘walks around with me’) \([du^{8}\cdot bo^{6}\cdot k^{5}\cdot a^{j^{4}}\cdot ka^{0}}\text{ 3M.SBJ}^{8}\cdot 1\text{Sg.OBJ}^{6}\)-with/atelic^{5}\cdot \text{NPST}^{4}\cdot \text{one.walks}^{0}], db\text{\textasciitilde}k\text{tajaŋgutn}^{\prime}\text{they lead me around}^{\prime} \([du^{8}\cdot bo^{6}\cdot k^{5}\cdot a^{j^{4}}\cdot a^{j^{1}}\cdot qu^{n}^{0}}\text{ 3SBJ}^{8}\cdot 1\text{Sg.OBJ}^{6}\)-with/atelic^{5}\cdot 3\text{AN.PL.SBJ}^{1}\)-many.walk.around^{0}].

3.2 Subject agreement in Ket intransitives

Productive intransitive stem types show one basic, unmarked pattern and four patterns associated with specific formal or semantic triggers. The unmarked pattern consists of a subject agreement marker in P3 for most inanimate-class subjects, and in P8 + P-1 for animate-class.

\[
\text{SBJ}^{8}\)-hanging\text{7}.\text{TH}^{3}\text{-NPST}^{4}\cdot \text{N.SBJ}^{3}\text{-be.extended}^{8}\text{-AN.PL.SBJ}^{1} \quad \text{‘S is hanging suspended’}
\]

\[
\begin{array}{lll}
1\text{s} & d^{8}\text{än}^{7}\text{g}^{5}(i)s^{4}t^{a^{0}} & 1\text{pl} & d^{8}\text{än}^{7}\text{g}^{5}(i)s^{4}t^{a^{n^{-1}}}
\end{array}
\]

\[
\begin{array}{lll}
2\text{s} & k^{8}\text{än}^{7}\text{g}^{5}(i)s^{4}t^{a^{0}} & 2\text{pl} & k^{8}\text{än}^{7}\text{g}^{5}(i)s^{4}t^{a^{n^{-1}}}
\end{array}
\]

\[
\begin{array}{lll}
3\text{m} & d^{8}\text{än}^{7}\text{g}^{5}(i)s^{4}t^{a^{0}} & 3\text{AN.PL} & d^{8}\text{än}^{7}\text{g}^{5}(i)s^{4}t^{a^{n^{-1}}}
\end{array}
\]

\[
\begin{array}{lll}
3\text{f} & d^{8}\text{än}^{7}\text{g}^{5}(i)s^{4}t^{a^{9}} & 3\text{n} (s, pl) & \text{än}^{7}\text{g}^{5}(i)p^{3}ta^{0}
\end{array}
\]

Figure 6. Sample paradigm of Ket Intransitive Configuration I (Basic Intransitive)

Intransitive verbs based on Russian loans normally conform to this pattern, with the borrowed element appearing as the verb’s semantic head in P7: \(dals\text{\textasciitilde}veravet ‘she works’ (< Russian \(ra\text{bota ‘work’) \([da^{8}\cdot lobet^{7}\cdot a^{4}\cdot bet^{0}}\text{ 3r.SBJ}^{8}\text{-work}^{7}\text{-NPST}^{4}\text{-atelic}^{0}].\)

Most intransitives without P7 also follow this pattern, though a significant number do not, as will be shown below.

Each of the remaining five productive inanimate patterns is associated with a particular morphological or semantic feature, at least for stems containing P7. Once again, root final verbs, which reflect a more ancient pattern in the language, show different morpholexical patterns with respect to agreement marking.

The second intransitive pattern is likewise extremely widespread. Several structurally distinct types of inanimate stems cross-reference their subject in P6. These include inceptives that name the verbal action with an infinitive in P7 (example 1 in Figure 7) and change-of-state verbs that name the resultant state by using a noun or infinitive in P7 (example 2). Change-of-state verbs containing an adjective root in P7, however, belong to Intransitive I: \(daq\text{\textasciitilde}ya\text{\textasciitilde}ya\text{\textasciitilde}n ‘she gets big’ \([da^{8}\cdot qa^{7}\cdot a^{4}\cdot qan^{0}}\text{ 3r.SBJ}^{8}\text{-big}^{7}\text{-NPST}^{4}\text{-incept}^{0}].\), qa\text{\textasciitilde}ya\text{\textasciitilde}va\text{\textasciitilde}n ‘it gets big’ \([qa^{7}\cdot a^{4}\cdot b^{3}\cdot qan^{0}}\text{ 3big}^{7}\text{-NPST}^{4}\text{-3N.SBJ}^{3}\text{-incept}^{0}].\)
Impersonal constructions in Ket

Figure 7. Sample paradigm of Ket Intransitive Configuration II

The third intransitive configuration involves multi-site subject agreement in P8, P6, and P-1. The marker bu appears in all forms with 3rd person subjects, just as in Transitive III. This pattern appears in certain auto-instrumental verbs, such as ‘to whistle using one’s lips’:

\[
\begin{align*}
1s & \quad q\dot{0}t^{7}b\dot{a}^{5}y^{a}a^{9}r\pi^{0} & \quad & d\dot{u}m^{7}\bar{b}a^{5}a^{9}q^{0} \\
2s & \quad q\dot{0}t^{7}\bar{k}u^{6}y^{a}a^{9}r\pi^{0} & \quad & d\dot{u}m^{7}\bar{k}u^{6}a^{9}q^{0} \\
3m & \quad q\dot{0}t^{7}a^{6}y^{5}a^{9}r\pi^{0} & \quad & d\dot{u}m^{7}a^{9}q^{0} \\
3f & \quad q\dot{0}t^{7}i^{6}5a^{6}a^{9}r\pi^{0} & \quad & d\dot{u}m^{7}i^{6}a^{9}q^{0} \\
5n (sg, pl) & \quad q\dot{0}t^{7}i^{6}5a^{6}a^{9}r\pi^{0} & \quad & d\dot{u}m^{7}i^{6}a^{9}q^{0} \\
1p & \quad q\dot{0}t^{7}d\dot{a}n^{6}y^{a}a^{9}r\pi^{0} & \quad & d\dot{u}m^{7}d\dot{a}n^{6}a^{9}q^{0} \\
2p & \quad q\dot{0}t^{7}d\dot{a}n^{6}y^{a}a^{9}r\pi^{0} & \quad & d\dot{u}m^{7}d\dot{a}n^{6}a^{9}q^{0} \\
3p & \quad q\dot{0}t^{7}d\dot{a}n^{6}y^{a}a^{9}r\pi^{0} & \quad & d\dot{u}m^{7}d\dot{a}n^{6}a^{9}q^{0} \\
\end{align*}
\]

Figure 8. Sample paradigm of Ket Intransitive Configuration III

This pattern also productively detransitivizes Transitive II stems, creating reflexives such as "datuyunitayit ‘she combs her’ [\(da^{8}-tukun^{7}-t^{6}-a^{4}-kit^{0}\) 3F.SBJ\(^{6}\)-comb\(^{7}-3F.SBJ\(^{6}\)-rub\(^{0}\)] \(\rightarrow\) datuyunbutayit ‘she combs herself’ [\(da^{8}-tukun^{7}-bu^{6}-t^{6}-a^{4}-kit^{0}\) 3F.SBJ\(^{6}\)-comb\(^{7}-3F.SBJ\(^{6}\)-TH\(^{5}\)-NPST\(^{4}\)-rub\(^{0}\)]. Other verbs of this type have no transitive equivalent and are not clearly linked to notions of reflexivity. Some Intransitive III verbs express unusually quick or intense motions, such as \(da^{8}i^{g}g^{a}d^{a}b^{u}^{6}-t^{6}-s^{4}-a^{q^{0}}\) 3F.SBJ\(^{6}\)-riverbank\(^{7}-3SBJ\(^{6}\)-TH\(^{5}\)-NPST\(^{4}\)-go.mom\(^{0}\)], or \(dab\u{u}nj\u{u}ss\u{u}p\u{u} \u{u} ‘she looks intently, searches’ [\(da^{8}-bu^{6}-n^{5}-s^{4}-q^{0}\) 3F.SBJ\(^{6}\)-3SBJ\(^{6}\)-TH\(^{5}\)-NPST\(^{4}\)-watch\(^{0}\)]. Synonyms without this nuance follow the basic intransitive pattern: \(da^{8}g^{a}d^{a}k^{a}q^{a}a^{q^{0}}\) ‘she makes a round trip to the river lasting several days’ [\(da^{8}-ig^{a}g^{a}d^{a}b^{u}^{6}-t^{6}-s^{4}-aq^{0}\) 3F.SBJ\(^{6}\)-riverbank\(^{7}-3SBJ\(^{6}\)-TH\(^{5}\)-NPST\(^{4}\)-go.mom\(^{0}\)], \(d\u{u}r\u{u}y\u{u}r\u{u} \u{u} ‘she looks, watches’ [\(da^{8}-k^{5}-a^{4}-do^{0}\) 3F.SBJ\(^{8}\)-TH\(^{5}\)-NPST\(^{4}\)-look\(^{0}\)]. Most reciprocal verbs also follow Intransitive Configuration III: that-\u{a}n\u{u}bu\u{u}ksiv\u{u}ten ‘they hug’ [\(du^{8}-h\u{a}ta\u{u}n^{7}-bu^{6}-k^{5}-s^{4}-b\u{u}t^{0}\) 3SBJ\(^{8}\)-close\(^{7}-3SBJ\(^{8}\)-TH\(^{5}\)-NPST\(^{4}\)-make\(^{0}\)] ‘they hug (each other)’. In all of its various meanings, Transitive III appears restricted to verbs that require a sentient subject.
The fourth intransitive pattern requires multiple-slot subject marking in P8 and P1 and can be used with inanimate-class as well as animate-class subjects. Unlike all other patterns that require P8, the P-1 animate-class plural suffix does not appear, so that subject number is expressed only in P1:

\[
\text{SBJ}^8 \text{-laugh}^7 \cdot \text{CAUS}^5 \cdot \text{NPST}^4 \cdot \text{SS}^1 \cdot \text{ITER}^0 \cdot \text{S starts laughing (repetitively)}
\]

1s \(d^8\text{dåq}^7q^a\text{a}^4\text{dij}^0\) (I start laughing) \hspace{1cm} 1pl \(d^8\text{dåq}^7q^a\text{\text{tan}}^1\text{dij}^0\) (we start laughing)

2s \(k^8\text{dåq}^7q^a\text{y}^u\text{rij}^0\) (you.S start laughing) \hspace{1cm} 2pl \(k^8\text{dåq}^7q^a\text{\text{tan}}^1\text{dij}^0\) (you.PL start laughing)

3m \(d^8\text{dåq}^7q^a\text{a}^4\text{ri}^j^0\) (he starts laughing) \hspace{1cm} 3AN.PL \(d^8\text{dåq}^7q^a\text{\text{an}}^1\text{dij}^0\) (they start laughing)

3f \(d^8\text{dåq}^7q^a\text{a}^4\text{ri}^j^0\) (she starts laughing)

Figure 9. Sample paradigm of Ket Intransitive Configuration IV

This pattern productively detransitivizes Transitive I stems, but many root-final verbs that follow this pattern have neither transitive stem analogs nor reflexive semantics and are instead basic intransitives. These include many basic verbs expressing the subject’s active change of state: \(\text{datájaraq} \text{‘she falls’} \) [\(d^8-\text{t}^5-\text{aj}^4-\text{daq}^0\) \(3S\text{.SBJ}^8-\text{TH}^5\text{-NPST}^4\text{-3S.SBJ}^1\text{-fall}^0\)], \(\text{daájatiq} \text{‘she grows’} \) [\(d^8-\text{aj}^4-\text{tij}^0\) \(3S\text{.SBJ}^8\text{-NPST}^4\text{-3S.SBJ}^1\text{-grow}^0\)].

The fifth and last productive intransitive type comprises mostly \text{habeo}-verbs that incorporate a monosyllabic possessum noun in P7 and express their subjects in P4/1.

\[
\text{dog}^7 \cdot \text{SBJ}^4/1 \cdot \text{have}^0 \quad \text{‘S has a dog’}
\]

1s \(\text{tip}^7\text{di}^1\text{vet}^0\) \hspace{1cm} 1pl \(\text{tip}^7\text{dan}^1\text{vet}^0\)

2s \(\text{tip}^7\text{ku}^1\text{vet}^0\) \hspace{1cm} 2pl \(\text{tip}^7\text{kan}^1\text{vet}^0\)

3m \(\text{tip}^7\text{a}^4\text{bet}^0\) \hspace{1cm} 3AN.PL \(\text{tip}^7\text{an}^4\text{vet}^0\)

3f \(\text{tip}^7\text{ij}^4\text{bet}^0\)

Figure 10. Sample paradigm of Ket Intransitive Configuration V

\text{Habeo}-verbs with polysyllabic incorporates follow Intransitive Configuration II rather than V: \(\text{uuntipbájbet} \text{‘own a puppy’} \) [\(\text{uuntip}^7\text{-\text{ba}^6\text{-bet}^0}\) \(\text{puppy}^7\text{-1.SG.SBJ}^6\text{-have}^0\) \text{‘own a puppy’}]. A few intransitives belonging to other semantic groups follow this agreement pattern, as well, so that it cannot be identified exclusively with possession: \(\text{iritet} \text{‘I spend the day’} \) [\(i^7\text{-di}^1\text{-bet}^0\) \(\text{day}^7\text{-1SG.SBJ}^1\text{-have}^0\)], \(\text{sitkaya} \text{‘you.PL wake up’} \) [\(s^7\text{-\text{ka}^1\text{-a}^0}\) \(\text{awake}^7\text{-2PL.SBJ}^4\text{-process.occurs}^0\)]. Similar to \text{habeo}-verbs, however, the remaining stems belonging to Intransitive Configuration V logically require sentient subjects; their conjugations are thus normally defective for the inanimate-class subject form.

Finally, as might be expected in a language where the agreement-marking strategy is lexically conditioned in idiosyncratic ways, some Ket intransitives have rare or
Impersonal constructions in Ket

Even unique agreement configurations. Several verbs use multi-site subject marking only in the plural forms: *diranddaŋŋa* ‘we fly’ [diŋ- day*-doq/ŋ0 1sbj*-fly/1pl.sbj*] (cf. *dir* ‘I fly’ [diŋ-(ji)-doq0 1sbj*-fly0]). The past tense forms of one intransitive verb require subject marking in P6 and P1: *éjbagbindy* ‘I jumped up’ [eŋ*-ba*-k5*-in2-di*-qos0 up7-1sg.sbj*-th3-pst3-1sg.sbj*-jump0]. A near exhaustive listing of these exceptional verb types can be found in Vajda (2004: 69–71).

Given the extremely complicated lexical nature of verb agreement in Ket, the identification of any pattern that simplifies the overall picture is worthy of attention, as it could shed light on the origin of this unusual system. The remainder of this article examines a phenomenon that suggests certain intransitives evolved out of impersonal uses of the two productive transitive conjugations.

4. *Da*-intransitives

Ket verbs possess another category of lexical morpheme in addition to the primary stem-building components occupying positions 7-5-0. The leftmost slot (P8) may contain an affix *da*- that expresses no grammatical agreement and thus remains in all conjugated verb forms. Interestingly, this morpheme is identical to the 3p inanimate-class subject marker used in Transitive Configurations I and II. Verbs containing this affix, which Vajda (2003) called an “involuntary causative” marker, can be called ‘*da*-intransitives’. In terms of the positioning of their subject agreement marker, they belong either to Intransitive Configuration II, with subject marking in P6, or Intransitive Configuration V, with subject marking in P4-3-1:

<table>
<thead>
<tr>
<th>Subject</th>
<th><em>da</em>-súlej7</th>
<th><em>da</em>-ét7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1S</td>
<td>db<em>ks5s</em>a0</td>
<td>d6et<em>di</em>rus0</td>
</tr>
<tr>
<td>2S</td>
<td>d6súlej7ku6ks5s*a0</td>
<td>d6et<em>ku</em>rus0</td>
</tr>
<tr>
<td>3M</td>
<td>d6súlej7g<em>ks5s</em>a0</td>
<td>d6et<em>aj</em>us0</td>
</tr>
<tr>
<td>3F</td>
<td>d6súlej7u<em>ks5s</em>a0</td>
<td>d6et<em>ii</em> Gus0</td>
</tr>
<tr>
<td>3N (SG, PL)</td>
<td>d6súlej7u<em>ks5s</em>a0</td>
<td>d6et<em>bi</em>rus0</td>
</tr>
<tr>
<td>1PL</td>
<td>d6súlej7dan6ks5s*a0</td>
<td>d6et<em>dan</em> Gus0</td>
</tr>
<tr>
<td>2PL</td>
<td>d6súlej7kan6ks5s*a0</td>
<td>d6et<em>kan</em> Gus0</td>
</tr>
<tr>
<td>3AN.PL</td>
<td>d6súlej7a<em>n</em>ks5s*a0</td>
<td>d6et<em>an</em> Gus0</td>
</tr>
</tbody>
</table>

*Figure 11. Da*-intransitives belonging to Intransitive Configurations II and V

*Da*-intransitives resembling the second example in Figure 11 are few in number and denote independent movements, such as ‘S jumps up,’ ‘S spins around.’ These
verbs, which mark their intransitive subject in positions 4, 3, or 1 (depending on person and gender, following the general pattern shown in Figure 2 above) belong to Intransitive Configuration V and sometimes correlate with transitive verbs belonging to Transitive Configuration I.

\[
\begin{array}{ll}
\text{'S lifts O up'} & \text{'S jumps up' (literally, 'it takes S up')} \\
\text{SBJ}^8\text{-up}^7\text{-OBJ}^{6/3/1}\text{-NPST}^4\text{-take}^0 & \text{R}^8\text{-up}^7\text{-NPST}^4\text{-SBJ}^{6/3/1}\text{-take}^0 \\
\text{k}^8\text{ét}^7\text{di}^1\text{rus}^0 & \text{da}^8\text{ét}^7\text{di}^1\text{rus}^0 (\text{I jump up}) \\
\text{d}^8\text{ét}^7\text{ku}^1\text{rus}^0 (\text{I lift up}) & \text{da}^8\text{ét}^7\text{ku}^1\text{rus}^0 (\text{you.S jump up}), \text{etc.} \\
\end{array}
\]

Figure 12. Transitive Configuration I and related \textit{da}-verb of Intransitive Configuration V

While it is difficult to assign a specific semantic function other than general detransitivization to the prefix \textit{da}- in such verbs, the \textit{da}-intransitives belonging to Intransitive Configuration II are involuntary causatives of various sorts. They mark their subject in P6 and contain an incorporate in P7 that names the effect produced. This pattern is productive with verbs denoting that the subject becomes a certain color, another example being \textit{daqilejoka} 'it turns yellow' \([\text{da}^8\text{-qalej}^7\text{-u}^6\text{-k}^5\text{-s}^4\text{-a}^0 \text{IC}^8\text{-yellow}^7\text{-3N.SBJ}^6\text{-TH}^2\text{-NPST}^4\text{-process.occurs}^0]\). A few of these verbs correlate with verbs belonging to Transitive Configuration II, as shown in Figure 13:

\[
\begin{array}{ll}
\text{'S wrinkles, rumples O'} & \text{'S gets wrinkled' (literally, 'it wrinkles S')} \\
\text{SBJ}^8\text{-wrinkles}^7\text{-OBJ}^6\text{-TH}^1\text{-NPST}^4\text{-make}^0 & \text{IC}^8\text{-wrinkles}^7\text{-SBJ}^6\text{-TH}^2\text{-NPST}^4\text{-make}^0 \\
\text{t}^6\text{kúran}^7\text{u}^6\text{k}^5\text{s}^4\text{bet}^0 (\text{I wrinkle it}) & \text{da}^8\text{kúran}^7\text{b}^6\text{k}^5\text{s}^4\text{bet}^0 (\text{I become wrinkled}) \\
\text{k}^8\text{kúran}^7\text{u}^6\text{k}^5\text{s}^4\text{bet}^0 (\text{you.S wrinkle it}) & \text{da}^8\text{kúran}^7\text{ku}^1\text{k}^5\text{s}^4\text{bet}^0 (\text{you.S become wrinkled}) \\
\text{t}^6\text{kúran}^7\text{u}^6\text{k}^5\text{s}^4\text{bet}^0 (\text{he wrinkles it}) & \text{da}^8\text{kúran}^7\text{f}^6\text{k}^5\text{s}^4\text{bet}^0 (\text{he becomes wrinkled}) \\
\text{da}^8\text{kúran}^7\text{u}^6\text{k}^5\text{s}^4\text{bet}^0 (\text{she/it wrinkles it}), \text{etc.} & \text{da}^8\text{kúran}^7\text{u}^6\text{k}^5\text{s}^4\text{bet}^0 (\text{it becomes wrinkled}), \text{etc.} \\
\end{array}
\]

Figure 13. Transitive Configuration II and related \textit{da}-verb of Intransitive Configuration II

Recent fieldwork undertaken by two of the authors during August 2008 (Andrey Nefedov and Edward Vajda) and by Andrei Nefedov during Jan.–Feb. 2009 analyzed the semantic and morphological categories of \textit{da}-intransitives, leading to the discovery of a number of additional facts relating to these verbs\(^1\). The findings largely confirm the hypothesis of the present article’s third author (Andrej Malchukov) that such verbs originated as detransitivizing derivations analogous to impersonal verb constructions cross linguistically.

---

\(^1\) The authors express their deepest gratitude to Bernard Comrie, Director of the Linguistics Department of Max Planck Institute for Evolutionary Anthropology (Leipzig), for funding all of the research that went into writing this article.
5. Typology and the evolution of da-intransitives

A comprehensive survey of 48 Ket verbs of impersonal semantics undertaken by Andrei Nefedov in January 2009 while working with Native Ket consultant V.A. Romanenkova revealed the following formal categories. Eight intransitives that obligatorily contain da- in all of their conjugated forms are associated with corresponding transitive verbs, while seven others have no corresponding transitive stem. In six others, the prefix da- is optional. The optional nature of da- with Ket verbs was first discovered during elicitation from V.A. Romanenkova by the three authors during September 2007 and is not reflected in any of the earlier published materials. In the standard work on the Ket verb by E. A. Kreynovich (1968) these verbs were given only without the prefix da-. The fact that da- can be optional suggests that in other verbs this prefix may have simply dropped altogether. For example, the remaining 20 impersonal verbs examined showed no ability to combine with the prefix da-. While these appear to be primary Intransitive Configuration II or V verbs, it is possible that they, too, were originally derived by intransitivizing da-, but have subsequently lost this morpheme.

The eight da-intransitives attested with corresponding transitive stems have the meanings WRINKLE (example 1), SWAY (example 2), WINK/SQUEEZE (example 3), COME.INTO.VIEW/BRING (example 4), BECOME.ENRAGED/FLY. INSIDE.OF (example 5), TURN (example 6), and JUMP UP/RAISE UP (example 7), SWAY.FROM.SIDE-TO-SIDE (example 8). In the example sets below, the da-intransitives appear under /a/, while the corresponding transitives follow under /b/ and /c/:

(1) a. dakuranuksivet
da^a^1-kudan^a^7-u^6^1-k^5-s^4-bed^0
ic^8^1-wrinkles^7^3n.sbj^6^th^5-npst^4-make^0

'it becomes wrinkled.' (literally, 'it wrinkles it')

b. āt ba'at tkuranuksivet
ād ba^2^d dīt^8^kudan^7^u^6^1-k^5-s^4-bed^0
1sg forehead 1sbj^8^-wrinkles^7^3n.obj^6^-th^5-npst^4-make^0

' I wrinkle my brow (repeatedly).'  

c. āt ba'at tkurapuksivet
ād ba'ād dīt[kudab^7^u^6^1-k^5-s^4-bed^0
1sg forehead 1sbj^8^-wrinkle^7^3n.obj^6^-th^5-npst^4-make^0

' I wrinkle my brow (once).'

(2) a. datinjboyabet
da^8^-tiŋ^7^bo^6^-k^5-a^4-bed^0
ic^8^-turn^7^1sg.sbj^6^-th^5-npst^4-iter^0

' I sway (back and forth while standing in place).'

(literally, 'it sways me')
b. bú át ttiŋboyabet
   bú ád du8-ttiŋ7-bo6-k5-a4-bet0
   3SG 1SG 3SBJ8-turn7-1SG.OBJ6-TH5-NPST4-make0
   ‘He makes me sway.’/‘He shakes me.’

(3) a. dalayejuksa
da8-lakej7-u6-k5-s4-a0
   1C8-press.close7-3N.SBJ6-TH5-NPST4-change.occurs0
   ‘It blinks.’ (literally, ‘it presses it close’)

b. át duu7 baya
   ád duu7 ba-ŋa
   1SG child 1SG.POSS-DAT
dlalayejuksa
di8-lakej7-o6-k5-s4-a0
   1SBJ8-press.close7-3M.OBJ6-TH5-NPST4-change.occurs0
   ‘I will press the child close to myself.’

(4) a. Ĩbot dauksaq
   iqot da8-u6-k5-s4-aq0
   sunshine 1C8-3N.SBJ6-TH5-NPST4-move.once0
   ‘The sun will come out.’ (literally, ‘it brings the sun out’)

b. tλ7ŋ ět builda
   tλ7ŋ ět builda
   money 1SG all
   lapkadiŋa duksaq
   lapka-di-ŋa di8-u6-k5-s4-aq0
   shop-3N.POSS-DAT 1SBJ8-3N.OBJ6-TH5-NPST4-move.once0
   ‘I will bring all the money to the shop.’

(5) a. dqabaddoq
da8-qa7-ba6-t5-doq0
   1C8-inside7-1SG.SBJ6-TH5-fly0
   ‘I become enraged.’ (literally, ‘it flies inside of me’)

b. ět ět luitis tqabaddoq
   ět ět luitis du8-qa7-ba6-t5-doq0
   as.if 1SG devil 3M.SBJ8-inside7-1SG.OBJ6-TH5-fly0
   ‘As if a devil possesses me.’ (literally, ‘will fly inside me’)

(6) a. ět dλṭtaŋ
   ěd da8-di1-tay0
   1SG 1C8-up7-1SG.SBJ1-turn0
   ‘I spin around.’ (literally, ‘it spins me’)


b.  bū āt duttan
bū ād du₈-di₁-тан⁰
3-PL 1SG 3SBJ⁶-1SG.OBJ¹-spin⁰
‘He spins me.’

(7) a.  āt daettosos
ād da₈-et²-di₁-qos⁰-in⁻¹
1SG 1C⁸-up⁷-1SG.SBJ¹-take⁰
‘I jump up.’ (literally, ‘it raises me up’)
b.  būŋ āt dettososin
būŋ ād du₈-et²-di₁-qos⁰-in⁻¹
3-PL 1SG 3SBJ⁸-up⁷-1SG.OBJ¹-take⁰-AN.PL.SBJⁱ
‘They lift me up.’

(8) a.  datatinboij
da₈-tatin⁷-bo⁶-k⁵-ij⁰
1C⁸-side.to.side⁷-1SG.SBJ⁶-TH⁵-ITER⁰
‘I stagger (sway from side to side while walking forward).’
b.  bū āt ttatinboij
bū ād du₈-tatin⁷-bo⁶-k⁵-ij⁰
3SG 1SG 3SBJ⁶-side.to.side⁷-TH⁵-ITER⁰
‘He sways (rocks) me from side to side.’

Example (8b) was provided by our consultant with some hesitation, possibly for the logical reason that the motion of swaying back and forth while walking forward cannot be readily caused by another person.

Note that sentences containing da-verbs in examples (1a–8a) do not permit inclusion of a subject noun denoting the causal agent. The verb’s sole agreement marker correlates with the affected person, so that these verbs must be considered intransitives. A noun or pronoun denoting the affected person (i.e. the subject) may be freely included, though its frequent omission serves as a backgrounding device in discourse.

A number of other da-intransitives have no corresponding transitives in the extant literature on the Ket verb. These include verbs denoting changes in color as well as several others:

(9)  āt dasulejboks
ād da₈-sulej²-bo⁶-k⁵-s⁴-a⁰
1SG 1C⁸-blood.colored⁷-1SG.SBJ⁶-TH⁵-NPST⁴-event.occurs⁰
‘I blush.’ (literally, ‘it reddens me’)

Verb phrases containing these da-intransitives, just like those in examples (1a–8a) that have corresponding transitives, do not permit inclusion of a subject noun expressing the causal agent.
Six other intransitive verbs were found to allow \textit{da}- optionally. The verbs in question are shown in examples (10–15). As mentioned above, this feature had not been recorded previously, the verbs in question having been recorded only without \textit{da}- As with the two groups of \textit{da}-intransitives mentioned above, the verb phrase containing them does not permit a 3p inanimate-class noun that would correlate with \textit{da}- as subject. Of the verbs in question, two of them (10–11) appear to be \textit{da}-intransitives that have begun to lose the involuntary causative morpheme, perhaps due to the opacity of their internal semantic structure.

(10) \textit{(da)}tatiŋbatava
\textit{(da)}-tatiŋ\textsuperscript{7}-ba\textsuperscript{6}-t\textsuperscript{5}-a\textsuperscript{4}-b\textsuperscript{3}-a\textsuperscript{0} \\
1sg (\textsuperscript{ic8}-)-straight\textsuperscript{7}-1sg.sbj\textsuperscript{6}-TH\textsuperscript{5}-NPST\textsuperscript{4}-TH\textsuperscript{3}-change.occurs\textsuperscript{0} \\
'I agree.' (? literally, 'it straightens me')

(11) \textit{ap} qaanas bu\textit{(ra)atabqo}
\textit{ap} qaan-as \textit{bū} \textit{(da)}-a\textsuperscript{6}-t\textsuperscript{5}-a\textsuperscript{4}-b\textsuperscript{3}-qo\textsuperscript{0} \\
1sg.poss words-instr 3 (\textsuperscript{ic8}-)3m.sbj\textsuperscript{6}-TH\textsuperscript{5}-NPST\textsuperscript{4}-TH\textsuperscript{3}-R\textsuperscript{0} \\
'He agrees with my words.' (internal structure of this verb is largely opaque)

(12) \textit{sanana} \textit{(da)}ayavut
\textit{sanana} \textit{(da)}-a\textsuperscript{6}-k\textsuperscript{5}-a\textsuperscript{4}-b\textsuperscript{3}-qut\textsuperscript{0} \\
\textit{spark} (\textsuperscript{3n.sbj}\textsuperscript{6}-)3m.sbj\textsuperscript{6}-TH\textsuperscript{5}-NPST\textsuperscript{4}-TH\textsuperscript{3}-rise.up\textsuperscript{0} \\
'A spark hits me.'/I get hit by a spark.'

In Kreynovich (1968:82) the verb in (12) was recorded consistently without the prefix \textit{da}-, though the nominal element \textit{sanana} 'spark' was shown as a separate word rather than as an element incorporated into the verb form. In the speech of our consultant, the presence or absence of \textit{da}- seems to correlate with whether \textit{sanana} 'spark' is incorporated into the verb form. The remaining three verbs in which \textit{da}- was recorded as optional (13–15) differ from examples (10–12) in that the \textit{da}- prefix is obligatorily present in transitive readings, but absent or optional when the causal factor behind the given action is not expressed in the verb phrase. These verbs have previously been recorded only in their intransitive sense, without the prefix \textit{da}- (cf. Kreynovich 1968:80–87):

(13) a. \textit{qo}^j\textit{deyna ira atabgit}
\textit{qo}^j\textit{ deyna} \textit{i} \textit{da}-a\textsuperscript{6}-t\textsuperscript{5}-a\textsuperscript{4}-b\textsuperscript{3}-kit\textsuperscript{0} \\
bear people-poss.an.pl scent \textsuperscript{ic8}-3m.sbj\textsuperscript{6}-TH\textsuperscript{5}-NPST\textsuperscript{4}-TH\textsuperscript{3}-rub\textsuperscript{0} \\
'The bear catches people's scent.' (literally, 'people's scent rubs the bear')

b. \textit{(da)}atabgit
\textit{(da)}-a\textsuperscript{6}-t\textsuperscript{5}-a\textsuperscript{4}-b\textsuperscript{3}-kit\textsuperscript{0} \\
\textit{(ic8)-}3m.sbj\textsuperscript{6}-TH\textsuperscript{5}-NPST\textsuperscript{4}-TH\textsuperscript{3}-rub\textsuperscript{0} \\
'He smells (something).' (literally, 'it rubs him')
a. \textit{dajëŋ āt daboltaqboksibet}
\textit{dajëŋ ād da^8-boltaq^7-bo^6-k^5-s^4-bed^0}
illness 1SG 3N.SBJ^8-rolled.up^7-1SG.OBJ^6-TH^5-NPST^4-make^0
‘I fall ill.’ (literally, ‘illness doubles me up’)

b. \textit{āt (da)boltaqboksibet}
\textit{ād (da^8-)boltaq^7-bo^6-k^5-s^4-bed^0}
1SG 3N.SBJ^8-rolled.up^7-1SG.OBJ^6-TH^5-NPST^4-make^0
‘I fall ill.’ (literally, ‘it doubles me up’)

(15) a. \textit{sës āt daboksivij}
\textit{sës ād da^8-bo^6-k^5-s^4-bej^0}
river 1SG 3N.SBJ^8-1SG.OBJ^6-TH^5-NPST^4-blow^0
‘The river carries me along.’

b. \textit{(da)boksivij}
\textit{(da^8-)bo^6-k^5-s^4-bej^0}
\textit{(ic^8-)1SG.OBJ^6-TH^5-NPST^4-blow^0}
‘I get carried along’ (literally, ‘it blows me along’)

Finally, twenty other intransitive verbs do not allow the presence of \textit{da-} even as an option. These include verbs with the meanings FEAR, BE ASHAMED, GET COOKED, SLIP, GET SNOWED ON, and the like.

The new data collected on \textit{da-}intransitives and their association with transitive stems suggest several interesting patterns in the distribution of non-agreement \textit{da-} in Modern Ket.

First, the issue of whether \textit{da-} is obligatory, optional, or not possible with intransitive verbs seems roughly to follow a gradient of transitivity ranging from canonical transitive verbs (whose associated intransitives tend to require retention of \textit{da-}) and endings with intransitive verbs (lack of \textit{da-}) with tranimpersonals as intermediate stages (\textit{da-}intransitives).

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
\textbf{TRANSITIVE} & \textbf{ACTIVE INTRANSITIVE} & \textbf{INACTIVE INTRANSITIVE} \\
\hline
Canonical transitives & \textit{da-}intransitives & intransitives lacking \textit{da-} \\
\hline
\end{tabular}
\caption{Transitivity and the affix \textit{da-}}
\end{table}

As far as concerns which of these verbs include and which omit the morpheme \textit{da-}, the observation can be made that many verbs that do not allow \textit{da-}insertion have a nominal element incorporated in P7. For example, verbs build with the P0 base -\textit{qan} ‘begin’ contain an infinitive in P7; none of these verbs allows \textit{da-}, regardless of whether the verbal semantics are active (\textit{ilbäyäkän} ‘I begin to sing’) or inactive (\textit{qotbäyäkän} ‘I begin to cry’). It is likely that this infinitive incorporant originally represented the
syntactic subject. Infinitives routinely serve as the syntactic subject of phasal verbs of completion, for example (cf. āpīl binut ‘My singing ended’). If this is correct, then loss of the original inanimate-class subject marker da- could have occurred upon the incorporation of the infinitive form into the verb complex. The example in (12) above, likewise suggests that the incorporation of a P7 nominal into the verb form is associated with deletion of da-. Ket intransitive verbs in which the causal agent or change of state is expressed by a nominal form may lack agreement morphology altogether (cf. sildayaan ‘it became summer’, ilvej ‘it was windy’).

Imperative formation from verbs that involve the morpheme da- suggest a second preliminary generalization. Most verbs that allow insertion of da-, as well as da-intransitives that have morphologically related transitive stems, form their imperatives in a regular fashion through deletion of any P8 subject agreement morpheme. Verbs lacking da- altogether, which fall generally lower on the transitivity gradient, seem to prefer an optative construction, corresponding to the technique typically used for Ket inactive intransitive verbs. The example set below shows imperative formation from a transitive verb (16a), da-intransitive verb associated with a transitive stem (16b), da-intransitive not associated with a transitive but in which da- is obligatory (16c), and finally from an intransitive that does not permit the inclusion of da- (16d).

Note that the imperatives of verbs in which da- is optional or not possible are formed periphrastically using the optative particle qān, while da-intransitives form imperatives using the pattern similar to transitives. Transitive verbs, on the other hand, form their imperatives synthetically, through deletion of the P8 morpheme and addition of a P2 affix /n/ or /l/, a lexically conditioned choice in which /n/ correlates generally with achievements and /l/ with activities, processes, or repetitions (Vajda 2004: 46):

(16)  a. ettonos (transitive verb whose indicative forms require P8 subject agreement)
     et7-n7-di1-qos0 (note metathesis of position 2 and 1 morphemes)
     up7-IMP2-1sg.sbj1-take0
     ‘Lift me up!’

     b. ettunos (P8 da- is obligatory in indicative forms)
     et7-n7-ku1-qos0 (note metathesis of position 2 and 1 morphemes)
     up7-IMP2-2sg.sbj1-take0
     ‘Jump up!’

     c. sulejula (P8 da- is obligatory in indicative forms)
     sulej7-ku6-k5-l2-a0
     red7-2sg.sbj6-th5-imp2-change.occurs0
     ‘Blush!’

     d.ū qān sigutaq (P8 da- not attested in any form of this verb)
     ū qān sī7-kw6-t5-aq0
     2sg OPTATIVE exist8-2sg6-th3-INCEPTIVE0
     ‘Be born!’
Intransitives so far identified as containing *da*- optionally are few in number. In the corpus of 48 verbs tested by A. Nefedov, none of these allowed formation of an imperative or optative construction of either type shown in (16a–d); this is in keeping with the fact that many intransitives do not form imperatives at all. However, the verb ‘understand’, which may be related etymologically to verbs denoting ‘sensing’ or ‘smelling’ discussed above (cf. examples 13a–b), perhaps provides an illustration of such a verb. As would be expected from a less active intransitive verb, this verb builds its imperative periphrastically by preposing the optative particle *qān* ‘let’ to the indicative form, and lacks the synthetic imperative characteristic of verbs whose indicative forms obligatorily contain *da*–.

(17)  
\[ū \, qān \, dakutavet\]  
\[ū \ \, qān \ \, da^8-ki^6-t^2-a^4-bet\]  
\[2\text{SG OPTATIVE IC}^8-\text{2SG}^6-\text{TH}^5-\text{NPST}^4-\text{R}^0\]  
‘Understand!’

Although imperative/optative forms cannot be elicited from many intransitive verbs, the attested patterns of imperative formation indicate that *da*- is linked to greater transitivity, since *da*-intransitives, like transitive verbs, tend to form their imperatives synthetically by a process involving the deletion of the P8 morpheme. This further suggests that P8 ‘involuntary causative’ *da*- derived from the homonymous inanimate-class subject marker *da*- found as a regular agreement marker in transitive verbs.

A final piece of evidence linking the tendency to retain *da*- in intransitive verbs to the notion of higher transitivity involves the semantic interpretation of deverbal nominalizations. In general, conjugated Ket verb forms can be nominalized by adding the universal nominalizing suffix –*s*, creating what Nefedov (in preparation) calls “headless relative clauses”. Nominalizations made from *da*-intransitives that are related to transitive verbs or which obligatorily retain *da*- favor an agentive interpretation (e.g. Example 18), while purely intransitive verbs favour rather a patientive interpretation (e.g. Example 19).

(18)  
\[dadāyajuskas\ (da- \text{is obligatory})\]  
\[da^8-dγ^8ku^8-u^6-k^5-s^4-\text{a}^0-\text{s}\]  
\[\text{IC}^8-\text{flare}^7-\text{3N}^6-\text{TH}^5-\text{NPST}^4-\text{EVENT}^0-\text{NOMINALIZER}\]  
‘Something that flares up.’

(19)  
\[anugbunsanš\ (da- \text{is never present})\]  
\[γn^7-u^6-k^5-b^3-n^2-sah\]  
\[\text{IC}^8-\text{cook}^7-\text{3N}^6-\text{TH}^5-\text{TH}^3-\text{occur}^0-\text{NOMINALIZER}\]  
‘Something that got boiled.’

To summarize, a number of formal and semantic patterns observable with *da*- intransitives suggest these verbs continue to align with notions of higher transitivity.
and likely derived historically from transitive verbs through a process in which the original inanimate-class subject marker *da-* became a non-agreement morpheme marking detransitivization.

6. Conclusions

The examination of *da*-intransitives undertaken above reveals two new facts about the famously idiosyncratic morphological system of verb agreement in Ket. First, most, if not all *da*-intransitives seem to have developed on the basis of impersonal uses of transitive verbs, in which *da-* was originally an inanimate-class subject marker. The subject markers in the resulting *da*-intransitives were originally object markers, which explains why they correspond perfectly to the object markers found in transitive verbs. This is suggested both by the impersonal semantics of many of these verbs, as well as by the formal morphological relationship of certain *da*-intransitives to corresponding transitives. Second, many intransitives belonging to Configurations II and V appear to have originated in the same way, with subsequent loss or partial loss of an earlier detransitivizing prefix *da-*. The resultant patterns simply contributed to the array of agreement patterns in the language rather than developed into a new, completely regular pattern of agreement.

Taking into account patterns in the original detransitivizing function of *da-*, together with the use of multi-subject marking associated with reflexive or reciprocal meaning, can greatly simplify our understanding of the lexical associations between Modern Ket verbs belonging to different Agreement Configurations. Many if not most Ket verbs that belong to productive patterns within Intransitive Configurations II, III, IV, V, as well as those belonging to the two unproductive Transitive Configurations that use multi-site subject marking, evince an etymological relationship with one or the other of the two productive Transitive Configurations in the following way:

- Transitive Configuration I + *da-* \(\rightarrow\) Intransitive Configuration V
- Transitive Configuration I + reflexive/reciprocal object \(\rightarrow\) Intransitive Configuration IV
- Transitive Configuration II + *da-* \(\rightarrow\) Intransitive Configuration II
- Transitive Configuration II + reflexive/reciprocal object \(\rightarrow\) Intransitive Configuration IV
- Transitive Configuration II + reflexive instrument \(\rightarrow\) Transitive Configuration III

**Figure 15.** Etymological relationships between the various Transitive and Intransitive Configurations in Modern Ket

Verbs belonging to Intransitive Configuration I represent the default class of verbs, which lacks any discernable etymological relationship with either of the productive
transitive conjugations. One can therefore posit a general diachronic rule stating that, at least as far as concerns productive patterns of Ket verb morphology, subject agreement is expressed in P8 for all verbs (both transitive and intransitive) except in the case of intransitives that derive from transitives via some sort of morphological process of detransitivization or intransitives that have undergone incorporation of the subject nominal (P7) into the verb form.

It must be stressed that the etymological patterns shown in Figure 15 still do not explain the origin of agreement marking in many basic verbs. These forms were presumably inherited into Modern Ket from an earlier time when the patterns shown in Figure 15 had not yet developed. Idiosyncrasies in the oldest layer of Ket vocabulary, taken together with comparisons of verb morphology from the extinct Yeniseian language Kott (Vajda 2008) reveal that the original subject position was P1 in ancient Yeniseian rather than P8. It is possible that surviving examples of P1 subject markers contributed to the conversion of P1 object markers in da-transimpersonals into subject markers. In any event, basic verbs in the Ket lexicon often do not follow the patterns outlined in Figure 15, which represent generalizations about the lexicon rather than grammatical rules pertaining to all verbs in the language.

The evolution of Ket verb agreement is a long and complex process, with different layers of the vocabulary conforming to different historical patterns in their selection of agreement marking strategy. It must be born in mind that the lexical associations shown in Figure 15 are based on the primacy of the P8 subject prefixes, a feature that itself is an innovation in the history of Ket. Even though the P8 subject markers are required by a majority of Ket verbs today, these morphemes, which are not found in the extinct southern Yeniseian languages (Vajda 2008), represent secondary innovations that seem to have arisen as Ket restructured itself from a prefixing language to one in which the verb’s semantic head was positioned as close as possible to the left edge of the verb complex (normally in P7). As Reshetnikov and Starostin (1995) first pointed out, the P8 morphemes that developed from the subject pronouns preceding the verb complex are special clitics that only sometimes join with the phonological verb (cf. Vajda 2000, 2004 for a full explication of the patterns involved). Most verb forms, even those containing a P8 subject marker, therefore begin phonologically with the lexical morpheme in P7, a fact that renders Modern Ket a largely suffixing language in terms of its verb morphology.

The diverse agreement patterns coexisting in Modern Ket must still be described with considerable reference to lexical idiosyncrasy. Although information from studying da-intransitives can contribute important diachronic insights into the evolution of Ket verb agreement, no obvious pattern exists on the synchronic level that might readily predict whether the originally detransitivizing prefix da- will be retained obligatorily, optionally, or deleted altogether, let alone Agreement Configuration membership overall.
Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AN</td>
<td>animate</td>
</tr>
<tr>
<td>N</td>
<td>inanimate (singular or plural)</td>
</tr>
<tr>
<td>CAUS</td>
<td>causative</td>
</tr>
<tr>
<td>NPST</td>
<td>non-past</td>
</tr>
<tr>
<td>F</td>
<td>feminine (singular)</td>
</tr>
<tr>
<td>PL</td>
<td>plural</td>
</tr>
<tr>
<td>DAT</td>
<td>dative</td>
</tr>
<tr>
<td>PST</td>
<td>past</td>
</tr>
<tr>
<td>IC</td>
<td>involuntary causative</td>
</tr>
<tr>
<td>R</td>
<td>semantics unclear</td>
</tr>
<tr>
<td>IMP</td>
<td>imperative</td>
</tr>
<tr>
<td>SG</td>
<td>singular</td>
</tr>
<tr>
<td>INTR</td>
<td>intransitive</td>
</tr>
<tr>
<td>SS</td>
<td>second subject agreement marker</td>
</tr>
<tr>
<td>ITER</td>
<td>iterative (repeated or habitual action)</td>
</tr>
<tr>
<td>TR</td>
<td>transitive</td>
</tr>
<tr>
<td>M</td>
<td>masculine (singular)</td>
</tr>
<tr>
<td>TH</td>
<td>thematic consonant (lexical element of usually opaque semantics)</td>
</tr>
<tr>
<td>MOM</td>
<td>momentaneous</td>
</tr>
</tbody>
</table>

References

Impersonal verbs in Central Alaskan Yupik (Eskimoan)

Osahito Miyaoka
Max Planck Institute of Evolutionary Anthropology, Leipzig (guest)

This paper is a description of impersonal verbs in Central Alaskan Yupik, an Eskimoan language. After providing minimum information on verb stem classification and valency modifications (including one by zero derivation), primary impersonals of natural force and derived impersonals of necessity (with valency-increasing suffix) are discussed. For both types, although the impersonal A (A_{IMP}) argument does not occur externally, it is coded in verb inflection (‘ending’) as transitive subject in the singular. Therefore the construction qualifies as a transitive impersonal (‘trans impersonal’), and allows for detransitivizing derivation (impersonal passives derived by impersonal A deletion). It will be argued that the polysemy of the impersonal marker may be regarded as a result of a (historically) secondary change of the impersonal suffix into a modality marker with no valency increase, coupled with deletion of either A or A_{IMP}.

**Keywords:** Central Alaskan Yupik; impersonals of natural force; impersonals of necessity; trans impersonal; modality marker

1. **Introduction**

Central Alaskan Yupik (an Eskimoan language; CAY), which is spoken in the southwestern part of Alaska (and its major cities) by ca. 10,000 people, is a language characterized by:

i. non-templatic polysynthesis (with recursive transcategorial conversion)

ii. almost exclusively predominant suffixation (but no “nominal suffixes” as in the Mosan languages, morphological “incorporation” as a type of V+N or N+V stem compounding)
iii. predominant agglutination (with an extent of fusion in inflections)
iv. double marking of grammatical relations.

This paper aims at describing two types of impersonals, primary and derived, as they are actually used by speakers. As we shall see in view of the current bilingualism of
CAY speakers and the decline of the language, there is considerable variation in how both types of impersonals are used.

The paper is structured as follows. §1 provides background information about CAY impersonals in general, primary stems and valency modification together with their case assignments.\(^1\) §2 centres on primary impersonals, while §3 on derived impersonals (with modality involved).

1.1 Basics of CAY impersonals

CAY has two types of impersonal verbs – (i) primary and (ii) derived:

i. Primary impersonals have an inherent S\(_{\text{IMP}}\) or A\(_{\text{IMP}}\) argument, that is, they may either be monovalent stems with an S\(_{\text{IMP}}\) or patientive (or non-agentive) bivalent stems with a P and A\(_{\text{IMP}}\). There are no impersonal ditransitive stems. (See §1.2 for three kinds of primary verb stems – monovalent, bivalent, and ditransitive). The S\(_{\text{IMP}}\) and A\(_{\text{IMP}}\) of primary impersonals correspond to some unidentifiable natural force or process.

A monovalent stem of primary impersonals is basically inflected intransitively with a marked S\(_{\text{IMP}}\). But it may become bivalent with a zero-derived P (by valency increase) and an A\(_{\text{IMP}}\) originating from the primary S\(_{\text{IMP}}\) (by valency rearrangement) and as such it may be inflected transitively. This is the transitive use of monovalent impersonal stems (2.3).

A bivalent stem of primary impersonals, on the other hand, may inflect intransitively by way of A\(_{\text{IMP}}\) deletion (one of de-transitivizing processes, cf. 1.2.2).

\(^1\) CAY data are all from my fieldwork. Representations of CAY examples are made in the practical orthography –a, i, u, e /i/; p, t, c, k, q; v, l, s /z/, g /γ/; r /γ./; vv /f/, ll /S/, ss /s/, gg /x/, rr /χ/ (written single next to a voiceless); w /xʷ/, y, m, n, ng /ŋ/ (with bar on top for a voiceless nasal). The minus sign (−), the equal sign (=) and the non-equal sign (≠) respectively mark a morpheme boundary, an enclitic and a non-enclitic boundary within a “bound phrase” as an articulus (or form – articulation unit at the content plane; Miyaoka 2010b §2). But the phonological representations in the modified international phonetic symbols are used to list underlying morphemes enclosed in | | (stems and suffixes) whose suffix-initial plus (+) indicates retention of a stem-final velar (back /γ/ and front /γ/) and the minus (−) its deletion.

In the glosses, “sg.” stands for ‘unpossessed singular nominal’ and “3sg.sg.” for ‘singular nominal with third person singular possessor’, while “3sg.” stands for ‘third person singular subject intransitive’ and “3sg.3sg.” for ‘third person singular subject and third person singular object transitive’. “1, 2, and 3R” stand for ‘first, second, reflexive third person’. “pl., du.” for ‘plural’ and ‘dual’. See the list at the end for other abbreviations.
Derived impersonals have an $A_{\text{IMP}}$ which is added to the stem by a valency-increasing suffix, namely $+\text{na'iqi-}$ or two other related suffixes; (47bc), (66), (67). They may be derived from any type of “simplex” verbs (1.3.1) by the suffix $+\text{na'iqi-}$, etc. The added $A_{\text{IMP}}$ implies necessity, obligation or destiny, and thus the resulting construction will be referred to as the “necessitative construction” (NEC). A derived impersonal (bivalent or trivalent) is basically inflected transitively, but it may also be inflected intransitively again by way of $A_{\text{IMP}}$ deletion as is the case of primary impersonals. When inflected intransitively, the suffix may only be taken as a modality marker (since $A_{\text{IMP}}$ is deleted).

This means that both the primary natural force impersonals and derived necessity impersonals can either occur with intransitive or transitive inflection. Thus we may speak of “impersonal passives” vs. “transimpersonals”.

It should, however, be mentioned beforehand that, for a trivalent stem (either primary or derived, either impersonal or not) to be inflected transitively, there should obligatorily be reduction of one argument since only two core arguments can be indexed in a verb inflection.

The $S_{\text{IMP}}$ or $A_{\text{IMP}}$ can only be indexed by means of third-person singular subject verb inflection. Even though an impersonal force may be perceived by speakers, it is never expressed externally by a free-standing NP – either a pronominal or dummy cognate subject (like ‘it’ or rain rains’) or such an entity as (c)ella ‘universe’ or (c)ella-m $yu-a$ ‘spirit of the universe’ (cf. WGR sila – Fortescue 1984:81).

It should be noted that, compared with derived impersonals, primary impersonals often imply a mirative sense of some discovery (with surprise) of change caused by a natural force or process (which had been unnoticed to the speaker or the hearer). Accordingly they frequently co-occur with attention-calling or interjectional particles (e.g. $atam$, 2.3.1).

2. But they are not “passives of intransitives” which have often been so called in the literature (like Latin $curritur$ ‘(it) is being, running is done’).

3. Haas (1940) and Malchukov (2008).

4. The noun (c)ella (~[Yukon] cella) (< |cila-|) is variously translated as ‘world, nature, weather, outdoors, awareness’, and its ‘owner’ ella-m (rel.sg.) $yu-a$ (person-abs.3sg.sg.; Eastern Eskimo inu-a) is taken as something like a highest-rank entity which, deeply imbedded in the Yupik culture, is reportedly believed to have its own personality or mind and is perceived as something to be treated with great awe and fear.
1.2 Primary verb stems

A CAY verb, either a “simplex” (1.3.1) or a “complex” (transitive with upper-clause subject; 1.3.2), is morphologically composed (just like a nominal) of one stem, derivational suffix(es), and one inflection, occurring in that order. It is the inflection that completes a verb (or a nominal, for that matter) as a minimal articulus (form) on the content plane (Note 1). Since a derivational suffix is not obligatory, a stem can be closed by an inflection. The stem and the inflection are obligatory constituents of a verb (or a nominal).

A verb inflection is a combination (occasionally fused) of two inflectional suffixes marking its obligatory grammatical categories – person (both subject and object) and mood – as opposed to case, number, and person (possessor) for a nominal inflection.

Inflected verbs are either intransitive with subject indexed or transitive with both subject and object.

Derivational suffixes which can follow a verb stem have a multifarious range of functions in CAY. They can provide:

i. deverbalizations (VN, including relativizations VNr and nominalizations VNn)
ii. valency modifications (VVs for simplex verbs and VVc for complex transitive verbs)
iii. various adverbial elaborations or grammatical modifications (VV like tense-aspect, modality-evidentiality, polarity, etc.).

Primary stems, i.e. stems with no valency modification, include three types, monovalent, bivalent, trivalent (ditransitive), with subtypes:

(1)  
   a. monovalent
      S (intransitive subject) – e.g. |ayay-| ‘to leave’; |kity-| ‘[SIMP] to summer’
   b. bivalent
      P (patient) and A (agent)
      i. agentive
         A = S type – e.g. |niy-| ‘to eat’
      ii. patientive
         P = S type – e.g. |navy-| ‘to break’
      iii. impersonal patientive
         P = S type – e.g. |ciku-| ‘[AIMP] to freeze’
   c. trivalent (ditransitive)
      T (theme), R (recipient), and A
      i. secundative
         R like P – e.g. |cikij-| ‘to supply (with)’
      ii. indirective
         T like P – e.g. |tuni-| ‘to give (to)’

A denominal verb, composed of a nominal stem followed by a denominalizing suffix (NV), may be monovalent or bivalent depending upon the suffix.

The three-way classification of primary stems as above into mono-, bi-, and trivalent, however, is not a hard and fast one. A certain extent of variety or discrepancy is observed among individuals, generations, and dialects in how speakers handle a specific stem or a certain group of stems. For instance, some stems are treated as
Impersonal verbs in Central Alaskan Yupik (Eskimoan)  463

intransitive by some speakers but as monotransitive by others, with or without slightly different implications.

Since core arguments (subject and object) are indexed in inflection, they need not be externally expressed by a free-standing NP (due to double marking of grammatical relations; 1-iv). If a core argument NP occurs externally, it is assigned a syntactic case, i.e. absolutive (on S, P, R or T) or relative\(^5\) (on A).

1.2.1 Monovalent stems

Monovalent stems have an S argument which may or may not be a volitional agent, that is, it may be ‘active’ or ‘inactive’. (There is no split intransitivity in CAY.) Some monovalent stems are semantically similar to predicative adjectives of other languages, but it would be difficult in the language to set up the category of adjectives as a separate word class.\(^6\)

Monovalent stems occur only with intransitive inflection with the S argument indexed. If externally expressed by a free-standing NP, the subject is marked with the absolutive case, as indicated by \(S_{\text{abs}}\).

However, with a small number of monovalent impersonal stems the singular subject \((S_{\text{IMP}})\) can only be inflectionally indexed but never be externally expressed.

1.2.2 Bivalent stems

Bivalent stems are monotransitive stems which occur with transitive P and A indexed. The P argument is the patient-like argument of a verbal event, while the A is the one that initiates the event toward P and is the more active participant with higher control. The argument A (inflectionally marked as transitive subject) presupposes P (marked as transitive object) but not \textit{vice versa}, that is, the case marking of A depends on P, which is represented here as P > A (stands higher than).

i. NPs corresponding to the P and A, if external, are marked respectively with the absolutive and the relative case, with alignment of \(P_{\text{abs}} > A_{\text{rel}}\) as illustrated in (2a) below. Given that monovalent \(S_{\text{abs}}\) also occurs with the absolutive (1.2.1), the language manifests the ergative pattern of S = P \((\neq A)\).

---

5. The term “relative” is used in Eskimology for “ergative”, covering the G (genitive) function as well as A (ergative). Non-syntactic cases include ablative-modalis, allative, locative (which may concern argument demotion), equalis and vialis (which are peripheral or adverbial).

6. The parameter of comparison, for instance, could not be deemed a criterial property to this category (Miyaoka 2009:88).
A transitive verb may be de-transitivized in different ways, with the single external NP (if any) marked with the absolutive case, i.e. $S_{\text{abs}}$. There are two types of de-transitivization to be distinguished:

ii. “agentive” bivalent stems, showing the pattern of $S=A$, are subject to zero-derived antipassivization by $P$ demotion to the ablative-modalis coupled with $A$ promotion to the absolutive status, hence alignment of $P_{\text{abs}} S (= A)_{\text{abs}}$, as illustrated in (2b).

iii. “patientive” bivalent stems, showing the pattern of $S=P$, are subject to either suffix-derived antipassivization or zero-derived medio-passivization, that is, “medialization” (neutralization between $A$ and $P$; $S = A_{\infty P}$) and passivization (through $A$ deletion; $S_{\text{abs}} = P$, $A = \emptyset$) – illustrated in (3b-i) and b-ii. See Bergsland (1955) for medialization.

Of patientive bivalent stems, a fair number are:

iv. “impersonal patientive” stems, with $P$ and $A_{\text{IMP}}$. The $A_{\text{IMP}}$ argument is inflectionally indexed as transitive subject, which affects case assignment on the other arguments but is never expressed externally by a free-standing NP, as illustrated in (4a). De-transitivization is only achieved by passivization (through $A_{\text{IMP}}$ deletion; $S_{\text{abs}} = P$, $A_{\text{IMP}} = \emptyset$), illustrated in (4b), but not by medialization (unlike iii above), which implies impersonal bivalent stems have no mediopassives – reasonably because there can naturally be no such a thing as medialization (or neutralization) between $A_{\text{IMP}}$ and $P$ – see Miyaoka (1984) and, more fully, Miyaoka 2010b (§ 39.5.2).

v. A transitive verb may also be de-transitivized by way of reflexivization/reciprocivization (e.g. between $P$ and $A$; indicated by $P \subset A$), except for impersonal patientive bivalent stems (again for the same reason as for no medialization above).

Agentive bivalent stems (ii above) are illustrated by transitive (2 a) and intransitive (2b), the latter involving de-transitivization by demotion of the $P$ coupled with $A$ changing to (derived) $S$.

\begin{itemize}
  \item (2) a. transitive:
    
    $\text{Angute-}m_{\text{A}} \text{ neqa}_{\text{P}} \text{ ner-aa.}$
    
    man-REL.SG. fish.ABS.SG. eat-IND.3SG.3SG.
    
    ‘The man is eating the fish.’
    
    – schematically with case alignment of $P_{\text{abs}} > A_{\text{rel}}$
\end{itemize}

7. The symbol $\infty$ stands for medialization between $A$ and $P$ or $A$ and $E$. 
b. intransitive:

\[ \text{Angun}_{S(A)} neq-mek(p) \text{ er'-uq.} \]

\[ \text{man ABS SG. fish ABM SG. eat-IND 3SG.} \]

‘The man is eating a fish.’

- schematically with case alignment of P_{abm} S(=A)_{abs}.

**Patientive bivalent stems** (iii above) may involve transitive verbs with P_{ABS} and A_{REL} as in (3a) or de-transitivized verbs as in (3b) or (3c), the latter involving antipassivization and its closely related adversative construction with the same marker:

3a. transitive:

\[ \text{Angute-m}_{A} \text{ sass’aq} \text{ avg-aa.} \]

\[ \text{man REL SG. watch ABS SG. break-IND 3SG 3SG.} \]

‘The man broke the watch.’

- schematically with case alignment of P_{abs}>A_{rel}.

b. intransitive (“medio-passive”):

\[ \text{Sass’aq}_{S} \text{ navig-tuq.} \]

\[ \text{watch ABS SG. break-IND 3SG.} \]

i. ‘The watch broke’ (much more common than ii)

- with alignment of S(P\Rightarrow A)_{abs}

ii. ‘The watch was broken.’

- with alignment of P_{abs} A(\Theta)

3c. intransitives of “experiencer verbs” with E_{ADV} [+yi-]:

\[ \text{Angun}_{S} \text{ sass’a-mek(p)} \text{ navg-i-uq.} \]

\[ \text{man ABS SG. watch ABM SG. break-E IND 3SG.} \]

i. ‘The man broke a watch.’

- with alignment of P_{abm} S(E\Rightarrow A)_{abs}

ii. ‘The man had a watch broken/a watch broke on (to the detriment of) the man; this reading accepted only by limited speakers

- with alignment of P_{abm} S(E)_{abs} A(\Theta).”

Two readings of (3b) depend upon (i) medialization (between P and A) or (ii) passivization (or A deletion). Two readings of (3c-i) antipassive and (ii) adversative – depend respectively upon medialization (between E and A)⁹ and passivization (or A deletion).

---

8. The term “experiencer” (E) is used in a broad sense, covering applicative (E_{APL}), benefactive (E_{BEF}), and adversative (E_{ADV}) – 1.3.

9. The E suffix -i- (+yi-) is not only adversative, but eventually serves to decrease valency because of the medialization for antipassivization, thus E_{APS} as in (3c-i). Adversative construction like (3c-ii) is found also in (30b) and (44).
It is important to note here the parallelism of processes (medialization and pas-
sivization) between (b) and (c) each with two readings (i, ii). This implies that CAY
derived) antipassives are closely related with (adversative) “experiencer” verbs. It
also explains the fact that impersonal verbs, either primary or derived, can, reason-
ably, never have antipassives corresponding to (3c) from a non-impersonal patientive
verb, given, again, that there can naturally be no such a thing as medialization between
A_{IMP} and E.

Impersonal patientive bivalent stem (iv above) are illustrated by (4a) for transi-
tive verbs with P_{abs} and A_{rel}, and (4b) for de-transitivization. For the same reason as
mentioned, no antipassivization or reflexivization/ reciprocization is possible with
impersonal verbs unlike in the case of non-impersonal patientives.

\[
\begin{align*}
(4) & \quad \text{a. transitive} \\
& \quad \text{Namvaq}_{P} \ ciku-a. \\
& \quad \text{lake.abs.sg. freeze-ind.3sg.3sg.} \\
& \quad \text{‘It [A_{IMP}] froze the lake, i.e. the lake is (now) frozen.’} \\
& \quad \text{– A_{IMP} argument is indexed as transitive subject but never be externally} \\
& \quad \text{expressed by a free NP in the relative case.} \\
& \quad \text{b. intransitive} \\
& \quad \text{Namvaq}_{S=P} \ ciku-uq. \\
& \quad \text{lake.abs.sg. freeze-ind.3sg.} \\
& \quad \text{‘The lake is (still) freezing.’} \\
& \quad \text{– de-transitivized through A (A_{IMP}) deletion as characteristic of} \\
& \quad \text{patientive stems.}
\end{align*}
\]

1.2.3 Trivalent (ditransitive) stems
Trivalent (ditransitive) stems have three arguments, i.e. T, R, and A. However, since CAY
verb inflection can only index up to two arguments, either T or R must be demoted in
order to be inflected transitively. The resulting stems are secundative if R behaves like
bivalent P but T is demoted to the ablative-modalis status, or indirective if T behaves
like P but R is demoted to the allative status. See Miyaoka (2010a; 2010b § 35.1) for
more details, and 1.3 for other trivalent and multivalent stems.

10. This interpretation of CAY antipassives as medialization (1.2.3 and Note 7) finds a
support in another E suffix (mainly applicative) \(+uc-\) which has both the roles of benefactive
(etc.), e.g. 2.1.1 as well as (58), (65), and antipassive, e.g. \(atra-ut-uq\) ‘he brings (s.t./s.o.)’ with
\(atjaj-\) ‘to go down’.

11. Thus, the \(ciku-\gammai-\) (freeze-E) > \(ciku-i-\) can be an adversative stem like (3c-ii) but not an
antipassive like (3c-ii).
Thus, intransitive and transitive verbs with primary stems, monovalent (5a), biva-
alent (5b), ditransitive (5c/d) show the case alignment below:

\[(5)\]

\[\begin{align*}
&\text{a. } S_{\text{abs}} & \text{intransitive} \\
&\text{b. } P_{\text{abs}} & A_{\text{rel}} & \text{transitive from bivalent} \\
&\text{c. } T_{\text{abm}} & R_{\text{abs}} & A_{\text{rel}} & \text{transitive from secundative ditransitive} \\
&\text{d. } R_{\text{all}} & T_{\text{abs}} & A_{\text{rel}} & \text{transitive from indirective ditransitive.}
\end{align*}\]

As (5) indicates, the absolutive case is assigned to the leftmost argument except when one argument T or R of ditransitives is demoted (from the absolutive) to the ablative-
modalis (c) or allative status (d).

1.3 Valency modifications

Primary stems (monovalent, bivalent, ditransitive) may be subject to valency modifica-
tion by a VVs or VVc suffix (1.2) which increases or decreases by one argument as well as rearranges arguments relative to each other. By way of valency increase, a monova-
 lent stem becomes a modified or derived bivalent one, and a bivalent stem becomes a trivalent one, and so on, since the increase can be recursive within a (morphologically) single verb.

1.3.1 Simplex verbs

Simple verbs involve increases relating to the A (agent addition only on monovalent stems, yielding a kind of causative verbs, but see 1.3.2 also), E (E_{\text{APL}}, E_{\text{ADV}}, E_{\text{BNF}} – Note 8), and A_{\text{IMP}} which is nothing but the increase responsible for derived (necessita-
tive) impersonals – §3.

Primary stems S, P, T|R, A with or without these extended (italicized) arguments E, A, A_{\text{IMP}} constitute a simplex verb (vs. a complex transitive verb; 1.3.2).

The causative A extension is formed by [+c] only on a monovalent stem with S – e.g. (27) maq-t-, 2.2 pek-t-, (56) tuqu-t- – and is distinct from the causative A’ for a complex transitive (cf. 1.3.2 and Note 14). It requires rearrangement of S into P (i.e. S + A → P + A), hence the same pattern as bivalent stems above (1.2.2) like P_{\text{abs}} > A_{\text{rel}}.

A whole simplex verb with or without (italicized) extensions follows the order (higher > lower) of:

\[(6)\]

\[S/P/T_{\text{indir}} \mid R_{\text{secun}} > E > A > A_{\text{IMP}}\]

generating to which the absolutive case status (intransitive subject or transitive object) is assigned to the leftmost (or highest) S, P, T or R, and the relative status (transitive
subject) to the next $E$ or $A$ (if no $E$). This is a general pattern of CAY case assignment, except for one caveat.\(^{12}\)

If a third argument is involved (as stated above), the absolutive argument is obligatorily demoted to oblique status – ablative-modalis for $T_{\text{indir}}$ but allative for $R_{\text{secun}}$ – and the relative argument is promoted to absolutive status so that the third argument can fill the vacated relative status. And if a fourth argument is involved, the (promoted) absolutive argument is demoted to ablative-modalis, followed by the promotion of the relative to absolutive and the fourth argument to relative, and so on. $A_{\text{IMP}}$ can never be an external NP (with case assigned), but can enjoy relative case status as the transitive subject (so indexed in the verb). Thus:

\[(7) \quad \begin{align*}
& a. \quad P_{\text{abm}} \quad E_{\text{abs}} \quad A_{\text{rel}} \\
& b. \quad T_{\text{abm}} / R_{\text{all}} \quad E_{\text{abs}} \quad A_{\text{rel}} \quad (\text{both secundative and indirective}) \\
& c. \quad P_{\text{abm}} \quad A_{\text{abs}} \quad A_{\text{IMP}(\text{rel})}.
\end{align*}\]

The ablative-modalis demotion is the most general pattern of demotion and is shared by other trivalent or multivalent verbs (incl. type 1 complex transitives; 1.3.2), while the allative demotion is limited to $R$, recipient-like applicative $E$ type 2 complex transitive $A'$ and its lower (embedded) $A$, and $A$ in transitive (type 2) nominalizations (details in Miyaoka 2010b). An illustration of the whole valency modification and case alignment is beyond the scope of this paper.

1.3.2 Complex transitive verbs

Complex transitive verbs have an upper-layer subject ($A'$, $A''$, ...; causative, directive, speculative, reportative, ignoteive, expectant) and morphologically embed a simplex verb (recursively and cumulatively, theoretically ad infinitum), like ‘she asks me to eat (s.t.)’ [directive], ‘she thinks (s.o.) asked me to eat (s.t.)’ [directive and speculative], and so on.\(^{13}\) An impersonal verb (primary and derived) may also be embedded into a complex verb. Complex transitive verbs may also be subject to valency modification.

Given the necessary increases in valency, a comple transitive verb is typically multivalent, up to six or seven arguments at least (e.g. ‘$A''$ says that $A'$ thinks that $A_{\text{IMP}}$ necessitates $A$ to sell $T$ to $R$ for the benefit of $E$ (that $A$ had to sell $T$ to $R$ for $E$)’). This topic is again excluded here for reasons of space – see Miyaoka 2010b (§40.3.3).

---

12. The only caveat is that an extended bivalent stem with $E_{\text{APL}}$ (not $E$ in general) requires $S > E_{\text{APL}} \rightarrow P(E_{\text{APL}}) > A(S)$ – e.g. ‘she (REL) went out for me (ABS)’, while $S > E_{\text{ADV}}$ follows the general pattern – e.g. ‘she (ABS) died on me (REL)’.

13. A CAY complex verb is not a compound verb or a serival verb but is morphologically a single transitive verb exclusively by means of suffixation.
Argument reduction necessary for complex transitive verb argument(s) may be achieved by demoting an argument inside the simplex verb into the appropriate case (a, b, c) or by demoting the upper-layer A’ into the allative-case if it is further followed by A” (d). Thus:

\[(8)\]
\[
\begin{align*}
\text{a. } & \{P_{abm} A_{abs}\} A’_{rel} & \text{- type 1} \\
\text{b. } & \{P_{abs} A_{all}\} A’_{rel} & \text{- type 2} \\
\text{c. } & \{P_{abm} A_{all}\} A’_{abs}, A”_{rel} \\
\text{d. } & \{P_{abm} A_{abs}\} A’_{all}, A”_{rel}
\end{align*}
\]

Only the two constructions of them (a) and (b) are exemplified below by giving a causative complex transitive \(nivii-vkaj’-\) ‘to make/let s.o. to eat s.t. (to feed)’. Compare with simplex verb construction (2) without valency increase:

\[(9)\]
\[
\begin{align*}
\text{a. transitive 1:} & \\
\text{Arna-} & A’ \quad \text{angun}_{P(A)} \quad \text{neq-mek}_{(P)} \quad \text{nere-vkar-aa.} \\
\text{woman-} & A_{rel}. \quad \text{man.abs.sg.} \quad \text{fish-} & \text{abm.sg.} \quad \text{eat-cau-} & \text{ind.3sg.3sg.} \\
\text{‘The woman let the man eat a fish.’} & \\
\text{– schematically with case alignment of } P_{abm} A_{abs} A’_{rel} \\
\text{b. transitive 2:} & \\
\text{Arna-} & A’ \quad \text{angut-mun}_{A} \quad \text{neqa}_{p} \quad \text{nere-vkar-aa.} \\
\text{woman-} & A_{rel}. \quad \text{man-all.sg.} \quad \text{fish.abs.sg.} \quad \text{eat-cau-} & \text{ind.3sg.3sg.} \\
\text{‘The woman let the man eat the fish.’} & \\
\text{– schematically with case alignment of } P_{abs} A_{all} A’_{rel}
\end{align*}
\]

The (c) and (d) have four arguments with recursive valency increase, with either A or A’ argument demoted to allative.

If a complex transitive verb has five or more arguments, two arguments may occur with one and the same case, in which case the two explicit NPs generally stand in mirror image order with the suffixes related to each argument. If the two agent arguments A and A’ (or A’ and A”) are coreferential, there is no need for demotion since coreferentiality itself does decrease one argument.

A complex transitive verb may be inflected intransitively. Of six kinds of complex verbs (1.3.2), all except the reportative \(+ni-\) ‘A’ to say that’ are patientive stems, so the antipassivizer E\(_{APS}\) \(+yi-\) does occur just as is the case within a simplex

\[14\] The causative (complex transitive) suffix has suppletive variants, i.e. \(-vkaj’-\) (after vowel) and \(+cic-\) (after consonant; /-cic-/), which latter should probably be distinct from denominizing NV \(+cic-\) ‘to have the quality of’ in e.g. (12c), (41), (42) (see Jacobson 1984:440–41, 445).
verb with patientive stems (e.g. *nere-vkar-i-uq* for (9)). On the other hand, zero-antipassivization (with no antipassivizer but only intransitive inflection) is sufficient for the reportative [+ni-] as this is agentive.

2. **Primary impersonals of natural force**

CAY has a considerable number of impersonal stems, both monovalent with $S_{\text{IMP}}$ argument and bivalent with $P$ and $A_{\text{IMP}}$. They mainly include verbs denoting changes in time (season, day/night)/weather and some other sorts of changes or process of things caused by a natural or supernatural force or process as an involuntary, uncontrollable, and invisible agent, though the $S_{\text{IMP}}$ or $A_{\text{IMP}}$ arguments are never expressed externally by a full NP (as stated).

Many of the impersonal monovalent stems are ambivalent and may possibly be used as nominal stems as well.

2.1 **Monovalent impersonal stems**

Monovalent impersonal stems include verbs of temporal change (season and day/night) and weather.

(10) *temporal change (season and day/night):* |iγc-| ‘to become morning (dawn)’, |kiaγ-| ‘to become summer’, |unuγ-| ‘to become night’, |unuaquγ-| ‘to dawn, become tomorrow’, etc.

Many monovalent stems may be subject to zero-derived transitive use with locational $P$ (§2.3). Example (11) may thus be (a) transitive or (b) in transitive, with quasi-equivalence:


summer-IND.3SG.3SG. here-EX.ABS.SG. land-ABS.1PL.SG.

‘It is/has become summer in our village.’

– *kiag-aa* lit. ‘it [A$_{\text{IMP}}$] summers on it/him [location].’

b. **Kiag-tuq** [ma-n’a nuna-vut]$_s$.

summer-IND.3SG. here-EX.ABS.SG. land-ABS.1PL.SG.

‘It is/become summer in our village.’

– de-transitivization of (a) by $A_{\text{IMP}}$ deletion.

15. Apart from this antipassivizer, valency modification after a complex suffix seems very rare if any.
There is another group of monovalent weather stems, which are denominal (that is, verbalized by a denominalizing NV suffix) and typically occur with intransitive inflection, even though the suffix generally yields bivalent stems (13):

weather:

(12) a. |anuq-liγ-| 'to be windy' – |anuqi-| 'wind'
|anuq-sa(y)aγ-| 'to be breezy' – NV/NN |sa(y)aγ-| 'slight(ly), little, dear'
|anuq'-vay-| 'to be very windy' – |+vay-| 'intensely'
b. |ivz-iγ-| 'to rain, drizzle' – |ivzuγ-| 'rain/drizzle'
|ivzu-ŋi-| 'to begin to rain, drizzle' – with NV |-ŋi-| 'to get'
c. |kiiγ-ic-| '(air, weather) to be hot' – |kiiγ-| 'heat', see

The |anuq-liγ-| and |ivz-iγ-| share the very productive NV suffix |liγ-| 'to have/provide plenty of', which is bivalent and the derived stem (from |atsaγ-| 'berry') can occur with either intransitive or intransitive inflection:

(13) a. atsa-lir-aa ~ ats-ir-aa (IND.3SG.3SG.) 'she added lots of berries to it'
b. atsa-lir-tuq ~ ats-ir-tuq (IND.3SG.) 'it (e.g. ice-cream) has lots of berries'.

However, transitive forms of a denominal weather verb (|qan-iγ-| 'to snow'; Note 16) may sometimes be encountered as in (14) which should probably be interpreted as a case of zero derived transitive use with locational P (2.3):17

(14) qanir-aa (IND.3SG.3SG.) 'it [A_IMP] is snowing on it/him [P; location]'
qanir-aanga (IND.3SG.1SG.) 'it [A_IMP] is snowing on me [P] (after I have left)'.

16. The variants -lir- vs. -ir- reflect a syncopation of a stem-final vowel and the suffix initial /l/ (together with velar deletion if any), which is a phonological adjustment idiosyncratic to NV suffixes beginning with |-l-|, e.g. |-li-| 'to make', |-liγ-| 'to work on, deal with' and NN suffix |-linγaγ-| 'trace/evidence of' as in (39).

As a matter of fact the suffix |-liγ-| is productive of many weather stems: |anuq-liγ-| 'to be windy' (|anuqi-| 'wind'), |qan-iγ-| 'to snow' (|qanuγ-| 'snow'; e.g. 15), |akijγ-iγ-| 'to be sunny' (|akijγ-| 'sun'), |taic-iγ-| 'to be foggy' (|taittuγ-| 'fog'), (c)|Ha-γy-| 'rain' from [(c)|Ha-γuy-| 'weather-bad'], |ivz-iγ-| 'to rain' (|ivzuγ-| 'rain, drizzle'; Nelson Island dialect, etc.), |kal-iγ-| 'to thunder' (|katuy-| 'thunder'), |amiγlu(γ)-iγ-| 'to be cloudy' (|amiγlu(γ)-| 'cloud').

17. There may still be a possibility that the transitive form reflects the original bivalency of a weather verb with the NV suffix, meaning something like 'it is supplying it (e.g. tundra) with lots of snow'.
2.1.1 Applicative |+uc-| However, the transitive use above seems to be very rare, and some speakers prefer an expanded stem with the applicative |+uc-|

(15) qani-ut-aanga
snow-E_{APL}-IND.3SG.1SG.
'it [A=S_{IMP}] is snowing on me [P=E_{APL}] (after I have left)'
– with S into A (below).

(16) a. ivsu-ng-ut-aa
    angun_{P=E}
    rain-INC-E_{APL}-IND.3SG.3SG. man.ABS.SG.
    'it started to rain on the man'
b. ivsu-ng-ut-aanga
    rain-INC-E_{APL}-IND.3SG.1SG.
    'it started to rain on me (after my departure)'.

(17) unuaqu-uc-iiq-aaten
    be.tomorrow-E_{APL}-FUT-IND.3SG.2SG.
    'it will be tomorrow before you(sg.) are done, lit. it (the dawn) will come on you' – from unuaquγ- in (10).

Note that the three examples above are the cases for the caveat of $S > E_{APL} \rightarrow E_{APL} > A(S)$ (caveat mentioned in Note 12) when the applicative |+uc-| occurs after a monovalent stem. See also attested with (58) |iqva-ut-| 'to pick berries for (s.o.)' which is no weather verb.

2.2 Bivalent impersonal stems

As stated (1.2.1), impersonal patientives can not have an antipassive, though the demarcation between the patientive proper and the impersonal may be blurred for some stems: Stems that are typically impersonal may nevertheless be found to occur with a personal agent, behaving as an ordinary patientive stem, but not vice versa.

The following list is representative of various semantic categories where $P$ argument is a thing subject to the change concerned:

freezing/burning:

(18) |ayu-| 'to spread (of fire)', |ciku-| 'to ice, freeze', |iha-| 'to weather, tan (of skin)', |ily-| 'to singe', |iyuj-| 'to jell', |kumlac-| 'to cool', |niñfiý-| 'to be cold', |nipi-| 'to go off, extinguish (of fire)', |qamí-| 'to die down', |qixu-| 'to freeze to death', |qijcu-| 'to freeze', |uñjuy-| 'to melt', |uu-| 'to cook', etc.

change in body parts:

(19) |cii-| 'to get chapped', |mami-| 'to heal, close in', |miqí-| 'to shed hair/fur', |pupiy-| 'to get infected sores', |qauj-| 'to get head sores', etc.
change in condition, shape, or position:

(20)  \[\text{aju-} \to \text{ripen, rot}, \text{iqa-} \to \text{get dirty}, \text{iqin-} \to \text{shrink},
\text{ki-} \to \text{come/peel off as a layer}, \text{kinj-} \to \text{dry}, \text{ninj-} \to \text{stretch},
\text{piki-} \to \text{move, stir}, \text{piji-} \to \text{bend}, \text{puvi-} \to \text{swell},
\text{qatj-} \to \text{get rusty}, \text{tini-} \to \text{fly}, \text{uki-} \to \text{get a hole}, \text{uli-} \to \text{flood}, \text{etc.}\]

The following example is a repetition of (4):

(21)  
\begin{enumerate}
\item \text{transitive}  
\begin{align*}
\text{Nanvaq}_p & \quad \text{ciku-}a. \\
\text{lake.abs.sg} & \quad \text{freeze-IND.3SG.3SG}. \\
\text{It \([\text{A_{IMP}}]\) froze the lake, i.e. the lake is (now) frozen.}
\end{align*}
\item \text{intransitive}  
\begin{align*}
\text{Nanvaq}_{S=P} & \quad \text{ciku-}uq. \\
\text{lake.abs.sg} & \quad \text{freeze-IND.3SG}. \\
\text{The lake is (still) freezing.}
\end{align*}
\end{enumerate}

The italicized ‘it’ in (21a) is meant to refer to an impersonal agent, i.e. some impersonal force or process of nature that can never be expressed externally by a free-standing NP. However, it is apparently perceived as such by some speakers at lease, who sometimes render it as ‘it is made \textit{that way}’ or even ‘someone is causing it’ (reasonably as it is inflectionally so marked).

The subject (\text{A_{IMP}}) in transitive inflection is always indexed as the third person singular, and never dual or plural or in the first or second person, while the P argument may be either singular, dual, or plural:

(22)  
\begin{enumerate}
\item \text{ciku-a}  \quad \text{(IND.3SG.3SG.)} \quad \text{it/he is frozen/freezing; lit. it \([\text{A_{IMP}}]\) freezes it/him \([\text{P}]\)}
\item \text{ciku-i}  \quad \text{(IND.3SG.3PL.)} \quad \text{they are frozen/freezing; lit. it \([\text{A_{IMP}}]\) freezes them \([\text{P}]\)}
\end{enumerate}

Despite the English gloss ‘they’ in (b), this is actually a P argument, which becomes S in the intransitive, with \text{A_{IMP}} (‘it’) deleted as the agent of patientive bivalent stems, i.e. as passives. The following pair is quasi-equivalent to the preceding (22ab):

(23)  
\begin{enumerate}
\item \text{ciku-uq}  \quad \text{(IND.3SG.)} \quad \text{it/he is (being) frozen}
\item \text{ciku-ut}  \quad \text{(IND.3PL.)} \quad \text{they are (being) frozen}
\end{enumerate}

Generally speaking, the impersonal transitive and intransitive constructions are semantically quasi-equivalent, but may display some (mainly aspectual) differences. The differences may be more noticeable with some stems and in some contexts than others. In the following pair, the transitive (a) may denote either completion (‘just frozen’) or a process (‘now freezing’) for at least some speakers, while the intransitive
(b) is more likely to denote process, each being particularly characterized by different aspect suffixes VV |+nauγ.-| ‘(therefore) now’ vs. |+niaγ.-| ‘(therefore) later’:

\begin{tabular}{l}
\text{lake.ABS.SG.} & \text{freeze-IND.3SG.3SG.} & \text{go-now-IND.1du.} \\
\end{tabular}
\begin{tabular}{l}
\text{‘The lake is (just, already) frozen, (so) let us(du.) go(now)!’} \\
\end{tabular}

b. Nanvaq\textsubscript{S} ciku-\textit{uq} cali
\begin{tabular}{l}
\text{lake.ABS.SG.} & \text{freeze-IND.3SG.} & \text{still} \\
\end{tabular}
\begin{tabular}{l}
\text{ayag-niar-tukuk unuaqu.} \\
\text{go-later-IND.1du. tomorrow} \\
\end{tabular}
\begin{tabular}{l}
\text{‘The lake is (still) freezing, (so) let us(du.) go tomorrow!’} \\
\end{tabular}

It is more likely that the former is perfective and the latter durative, but, in the next pair, the distinction is neutralized since the perfective (-continuative) marker VV |+uma-| emphasizes the completion in the latter (b) as well:

(25) a. Nanvaq\textsubscript{P} ciku-ma-\textit{a}.
\begin{tabular}{l}
\text{lake.ABS.SG.} & \text{freeze-PRF-IND.3SG.3SG.} \\
\end{tabular}
\begin{tabular}{l}
\text{‘The lake is frozen.’} \\
\end{tabular}

b. Nanvaq\textsubscript{S=P} ciku-ma-uq.
\begin{tabular}{l}
\text{lake.ABS.SG.} & \text{freeze-PRF-IND.3SG.} \\
\end{tabular}
\begin{tabular}{l}
\text{‘The lake is frozen.’} \\
\end{tabular}

As is the case with the ‘lake (water)’, the P argument for impersonal verbs is something that is susceptible to – or has potentiality for – changing or moving by itself, i.e. ice (above), weather, time, animate things, etc. Thus, the stem |pikî-| ‘to move, stir’ can occur with \textit{yaqulek} ‘bird’ as in the following but not with, e.g. \textit{ena} ‘house’ or \textit{nuna} ‘land’: (a) and (b) are quasi-equivalent:

(26) a. Pek'-\textit{uq} (IND.3SG.) yaqulek\textsubscript{S=P} (ABS.SG.)
\begin{tabular}{l}
\text{‘the bird stirred (made a movement when hatching), is stirring’} \\
\end{tabular}

b. Pek'-\textit{aa} (IND.3SG.3SG.) yaqulek\textsubscript{P} (ABS.SG.)
\begin{tabular}{l}
\text{‘the bird stirred, made a movement!’} \\
\end{tabular}
\begin{tabular}{l}
\text{– with a mirative (surprise) sense, naturally accompanying an attention-calling particle atam (‘look!’) or phrase tang yaa=i (‘look, there!’) in particular.} \\
\end{tabular}

Here, the intransitive (26a) is much more frequent than the transitive (26b).

In general, while many speakers seem to accept the use of both transitive and intransitive forms with impersonal patientive stems, some speakers feel it difficult or impossible to use the transitive. The acceptability differs, at least partly, from one stem to another. The transitive use may conceivably be one of the areas which have been affected by rather recent linguistic erosion.
If the P argument for the verb \[piki\] is a ‘house’ or ‘land’ (instead of ‘bird’), some causation has to be added by personal A (\(pek-t-aa\); with \(+c\), 1.3.1; one event) or by a complex transitive verb (upper-layer) agent \(A' \ (pek-cet-aa; \text{ with } +cic-\neg+vka\neg) \ 'to let/cause' \); two events) – see also (27).

2.3 Transitive use of monovalent stems (with zero derivation)

Monovalent stems may be expanded by a valency-increasing suffix to become bivalent so as to occur with transitive inflection. But, even without such valency modification, a fair amount of monovalent stems are used transitively, that is, by zero derivation, as if they were primarily bivalent stems.

The transitive use of monovalent stems with zero derivation, however, is far from anything very common in CAY and is more apt to be accepted by elders (with acceptability differing even among them) than younger speakers who generally use only the intransitive construction. Here is again an interesting area of fluidity, obscurity, and fuzziness which will certainly require extensive exploration in view of language erosion.

The arguments added on monovalent stems by zero derivation include locational P, e.g. (27), though not directly relevant to this paper, and personal/impersonal A (2.3.1).

(27) a. \(maq'-uq \ \text{(IND.3SG.)} \ 'it [S: liquid] flows out/leaks'\)
   b. \(maq-aa \ \text{(IND.3SG.3SG.)} \ 'it [A=S] flows out on it [P: place]'\)

   transitive (b) with zero-added locational P has very limited acceptance, as compared with personal A added:

   cf. \(maq-t-aa \ 'he [A] is making it [P(=S): liquid] flow out'\)

   flow-A-IND.3SG.

2.3.1 Impersonal and personal A

Impersonal \(A_{\text{IMP}}\) and/or personal A is supplied by zero derivation for a wide range of verbs that behave like patientive stems (2.2). The primary S takes the role of P, i.e. with P (=S) and \(A/A_{\text{IMP}}\).

The acceptability of transitive verbs with zero-added impersonal A is generally very low (if ever), while a personal A reading seems to be accepted by more speakers and may be the only one that is acceptable to them. For some stems, however, two readings may be possible.

Some of the relatively few speakers who accept the transitive construction seem to perceive little or no difference from its intransitive counterpart, but others often feel

---

18. But apparently not as freely as in Siberian Yupik where Malchukov (2008:79), referring to Emeljanova and Vakxtin, concludes ‘transimpersonal inflection can be used with any intransitive verb to indicate the lack of control on the part of the subject’.
that the transitive has a mirative implication of (the speaker noticing) some natural or supernatural change or force which is unseen or unnoticed by the hearer, with frequent co-occurrence of the attention-calling particle or phrase.

Whether the zero-added argument is a personal A and/or impersonal A_{IMP} seem to depend largely upon the semantics of the stems as well as the speakers.

color stems:

(28) a. Tungu-a ∼ tungu-uq   kelipaq_{P/S}
    black-IND.3SG.3SG./IND.3SG.  bread.ABS.SG.
    ‘The bread is black, it has blackened the bread.’
b. Tungu-a   mingug-a-nip
    black-IND.3SG.3SG.  painting-VN R-ABS.3RS G.SG.
    ‘He made his (own) painting (too) black.’

(29) a. Qater-pag-taten! ∼ qater-pag-cit!
    whiten-AUG-INT.3SG.2SG./INT.2SG.
    ‘How white you(sg.) are!, How white it [A_{IMP}] has made you(sg.)!’
    – the exclamatory use of an interrogative-mood verb.
b. qat-siyaag-an
    white-AUG-IND.2SG.3SG.
    ‘You (sg.) are making it too white.’

(30) a. Qiu-gaa ∼ Qiu-guq   qeggina-a_{P/S}
    discolor-IND.3SG.3SG./IND.3SG.  face-A BS.3SG.SG.
    ‘His face is discolored.’
b. Ilunga-ma_{A}  qiu-gi-anga   unate-mnek_{P/}.
    cousin-REL.SG.  discolor-E_{ADV}-IND.3SG.1SG.  hand-ABM.1SG.SG.
    ‘My (female cross-)cousin bruised me (E) on my hand.’
    – (b) with personal A can have the adversative |-γi-|. Together with zero-added A, it constitutes a trivalent verb with arguments P(S) > E_{ADV} > A; cf. e.g. (44).

weather stems:

(31) a. ella-anga ∼ ella-unga
    weather-IND.3SG.1SG./IND.1SG.
    ‘It has weathered (tanned) me.’
b. Akerte-m_{A}  ella-anga.
    sun-REL.SG.  weather-IND.3SG.1SG.
    ‘The sun has weathered (tanned) me.’

adjectival stems: The following (a) examples with an impersonal A reading may be used in speaking to someone who has not noticed the change, thus the frequent co-occurrence of the attention-calling atam. Note also the asp ectual implication in the glosses.
Impersonal verbs in Central Alaskan Yupik (Eskimoan)

    small-IND.3SG.SG. ATN cap.ABS.SG.
   ‘See, the cap is small!’

    b. *Mik’-ak iakma-li-a-gniₚ.*
    small-IND.3SG.3du. sled-make-VNr-ABS.3RSG.du.
   ‘He has made his own sled small, lit. he has made the one
   (which he made) small.’

(33)  *Puqig-aa atam Nace-aqₚ.*
    articulate-IND.3SG.3SG. ATN name-LNK.ABS.SG.
   ‘See, Nace is articulate (clever)!’

**Monovalent stems of motion:**

(34)  *Akag-aa angqa-aₚ.*
    roll-IND.3SG.3SG. ball-ABS.3SG.SG.
   ‘His ball is rolling (it rolls his ball).’


(35)  *tep-liq-aa ≌ tep-liq-uq*
    smell-have.bad-IND.3SG.3SG./IND.3SG.
   ‘it smells bad’

    cold-get-not.yet-IND.3SG.3SG. bowl.ABS.SG. food-ABS.1SG.SG.
   ‘My bowl of food is not cold yet (lit. it [AIMP] has not get my bowl of food cold).’

   cf. *nenglla-nge- ≌ nengllir- (non-impersonal) [niŋiŋ-] ‘to cold’.*

(37)  *Uksu-urte-ng-aa ≌ Uksu-urte-ng-uq nuna-vutₚ/ₚS.*
    winter-become-INC-IND.3SG.3SG./IND.3SG. land-ABS.1PL.SG.
   ‘Our village is beginning to become wintery.’ ct. (16)

(38)  *Ciul-va-u-gaa ≌ Ciul-va-u-guq atam Nace-aqₚ/ₚS.*
    ear-big-be-IND.3SG.3SG./IND.3SG. ATN name-LNK.ABS.SG.
   ‘See, Nace is big-eared.’

2.3.2 Denominal verbs – with certain verbalizing suffixes (including copula-like relational verb NV like ‘to be, become’) may behave the same way with zero-added impersonal A, although the acceptability may vary among speakers.

(35)  *tep-liq-aa ≌ tep-liq-uq*
    smell-have.bad-IND.3SG.3SG./IND.3SG.
   ‘it smells bad’

    cold-get-not.yet-IND.3SG.3SG. bowl.ABS.SG. food-ABS.1SG.SG.
   ‘My bowl of food is not cold yet (lit. it [AIMP] has not get my bowl of food cold).’

   cf. *nenglla-nge- ≌ nengllir- (non-impersonal) [niŋiŋ-] ‘to cold’.*

(37)  *Uksu-urte-ng-aa ≌ Uksu-urte-ng-uq nuna-vutₚ/ₚS.*
    winter-become-INC-IND.3SG.3SG./IND.3SG. land-ABS.1PL.SG.
   ‘Our village is beginning to become wintery.’ ct. (16)

(38)  *Ciul-va-u-gaa ≌ Ciul-va-u-guq atam Nace-aqₚ/ₚS.*
    ear-big-be-IND.3SG.3SG./IND.3SG. ATN name-LNK.ABS.SG.
   ‘See, Nace is big-eared.’

---

19. The dual form for a single sled is a non-singular used for a single but composite object.
(39)  Ciru-ar-u-i  \(\equiv\)  Ciru-ar-u-ut  \(\{\text{kayangu-t}\)  \\
cover-VNr-be-IND.3SG.3PL./IND.3PL.  egg-ABS.PL.  \\
tua-ni  tuutangay(a-\{\text{\text{-}}\}imra-a-t\}\_\text{PS}  \\
there-LOC  goose-product-EV-ABS.PL.  \\
'The Canadian goose eggs are covered (like a nest) over there.'  \\
\quad \text{– with the very common deletion of } /\text{VOWEL-}l/ \text{ mentioned in Note 16.}

Both personal and impersonal A can be supplied at least in the following:

(40)  a.  Iqtu-an  qasp-e-li-a-n\text{\_}\text{p}.  \\
wide-IND.2SG.3SG.  parka-EV-make-VNr-ABS.2SG.SG.  \\
'You(sg.) made your parka (lit. your made parka) too wide.'  \\
\quad |iqtu-| 'to be wide'

b.  Iqtu-a  kuik\text{\_}\text{p}  atam.  \\
wide-IND.3SG.3SG.  river.ABS.SG.  ATN  \\
'Hey, the river is (has become) wide, lit. \text{it } [A_{\text{IMP}}] \text{ has widened it!}'  \\
\quad \text{cf. } iqtu-uq \text{ (3SG.) } \text{‘it is wide’ } \text{– a mere statement instead of mirative.}

The -cet- in the following two examples should be a denominalizing suffix |+cic-| which is probably distinct from post-consonantal variant of the causative complex suffix (cf. Note 14):

(41)  Qiug-cet-aa  \(\equiv\)  Qiug-cet-uq  qilak\text{\_}\text{PS}.  \\
blue-EX-IND.3SG.3SG./IND.3SG.  sky.ABS.3SG.SG.  \\
'The sky is (has become) blue.'  \\
\quad \text{– cf. (30).}

(42)  Kavir-cet-aa  \(\equiv\)  Kavir-cet-uq  qen ga-a\text{\_}\text{PS}.  \\
red-VVc-IND.3SG.3SG./IND.3SG.  nose-ABS.3SG.SG.  \\
'His nose is red.'

2.3.3 Primary vs. derived bivalent impersonals  One may naturally question whether an impersonal bivalent (§2.2) is actually the same as a zero-derived transitive use of a monotransitive (§2.3ii-b). But they are utterly distinct.

The often-cited impersonal bivalent [ciku-| ‘to freeze’ is compared below with two monovalent stems, i.e. the impersonal monovalent [kiay-| ‘to summer’ and the monovalent [ani-| ‘to go out’. The second argument in a transitive verb is not an added one for the former, but it is for the latter:

(43)  a.  ciku-ut  \(\equiv\)  ciku-i  \\
freeze-IND.3PL./IND.3SG.3PL.  \(S_{\text{abs}}(P) \equiv P_{\text{abs}} A_{\text{IMP}(\text{rel})}\)  \\
'rethey are freezing (frozen)'

b.  kiag-tuq/kiag-aakut  cf. (11ab)  \\
summer-IND.3SG./IND.3SG.1PL.  \(S_{\text{IMP}(\text{abs})} \equiv P_{\text{abs}} A(S)_{\text{IMP}(\text{rel})}\)  \\
'\text{it is (has become) summer/it is (has become) summer on us}.'
It should be clear that ‘to freeze (s.t.)’ and ‘to become summer (on s.o./s.t.)’ are “grooved” or conceptualized differently, that is, are either primarily bivalent or monovalent.

The primarily mono- vs. bi-valent contrast is also made clear in adversative constructions by VVs |+γi-| ‘to the detriment of; on’ See the transitive inflections with different person relations, 3sg.1sg. vs. 1sg.3sg., depending upon the argument ordering:

(44) a. *ciku-i*-gaanga *Pабm* EADVabs A_Imp(rel)
freeze-EADV-IND.3SG.1SG.
‘it (A_Imp) froze (s.t.) on me’

b. *an-i-aqa* Sabs EADVrel
go.out-EADV-IND.1SG.3SG.
‘it went out (somewhere) on me, it escaped out from me’

Like (44b), the other adversative verb (30b) *qiu-gi-anga* cited above begins with a monovalent stem (‘to discolor’), but it is trivalent together with EADV and the zero-derived agent, hence the same person relation of 3SG.1SG. as (a).

3. Derived impersonals (necessitative) with |+nayqi-|

While the A_Imp argument of primary impersonal patientive verbs (2.2) is some natural power or process, as in (22a) *ciku-a* ‘it (A_Imp) froze it, i.e. it is frozen’, the added impersonal agent by means of the VVs |+nayqi-| group suffixes (47) below is something along the line of necessity or destiny (beyond personal or human control). An extended stem with |+nayqi-|, etc. is patientive.

Just as the primary impersonal *ciku-*, derived impersonal verbs are patientive and occur with either transitive or intransitive inflection. That is, the added impersonal A argument is inflectionally indexed as the transitive subject in the third person singular (though it never occurs as a full NP, hence no need of case assignment). The following two examples come respectively from a monovalent and an (agentive) bivalent stem:

(45) *ayag-narq-aaten*  
go-NEC-IND.3SG.2SG.
‘you(sg.) have to go; lit. it necessitates you to go’.

(46) *Ner-narq-aanga*  
saq-mek_Pf  
eat-NEC-IND.3SG.1SG. fish-ABM.SG.
‘I have to eat fish; lit. it necessitates/destines me to eat fish.’

The A_Imp argument in (45) and (46) may be deleted in de-transitivization (i.e. passivization) as the extended stems are patientive (as stated above), and their intransitive forms (48b) and (50b) are again more common than transitive.
The three $A_{IMP}$ adders are:

(47) a. $[+naγi-]$ ‘to necessitate, to be destined to’
   b. $[+naγi-]$ ‘to be the time to, to be time-wise destined’
      – compositive with inchoative $[γi-]$ (2.3.2.1)
   c. $[-kixnaγi-]$ ‘to be a good time to’ (rather rare)
      – with $[-kγc(i)-]$ ‘to have a good/nice’.

which are clearly related to NN/NV $[+naγi-]$ ‘one that causes/to cause’ (e.g. $tuqu-naq$ ‘poison’ from $tuqu-|$ ‘to die’, $aling-naq$ ‘hazardous thing’ from $aliŋi-|$ ‘afraid, scared’) expanded by transitive relational verb NVr $[ki-]$ ‘to have as’ into $[+naγi-]$. This necessitative suffix $[+naγi-]$ is incidentally one of the most productive and extensive derivative suffixes in CAY.

It is important to note that the suffix is not only valency-increasing (responsible for impersonal A of necessity/destiny) but may also behave merely like a modality marker which does not affect valency, inflecting with any person and number (first and second, dual and plural as well). Accordingly (47a). $[+naγi-]$ verb with third person singular inflection may have ambivalent readings (A ‘he’ or $A_{IMP}$ ‘it’) as mainly illustrated in 3.2.

The suffix (whether valency-increasing or not) may not only occur with primary stems (mono-, bi-, tri-valent) but also with expanded multi-valent stems, thereby rendering case alignment complicated, though it regularly follows the same general pattern of reduction.

Discussion below are made of $[+naγi-]$, with the other two suffixes (b and c) being illustrated only at the end in (66) and (67).

3.1 Impersonal A

Illustrations are made of impersonal transitives and their de-transitivizations, with case alignment shown at the very end. In the following pairs, (a) is transitive or transimpersonal (though there may be another reading), while (b) is intransitive or an impersonal passive, with $A_{IMP}$ deleted as is the case with any patientive verbs. The transitive subject in (a) is always in the third person singular (while the transitive object can be any person and number).

3.1.1 Monovalent stems – The $A_{IMP}$ suffix occurs with a monovalent stem (‘to go’) in the following, yielding a derived necessitative impersonal bivalent stem with

20. The suffix needs a particular phonological adjustment of $iγl$ being deleted before $qC$ triggered by final $iγl$ deletion, before certain consonant-initial suffixes: $aling-naq-luni$ ‘(he/it) frightening’ $< aling-naqγe-luni (aliŋi-|-‘fearful-get’). But no $iγl$ deletion occurs if final $iγl$ deletion does not occur before a consonant cluster stays: $tag-naqγe-nrγe-ami$ (go.up-NEG-NEG-CNNGG.$γc.3Rsg.$) ‘as it was not far to go up’.
P(S) and $A_{IMP}^{21}$ which, in de-transitivization, requires $A_{IMP}$ deletion coupled with P demotion. The $A_{IMP}$, indexed in the transitive (a) as its subject, can never be external. In the corresponding intransitive (b), with $A_{IMP}$ deleted, the subject S from P(S) is personal and the suffix is little more than a modality marker:

(48)  a. **ampi ajag-*narq-aaten** $P_{abm}(S) A_{IMP(rel)}$
      hurry go-NEC-IND.3SG.2SG. = (45)
      'you(sg.) have to (i.e. it necessitates you to) hurry up and go'.

b. **ampi ajag-*narq-uten** $S_{abs} A_{IMP(O)}$
      hurry go-NEC-IND.2SG.
      'you(sg.) must hurry up and go'.

(49) **apiatar-narq-aakuk**
      eat.lunch-NEC-IND.3SG.1du.
      'we(du.) have to (i.e. it necessitates us to) eat our noon meal'.

3.1.2 **Agentive bivalent stems** – The $A_{IMP}$ suffix, occurring with a bivalent stem, derives a trivalent stem which requires argument reduction (P demotion), yielding the alignment $P_{abm} A_{abs} A_{IMP(rel)}$. If the bivalent stem is agentive (‘to eat’) as in the following example, the reduction is made by demoting the P argument to ablative-modalis status, thereby promoting the A to the absolutive and promoting $A_{IMP}$ to the relative case (but with no full NP). Note that ‘I’ (‘eater’) is the transitive object:

(50)  a. **Ner-narq-aanga neq-mek** $P_{abm} A_{abs} A_{IMP(rel)}$
      eat-NEC-IND.3SG.1SG. fish-ABM.SG. = (46)
      'I have to eat fish’ (as if ‘it necessitates me to eat fish’)
      cf. ner-aanga ‘he is eating me’, but not *ner-aanga neq-mek

b. **Ner-narq-ua neq-mek** $P_{abm} S_{abs}(A) A_{IMP(O)}$
      eat-NEC-IND.1SG. fish-ABM.SG.
      'I must eat fish'.
      cf. ner-ua neq-mek ‘I am eating fish’ – see (2b).

21. The suffix also has produced more or less lexicalised monovalent stems, especially from many a-valent roots: **aling-narq-uq** ‘it being frightening’ (footnote 20), **ayuqnia-narq-uq** ‘he is enviable’ ([ayuqniaj| ‘to envy’), **cakenr-narq-uq** ‘it is difficult, makes it hard (on you)’ ([caŋkni| ‘to struggle to, have a hard time’), **caperr-narq-uq** ‘it is difficult’ ([capx| ‘difficult’), **qacig-narq-uq** ‘it is easy’ ([qaciy| ‘easy’), **uumi-narq-uq** ‘it is infuriating’ ([uumi| ‘irritate’), **takar-narq-uq** ‘he is intimidating, makes one (respectfully) shy’ ([takaj| ‘shy’), **tanger-narq-uq** ‘it is visible’ ([taŋx| ‘to see’), **niit-narq-uq** ‘it can be heard’ ([niic| ‘to hear’), **nallu-narq-uq** ‘it is hard to know, understand’ ([natu| ‘not to know’), **akngir-narq-uq** ‘it (body part) hurts’ ([akŋiy| ‘in pain’), **akngir-r-narq-uq** (it; A_{imp}) causes him to get hurt, he can hurt’ ([akŋiy-c-|) with A (causative) adder [+c-], **tuqu-narq-uq** ‘it is deadly, poisonous’ ([tuqu| ‘to die’), **cang-narq-uq** ‘it has lots of fish’ ([ca-ŋ | ‘to catch fish/game’ from ‘something-get’).
The impersonal A is deleted in the intransitive (50b), since the extended *ner-narq-* is patientive.

The (a) clause in the following example has the same pattern as the (a) in the preceding ones, the only difference being in the person (3sg.) of the P argument (which is the ‘eater’). But note that (b) is ambivalent:

(51) a. *ner-narq*-aa  neq-mek$_{(P)}$  P$_{abm}$  A$_{ab}$  A$_{IMP(rel)}$
   eat-NEC-IND.3SG.3SG.  fish-ABM.SG.
   ‘he has to eat a fish’

b. *ner-narq*-aa  neq$_{p}$
   eat-NEC-IND.3SG.3SG.  fish-ABS.SG.
   i. ‘the fish must be eaten’  P$_{ab}$  A$_{(o)}$  A$_{IMP(rel)}$
   ii. ‘he must eat the fish’  P$_{ab}$  A$_{rel}$
   – in (i) the A argument is deleted while the impersonal A is the transitive subject, while in (ii) the suffix has no valency increase and the transitive A is personal as will be illustrated more in 3.2.

In the following intransitive construction, both A and A$_{IMP}$ are deleted:

(52) Qaillun  u-na$_{SgP}$  atur-narq-a
   how  this-EX.ABS.SG.  use-NEC-INT.3SG.
   ‘How is this supposed to be used?’  S$_{ab}$  (P)  A$_{(o)}$  A$_{IMP(o)}$

3.1.3 Patientive bivalent stems – On the other hand, if the bivalent stem is patientive (e.g. ‘to help’) as in the following, the reduction is made by deleting the A (‘helper’) in (a) and further by deleting A$_{IMP}$ in (b):

(53) a. *ikayur-narq-aaten*  S(P)$_{ab}$  A$_{(o)}$  A$_{IMP(rel)}$
    help-NEC-IND.3SG.2SG.
    ‘you(sg.) have to be helped, need help (lit. it necessitates you to be helped)’
    cf. *ikayur-aaten* (help-IND.3SG.2SG.)  ‘he is helping you’

b. *ikayur-narq-uten*
    S$_{ab}$  (P)  A$_{(o)}$  A$_{IMP(rel)}$
    help-NEC-IND.2SG.
    ‘you(sg.) have to be helped, you need help’

In the following example with the antipassive stem (with E$_{APS}$ |γi-|), reduction involves P demotion like (50b) above in 3.1.2:

(54) a. *ikayur-i-narq-aanga*  naulluu-ria-mek$_{(P)}$
    help-E$_{APS}$-IND.3SG.1SG.
    ‘I must help a patient.’

b. *ikayur-i-narq-ua*  naulluu-ria-mek$_{(P)}$
    help-E$_{APS}$-IND.1SG.
    ‘I must help a patient.’
Impersonal verbs in Central Alaskan Yupik (Eskimoan) 483

(55) Pissur-yar-pailemta casku-put\textsubscript{p} P\textsubscript{abs} A\textsubscript{rel}
hunt-go.to-CNNbc.1pl. weapon-ABS.1PL.1PL.

kitugg-narqe-ciq-aput fix-NEC-FUT-IND.1PL.3PL.

‘Before we go out hunting we will have to fix our weapons.’

Since impersonal A\textsubscript{IMP} and personal A arguments are involved, bivalency may arise as in (56a), depending upon which of the two As is deleted or in what order they are deleted (if both are deleted). Some evidence shows that the recent tendency is more toward A\textsubscript{IMP} deletion rather than personal, if the transitive subject is in the third person.

In the following example the stem [tuqu-\textsubscript{c-}] (tuqu-\textsubscript{t-}) is a patientive bivalent derived from monovalent [tuqu-\textsubscript{-}] ‘to die’ with causative A |+c-| (cf. 1.3.1):

(56) a. Tuqu-t-narq-aa qimugta\textsubscript{p}
die-A-NEC-IND.3SG.3SG. dog.ABS.SG.
 i. ‘The dog has been killed.’ – with A deleted
 ii. ‘He must kill the dog.’ – with A\textsubscript{IMP} deleted, in which a personal A (‘killer’) may be expressed by a relative-case NP, e.g. angute-m (man-REL.SG.)
 b. Tuqu-t-narq-uq qimugta\textsubscript{s}
die-A-NEC-IND.3SG. dog.ABS.SG.
 ‘The dog must be killed.’ – with both A and A\textsubscript{IMP} deleted.

(57) ellet-narq-uq (~[Yukon] ell’-narq-uq)
 ‘it (air) must be squeezed out’
 – patientive |ilc-| ‘to deflate, squeeze’.

Bivalent applicative stems with a monovalent stem are also patientive:

(58) Iqva-ut-narq-aa = iqva-ut-narq-uq arnaq(E)
pick.berry-E\textsubscript{APL}-NEC-IND.3SG.3SG./-IND.3SG. woman.ABS.SG.
 ‘The woman requires (s.o.) to pick berries for the woman, i.e. s.o. has to pick berries for the woman.’
 – agentive |iqvaγ-| ‘to pick (berries).’

3.1.4 Trivalent (ditransitive) stems – Secundative (59) is compared with indirective (60): see (5cd) for the different case alignments.

(59) a. Angun cikir-narq-aa akuta-mek.
 man.ABS.SG. give-NEC-IND.3SG.3SG. ice.cream-ABM.SG.
 i. ‘The man has to be given ice cream.’
 – with A\textsubscript{IMP} as the subject (less acceptable)
 ii. ‘She must give ice cream to the man.’
 – with e.g. arna-m (woman-REL.SG.) for the subject.
 b. Angun cikir-narq-uq akuta-mek.
 man.ABS.SG. give-NEC-IND.3SG. ice.cream-ABM.SG.
 ‘The man must be given ice cream.’ – with both A\textsubscript{IMP} and A deleted.
(60) a. Akutaq tune-narq-aa angut-mun.  
        ice.cream.abs.sg. give-NEC-IND.3SG.3SG. man-ALL.SG.  
        i. ‘The ice cream has to be given to the man.’ – with A\textsubscript{\textsc{imp}} as the subject.  
        ii. ‘She (arna-m; (59a-ii)) must give ice cream to the man.’

        ice.cream.abs.sg. give-NEC-IND.3SG. man-ALL.SG.  
        ‘The ice cream must be given to the man.’ – with both A\textsubscript{\textsc{imp}} and A deleted.

3.2 Modality marker

The impersonal [+\textsc{naγi}]-, however, may not be a valency-increasing suffix (unlike 3.1) but is a modality marker, when the S or A argument of primary verbs may be in any person and number (as opposed to the third person singular for A\textsubscript{\textsc{imp}}; 3.1).

Compare with (50a) the following which has the reversed personal relation inflection of 1SG.3SG. as opposed to 3SG.1SG., though both being ‘I’ as the eater. Now not being an impersonal patientive verb, an antipassive form is possible with P demotion (b) or with A deletion (62a), which can have another reading (62b).

(61) a. Ner-narq-aqa neqa\textsubscript{\textsc{p}} P\textsubscript{\textsc{abs}} A\textsubscript{\textsc{rel}}  
        eat-NEC-IND.1SG.3SG. fish.abs.sg.  
        ‘I must eat the fish (e.g. before a dog eat it).’
        cf. ner’-aqa neqa ‘I am eating the fish’ – see (2a)

b. Ner-narq-ua neq-mek\textsuperscript{(p)} P\textsubscript{\textsc{abm}} S(A)\textsubscript{\textsc{abs}}  
        eat-NEC-IND.1SG. fish-ABM.SG.  
        ‘I must eat fish (right away).’
        – zero-derived antipassivization (with demoted P in the ablative -modalis) as agentive stem (1.2.2-i) since it does not change the stem type as a modality suffix:
        cf. ner’-ua neq-mek ‘I am eating fish’ – see (2b).

In the following pair with the third person subject intransitive forms, (a) shows passivization with A (‘eater’) deleted, while (b) has the absolutive neqa as the subject of the (zero-derived) antipassive construction like (61b):

(62) a. Ner-narq-uq neqa\textsubscript{s} S(P)\textsubscript{\textsc{abs}} A\textsubscript{\textsc{(o)}}  
        eat-NEC-IND.3SG. fish.abs.sg.  
        ‘the fish must be eaten (right away)’
        – with A (‘eater’) deleted in de-transitivization\textsuperscript{22}

\textsuperscript{22} An agentive stem may be subject to passivization if expansion (esp. aspect) is made, e.g. ner-uma-uq ‘it has been eaten.’
b. \textit{Ner-narq-uq neqa}s \textit{P}\textsubscript{abm} S(A)\textsubscript{abs}  
\textit{eat-NEC-IND.3SG. fish.ABS.SG.}  
‘the fish must eat something (right away)’  
– ‘something’ (P) may be added by an ablative-modalis NP (e.g. \textit{ciissi-mek} ‘bugs’).

The transitive subject in the third person singular in the following example may be personal A or impersonal A\textsubscript{IMP}, hence two readings:

\begin{enumerate}[label=(\arabic*)]
\item \textit{ner-narq-aa neqa}\textsubscript{p}  
\textit{eat-NEC-IND.3SG.3SG. fish.ABS.SG.}  
\begin{enumerate}[label=(\roman*)]
\item ‘he (e.g. \textit{Maya}-m REL.sg.) must eat the fish’, cf. (61a)  
– with \textit{P}\textsubscript{abs} A\textsubscript{rel}
\item ‘the fish has to be eaten’  
– with \textit{P}\textsubscript{abs} A(\textsubscript{Ø}) A\textsubscript{IMP(rel)}  
which is equal to (51b) as opposed to (51a) \textit{ner-narq-aa neq-mek} ‘he has to eat a fish’ with \textit{P}\textsubscript{abm} A\textsubscript{abs} A\textsubscript{IMP(rel)}.
\end{enumerate}
\end{enumerate}

3.3 Further derivations of impersonal verbs

Finally, impersonal verbs, either primary or derived, are subject to further verbal modifications like intensification, time-aspect, negation, modality, evidentiality. In particular, they are very often found to occur with |\textit{+caaq} G\textperiodcentered| (‘but, without success, in vain’) to imply ‘it is possible to –’:

\begin{enumerate}[label=(\arabic*)]
\item \textit{ayag-naq-saaq-ua}  
\textit{go-NEC-but-IND.1SG.}  
‘it is possible for me to go’
\item \textit{ayag-naq-saaq-aanga}  
\textit{go-NEC-but-IND.3SG.1SG.} – cf. (48).  
‘it (A\textsubscript{IMP}) makes it possible for me to go’
\end{enumerate}

An impersonal verb clause may be put into a complex transitive construction:

\begin{enumerate}[label=(\arabic*)]
\item \textit{Apy-ut-naq-ni-anga mulia-vnun}  
\textit{ask-\textit{E}\textsubscript{APL}-NEC-A’.say-IND.3SG.1SG. wife-\textit{ALL}.2SG.SG.}  
\textit{akuta-li-yara-mek. ice.cream-make-VNn-ABM.SG.}  
‘She said that she herself (or s.o. else) must ask your wife for me about how to make ice cream.’  
– with alignment of T\textsubscript{abm} R\textsubscript{all} E\textsubscript{abs} A\textsubscript{rel/all} A\textsubscript{IMP(Ø)} A\textsubscript{rel'} in which here A may be coreferential with A\textsubscript{rel'} (relative) or can be another person (allative).
\end{enumerate}
Finally the two other $A_{\text{IMP}}$ suffixes are illustrated, though they are much less common than $|+\text{na} \bar{y}qi\text{-}|$.

(66) \textit{Pitgar-nari-qatar-qani=gguq ayakar-naur-tuq.} \\
shoot-time.to-IMN-CNNif.3sg.3Rsg.=RPT flee-would-IND.3sg. \\
‘When he was about to shoot (it), it is said, it would fly.’

(67) \textit{Iqa-ir-i-keggnarq-aaten unuamek.} \\
\textit{dirt-PRV-$E_{\text{APS}}$ good.time.to-IND.3sg.2sg. today} \\
‘Today is a nice day for you(sg.) to wash (clothes).’ \\
\cong \textit{Iqa-ir-i-keggnarq-ut-aaten unuamek} \\
– with applicative -ut- (2.1.1).

4. Summary

As described, CAY impersonal verbs, of which there are two kinds, (i) primary and (ii) derived, occur very commonly and are still a productive feature of the language.

i. A primary impersonal belongs either to monovalent stems with an $S_{\text{IMP}}$ argument or to patientive bivalent stems with $A_{\text{IMP}}$ and $P$. The impersonal argument involved is some unidentifiable natural force or process. There are no primary trivalent (i.e. ditransitive) impersonal stems with $A_{\text{IMP}}$, $R$, and $T$.

Primary impersonal stems as listed in (2.1 and 2.2) are not necessarily exhaustive. The transitive use of monovalent stems (2.3) varies widely depending upon the stems and the speakers. Some of the bivalent impersonal stems may rather be used by some speakers with a personal agent like ordinary patientive stems (2.2).

ii. The derived impersonals are characterized by the $A_{\text{IMP}}$ adding suffix $|+\text{na} \bar{y}qi\text{-}|$ (and its related two suffixes), and the argument $A_{\text{IMP}}$ is some force of necessity or destiny (thus necessitative impersonal). It is productively and freely used by any speakers with any type of primary stem – monovalent, bivalent, or trivalent – or with a valency increased stem (3.1).

The suffix may only be a modality marker with no valency increase, thereby causing some ambivalency, with different person inflection (3.2).

iii. Like many other CAY verbs, formation of impersonal verbs is a result of the whole dynamic process with possibly recursive valency modification (increase, decrease, arrangement) by suffix- or zero-derivation. They may also be embedded into a complex transitive verb with an upper-layer subject. Hence complicated patterns in case alignments (possibly with ambivalency), which, however, are regularly explainable by the general case assignment rule (1.3).

iv. Whatever modifications may be involved in formation, both primary and derived impersonals may be inflected intransitively (with a hardly perceived impersonal...
force) and transitively, though transitive forms are generally found to be used much less commonly by the younger generations and are seemingly in gradual decline in the language.

Striking variability in acceptability of many impersonal verbs – not only among generations and speakers but also from one verb stem to another – may at least partly be a phenomenon due to a very rapidly changing and vanishing language in variable bilingual situations with the encroaching of English.

**Abbreviations**

<table>
<thead>
<tr>
<th>Abbr.</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>agent/transitive subject</td>
</tr>
<tr>
<td>A_imp</td>
<td>transitive impersonal agent</td>
</tr>
<tr>
<td>A′ (A″,..)</td>
<td>upper-layer agent (complex transitives)</td>
</tr>
<tr>
<td>ABM</td>
<td>ablative-modalis case</td>
</tr>
<tr>
<td>ABS</td>
<td>absolutive case</td>
</tr>
<tr>
<td>ADV</td>
<td>adversative</td>
</tr>
<tr>
<td>ALL</td>
<td>allative case</td>
</tr>
<tr>
<td>APL</td>
<td>applicative</td>
</tr>
<tr>
<td>APP</td>
<td>appositional mood</td>
</tr>
<tr>
<td>APS</td>
<td>antipassive</td>
</tr>
<tr>
<td>ATN</td>
<td>attention</td>
</tr>
<tr>
<td>BNF</td>
<td>benefactive</td>
</tr>
<tr>
<td>CAU</td>
<td>causative</td>
</tr>
<tr>
<td>CAY</td>
<td>Central Alaskan Yupik</td>
</tr>
<tr>
<td>CNN</td>
<td>connective mood</td>
</tr>
<tr>
<td>CNS</td>
<td>constancy</td>
</tr>
<tr>
<td>DEM</td>
<td>demonstrative</td>
</tr>
<tr>
<td>DES</td>
<td>desiderative</td>
</tr>
<tr>
<td>E</td>
<td>experiencer</td>
</tr>
<tr>
<td>E_ADV</td>
<td>adversative</td>
</tr>
<tr>
<td>E_APL</td>
<td>applicative</td>
</tr>
<tr>
<td>E_APS</td>
<td>antipassive</td>
</tr>
<tr>
<td>E_BNF</td>
<td>benefactive</td>
</tr>
<tr>
<td>EV</td>
<td>epenthetic vowel</td>
</tr>
<tr>
<td>EX</td>
<td>(root) expander</td>
</tr>
<tr>
<td>FUT</td>
<td>future</td>
</tr>
<tr>
<td>IMP</td>
<td>impersonal</td>
</tr>
<tr>
<td>INC</td>
<td>inchoative</td>
</tr>
<tr>
<td>IND</td>
<td>indicative mood</td>
</tr>
<tr>
<td>INF</td>
<td>inferential</td>
</tr>
<tr>
<td>INJ</td>
<td>interjection</td>
</tr>
<tr>
<td>INS</td>
<td>instrumental</td>
</tr>
<tr>
<td>INT</td>
<td>interrogative mood</td>
</tr>
<tr>
<td>LNK</td>
<td>linker</td>
</tr>
<tr>
<td>LOC</td>
<td>locative case</td>
</tr>
<tr>
<td>NEC</td>
<td>necessitative</td>
</tr>
<tr>
<td>NN</td>
<td>nominal-elaborating suffix</td>
</tr>
<tr>
<td>NP</td>
<td>nominal phrase</td>
</tr>
<tr>
<td>NV</td>
<td>denominizing suffix</td>
</tr>
<tr>
<td>NVR</td>
<td>relational verb</td>
</tr>
<tr>
<td>OPT</td>
<td>optative mood</td>
</tr>
<tr>
<td>P</td>
<td>patient</td>
</tr>
<tr>
<td>PL</td>
<td>plural</td>
</tr>
<tr>
<td>PPL</td>
<td>participial mood</td>
</tr>
<tr>
<td>PRF</td>
<td>perfective</td>
</tr>
<tr>
<td>PRL</td>
<td>perative case</td>
</tr>
<tr>
<td>PRV</td>
<td>privative</td>
</tr>
<tr>
<td>PST</td>
<td>past</td>
</tr>
<tr>
<td>R</td>
<td>recipient</td>
</tr>
<tr>
<td>RCL</td>
<td>relative clause</td>
</tr>
<tr>
<td>REL</td>
<td>relative case</td>
</tr>
<tr>
<td>RPT</td>
<td>reportative</td>
</tr>
<tr>
<td>S</td>
<td>intransitive subject</td>
</tr>
<tr>
<td>S_IMP</td>
<td>intransitive impersonal</td>
</tr>
<tr>
<td>S.O.</td>
<td>someone</td>
</tr>
</tbody>
</table>
s.t.  something
T    theme
VN   deverbalingizing suffix
VNN  nominalizer
VNR  relativizer

References


Impersonals in Innu

Lynn Drapeau
Département de linguistique

This paper analyzes impersonal constructions in Innu, an Algonquian language, closely related to Cree and spoken in North Eastern Quebec. Impersonalization is determined by the ranking of the suppressed entity on the animacy hierarchy. Suppressing an Animate ‘subject’ requires additional verbal derivation. By contrast, Inanimate subject verbs may have both a personal and an impersonal reading without further derivation. The language also exhibits a large number of lexical impersonal verbs since natural forces and elements of the landscape are lexicalized as verbs, or verb parts. Impersonal verbs belong to the class of Inanimate Intransitives and their suppressed ‘actor’ argument may never find syntactic realization as a lexical NP. Attention is drawn to the complex interaction between lexicalization patterns, constraints on the expression of subjecthood and information structure.

Keywords: Algonquian; Innu; Cree; Impersonals; Animacy Hierarchy; Natural forces

1. Introduction

Impersonals are defined cross-linguistically as constructions in which the element that normally assumes the ‘subject’ function in topic-comment statements is either downgraded or completely suppressed. In this paper, I adopt the subject-centered view of impersonals focusing on constructions which lack a referential subject either intrinsically (lexical impersonals) or through derivation (derived impersonals). The paper describes and analyzes this type of construction in Innu, an Algonquian language.

---

1. Funding by the Canada Social Sciences and Humanities Research Council (grant no 856-2004-1068) for the Innu Grammar Project is gratefully acknowledged. I also wish to thank Jimena Terraza, Nicholas Vaughan and Marguerite MacKenzie for their comments on an earlier draft and the editors and reviewers of this volume for many helpful comments. Due thanks also to Magali Lachapelle for her help in formatting the paper. Usual disclaimers apply.
spoken in Northeastern Quebec. It will show that impersonalization in Innu is determined by the ranking of the suppressed entity on the animacy hierarchy (Siewierska 2004; Silverstein 1976). There is a cut-off point on the animacy hierarchy beyond which inanimates are not encoded as arguments in the subject function and are not even lexicalized as nouns. The result is a large class of lexical impersonal verbs. As for derived impersonals, whereas suppressing a human subject requires special derivation on the verb, suppressing an inanimate S need not require special morphology since verbs with inanimate subjects can have both a personal and an impersonal reading. The analysis reveals that impersonals fit into the Intransitive class of verbs and have a dummy Inanimate subject index. Second, impersonalization requires the total suppression of the downgraded argument, hence the impossibility of its finding syntactic realization as a lexical NP in the clause. Third, in information structure terms, impersonal verbs represent thetic (sentence-focus) constructions. For breadth of coverage, and since they are so closely related, I also include data on other instances where the Agent is unspecified (so-called “transimpersonals”) which are not strictly speaking impersonal under the subject-centered view.

I begin by a presentation of the relevant properties of Innu (§2), and then move on to the description of lexical impersonal verbs (§3). §4 is devoted to the analysis of impersonals derived from the suppression of an animate subject. In §5 I turn to a residual group of transimpersonal forms that lack an inanimate A despite their transitive morphology.

2. Typological and language-specific features of Innu

This section is devoted to the presentation of the features of Innu that have a bearing on the topic under discussion. I begin by looking at some broad typological parameters and then go into more specific information on the language. Lastly, I address the issue of the identification of impersonal forms.

2. The language was formerly known as Montagnais, a term replaced by the autonym Innu in the last decade. It should be underscored that the language is part of the chain of Cree dialects that ranges from the Canadian Rockies to the easternmost Innu dialects spoken in North Eastern Quebec.

3. The data presented here come from the author’s texts, dictionary (Drapeau 1991) and field notes gathered over the years in the Innu community of Pessamit. Given the numerous phonetic and morphophonological changes in that dialect of Innu, the system of transcription used in the morphemic glossing is quite abstract; it is a practical compromise designed both to unpack the surface fusion of segments and morphemes and to facilitate comparison with other dialect groups and sister languages.
2.1 Impersonals in Innu

Innu is a consistently head-marking language in the sense proposed by Nichols (1986). As a polysynthetic language, it typically displays a high ratio of morphemes per word, especially on the verb. The verb bears a pronominal index for its higher-ranking arguments, albeit with restrictions related to the animacy hierarchy as shown in the next section. There is no case-marking on nouns. As in many polysynthetic languages, Innu exhibits productive noun incorporation. Incorporated elements are not restricted to Patients but correspond to a wide array of possible semantic roles (Drapeau 2008).

2.2 Hierarchical alignment

Algonquian languages are described on current accounts of alignment types as 'hierarchical alignment' languages (see Siewierska 2005). Languages with hierarchical alignment typically exhibit an 'inverse' system (Klaiman 1991; Siewierska 1984; Thompson 1994; Zúñiga 2006) alongside a 'direct' system of morphological encoding with two variables coming into play: the semantic role of the argument and its rank on the person hierarchy. The person hierarchy in Innu is 2>1>3>3'>Inanimate (see also Dahlstrom 1991; Wolfart 1973). The logic that underlies the active-inverse system is the following: in transitive clauses the verb encodes with direct morphology all instances in which an argument higher in the hierarchy assumes the Agent role and with inverse morphology all instances in which an argument lower on the hierarchy assumes that role. When both arguments are third person, the more topical argument, called the proximate (symbolized as 3), is higher on the hierarchy than the less topical one, called the obviative (symbolized as 3'). When only third person referents are involved, the verb displays forms with direct markers when the topic is Agent and inverse marking when it is Patient. The efficiency of the system rests crucially on the possibility of disambiguating third person referents. Obviation marking on both nouns and verbs serves as a reference-tracking device (Comrie 1989b) by marking the less topical argument as obviative, while the proximate remains unmarked.

2.3 Gender and verbal morphology

The language groups nouns into two grammatical genders. Animate nouns cover both semantic animates and a residual class of semantically inanimate entities (such as 'sun';

---

4. In this respect, the language deviates from the normal ranking (1>2>3) found in typological surveys (Siewierska 2004).

The Algonquianist tradition distinguishes primary and secondary derivation (see Bloomfield 1946; Goddard 1990). In Goddard’s model, the derivation of the primary verb stem exhibits a tripartite structure: the initial, followed optionally by a medial and lastly, the final complex as exemplified in (1). Inflectional material may be added on both sides of the stem. The initial is a bound root, an already existing stem or part of a stem. The medial slot is filled by morphemes denoting entities: incorporated nouns, classifiers or other bound morphemes denoting entities. The last slot is devoted to a closed class of ‘final’ elements (identified as FIN in the morphemic glossing) that have concrete or more or less abstract meaning. Concrete finals are mostly manner-of-action morphemes expressing either the manner of human action (by hitting, with an instrument, by pulling, walking, running, etc.) or the action of natural forces (wind, heat, frost, etc.). Abstract finals mostly mark verb class but they may also bear abstract semantic properties (see Denny 1984).

(1) primary verb stem: [initial + (medial) + final]

As for secondary derivation, in Goddard’s model, it is analogous to transposition. Transposition to a different category is carried out by a number of secondary finals that change the category of the stem or its valence.

Algonquian languages display an outstanding system of verb stem alternation. For every transitive verb, there is a stem form in case the P is Animate and an alternate one if it is Inanimate. Likewise, for each intransitive verb, there is a stem form in case the S is Animate and an alternate one for an Inanimate S. Stem alternation is located in the ‘final’, as it is the finals that change form depending on the gender of the argument. Since these finals also convey lexical information, they are not part of the inflectional system. All in all, a system of four morphological verb classes emerges identifiable on the basis of the shape of their final morpheme.6

- Transitive Animate (TA) verbs are morphologically transitive and take an Animate P
- Transitive Inanimate (TI) verbs are morphologically transitive and take an Inanimate P

5. It is common in Algonquian studies to symbolize Animate third person inflection as 3 and Inanimate third person inflection as 0. This practice however obscures the fact that Animates and Inanimates are indexed on verbs by the same set of suffixes and it is therefore not followed in the present paper.

6. Because the finals take different forms depending on the morphosyntactic verb class to which they belong, the class is marked in the morphemic glossing on the FIN morpheme.
- Animate Intransitive (AI) verbs are morphologically intransitive and take an Animate S
- Inanimate Intransitive (II) verbs are morphologically intransitive and take an Inanimate S

The system of stem alternation is exemplified in the following examples all of which are derived with the root wâp-. The transitive final for verbs of ‘seeing’ has the shape \(-am\)- when the object is Animate (2) and \(-at\)- when it is Inanimate (3). The stative final is \(-išî\) when the S is Animate (4) and \(-â\)- when it is Inanimate (5).

\[(2)\] TA: \(n\- wâp\- am\- â\- w\)
\[1\- \text{ROOT}\- \text{FIN.TA}\- \text{DIR}\- 3\]
'I see him'

\[(3)\] TI: \(n\- wâp\- at\- ê\- n\)
\[1\- \text{ROOT}\- \text{FIN.TI}\- \text{TIT}\- 1\]
'I see it'

\[(4)\] AI: \(n\- wâp\- išî\- n\)
\[1\- \text{ROOT}\- \text{FIN.AI}\- 1\]
'I am white'

\[(5)\] II: \(wâp\- â\- w\)
\[\text{ROOT}\- \text{FIN.II}\- 3\]
'it is white'

It is worth noting that the isomorphism with syntactic transitivity is not perfect because (a) some TI verbs are objectless, and (b) a number of verbs of the AI class are ambitransitive.

In addition, the language exhibits mutually exclusive sets of verbal inflectional suffixes called *orders* in the Algonquianist tradition: the independent, the conjunct and the imperative orders. The motivation for the orders is by and large syntactic. Roughly, independent order forms occur in main and independent clauses and conjunct order forms in WH-questions and embedded clauses.\(^7\) In our glosses only the conjunct order is explicitly mentioned in the morphemic glossing, such that unless otherwise specified, a verbal inflectional suffix should be interpreted as belonging to the independent order set.

\(^7\) Some tenses and modes only have independent order forms however, and this fact skews the syntactic distribution.
2.4 The coding of grammatical functions

Since impersonals lack a referential subject, it is crucial to define the notion in the language under study. In that respect, we need to look first at how grammatical functions are encoded.

The distribution of external NPs and pronouns leaves the verb as the central, often stand-alone, element in the clause. Indeed, external NP arguments are never compulsory and they surface mainly when the referent is not otherwise recoverable from context or for contrastive purposes. Moreover, external personal pronouns are used for contrastive emphasis and must bear contrastive stress. Verb inflection is compulsory in the form of inflectional suffixes and prefixes (the latter only in independent order forms). Core arguments (i.e. those entailed by the verb’s meaning) are indexed on the verb through inflectional affixes which in turn constitute the main cue in determining the status of a verb’s arguments. For this reason, Innu may be defined as belonging to the ‘pronominal argument’ type, frequent in nonconfigurational languages, in which the indexes on the verbs are pronouns taken as the de facto ‘arguments’ (see Jelinek 1984; Van Valin & LaPolla 1997). On this view, the only way to determine the function of an argument is to look at how it is indexed on the verb. Let us now turn to how arguments are actually selected for indexation on the verb.

The indexing of core arguments on the verb is constrained by the person hierarchy and by the animate/inanimate distinction as summarized in the hierarchy presented in §2.2. In transitive verbs, only Animate gender arguments (corresponding to A or P) are indexed according to the following constraints.

- when both the A and the P are Animate (as in Paul feeds the baby), both are indexed on the verb
- when the A is Animate and the P is Inanimate (as in Paul fills the tank), the verb only indexes the Animate A

---

8. This view has however been challenged for Maliseet-Passamaquoddy by LeSourd (2006).
9. Siewierska (2004) decomposes the hierarchies into sub-hierarchies: the person hierarchy (1>2>3), the animacy hierarchy (human>animate>inanimate>abstract), the referentiality hierarchy, etc. Only the animacy hierarchy need concern us here.
10. I use the labels proposed by Comrie (1989a): S symbolizes the single argument of intransitive predicates, A stands for the more agent-like argument of a transitive predicate and P represents the more patient-like argument of a transitive predicate.
11. There are only minor exceptions to this rule, namely cognate objects.
– when the A is Inanimate and the P is Animate (so-called *Inanimate actor forms* as in *the food makes Mary ill*), the verb only inflects for the Animate patient but the Inanimate Agent may be realized syntactically as an external NP\(^{12}\).

– when both arguments are Inanimate, the Inanimate Agent is encoded in some other fashion, either as an incorporated noun or as a manner-of-action morpheme on the verb stem.

In other words, both arguments of a transitive verb are indexed on it only if they are both Animate. If only one is Animate, it is the only one to become indexed. It follows that AI, II and TI verbs only index one argument, while TA verbs index (at least) two arguments. Hence in AI, II and TI verbs, the sole indexed argument on the verb automatically gets identified as the ‘subject’ (S). It also follows from these stringent animacy-related restrictions on argument indexing that one cannot deduce the syntactic valence of a verb only by looking at its inflectional markers.

We are now ready to look at how the arguments are actually encoded on the verb. Such encoding is achieved through prefixes and suffixes in the independent order but uniquely through suffixes in the conjunct and the imperative order. The following illustrates the system as it operates in the independent order.

Nouns and verbs in the independent order share essentially the same set of personal prefixes presented in Table 1. Observe that verbs do not use third person prefixes.

**Table 1. Personal prefixes**

<table>
<thead>
<tr>
<th>Nouns</th>
<th>Verbs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 ni(t)-(^{13})</td>
<td>ni(t)-</td>
</tr>
<tr>
<td>2 ci(t)-</td>
<td>ci(t)-</td>
</tr>
<tr>
<td>3 u(t)-</td>
<td>Ø</td>
</tr>
</tbody>
</table>

If the verb has only one indexed argument, it will be indexed by both a person prefix and by a person suffix. Single argument verbal suffixes are provided in Table 2.

---

12. See Rhodes (1990: 402) on “the Inanimate Subject Ban: no clause containing a final object can have an inanimate final subject”.

13. Furthermore, in this set of personal prefixes, an epenthetic /t/ is added to the prefix when the stem begins by a vowel (with exceptions when the stem begins with /u/).
When two arguments are indexed (i.e. verbs with two Animate arguments), the mapping between the prefix and the suffixes on the one hand, and their role or function on the other is constrained by the person hierarchy as explained in §2.2. The system uses the prefixes presented in Table 1 and a number of suffixes. The first set of suffix to appear after the stem is a direction marker that indicates whether the Actor/Agent is higher than the Patient on the person hierarchy (2>1>3>3’). There is a limited set of these direction markers: a pair for so-called ‘local’ speech act participant (SAP) forms, a pair for mixed SAP and non-SAP forms and a pair for exclusively non-SAP forms. The direction markers are followed by other suffixes that further specify the person involved (first person plural inclusive, first person plural exclusive, second person plural), obviation, plurality of third person referents, etc. The main feature of the system is that it uses a limited number of suffixes indeterminate for function (subject or object). It is the direction markers that determine the function of the arguments. Thus, the person prefix indexes the A in direct forms (i.e. in forms that host a direct marker) and the P in inverse forms (forms that host an inverse marker). The reader is referred to Wolfart (1973) or Ellis (1971) for full sets of TA paradigms. The following examples illustrate how the system operates. In (6), the person prefix refers to the A and the plural suffix refers to the P, with the -â- being the direct marker for mixed (SAP and non-SAP) forms. In (7) the same affixes refer respectively to the P and the A (the inverse situation), -ikw- being the inverse direction marker for mixed forms. Likewise, in (8) the plural suffix indexes the A but in (9) it refers to the P. Recall that third person arguments are not indexed by prefixes on verbs.

(6) n- wâpam- â- w -at
   1- see.TA -DIR- 3- PL
   'I see them'

(7) n- wâpam- ikw- at
   1- see.TA -INV- PL
   'they see me'
In the TA paradigm illustrated in the above examples, inverse forms should not be confused with transitive passives. Inverse forms are active and index both participants in the manner specified in the above paragraphs. By contrast, passives only index the P participant as will be shown in §4.2.

To sum up, Innu is a language with relatively free word order, no case marking on nominals, optional use of lexical NPs and exclusively contrastive use of full pronouns. Core arguments are indexed on the verb following the animacy hierarchy. Although the lexical stem of the verb exhibits gender alternation, argument indexing proper (in person and number) is carried out by the verb’s inflectional prefixes and suffixes. If a verb has only one core argument (S), that argument gets indexed on the verb. If it has two core arguments, the P is indexed only in case it is Animate, otherwise only the A gets indexed with the same markers that index an S. In those paradigms where the verb indexes two (Animate) arguments, one cannot in isolation match an affix with a syntactic function. A given inflectional affix can serve in the A or the P function and takes its functional value only when matched with the direction markers.

There has been disagreement among Algonquianists about the identification of grammatical functions for the third person arguments of TA verbs (see on this topic Dahlstrom 1991; Rhodes 1976). Wolvengrey (2005) argues that Algonquian languages lack the syntactic functions ‘subject’ and ‘object’ altogether, a view which is also espoused by Dryer (1997). In his (1946) description of Algonquian languages, Bloomfield does not use the notions ‘subject’ and ‘object’ but refers instead to concepts such as ‘actor’ and ‘goal’. Wolfart (1973) follows the same practice. Rhodes in his many papers on the description of Ojibwa, a closely related language of the central branch of Algonquian languages, uses the primitive notions of ‘subject’ and ‘object’ as part of the apparatus of Relational grammar in which he casts his description. Grammatical relations are also basic to the model used by Dahlstrom (1991) in her description of Plains Cree. However, the tests adduced by Dahlstrom for Plains Cree to justify the syntactic functions of S and O do not apply well in Innu and the evidence for the syntactic notions of subject and object in the language is not very persuasive. The data however do support a view of ‘subject’ as a neutralization of semantic roles for the purpose of argument indexing on the verb. The labels A, P and S are therefore meant here as symbols for the language particular phenomena described above without committing to
the view that they actually represent crosslinguistically universal syntactic relations in the ‘Platonic’ sense expounded by Dryer.

2.5 The identification of impersonal forms

While all verbs carry a person marker, not all of these markers actually refer. Impersonal verbs in Innù are therefore formally indistinguishable from personal verbs since they always carry a dummy inflectional third person singular marker. The presence of the inflectional person marker is therefore not \textit{prima facie} an indication of the status of the verb and we must look elsewhere for confirmation of its impersonal status. As we have seen above, in many head-marking languages, a core external (as opposed to incorporated) lexical NP argument may be omitted if it is recoverable from context. However, such an NP can of course be inserted. Consequently, it is the impossibility of occurrence of an external NP that is indicative of impersonal status. Hence our working definition: a third person verb is impersonal if it cannot occur with an external lexical NP fulfilling the subject (or highest-ranking argument) function.

In the identification of the person index in the examples presented in this paper, it should be borne in mind that throughout the Cree dialect continuum (Ellis 1971; Wolfart 1973), independent order third person -\textit{w} is deleted after Inanimate Intransitive verb stems that end in /n/ so that these verbs seem on surface not to bear inflection. The person marker however is recoverable in the conjunct order forms. But in the latter case, the stem drops its final /n/. The following examples contrast the behavior of both types of stem, vowel final or /n/ final, in the independent order (10) and in the conjunct order (11). Accordingly, both the inflectional marker and its \textit{Ø} allomorph will be glossed as (3).

\begin{align*}
\text{(10)} & \quad \text{miluškami-} & \text{w} & \text{šikun-} & \text{Ø} \\
& \text{be\_late\_spring.II- 3} & \text{be\_early\_spring.II- 3} & \text{Ø} \\
& \text{‘it is late spring’} & \text{‘it is early spring’} & \\
\text{(11)} & \quad \text{miluškami-} & \text{t} & \text{šiku-} & \text{t} \\
& \text{be\_late\_spring.II- 3Cj} & \text{be\_early\_spring.II- 3Cj} & \text{t} \\
& \text{‘that it is late spring’} & \text{‘that it is early spring’} & \\
\end{align*}

3. Lexical impersonal verbs

This section is devoted to lexical impersonal verbs which all display an Inanimate Intransitive (II) stem and never occur with an external lexical subject. I mentioned in the previous section that they bear non-referential third person singular ‘dummy’ inflection.
Lexical impersonal verbs include meteorological verbs and others that form part of a larger class of verbs of natural events. Among these we find a subclass of verbs in which a natural force is expressed as a manner-of-action morpheme on the verb and not as a lexical external NP. A closer look at the latter helps to clarify the status of impersonal verbs as presentational thetic constructions. After recapitulating the salient characteristics of lexical impersonal verbs, I conclude that many verbs of the II class are in fact ambivalent in that they may have both a personal reading with either predicate or argument focus and an impersonal reading with sentence (thetic) focus (Lambrecht 1994; Lambrecht 2000).

3.1 Verbs of natural events

Since they entail no argument, meteorological verbs form part of a restricted class of avalent verbs that have no logical subject and assign no semantic roles. In many languages, avalent verbs surface with dummy pronouns and/or a dummy person inflection, usually third person pronominal forms. There is therefore a mismatch between their zero-argument conceptual structure and their surface realization.

Meteorological verbs in Innu include far more than the habitual list of weather verbs. Some are simple stems (i.e. a root and an abstract ‘final’) as in (12) and (13), others are more complex involving manner of action ‘final’ morphemes as in (14) through (16). They include verbs depicting the time of day and the seasons and form a large class of verbs of natural events as in (17) through (19).

(12) cimwan-Ø
to_rain.ii- 3‘it rains’

(13) lûtin-Ø
be_windy.ii- 3‘there is wind, it is windy’

(14) mamên-ûtan-Ø
on_and_off- to_rain.ii- 3‘there are scattered showers’

(15) cînikwân- pîputwê- ñtan-Ø
round- blurred- storm_force.ii- 3‘there are large spinning gusts of blizzard’

(16) nâmûn-âtân-Ø
downwind- wind_force.ii- 3‘the wind is favourable’
Formally they belong to the II class of verbs: (a) their stem shape is that of II verbs; (b) they bear third person singular inflection and may inflect for categories such as tense, mood and obviation as in (20). However, they may not inflect for plurality and, unlike other II verbs, they can never appear with a lexical NP subject (or pronoun).

(20) lütin- li- pan mêšikâ- t
    be_windy.II- OBV- PST CC.arrive.AI- 3CJ
    ‘it was windy when he arrived’

Notice that the concepts of ‘wind’, ‘rain’ and ‘storm’ are lexicalized either as full verbs as in (12) and (13) or as manner-of-action finals as in (14) to (16) but never as nouns.

There exist a number of manner-of-action finals of the type found in (14) to (16), that express the action of natural forces such as ‘the tide’, ‘the current’, ‘the wind’, ‘the frost’, or ‘the heat’. These natural forces are not lexicalized as nouns, but are only expressed as verbs or verb parts. The resulting verb sometimes depicts the action of the natural force on a Patient, in which case the verb inflects for the Patient argument as in (21) where it is Inanimate and (22) where the S is Animate. The salient point is that the same type of verb structure also depicts the action of the natural force itself as in (15) and (16) where there is no S. As a result, several impersonal verbs express predications about natural forces that cannot appear as external nouns because they are not lexicalized as such.

(21) wêp- âštan- Ø
    away- wind_force.II- 3
    ‘It (Inanimate) is blown away by the wind’

(22) wêp- âšu- w
    away- wind_force.AI- 3
    ‘he/it (Animate) is blown away by the wind’

Verbs of natural events, including those describing natural forces, beg the question of lexicalization patterns and the exploration of meaning-to-surface relations (see Talmy 1985). Languages have a fixed and quite restricted number of linguistic categories
(N, V, ADJ, ADV, morphemes, etc.) which map onto a wide range of ontological categories with variable results depending on the language. In Innu, natural forces are expressed through morphemes on the verb stem. Recall also that inanimates cannot surface as subjects of transitive predications by virtue of the hierarchical restriction that bans inanimate agents from acting as A. The interaction of these constraints results in a large number of impersonal verbs with no possibility of an overt lexical subject NP. The next section further investigates this topic in connection with topographic verbs.

3.2 Topographic verbs

The components of the landscape mass are seldom conveyed by a noun in Innu but rather through verbs denoting a combination of properties, such as shape, expressed in the root, and substance, expressed in the medial followed by an abstract final, in this case -â- ‘extension in space’. As in the case of verbs of natural events, the concepts of ‘bay’ or ‘long point of land’ for instance, do not exist as nouns in the language. Consequently, the statements in (23) and (24) are impersonal verbs that cannot have a lexical NP subject. Here, the verb merely asserts the existence of a component of the landscape that fits its description.

(23) wâš- á- w
    bay_shape- FIN.II- 3
    ‘it is a bay’

(24) macitêwê- y- á- w
    long_point- LK- FIN.II - 3
    ‘it is a long point of land’

We observe once again that impersonalization is closely linked to the unavailability of some nominals in the language.

Similarly, some medials that express landscape elements do not have a ‘noun’ equivalent. Recall that Innu possesses a specific slot, the medial position, for incorporating noun-like elements. The medials for ‘cliff’, ‘pebbles’, ‘hill’, ‘area of burnt forest’ do not have an autonomous noun form. As shown in the examples (25) to (28), the verbs that contain these medials are predications about entities that do not exist as lexical nouns, hence the impossibility of an external S. Such verbs are hence impersonal.

(25) âlim- išêk- á- w
    difficult- cliff- FIN.II - 3
    ‘it is a difficult cliff’

14. Abstract finals are difficult to gloss and their precise meaning is often obscure. They are simply labeled as FIN in the morphemic glossing.
Summing up, some of the concepts identified ontologically as THINGS are not lexicalized as nouns in the language but as verbal components: roots, medials or even finals, as in the case of natural forces. This situation gives rise to impersonal verbs that cannot have an external lexical S. On this topic, Comrie (1989a: 189) aptly observed that “entities of lower animacy are more readily perceived as an indeterminate mass”, a fact which decreases their likelihood of being lexicalized as nouns.

### 3.3 Impersonals in sentence-focus constructions

We have just seen that impersonality was linked to the absence in the language of a lexicalized nominal that could occupy the S function. In the next set of examples however, such a noun exists and impersonalization is not coerced by the absence of a lexical noun.

The first set of examples involve the manner-of-action final that expresses the 'force of waves' (29), as opposed to a similar predication about the 'wave' lexicalized as a noun kaškan in (32). Both examples display the same stative root ‘big’.15

In (29), the verb makâyn is a complex Inanimate Intransitive verb formed with a root mak- ‘big’ and a manner-of-action final complex -âyn- denoting an event that involves the waves. This verb is ungrammatical as a predication about an inanimate noun (30). It is worth noting that impersonality is not a property of the final -âyn- itself since, as shown in (31), it may occur in monovalent personal verbs. By contrast, example (32) is a predication about the waves, kaškan, which is an Animate noun. The verb mišištiwat ‘they are big’ exhibits an Animate Intransitive stem form and Animate plural inflection since the verb agrees with its Animate subject. This set of examples allows us to contrast two different types of verb. Personal verbs, as in (31) and (32), are predications about a nominal S which triggers stem and verb agreement. By contrast, an impersonal verb, as in (29), is not a predication about an entity, but a general

---

15. There are several roots for 'big'. The root -mak- appears in (29), whereas (32) uses the root miš-. 
depiction of a situation, an event-reporting statement. It cannot appear with an external lexical NP and if it did, it would yield a personal verb as in (32).

(29)  \[mak-\text{\`ayn-} \emptyset \text{big-} \text{wave\_force.}11- \ 3\]
      ‘there are big waves’ (i.e. ‘there is big wave action’)

(30)  \ *[mak-\text{\`ayn-} \emptyset \text{\`a}kaykan \text{big-} \text{wave\_force.}11- \ 3 \text{lake} \]
      ‘the lake has big waves’

(31)  \[pikw-\text{\`ayn-} \emptyset \text{nit-} \text{\`a}sh \text{broken-} \text{wave\_force.}11- \ 3 \text{1-} \text{canoe} \]
      ‘my canoe has been damaged by the waves’

(32)  \[mi\text{\`i}shi\text{\`i}-\text{w-} \text{at k\`askan-} \text{at} \text{big.}11- \ 3 \text{pl wave-} \text{pl} \]
      ‘the waves are big’

Examples (29) and (32) correspond to different construals of the situation. I conclude that (29) and (32) are allosentences in Lambrecht’s (1994:35) terms: “semantically equivalent but formally and pragmatically divergent surface manifestations of given propositions”. In (32) the verb is personal, its ‘wave’ argument is topical, occupies the S function and, in terms of information structure, the statement is of the topic-comment type. By contrast, the impersonal construction in (29) has the ‘wave’ argument lexicalized as a verb part and not as S. The statement is purely event-centered and the ‘waves’ are not topical, which corresponds to a thetic statement, (sentence-focus also called all-focus statement). Lambrecht (1994:140) points out that such sentences “do not predicate a property of some entity but they simply assert or “pose” (hence “thetic”) a fact or state of affairs”. Topicality of the subject (or highest-ranking argument) is the defining criterion for the unmarked topic-comment (or predicate-focus) structure. Sentence-focus constructions by contrast are characterized by the absence of predicate-focus structure (Lambrecht 1994:235), as reflected by the lack of an external lexical NP as highest ranking argument. Here the focus domain is the entire verb/sentence.

The second set of examples involves verbs that incorporate an inanimate S denoting a mass entity (such as ice, snow, mud, water, sand, etc.) in the medial position. Sentence (33) exhibits a simple stem: a root \text{pakune-} ‘have a hole’ followed by abstract final -\text{\`a}- ‘extension in space’. In (34) the same stem incorporates the medial -\text{\`a}k-, a classifier for \text{dry wood}. In both sentences, the external NP is omitted if the referent is recoverable from context and the pronominal index on the verb cross-references the S.

(33)  \[pakun\text{\`e-} \text{\`a} \text{w (n\`e mici\text{\`i}sw\`akan)} \text{have\_a\_hole-} \text{LK-} \text{FIN.}11- \text{3 DEM table} \]
      ‘it/(the table) has a hole in it, there is a hole in it’
Of interest to us here are the cases in which the incorporation of a nominal entity in medial position results in an impersonal verb, as in the following two examples that involve the presence of the medials for ‘snow’ (35)\textsuperscript{16} and ‘ice’ (36). Both verb forms involve an Inanimate Intransitive stem. The same verb forms would yield ungrammatical sentences if the lexical subject NP (the noun for ‘snow’ and the noun for ‘ice’) were present in the sentence as external lexical nouns as in (37) and (38). Moreover, the incorporated verbs in (35)–(36) may not occur with any other lexical NP in subject position. As a result, the latter are syntactically avalent and impersonal.

\begin{align*}
(35) & \text{milamw}- \text{âkunak}- \text{â}- w \\
& \text{damp- snow- FIN.II- 3} \\
& \text{‘it is an expanse of damp snow’}
\end{align*}

\begin{align*}
(36) & \text{papakunê- šikw- â- w} \\
& \text{have_a_hole- ice- FIN.II- 3} \\
& \text{‘it is an expanse of ice with a hole in it’}
\end{align*}

\begin{align*}
(37) & \text{milamwàkunakàw *kûn} \\
& \text{*‘the snow is damp snow’}
\end{align*}

\begin{align*}
(38) & \text{papakunëšikwàw *miškumî} \\
& \text{*‘the ice is an ice with a hole’}
\end{align*}

Considering that the nouns for ‘snow’ kûn and ‘ice’ miskumî are both of Animate gender, the ungrammaticality of (37) and (38) is due to the mismatch between the gender of the N and the gender of the verb stem. Adding the external NP yields perfectly grammatical sentences in (39) and (40), in which the verb stem corresponds to verbs that take an Animate subject.

\begin{align*}
(39) & \text{milamw- âkunac- iši- w (kûn)} \\
& \text{damp- snow- FIN.AI- 3 (snow)} \\
& \text{‘the snow is damp snow’}
\end{align*}

\begin{align*}
(40) & \text{pakunê- šiku- ū- w (miskumî)} \\
& \text{have_a_hole- ice- FIN.AI- 3 (ice)} \\
& \text{‘a piece of ice has holes in it’}
\end{align*}

These examples warrant the conclusion that incorporating a nominal into a verb does not automatically yield an impersonal verb and that the sentences with impersonal verb forms in (35) and (36) are not equivalent to those in (39) and (40). The latter are

\textsuperscript{16} Notice that the verb ‘to snow’ mišpun is unrelated to the noun ‘snow’ kûn. The medial for snow is -âkun(ak)-.
referential; they are predications about individuated things: a piece or a cube of ice, a
patch of snow on one’s coat or under one’s boots, etc. By contrast, the impersonal verb
forms in (35) and (36) are depictions of areas in which a certain substance (‘snow’, ‘ice’)
is found. Both sets of sentences reflect different construals of the situation. On the one
hand, the snow or the ice is individuated and topical, as in (39) and (40), and the sen-
tence is of the topic-comment type with predicate focus. On the other hand, the ‘entity’
incorporated is construed as mass and is non-topical, the result being a presentational
thetic construction as in (35) and (36).17

Noun incorporation is not per se an impersonalizing operation. However, the
overwhelming number of verbs with an incorporated inanimate non-count noun (ice,
snow, mud, water, sand, etc.) yield impersonal verb forms corresponding to prese-
tational (sentence-focus) statements. The cases put forth in the previous sections, in
which impersonalization was coerced by the lack of an available nominal, can also be
described in information structure terms as event-centered sentence focus construc-
tions (Lambrecht 2000).

3.4 Pragmatic impersonals and the ambivalence of Inanimate Intransitives

We have seen in the previous section that there exist potential allo sentences:
– a statement of the topic-comment type with an Animate Intransitive verb and an
Animate S
– an alternative ‘impersonal’ sentence-focus statement in Inanimate Intransitive
form without any possibility of a lexical S

In several cases however, there is no formal variation between the ‘personal’ and the
‘impersonal’ reading of the verb. Many II verb forms are simply ambivalent: they may
take a personal (topic-comment) interpretation or an impersonal purely event-report-
ing interpretation. On the personal reading, an external NP may appear as S, but not
on the impersonal reading. In any case, the form of the verb remains unchanged.

For example, a verb like takâw ‘it is cold’ may be a predication about a specific
entity as in ‘the milk is cold’ or be an impersonal statement about the climate. Likewise,
evaluative verbs like mwestâtelitâkwan ‘it is boring’ can be a predication about a spe-
cific thing (movie, book, song, etc.) or just an impersonal comment on a general situa-
tion. To illustrate this point further, example (41) gives an idea of the range of uses for

17. It is noticeable that the ‘impersonal’ interpretation in (35) and (36) is not attributable to
the presence of the abstract final -â- since the same final occurs in many instances of referen-
tial verbs as in (33) and (34). There are no finals that are specific to impersonal verbs. Indeed,
impersonal verbs are formally indistinguishable in form from other Intransitive verbs with an
Inanimate subject.
verbs with the manner-of-action final ‘by heat’. In a topic-comment frame, the S may be (a) the Source, (b) the Undergoer, or a cognate subject as in (c). The same verb may also denote an impersonal event-reporting thetic statement as in (d). On the impersonal reading the verb is a thetic statement. On the personal reading, it has predicate focus. Examples such as (41) are pragmatic impersonals in the sense that, with no change in the verb form, the sentence may be interpreted with or without a subject.

(41) milu- katê- w
    good- heat_force.Π- 3
    ‘it heats/cooks well’
    a. source as S: ‘the stove/the wood provides good heat’
    b. undergoer as S: ‘the food cooks well’
    c. cognate S: ‘the fire heats well’
    d. zero-argument: ‘there is good heating’

Some Intransitive Inanimate verbs are polysemic: they have one meaning when ‘personal’ (in topic-comment type statements), but take a distinct meaning in impersonal thetic statements as in the following examples.

(42) cikâ- nákwan- Ø
    clear- appear.Π- 3
    a. predicational: ‘x appears clearly’
    b. meteorological impersonal: ‘there is clear weather’

(43) cikâ- y- ápišk- à- w
    cleft- LK- mineral- FIN.Π- 3
    a. predicational: ‘x (mineral) is cleft’
    b. place impersonal: ‘there is a cleft, a passage between two rocky peaks’

(44) usê- y- à- w
    ridge_shape- LK- FIN.Π- 3
    a. predicational: ‘x (bound object) is ridge shaped’
    b. place impersonal: ‘x (unbounded) has a long ridge, crest’
    c. place impersonal: ‘it is the watershed’

The following sentences further exemplify the ambivalence of Π verbs as the same verb surfaces in two statements that convey different information. In (45) the muskeg (S) is the topic of the statement with predicate focus. In example (46), the muskeg is non-topical and incorporated into the verb in a presentational sentence-focus statement. Having no topical S, the verb is impersonal.

(45) álim- assêk- à- w né massêkw
    difficult- muskeg- FIN.Π- 3 DEM muskeg
    ‘this muskeg is a difficult muskeg’
The potential for impersonal readings applies only to verbs that have a semantically inanimate entity as their highest-ranking argument. A verb with an animate S cannot take an impersonal reading through incorporation of the animate S. The question therefore arises whether all Intransitive Inanimate verb forms are ambivalent? Are there II verbs that do not allow an impersonal interpretation, and if so how should they be characterized? Indeed, some verbs are not conducive to an impersonal interpretation. A verb like *atitêw* ‘it changes color’ depicts berries that, as they ripen, change their color: the verb’s selectional restrictions are such that it requires a ‘fruit’ argument. Likewise the verb *âytimêmuw* ‘it is installed on the wrong side’ is clearly a predication about a specific ‘installed’ entity. Thus, some verbs are less likely to get an impersonal reading, but overall the potentiality is inherent to the class of verbs that describe inanimate entities.

3.5 A recapitulation of the formal and semantic properties of lexical impersonal verbs

From a formal point of view, lexical impersonal constructions, as discussed so far, are centered on verbs, simple or complex, that correspond to Inanimate Intransitive stems inflecting for a third person singular dummy S and never occur with an external lexical NP subject. Some stative verbal constructions in Innu incorporate their highest-ranking argument. If that argument is non topical, the verb is impersonal and corresponds to a presentational statement. Thus impersonal verbs form a covert category often indistinguishable from personal verbs, excepting the fact that they cannot appear with an external lexical NP as S.

Semantically, lexical impersonal constructions include (a) meteorological verbs in the broad sense, including verbs of natural forces; (b) verbs that describe elements of the landscape; (c) verbs that describe entities construed as mass, non-individuated, non-count elements. It is important to underscore that all the impersonal forms examined in this section relate to ‘inanimate’ arguments at the lowest end of the animacy hierarchy. The animacy hierarchy is also often described as a topicality hierarchy (Siewierska 2004): a ranking of entities as per their capacity to be topics. Our data show that the lower the element on the animacy hierarchy, the less likely it is to enter into topic-comment construction and even to be lexicalized as a noun. If an argument is not lexicalized as a noun, it cannot be a suitable topic and enter in topic-comment constructions, in which case, the statement is ‘event-centered’ and has sentence-focus. Entities located at the far end of the animacy hierarchy, i.e. unbounded, non-individuated and abstract, are therefore more conducive to impersonal sentence-focus statements.
4. Derived impersonals: Suppressing an animate argument

We have just seen that verbs that encode a semantically inanimate S are mostly ambivalent since they may have both a 'personal' and an impersonal reading. There is therefore no need to employ special morphology to ‘impersonalize’ them. By contrast, in verbs that index a semantically animate core argument, that argument is always understood as specific, (definite or indefinite), and its suppression requires specific derivation. This section will examine the operations involved in the suppression of an Animate S or A.

There are two main ways to suppress the Animate Actor in Innu depending on the transitivity of the clause. Suppressing an Animate Actor in intransitive clauses yields an impersonal verb. Suppressing the Animate Actor of a transitive clause yields a personal passive (as in Siewierska 1984). Both types of derivations are fully productive and I analyze them both as types of passive. The metaphor of suppression expresses the shift from a basic voice construction (Klaiman 1988) where the A/S is fully specified and topical to an alternate construction, the passive, where it is non-topical and left unspecified. Both types of passive (the personal passives of transitives and the impersonal passives of intransitives) decrease the valence of the verb and in both cases the unexpressed argument remains implicit in the conceptual frame. In passives of transitives, the implicit A is understood as human or animal while in impersonal passives the implicit S is always understood as human. Moreover, in each type of passive the implicit argument may include the speaker. Indeed, speakers often draw on unspecified actor constructions to leave a logical we subject non topical while focusing exclusively on the event under consideration.

The major part of this section is devoted to constructions in which an animate S is suppressed yielding an impersonal verb (4.1). Transitive passives (4.2) are presented also mainly because the distribution of the passive suffixes overlaps, highlighting the connection between the two operations. From a morphosyntactic viewpoint, only the suppression of an animate S results in an impersonal construction since, as will become clear below, transitive passives are promotional and 'personal'.

4.1 Suppressing an animate S: Derived impersonal verbs

It is possible to suppress the Animate S of an intransitive clause by adding the portmanteau suffix -nâniwi- to a morphologically Intransitive Animate verb stem. This

---

18. For more details on passives in Innu, see Drapeau (Forthcoming).

19. The sequence /iwi/ is contracted to /u:/ in some Innu dialects, and the /na:/ is deleted after verb stems ending with /e:/ or /a:/ . These dialectological idiosyncrasies have no bearing on the argument.
suffix which denotes “loosely collective human actors”, will be glossed as the ‘passive of intransitive’ suffix (PI).\textsuperscript{20} The resulting verb has a syntactically inert and implicit unspecified argument that never finds syntactic realization as a lexical NP.

The following example (47) narrates a sequence of events in impersonal form and the PI suffix is underlined. Notice that the speaker is semantically included in the action in this narration. The translation of the whole Innu verb is underlined in the English gloss.

\begin{enumerate}
\item \textit{êkwan mâ nê šâš êyêškupinânût tshê nâšipânût nûtim kâci aštacikânânût, êkwan cê pûšinânût.}
\item \textit{‘And then \textbf{there is being ready} to going back to the coast after \textbf{there is hiding the goods on platforms}. That is when \textbf{there is embarking on a journey’}.}
\end{enumerate}

In morphological terms, the derived impersonal verbs with the -nâniwi- (PI) suffix in (47) belong, by default, to the class of Intransitive Inanimate verbs. Evidence for this claim is adduced by the distribution of proximate and obviative marking between clauses. An Animate proximate argument triggers obviative marking on a non-coreferential dependent clause verb, but Inanimates do not trigger obviation on other Inanimates. In (48) the impersonal verb in the main clause nâšipânût does not trigger obviative marking on the dependent clause verb meluškamiti (the ungrammatical form is in the obviative).

\begin{enumerate}
\item \textit{wîpat mëlûškami- t-i êkêwê nâšipâ- nû- t}
\item \textit{‘as soon as it is spring, there is going back to the coast’}
\end{enumerate}

The derived impersonal verbs discussed in this section thus exhibit the properties of Inanimate Intransitives. Moreover, although the impersonal verb may inflect for obviation, mood and tense, it cannot be pluralized and its inflectional third person index is non-referential.

Impersonal verbs are reported crosslinguistically mainly in connection with activity verbs, hence the conception that impersonalization crucially involves ‘actor’ defocusing. It is worth noticing that in Innu the PI suffix is also compatible with stative verbs (as in example (47) ‘\textit{there is being ready}’) and may be added to all manner of AI stems, irrespective of the lexical-aspectual (\textit{aktionsart}) class of the verb.

These impersonal constructions are analyzed as ‘passive’ rather than active impersonals for the following reasons. In the active voice, Animate arguments are always referential and indexed for person and number on the verb as explained in §2.4. In derived impersonal forms, not only does the index on the verb not ‘refer’, the verb is

\textsuperscript{20} The meaning of this construction is comparable to the Mann impersonals in German and similar constructions in other European languages.
shifted to a different class, the II class, since, as was shown in (48), the verb is treated as if it had an Inanimate S. Treating semantic animates as morphosyntactic Inanimates is a major departure from the basic parameters of the active voice. The distribution of the ‘passive’ suffixes offers another reason for treating the impersonals of intransitives as ‘passives’. We shall see in §4.2 that the distribution of the two suffixes that derive ‘passive’ forms, the PI above and the PT presented in the next section, is not governed by purely syntactic (transitive vs. intransitive) factors but is sensitive to the morphological class of the verb. The two suffixes share the common function of signaling that the argument occupying the highest function is left unspecified, an operation which, in transitive clauses at least, is generally recognized as passivization. A detailed explanation of the overlap in the distribution of the two passive suffixes is presented in the coming section.

4.2 Suppressing an animate A: Personal passives

Keeping in mind that in the language a basic transitive verb always indexes a referential Animate Agent, suppressing such an Agent yields a passive in which the topical Patient gets promoted to S as the sole indexed argument. Passives of transitives are thus ‘personal’ in the sense that the verb inflects for the morphosyntactic properties of the derived S. They are also agentless because it is impossible for the Agent to appear as an NP in the sentence.

Passive forms are derived from both Transitive Animate object (TA) stems or from Transitive Inanimate object (TI) stems depending on the gender of the object (ultimately, the derived S). The suffix -ikaw- is added to a TA stem when the Patient is first and second person and -âkani- when it is third person. The latter suffix is also used to form passives of TI stems. The -âkani- passive suffix is glossed as the passive of transitive (PT) suffix.

Authors such as Keenan (Keenan 1985; Siewierska 1984) have noted that passives of transitive and passives of intransitives often share common morphology. Likewise in Innu, both suffixes (the PI presented in the previous section and the PT discussed here) overlap in their distribution, a fact which is congruent with their sharing a common function and warrants treating them both as ‘passive suffixes’ as the rest of the section will show.

Some syntactically intransitive verbs use the PT suffix normally used to form passives of transitive clauses. There is indeed a residual class of intransitive verbs that do

---

21. This form combines two morphemes, the -âkan which forms deverbal patientive nominals and the abstract final -i- which turns it into an Animate Intransitive verb stem. The PT suffix has a -kan allomorph in passives of TI verbs. Note that the Western dialects of Cree have different forms for passives of transitives (see Dahlstrom 1991; Wolfart 1991).
not form their passive with the expected PI suffix. They belong to a group of objectless Transitive Inanimate verbs (Wolfart 1973). These morphologically transitive, but syntactically intransitive verbs form their passive with the PT suffix in Innu, that is, with the suffix normally used in deriving passives of transitive verbs. The resulting verb is an impersonal passive with a non-referential third person inflectional index and no possibility of a lexical NP as in (49).

(49) mêluškami- t- i mîšt-wîpat mûšwâway- kani- w
    cc.be_spring.II- 3cj- sub very-early go_to_sea.TI- PT- 3
    'in the spring very early there is going out on the open sea'

Likewise, there is a subclass of ambitransitive verbs in the Animate Intransitive class, illustrated in (50), that form their passive with the PI suffix, even when syntactically transitive. As the non-topical A is left unspecified, the P is promoted as sole argument (as in the case of transitive personal passives) and the derived verb is intransitive.

(50) ašâm- at êssimâ- nû- tâw- i
    snowshoe- PL cc.weave.AI- PI- 3PL.CJ- sub
    'when the snowshoes were woven'

In sum, some personal passive forms exhibit the PI suffix normally found in passives of intransitives. Likewise, some derived impersonal verbs employ the personal passive PT suffix. Such misalignment between the morphology of the verb stem and its syntax shows that the choice between the two suffixes does not depend strictly on syntactic criteria, but is sensitive to the morphological class of the base stem. The shared distribution of the passive suffixes highlights the connection between the operation of suppressing the A of transitives and the S of intransitives.22

4.3 Conclusion on impersonal passives

In the following, I will both recapitulate the derivation of passive forms in Innu and situate it with respect to the normal parameters of argument encoding in active clauses. We have seen in §2.2, in connection with hierarchical alignment, that Innu possesses a morphological device for encoding on transitive verbs with third person arguments the relative topicality of the A and the P. In direct clauses, the Agent is topical, whereas in inverse clauses, it is the Patient that is topical. In both cases, the verb is transitive and active. In active clauses, direct and inverse alike, the Agent can be realized syntactically as an external lexical NP (but does not have to be overt if it is contextually recoverable). As a consequence, merely reversing the topicality status of the (third person) A

22. Both suffixes presented in this section also appear as nominal suffixes after the removal of the verbal finals. This is discussed in greater detail in Drapeau (Forthcoming).
or the (third person) P argument does not yield a passive. Passives arise when the A or the S of an active clause is left unspecified, an operation which requires that a new verb stem be derived by adding a passive suffix, either the PI or the PT. In such cases, the result yields an agentless construction and the derived verb is intransitive. In case the Agent of a transitive verb is left unspecified, the Patient gets promoted to highest-ranked argument indexed on the verb (the derived S). When the Actor of an intransitive clause is left unspecified, the derived verb is *impersonal* bearing a non-referential third person index.

In morphosyntactic terms, impersonals in Innu are a by-product of leaving the S unspecified. Derived impersonals show formal similarities with the lexical impersonals seen in §3. From the point of view of form, all impersonal verbs, be they lexical or derived, form Inanimate Intransitive stems that bear a dummy third person singular inflection. We also saw that lexical impersonal verbs are either avalent, as in the case of meteorological verbs, or they are understood as having an abstract non-lexicalized argument. As a result, the category of *impersonals* covers the constructions where the S is left unspecified by virtue of being both non-topical and out of focus or, by default, of not being expressed as a nominal.

5. **Suppressing an Inanimate Agent**

Innu displays a residual class of verbs which appear with transitive morphology despite being syntactically monovalent, their Inanimate Agent being left unspecified. I use the label of trans impersonals by analogy to a type identified by Malchukov (2008:77) as “constructions in which the verb is transitive and takes an experiencer as its object” which have, over time, become intransitive with an experiencer S. Innu trans impersonals are formally built as Inanimate Agent forms, but no Inanimate Agent can be realized syntactically. Such trans impersonals are therefore intransitive and have an experiencer S as sole possible argument. As was the case for the constructions examined by Malchukov, Innu trans impersonals do not correspond to a productive operation in the language, they are mostly lexicalized and restricted to verbs of bodily experience.

The following will describe the canonical transitive construction of Inanimate agency (so-called *Inanimate Actor forms*) in Innu in order to ascertain how trans impersonals depart from this normal construction. The canonical transitive construction of Inanimate agency is one which the verb (a) is built on a TA stem, (b) has two core arguments: an Animate Patient and an Inanimate Agent, (c) indexes the Patient only by person and number inflection. The Inanimate Agent, although not indexed on the verb as such, may appear as an external NP. Examples (51) and (52) illustrate
Impersonals in Innu

This type of construction. The stem is derived from a TA stem, followed by the inverse suffix -iku- (see also Wolfart 1973). Whereas normal TA verbs index both the Animate A and the Animate P (i.e. are both syntactically and morphologically Transitive), transitive verbs of Inanimate Agency are morphologically monovalent but syntactically bivalent but only index the P.

(51) \[ni- [cîskwêskâ]- ku- ni- micim\]
    \[1- cause_dizziness.TA- \] \[INV- 1- food\]
    'I am made dizzy by my food'

(52) \[Ø- [cîskwêskâ]- ku- w ni- kâwi nê- lu micimi- liw\]
    \[3- cause_dizziness.TA- \] \[INV- 3- mother DEM- OBV food- OBV\]
    'my mother is made dizzy by that food'

Transimpersonals are formally indistinguishable from normal transitive constructions of Inanimate Agency, except for the fact that they are syntactically monovalent. No Inanimate Agent may be expressed overtly, either because (a) there is no identifiable Agent, as in the case of bodily sensations and experiences, or (b) there is no nominal available in the language to appear in the NP slot. The result is a syntactically intransitive verb with an experiencer S with an unfilled slot for the Inanimate Agent NP as illustrated in (53) to (56). Presumably, the choice of a Transitive stem stresses the lack of control of the experiencer over an unpleasant situation that is forced onto him.

(53) \[n- ucipit- iku-n\]
    \[1- pull.TA- \] \[INV- 1\]
    'I have an epileptic seizure' (unspecified object pulls at me)

(54) \[ci- cîskwêkustim- iku-n\]
    \[2- make_dizziness_by_sleep.TA- \] \[INV- 2\]
    'you are sleep-dizzy' (unspecified object makes you sleep-dizzy)

(55) \[ni- pûtât- iku- n\]
    \[1- blow.TA- \] \[INV- 1\]
    'I get a steam-burn' (unspecified object blows at me)

(56) \[alây- ku- w\]
    \[overtake.TA- \] \[INV- 3\]
    'he gets ahead of himself' (unspecified object gets ahead of him)

The unavailability of an existing nominal to name the inanimate causer is the common denominator. Such transimpersonals share with the derived impersonals discussed in §4 the property that the Agent (or argument otherwise occupying the higher semantic role) is left unspecified. In Section 4, the unspecified argument
was Animate, in the case of transimpersonals it is Inanimate. Since they have an expressed experiencer subject these “transimpersonals” are not impersonals in the subject-centered view. They share with impersonals, as well as with personal passives, the property of having an unspecified Agent. As mentioned above, these forms are lexicalized and do not stem from a productive operation.

6. Concluding remarks

In morphosyntactic terms, Innu impersonals are constructions in which the S argument is left unexpressed and the verb bears a default non-referential third person singular index added to an Inanimate Intransitive verb stem. The paper has shown the impact of the animacy hierarchy on the phenomenon of impersonals in Innu. The examination of lexical impersonals has revealed that there is a cut-off point on the animacy hierarchy beyond which inanimates are not encoded as arguments in the subject function and are not even lexicalized as nouns. The result is a large class of lexical impersonal verbs expressing events originating from natural forces or depictions of the landscape. As for the suppression of Animate ‘subjects’, it was shown to require special derivation on the verb and to yield, in intransitive clauses, an impersonal verb, i.e. a derived intransitive stem with a dummy Inanimate subject. I have argued that these intransitive impersonals share the properties of passives.

It is also worth noting that, as in the case of passives of transitive clauses, impersonalization requires the total suppression of the downgraded argument, hence the impossibility of its finding syntactic realization as a lexical NP in the clause. Moreover, alternations between constructions with and without an external NP ‘subject’ has shown impersonal constructions to be thetic with respect to their focus structure type.

The paper has underscored the fact that the investigation of impersonals ought to focus not only on the nature of the predicates involved, but also, and perhaps more importantly, on the nature of the entities under predication. It also draws attention to the complex interaction between lexicalization patterns, language-specific constraints on the expression of core arguments, and information structure.

23. Plains Cree TA third person passive forms share the characteristics of transimpersonals: they exhibit regular Transitive active morphology without additional derivation, cannot take an overt Agent and index only the Patient (see Dahlstrom 1991). It is not altogether clear which pattern is more archaic, the Plains Cree pattern (wápmáw ‘he is seen’) or the -ákaní- pattern found in the dialects east of Plains Cree including Innu (wápmákaníw ‘he is seen’).
Abbreviations

The following abbreviations are used in the morphemic glossing:

- **AI**: Animate Intransitive verb
- **CC**: changed conjunct
- **CI**: conjunct order
- **DEM**: demonstrative
- **DIR**: direct marker
- **FIN**: final marker
- **II**: Inanimate Intransitive verb
- **INV**: inverse marker
- **LK**: phonetic linker
- **OBV**: obvation
- **PI**: passive of intransitive
- **PL**: plural
- **PST**: past
- **PT**: passive of transitive
- **SUB**: subjunctive
- **TA**: Transitive Animate verb
- **TI**: Transitive Inanimate verb
- **TTIT**: Transitive Inanimate Theme

Person categories on verbs and nouns are encoded as:

1  first person  
2  second person  
1PE first person plural excluding the addressee  
1PI first person plural including the addressee
2PL second person plural  
3  proximate third person  
3PL third person plural  
3’  obviative third person

I use capitals for grammatical properties and low case for semantic categories. Thus Animate refers to animate gender and animates refer to the semantic class of animates.

References


A diachronic study of the impersonal passive in Ainu*

Anna Bugaeva
Waseda Institute for Advanced Study

This paper provides a diachronic account for the impersonal passive in Ainu which is marked with $a(n)=$ in Saru (Southwestern Hokkaido) and with $a(n)=$ for 3SG/PL.O and $ci=$ for 1SG/PL.O in Ishikari (Northeastern Hokkaido). I have argued that in spite of clearly impersonal origins, neither the $a(n)=$ -marked nor the $ci=$ -marked construction can be regarded as impersonal because of the possibility of the inclusion of an oblique actor phrase and that none of the constructions can be regarded as personal passive because reanalysis of O to S is incomplete: only some topic-related subject properties (word order etc.) have been acquired by O but no single agent-related subject property has been acquired yet. In the context of the discussion of grammaticalization pathway IMPERSONAL > IMPERSONAL PASSIVE > PASSIVE, of special interest is the action nominalization origin of the $a(n)=$ -marked construction.

Keywords: Ainu; Saru; Ishikari; impersonal passive; grammaticalization; action nominalization

One great difference between the Ainu and Japanese languages is that while the Japanese tongue abhors passives the Ainu use them wherever possible and the Japanese only when forced to do so. (Batchelor 1938:2)

---

* The present research was supported by Global COE Programme “Corpus-based Linguistics and Language Education” of Tokyo Institute of Foreign Languages (SLE 41 workshop attendance), by La Trobe University, RCLT (SLE 41 workshop application stage), and by ELDP IPF0128 (SOAS, University of London). I would like to thank Andrej Malchukov, Anna Siewierska, Alexandra Aikhenvald, Emma Geniušienė, Hiroshi Nakagawa, Simon Overall, Birgit Hellwig and two anonymous referees for valuable comments on the earlier versions of this paper. None of them necessarily shares my views expressed here or is responsible for any errors.
1. **Introduction**

As is well-known nowadays, Ainu has no personal passive corresponding to that of the Japanese construction. It, does, however have an impersonal passive, which may, more or less, be regarded as the functional equivalent of the passive in other languages, and is, indeed, highly-frequent in texts. The impersonal passive in Ainu (Southern Hokkaido) is a subjectless construction with a retained direct object which is formed by the prefix *a* on transitive action predicates (1a). An actor may be overtly expressed and is then marked via an oblique phrase, viz. *or-o wa* 'from the place of' for animate actors (1b) or *ani* 'by' for inanimate actors (10), which is rather uncommon cross-linguistically (Siewierska 1984: 174).

1. neno *e=iki yak a=e=koyki na*
   
   like.this 2SG.S=do if INDE.A=2SG.O=scold SGST
   
   'If you do that, you will be scolded.'
   
   lit. ‘…someone will scold you.’ (T2 71)

2. *hapo or-o wa a=en=koyki*
   
   mother place-poss from INDE.A=1SG.O=scold
   
   'I was scolded by mother.'
   
   lit. ‘Someone scolded me by mother.’ (T2 72)

An impersonal passive is impersonal under both the subject-based (formal) and instigator-based (functional) views of impersonalization (for details see Siewierska (2008: 124)) since it lacks a subject and the agent has been demoted.

In Ainu studies, the construction in (1) has traditionally been regarded as active impersonal (Kindaichi 1931: 132–3; Chiri 1936: 63–4; Tamura 1988: 30) due to its active morphology. However as shown by Shibatani (1985, 1990), it is actually passive as evidenced by the possibility of expressing the demoted actor in (1b) which is typical of passives but impossible with active impersonals. However, there is no doubt that diachronically the impersonal passive1 in Ainu originated from an active impersonal: the impersonal passive (Southern Hokkaido) is marked with the indefinite transitive subject prefix *a*= (1), which may possibly be traced back to the indefinite intransitive subject suffix *=an* (2), the latter must have, in its turn, developed from the existential verb *an* 'exist', cf. the second clause in (2).

---

1. The impersonal passive analysis accepted in this paper was adduced by Satoo (1995)
A diachronic study of the impersonal passive in Ainu

A development leading from an indefinite subject (=impersonal) construction to an impersonal passive and ultimately to a personal passive appears to be a commonly attested grammaticalization path (Shibatani 1985; Malchukov 2008:96; Siewierska 2010), but indefinite subject marking originating in an existential verb has hardly been attested elsewhere.

The impersonal passive in Ainu displays a number of cross-linguistically unusual properties, and thus presents a challenge to a comprehensive typology of impersonal and passive constructions. I will offer a comparative analysis of the impersonal passive constructions in the relatively well-described Southern Hokkaido dialects (Southwestern group), viz. Saru (Tamura 1988/2000) and Chitose (Nakagawa 1995; Bugaeva 2004), see §3, and in a less described Central Hokkaido dialect (Northeastern group), viz. Ishikari (Asai 1969; Tamura 1970a), see §4, in order to provide a tentative grammaticalization scenario for these constructions and to discuss their typological implications, see §6 and §7. Since the impersonal passive marking originates in personal marking, a word on the Ainu person system as a whole is in order (§2.2. and §2.3). In §5, I will additionally scrutinize the coding and syntactic properties of the patients in the relevant constructions to find out whether there are any signs of the promotion of the O to subject position and of reanalysis of the impersonal passive into a personal passive. I will use such syntactic tests as equi-deletion and reflexive binding.2

The analysis is based on the data of colloquial Ainu only from Saru, Chitose, and Ishikari dialects (see list of Sources in the end: T1, T5, KJ for Saru; B for Chitose, and SK for Ishikari). Additionally, the data from specialist literature are also used (see T2 and T4 for Saru and T3 for Ishikari). The reason for not using numerous folklore texts in these dialects is that impersonal marking is materially identical with ‘logophoric’ marking (the person of the protagonist; see Tables 1 and 2) and its use here would have significantly complicated the study.

2. The coordination test which is commonly employed for testing subjecthood in other languages is inapplicable in Ainu because of the general tendency to omit topical arguments in both subject and object positions.
2. General information on Ainu

2.1 Genetic, dialectal, and typological profile of Ainu

Ainu is a critically endangered language of unknown genetic affiliation which shows considerable dialectal variation. The three primary divisions are geographically based, and distinguish between the dialects once spoken on Hokkaido, Sakhalin, and the Kurile Islands. Sakhalin and the Kuriles form part of the Russian Federation today, with Hokkaido being the last autochthonous location of native speakers. The Hokkaido dialects can be roughly divided into Northeastern (Northern, Eastern, and Central) and Southwestern (Southern and Southwestern) groups, which are further subdivided into local sub-dialectal forms (see Hattori 1964: 18).

Ainu is polysynthetic and agglutinating, with SV/AOV\(^3\) constituent order. It is predominantly head-marking and prefixing. Possession is the only category obligatorily marked on nouns; inalienable possession is expressed by the “izafet” construction. The difference between transitive and intransitive verbs is clear-cut. Some verbs employ different stems for singular and plural. In the case of intransitive verbs, plurality refers to the number of subject referents, while in the case of transitive verbs plurality refers to the number of object referents. Tenses proper are absent; there is only one tense-aspect marker – the perfect auxiliary \(a\). The verb has extensive aspectual, modal, and evidential systems which are expressed by suffixes, auxiliaries, and particles. There are many affixes to mark different types of actant relations.

2.2 Grammatical relations in Ainu

Arguments do not inflect for case in Ainu: A and O are distinguished by their relative position in clause structure and by verbal cross-referencing (=agreement) markers; obliques are marked by postpositions. Personal pronouns in the subject (A/S) and object (O) position are often omitted: Ainu is a so-called pro-drop language but verbal cross-referencing markers are obligatory.

The system of grammatical relations in Ainu may be characterized as basically tripartite since there is a distinct marking for S, A, and O in 1pl. verbal cross-referencing, viz. the prefix \(ci=\) marks A, the suffix \(as\) S, and \(un\) O for the first person exclusive, and \(a=\) A, \(an\) S, and \(i=\) O for indefinite person, see Table 1. Indefinite is used not only to refer to the indefinite speaker or addressee (see Tables 1 and 2), but also as the first person plural inclusive, second person singular/plural honorific, and

---

3. The abbreviation S stands for intransitive subject, A for transitive subject, and O for object of transitives (Dixon & Aikhenvald 1997:72). None of them indicates any particular semantic role but rather a group of roles that are typically expressed as subjects of intransitive/transitive predicates and to those that are typically expressed as objects.
logophoric (person of the protagonist). The latter is common in folktales because they have the structure of reported discourse with the whole story being a quote, for details see Bugaeva (2008). However, there are also some elements of a neutral system: no inflectional difference between the second and third person verbal cross-referencing, viz. A, S, and O are marked by e= in 2sg and by eci= in 2pl, and the third person is always zero-marked; personal pronouns are the same for A, S, and O. There is also one feature of the nominative/accusative system: A=S, and O are distinct in the 1sg verbal cross-referencing, viz. 1sg ku= marks A, S, and en= O, see Table 1. Although most forms are cognate, person marking varies considerably depending on a group of dialects and on a particular dialect, cf. Table 1 showing person marking in the Saru dialect of Ainu (Southern Hokkaido, Southwestern group) and Table 2 showing person marking in the Ishikari dialect of Ainu (Central Hokkaido, Northeastern group).

In verbal cross-referencing of transitive verbs, A affixes are, in most cases, placed before O markers (see Tables 3, 4). However, when it comes to the interaction of first and second person participants, person marking on verbs is not always analyzable as ARGa-ARGo-VERB and it is, indeed subject to considerable dialectal variation. For instance, in Southwestern dialects of Ainu, the combination of first person singular/plural subject (A) and the second person singular/plural object (O) is never marked as *ku=e=, *ku=eci=, *ci=e=, and *ci=eci=, as would be expected from Table 1, but is invariably marked with the portmanteau prefix eci= (1sg/pl.A+2sg/pl.O) (see Table 3) which has a key function of the second person plural marker for A, S, or O (cf. Table 1). The combinations of the second person singular subject and first person singular/plural object are marked with e=en= (2sg.A=1sg.o=) and e=un= (2sg.A=1pl.o=) in Northeastern dialects (see Table 4), just as expected from Table 2, but in Southwestern dialects they are marked with en= (2sg.A+1sg.o=) and un= (2sg.A+1pl.o=) (see Table 3) having the key functions of the first person singular and plural O markers respectively (cf. Table 1).

The A set of prefixes (Tables 1 and 2) is also employed in the inalienable possessive construction, where they are attached to the head noun to mark the person and number of the possessor. The head noun is additionally marked with the allomorphic possessive suffixes -V or -(V)hV which copy a root-final vowel once or twice with the epenthetic /h/ being inserted, as in ku=sapa-ha (1sg.A=) ‘my head’ or ci=setur-u/ci=setur-uhu (1pl.(exc).A=) ‘our backs’.

4. The allomorphic possessive suffix -ha/-hu/-ho/-he/-hi is employed for vowel-final roots and -a(ha)/-u(hu)/-o(ho)/-e(he)/-i(hi) for consonant-final roots.

5. The difference between “short” (ci=setur-u) and “long” (ci=setur-uhu) possessive forms is not clear.
Table 1. Person marking in the Saru dialect of Ainu (Southern Hokkaido, Southwestern group), adapted from Tamura (2000:49) with a modification

<table>
<thead>
<tr>
<th>Person-number</th>
<th>A/S/O pronouns</th>
<th>A markers</th>
<th>S markers</th>
<th>O markers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>kani ‘I’</td>
<td>ku=</td>
<td>ku=</td>
<td>en=</td>
</tr>
<tr>
<td>1PL.(EXC)</td>
<td>cóka ‘we (I and he/she/them)’</td>
<td>ci=</td>
<td>=as</td>
<td>un=</td>
</tr>
<tr>
<td>2SG</td>
<td>eani ‘you.SG’</td>
<td>e=</td>
<td>e=</td>
<td>e=</td>
</tr>
<tr>
<td>2PL</td>
<td>ecioká ‘you.PL’</td>
<td>eci=</td>
<td>eci=</td>
<td>eci=</td>
</tr>
<tr>
<td>3SG</td>
<td>sinuma ‘he/she’</td>
<td>Ø</td>
<td>Ø</td>
<td>Ø</td>
</tr>
<tr>
<td>3PL</td>
<td>oka ‘they’</td>
<td>Ø</td>
<td>Ø</td>
<td>Ø</td>
</tr>
<tr>
<td>INDF</td>
<td>aoka</td>
<td>a=</td>
<td>=an</td>
<td>i=</td>
</tr>
</tbody>
</table>

also has the functions of:
1. 1PL.INC ‘we (I and you)’
2. 2SG/PL honorific
3. logophoric

Table 2. Person marking in the Ishikari dialect of Ainu (Central Hokkaido, Northeastern group), adapted from Tamura (2000:49) with a modification

<table>
<thead>
<tr>
<th>Person-number</th>
<th>A/S/O pronouns</th>
<th>A markers</th>
<th>S markers</th>
<th>O markers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>kuani ‘I’</td>
<td>ku=</td>
<td>ku=</td>
<td>en=</td>
</tr>
<tr>
<td>1PL.(EXC)</td>
<td>ciokay ‘we (I and he/she/them)’</td>
<td>ci=</td>
<td>=as</td>
<td>un=</td>
</tr>
<tr>
<td>2SG</td>
<td>eani ‘you.SG’</td>
<td>e=</td>
<td>e=</td>
<td>e=</td>
</tr>
<tr>
<td>2PL</td>
<td>esokay ‘you.PL’</td>
<td>es=</td>
<td>es=</td>
<td>es=</td>
</tr>
<tr>
<td>3SG</td>
<td>anihi ‘he/she’</td>
<td>Ø</td>
<td>Ø</td>
<td>Ø</td>
</tr>
<tr>
<td>3PL</td>
<td>okay ‘they’</td>
<td>Ø</td>
<td>Ø</td>
<td>Ø</td>
</tr>
<tr>
<td>INDF</td>
<td>anokay</td>
<td>an=</td>
<td>=an</td>
<td>i=</td>
</tr>
</tbody>
</table>

also has the functions of:
1. 1PL.INC ‘we (I and you)’
2. 2SG/PL honorific
3. logophoric

Table 3. Subject-object cross-referencing of transitive verbs in the Saru dialect of Ainu (Southern Hokkaido, Southwestern group), adapted from Tamura (2000:59) with a modification

<table>
<thead>
<tr>
<th>O</th>
<th>1SG</th>
<th>1PL</th>
<th>2SG</th>
<th>2PL</th>
<th>3SG</th>
<th>3PL</th>
<th>INDF</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>–</td>
<td>–</td>
<td>eci=</td>
<td>–</td>
<td>ku=Ø=</td>
<td>ku=i=</td>
<td></td>
</tr>
<tr>
<td>1PL</td>
<td>–</td>
<td>–</td>
<td></td>
<td>ci=Ø=</td>
<td>a=i=</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2SG</td>
<td>en=</td>
<td>un=</td>
<td>–</td>
<td>–</td>
<td>e=Ø=</td>
<td>e=i=</td>
<td></td>
</tr>
<tr>
<td>2PL</td>
<td>eci=en=</td>
<td>eci=un=</td>
<td>–</td>
<td>–</td>
<td>eci=Ø=</td>
<td>eci=i=</td>
<td></td>
</tr>
<tr>
<td>3SG/PL</td>
<td>Ø=en=</td>
<td>Ø=un=</td>
<td>Ø=e=</td>
<td>Ø=eci=</td>
<td>Ø=Ø=</td>
<td>Ø=i=</td>
<td></td>
</tr>
<tr>
<td>INDF</td>
<td>a=e=</td>
<td>a=un=</td>
<td>a=e=</td>
<td>a=eci=</td>
<td>a=Ø=</td>
<td>a=i=</td>
<td></td>
</tr>
</tbody>
</table>
Table 4. Subject-object cross-referencing of transitive verbs in the Ishikari dialect of Ainu (Central Hokkaido, Northeastern group), adapted from Tamura (2000: 59) with a modification

<table>
<thead>
<tr>
<th></th>
<th>O</th>
<th>1SG</th>
<th>1PL</th>
<th>2SG</th>
<th>2PL</th>
<th>3SG</th>
<th>3PL</th>
<th>INDF</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>–</td>
<td>–</td>
<td>e=...=an,</td>
<td>es=...=an</td>
<td>ku=Ø=</td>
<td>?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1PL</td>
<td>–</td>
<td>–</td>
<td>e=...=as</td>
<td>es=...=as</td>
<td>ci=Ø=</td>
<td>?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2SG</td>
<td>e=en=</td>
<td>e=un=</td>
<td>–</td>
<td>–</td>
<td>e=Ø=</td>
<td>?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2PL</td>
<td>es=en=</td>
<td>es=un=</td>
<td>–</td>
<td>–</td>
<td>esi=Ø=</td>
<td>?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3SG/PL</td>
<td>Ø=en=</td>
<td>Ø=un=</td>
<td>Ø=e=</td>
<td>Ø=esi=</td>
<td>Ø=Ø=</td>
<td>Ø=i=</td>
<td></td>
<td></td>
</tr>
<tr>
<td>INDF</td>
<td>en=...=an, un=...=an</td>
<td>e=...=an,</td>
<td>es=...=an</td>
<td>an=Ø=</td>
<td>an=Ø=</td>
<td>i=...=an</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*The third person subjects and objects are marked with “Ø=” to facilitate distinguishing the respective forms from portmanteau morphemes.

2.3 Notes on grammaticalization of person marking in Ainu

Verbal cross-referencing person markers in Ainu have varying morphological status: some of them exhibit properties of words, others are closer to clitics or affixes. For instance, the person markers =an, =as may be separated from the stem by other words and are autonomous prosodic units, therefore they can be classified as words. The person markers a=, eci= cannot be separated from the verb and do not have an accent of their own, which are the characteristics of bound morphemes. On the other hand, they do not change the position of the accent kernel in the stem, which distinguishes them from regular prefixes. Thus a= and eci= should rather be classified as clitics. The personal markers ku=, ci=, e=, en=, un=, i= are full-fledged prefixes since they may not be separated from the root and do affect the position of the accent kernel in the stem in accordance with the general accentuation rule (Tamura 1970b: 605; 2001). The varying morphological statuses of Ainu personal markers are indicative of their being at different stages in the process of grammaticalization from independent pronouns to obligatory cross-referencing affixes, which is a commonly attested grammaticalization scenario. Moreover, it has been pointed out that “existing personal pronouns can be replaced by a set of new personal pronouns but survive as verbal clitics or affixes” (Heine 2007: 95). This is exactly what has happened

---

6. Due to the varying morphological statuses of personal markers in Ainu, there is no agreement among scholars of Ainu about their spelling, i.e. Kirikae (2003) prefers to write all personal markers separately as words, Tamura in her early works and texts (1984) writes them in one word as affixes and in her later works with a double hyphen (1996) which is probably indicative of the clitic-like analysis, the latter approach is accepted in this paper.
in Ainu. In terms of morphology, respective sets of “new” personal pronouns in all Ainu dialects (Tables 1, 2: A/S/O pronouns) exhibit clear features of being the result of secondary development. Each of them contains a respective “old pronoun”, based on an existential/locative verb an (sg) or oka (pl.) and a multifunctional nominalizer -i/-y (< hi ‘thing/place/time’), e.g. eani ‘you (sg)’ < e=an-i (2sg. A/S/O=exist.sg-thing. NR) lit. ‘the place where you (sg) exist’.

In their turn, most “old” personal pronouns functioning at present as verbal cross-referencing markers also have various lexical sources, except for the prefixes ku (1sg.a/s) ‘I’ and e= (2sg.a/s/o) ‘you (sg)’. At present, Ainu (Southern Hokkaido) has seven materially different personal prefixes (A or O) and two suffixes (S only). However, there is evidence that at least three existing personal enclitics originate in proclitic sources and have been rearranged as enclitics in the process of grammaticalization, viz. a(n)= (indf.a) < an ‘exist.sg’, i= (indf.o) < hi ‘thing’, and ci= (1pl. exc) < -hci (impersonal) (< -hci (3pl), because of a general tendency for prefixation in Ainu.

3. Impersonal passive in the Saru and Chitose dialects of Ainu  
(Southern Hokkaido, Southwestern group)

In this section, I will concentrate on the data from the well-documented Saru and Chitose dialects of Ainu (Southern Hokkaido, Southwestern group) and in the next section I will present additional data from the less known Ishikari dialect of Ainu (Central Hokkaido, Northeastern group) to support the plausibility of my proposed grammaticalization scenario.

In the present paper, I will claim that the impersonal passive construction in Ainu has developed from the impersonal (active) construction and the latter has, in its turn, developed from a nominalized verb phrase construction (cf. “the VP-nominalization passive” attested in Ute in Givón (1990: 610)). As mentioned in §1, the development of a nominalization into an impersonal construction in Ainu involves an existential construction, i.e. the impersonal markers an= and =an may be traced back to the existential verb an (vi).

(3)  an ‘exist.sg’ (vi) > an (Indefinite pronoun) > =an (Indefinite s) > a(n)= (Indefinite A).

– Stage I: Existential construction with nouns in S function

(4) a. pet enka ta [cikap]S [Ø=an]V 
   river above at bird 3.s=exist.sg
   ‘There is a bird over the river.’  (KJ); Saru
A diachronic study of the impersonal passive in Ainu

(b) \[a=unu-hu\]$_S$ \[\theta=an\]$_V$

INDF.A=mother-POSS 3.s=exist.SG

\[a=ak-ihi\]$_S$ \[\theta=an\]$_V$

INDF.A=younger.brother-POSS 3.s=exist.SG

‘There was my mother, there was my younger brother.’ (B 121); Chitose

(\text{In the beginning of a folktale.})

– Stage II: Existential construction with action nominalization in S function

> Impersonal construction with intransitive predicates

(5) a. \(rok=an\) yak-ka pirka ya?

sit.pl=INDEF.S if-even be.good Q

‘May one sit down?’

lit. ‘Is it good, if there is sitting down?’ (KJ)

b. \(to\) \(ka\) sat \(kane\) cip-ta=7

lake top be.dry as.if boat-dig=INDEF.S

‘People (in old days) made (so many) boats [lit. ‘people boat-made’] that the lake (looked) dry (when all the boats were there).’ (T1 26)

lit. ‘there was boat-making…”

In (5), all the impersonal constructions may equally be interpreted as action nominalizations, cf. the literal translations. Generally, most action nominalizations in Hokkaido Ainu are zero-marked,\(^8\) which probably correlates with the above-mentioned ability of intransitive stems in Ainu to function as nouns without any change in the morphology. For instance, the stem \(ukoyki\) may equally be interpreted as the intransitive verb ‘to fight’ and as the noun ‘fight’. The fact that the indefinite (impersonal) marker \(=an\) originated in the existential verb has been pointed out by Tamura (2001 (1970b): (217) who notes that a statement about an action by an indefinite person /\(ukoyki=an/\) can be interpreted as an impersonal as in the case of (6a) or as an action nominalization, as in (6b).

(6) a. \(u-koyki=an\)

rec-fight=INDEF.S

‘They (indefinite person) fight.’

b. \([u-koyki]_S\) \([\emptyset=an]_V\)

rec-fight 3.s=exist

‘There is a fight.’ (T4 217)

7. This verb contains the incorporation of O.

8. In Sakhalin Ainu dialects, action nominalizations are marked with possessive suffixes which are also attached to a possessee in possessive noun phrase (Murasaki 1979:95), recall the end of §2.2.
Moreover, even synchronically it is rather difficult to draw a borderline between (6a) and (6b) because, as Tamura (2001; 1970b: 217) shows, \textit{an} can be “separated from the stem by certain particles and/or certain other verb-modifying forms”, as in (7b).

(7) a. \textit{tokap-mokor=an kor…}
\text{noon-sleep=INDF.S} \text{ if}
‘When \textit{one} takes a long nap…’ (constructed example – A.B.)

b. \text{[tokap-mokor]}\textit{ eytasa [Ø=an]}\text{ kor…}
\text{noon-sleep too.much [3.s=exist.sg]} \text{ if}
‘if \textit{one} takes too long a nap…’
\cite{Saru}

Importantly, under the subject-based view of impersonalization, action nominalizations are not considered to be impersonal constructions because nominalization is a formal subject of the existential verb \textit{an} (Siewierska 2008: 124–125); however they may be classed as such under instigator-based view of impersonalization since, in (5)–(7), the focus is clearly on a process at the cost of the participants involved in that process.

– Stage III: Impersonal construction with intransitive predicates > Impersonal construction with transitive predicates

At the next stage, the class of predicate involved in the impersonal construction has been extended to transitive predicates, which can be regarded as a very start of passive reanalysis (see Siewierska 2010: 103). Consequently, the intransitive impersonal enclitic =\textit{an} has grammaticalized into the transitive impersonal proclitic \textit{an=} which is registered in the Northeastern (including Central) Hokkaido and Sakhalin dialect groups. In the Saru and Chitose dialects (Southern Hokkaido), one can also occasionally find \textit{an=} in a number of fossilized verbal forms but a commonly used form is the cognate prefix \textit{a=} (8) which appeared as a result of erosion, i.e. the loss in phonetic substance, usually accompanying the process of grammaticalization. Note the overt O marker in (8e) which additionally shows that the construction in question originally emerged as the impersonal active.

(8) a. \textit{num-i a=Ø=kar wa a=Ø=sakanke}
\text{nut-poss INDF.A=3.o=make and INDF.A=3.o=boil.and.dry}
‘They (people in old days) took the nuts of (water caltrops) and preserved them by boiling and sun drying.’ \cite{T1 26}; Saru

b. \textit{cip a=Ø=wente ruwe ne}
\text{boat INDF.A=3.o=be.bad-caus INF.EV COP}
‘The boat was shipwrecked.’ \cite{KJ}; Saru

c. \textit{kippu hunak ta a=Ø=hok pe an?}
\text{ticket where at INDF.A=3.o=buy NR exist.sg}
‘Where does one buy the tickets?’ \cite{KJ}; Saru
A diachronic study of the impersonal passive in Ainu

527

\[d. \text{Hokkaido tā ku } a=Ø=eywanke ruwe an? \]

\[\text{Are bows used in Hokkaido?} \]

(KJ); Saru

\[e. a=en=ko-pisi p anak-ne opitta ku=Ø=ye wa} \]

\[\text{I'll say everything I am asked about. (lit. ‘The things one asks me about, I'll say all of them.’)} \]

(T1 12); Saru

Regarding the origin of \(a=\) Tamura (2001;1970b: 217) notes the following:

It is probably all right to infer from this correspondence and the prosodic characteristics of the ‘\(a\)-’ and the ‘\(an\)-’ [see §3], that these prefixes originate from one and the same form as the suffix ‘\(-an\)’, which seems to originate from the same form as the [existential] verb ‘\(ān\)’.

Although the relationship between the suffix ‘=an’ and the existential verb ‘an’ is beyond doubt, some researchers of Ainu (Tomomi Satoo, p.c.) are rather sceptical about tracing ‘\(a(n)=\)’ (indf.a) back to the suffix ‘=an’ (indf.s) because of the drastic change in constituent order that this involves: a postverbal clitic becomes preverbal. However, if we consider the personal system as a whole (see §2.3), it becomes clear that similar changes in constituent order in the process of grammaticalization are not uncommon because of a general tendency for prefixation in Ainu.

– Stage IV: Impersonal construction with transitive predicates > Impersonal passive construction (with actor extention)

As a further grammaticalization step, the impersonal came to allow for actor extension\(^9\) via an oblique phrase, viz. ablative or-o wa ‘from the place of’ with animate actors (9), or instrumental ani ‘by’ with inanimate actors (10), which means that the construction cannot be any longer analyzed as the impersonal (active). On the other hand, the presence of the overt O marker in (9) indicates that the construction is still non-promotional and should be best analyzed as the impersonal passive. It has recently been argued that the emergence of an agent phrase is likely to precede the full subjectivization of the patient and that the presence of an overt agent is also likely to speed up the acquisition of more subject properties on the part of the patient, i.e. the development of the non-promotion passive into a promotional one (Siewierska 2010: 82); a detailed account the syntactic properties of the patients in the relevant constructions in the process of reanalysis will be provided in §5.

\[\text{Siewierska 2010: 82; a detailed account of the syntactic properties of the patients in the relevant constructions in the process of reanalysis will be provided in §5.}\]

\[\text{9. Similar actor extention in the impersonal “passive” has been reported for a number of languages, i.e. Kimbundu (Bantu) (Givón 1990:606)), Trukic (Austronesian), Indonesian (Austronesian) (Shibatani 1985:845), and Estonian (Torn-Leesik 2009).}\]
The marker \textit{a(n)}- is also used with a group of so-called transitive ‘psych predicates’ denoting perception, viz. \textit{nukar} ‘see sth/sb’ and \textit{mu} ‘hear sth’, and cognition, viz. \textit{ramu} ‘think about sth’. In this case, \textit{a(n)}- probably has a detransitivizing effect and triggers a decausative\textsuperscript{10} interpretation which may be regarded as an extension of its original function. Unfortunately, at the present stage, additional syntactic tests on the detransitivization of the base construction (11a) are not available. It is generally clear that the morphology employed is transitive (\textit{a=}), at least originally, but the expression of the actor with an NP in the subject function is not possible, which is a characteristic of the impersonal, and the expression of the actor with an oblique phrase, as in the case of the impersonal passive, is not possible either, which is suggestive of the decausative interpretation: decausatives do not allow for actor extension and are detransitivizing by definition. Thus, in this case, \textit{a(n)}- is treated as a derivational prefix and is glossed differently.

\begin{enumerate}[a.]
\item \texttt{heasi cise \textit{Ø=Ø=nukar}}
\hfill \texttt{at.the.sea.shore house \textit{3.A=3.O=see}}
\hfill (constructed example – A.B.)
\item \texttt{heasi cise \textit{a=nukar}}
\hfill \texttt{at.the.sea.shore house \textit{DEC=see}}
\hfill (KJ); Saru
\end{enumerate}

\begin{enumerate}[a.]
\item \texttt{cep hur-a-ha \textit{a=nu}}
\hfill \texttt{fish smell-poss \textit{DEC=hear}}
\hfill \texttt{I’ve heard the fish smells.}
\hfill (KJ); Saru
\end{enumerate}

4. Impersonal passive in the Ishikari dialect of Ainu
(Central Hokkaido, Northeastern group)

In this section, I will focus on the impersonal passive (Stage IV) in the Ishikari dialect of Ainu (Central Hokkaido, Northeastern group). Ishikari and other dialects of the

\textsuperscript{10} A similar analysis is provided in Shibatani (1990:56) who calls the construction in question “spontaneous” and by Kindaichi ((1931) 1993) who calls it “middle passive”.

\[\text{\textit{a(n)-}}\]
Northeastern group have retained certain constructions that may help us to complete the grammaticalization scenario of the impersonal passive in Ainu. Unlike the Saru dialect of Ainu (Southern Hokkaido, Southwestern group), the Ishikari dialect of Ainu (Central Hokkaido, Northeastern group) employs different indefinite markers for transitives depending on the person of O (see Table 5). Moreover in some persons the respective constructions cannot even be regarded either as impersonal or impersonal passive from a purely structural point of view: they are action nominalizations of the type described for intransitives11 in Saru (Stage II).

Table 5. Indefinite marking in Ainu.

<table>
<thead>
<tr>
<th></th>
<th>1SG O</th>
<th>1PL.(EXC) O</th>
<th>INDF O</th>
<th>2SG O</th>
<th>2PL O</th>
<th>3SG/PL O</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ishikari</td>
<td>en=c</td>
<td>un=c</td>
<td>i=...=an</td>
<td>e=...=an</td>
<td>es=...=an</td>
<td>an=Ø=</td>
</tr>
<tr>
<td>Saru</td>
<td>a=en=</td>
<td>a=un=</td>
<td>a=i=</td>
<td>a=e=</td>
<td>a=eci=</td>
<td>a=Ø=</td>
</tr>
</tbody>
</table>

Thus in the case of third person O, there is no difference between the two representative dialects from the major dialectal groups: Ishikari employs for the impersonal passive the transitive prefix \( an= \) (13) which is a cognate of the transitive prefix \( a= \) of Saru and other Southern Hokkaido dialects, cf. (8). In Ishikari as in Saru, the impersonal passive allows for actor extension via an oblique phrase, viz. ablative or -(o) wa ‘from the place of’ with animate actors (13a), cf. (9); I have no examples of the inanimate actor extension with the oblique instrumental \( ani ‘by’ \), corresponding to that in (10).

(13) With 3SG.O: Impersonal passive

a. macya ekota Ø=san a wa wen sisam or
town to 3.s=go.down.SG PERF and be.bad the.Japanese place
wa an=Ø=kisma, Ø=isitoma, Ø=paraparak
‘She went to the town and was grabbed by the Japanese (man),
she was scared and cried aloud.’ (SK 177); Ishikari

11. The impersonal construction of intransitives which is, strictly speaking, a nominalization also exists in the Northeastern group of Ainu dialects, it will not be discussed here for its structural similarity with that of Southwestern group discussed in §3.

12. Both Tamura (1970a: 261) and Asai (1969:779) treat the impersonal forms with 1sg/pl.o as portmanteau morphemes, viz. enci= and unci=, see. Table 4, but here I prefer a more discrete analysis, cf.Table 5.
b. $ci=Ø=kor$ $pon$ $peko$ $onuman-an$ $a=Ø=ahun-ke$

1PL.A=3.O=have young cow evening-exist.SG INDF.A=3.O=enter-CAUS

$oyra$ $soy$ $ta$ $Ø=hotke$ $wa$ $Ø=an$

forget outside at 3.S=sleep and 3.S=exist.SG

‘Last evening, we (lit. ‘someone’) forgot to put our calf (inside) and it slept outside.’ (The children are confessing to their mother.)

(SK 177); Ishikari

c. $Kindaichi$ $tono$ $nispa$ $ku=Ø=nukar$

Kindaichi lord rich.man 1SG.A=3.O=see

$tono$ $nispa$ $tura$ $ci=nokai$ $a=Ø=uk$

lord rich.man with 1PL.Poss=image.Poss INDF.A=3.O=take.SG

‘We met Prof. Kindaichi. We had a picture taken with him.’

(SK 185); Ishikari

According to Asai (1969:779) and Tamura (1970a: 261), in the case of the second person and indefinite O (also used to refer to 1PL.INC, see Table 2), the impersonal marker employed is the intransitive indefinite $=an$ (INDF.S) which, however, co-occurs with the respective verbal cross-referencing prefixes for O, viz. $e=...=an$ (2SG.O...INDF.S), $es=...=an$ (2PL.O...INDF.S), and $i=...=an$ (INDF.O...INDF.S). Here I argue that this traditional circumfixal interpretation raises several questions as to the transitivity of the construction due to the co-occurrence of intransitive (S) and transitive (O) morphology on the predicate. One might be tempted to hypothesise that the object prefixes $e=/es=$ and $i=$, have been reanalyzed as S markers and $=an$ as a passive marker and that the construction itself has become a promotional passive. However, it appears that there is no evidence to support the above hypothesis. In fact there is evidence strongly indicative of no promotion having taken place. In (14b), the base verb $kore$ ‘give sth to sb’ is ditransitive (here conceived of as a three-argument verb). So if it were to be a passive we would expect a reduction of valency by one, i.e. the verb should be a (two-argument) transitive still requiring the same transitive morphology, viz. $an=$ (INDF.A). However, what we find is the intransitive $=an$ (INDF.S).

I suggest that this marker $=an$ (INDF.S) should be analyzed here as the existential verb $an$ ‘exist.SG’ and that the construction has the argument structure of an action nominalization which functions as an S argument of the verb $an$ ‘be/ exist’, just like the action nominalization of intransitives in §3 (Stage II). Thus, (14a) may be interpreted along the lines of $[e=rayke]_s$ $[an]_v$ lit. ‘there will be killing you’. In the action nominalization, the object cannot be omitted, as in (14)–(16), while the subject should be omitted, as in (14a), or demoted to an oblique constituent, viz. the ablative $or-o$ $wa$ ‘from the place of’, as in (14b), (15), and (16). In terms of Koptjevskaja-Tamm (1994, 2005), this is a so-called Restricted or Argument-Reducing ANC in which both A and P cannot appear in the same construction.
As mentioned, the nominalization cannot be regarded as the impersonal from the purely structural point of view.

(14) With 2sg.o: Existential construction with action nominalization in S function

a.  
\[ e=po \quad oya-ke \quad ta \quad e=an \quad yak \]
\[ 2sg.a=son.poss \quad different-place \quad 2sg.s=exist.sg \]
\[ [e=ray-ke]_S \quad [\emptyset=an]_V \]
\[ 2sg.o=die-caus \quad 3.s=exist.sg \]

‘If you (SG) move out of your son’s place, you will be killed.’
lit. ‘…there will be killing you.’ (of vt)  (SK 168); Ishikari

b.  
\[ e=yup-i \quad or-o \quad wa \quad [e=kor-e]_S \]
\[ 2sg.a=elder.brother-poss \quad place-poss \quad from \quad 2sg.o=have-caus \]
\[ [\emptyset=an]_V \quad pe \quad \emptyset=ne \quad ruwe? \]
\[ 3.s=exist.sg \quad thing \quad 3.a=cop \quad perf \quad inf.ev \]

‘Is it the thing you were given by your elder brother?’  (T3 247); Ishikari
lit. ‘Is it the thing that there was giving to you from your elder brother?’ (of vd)

(15) With 2pl.o: Existential construction with action nominalization in S function

\[ es=aca-ha \quad or-o \quad wa \quad [es=koyki]_S \quad [\emptyset=an]_V \quad a \quad ruwe? \]
\[ 2pl.a=uncle-poss \quad place-poss \quad from \quad 2pl.o=scold \quad 3.s=exist.sg \quad perf \quad inf.ev \]

‘Have you (PL) been scolded by your uncle?’  (T3 248); Ishikari
lit. ‘Was there scolding you by your uncle?’

(16) With indf(or 1pl.inc).o: Existential construction with action nominalization in S function

\"okkayo \quad or-o \quad wa \quad [i=koyki]_S \quad [\emptyset=an]_V \quad kor \]
\[ man \quad place-poss \quad from \quad indf.o=abuse \quad 3.s=exist.sg \quad and \]
\[ tan-pe \quad neno-an \quad tusseka \quad ne \quad na. \]
\[ sekor \quad \emptyset=en=epakasnu \]
\[ this-thing \quad like.this-exist.sg \quad kick \quad cop \quad fin \quad quot \quad 3.a=1sg.o=teach \]

‘(The man) taught me: “If one (or ‘you and I’(1PL.INC)) are abused by a man, kick him like this.’ lit. ‘if there is abusing one by a man…”’  (SK 133); Ishikari

And finally, in the case of a first person O, Ishikari employs an entirely different impersonal marker ci= which is related to the first person plural exclusive transitive subject prefix ci= (1PL.(EXC).a=), cf. Tables 2, 4, and 5.

(17) With 1sg/pl.o: Impersonal passive

a.  
\[ huci \quad or-o \quad wa \quad sayo \quad ney \]
\[ grandmother \quad place-poss \quad from \quad gruel \quad always \]
\[ pak-no \quad en=ci=e-re \]
\[ till-ADV \quad 1sg.o=1pl.a=eat-caus \]

‘My grandmother always fed gruel to me (and ate less expensive food (fish) herself).’  (SK 63); Ishikari
b.  

tumi  kes  ta  Upun  ta  toy

war  end  at  Upun  at  land

-tono  or-o  wa  un=ci=kor-e

lord  place-poss  from  1PL.O=1PL.A=have-caus

‘In the end of the war, we were given land at Upun by the Japanese authorities.’  

(SK 154); Ishikari

As we can see, in the Ishikari dialect, in the case of first person O, the verbal morphology is transitive, viz.  *ci*  =  (1PL.(exc).A=), and 1SG/PL object markers  *en*=/*un*=  are retained, but they are attached in the reverse order, viz.  *en=ci*  =  (1SG.O=1PL.A=) and  *un=ci*  =  (1PL.O=1PL.A=); cf.  *a=en*  =  (INDEF.A=1SG.O=) of Saru (Table 5) with the regular order of prefixes, which iconically reflects the basic constituent order SV/AOV of Ainu. The inversion of markers may be taken as an evidence for reanalysis of the impersonal construction into passive, however there is no evidence that  *ci*  in the combinations of  *en=ci*  /  *un=ci*  has completely been reanalyzed into a personal passive marker, so the construction in  *ci*  should be best regarded as one more impersonal passive, to be discussed in detail in §5.2.

*Ci*  as the first person plural exclusive  *ci*  =  (1PL.(exc).A=) is found in all the Hokkaido Ainu dialects. The Sakhalin dialects of Ainu have retained the third person plural subject/object verbal suffix  *-hci*  (<hi ‘thing’ + ci ‘pl’) (Murasaki 1979:15) and the possessive nominal plural suffix  *-hci-n*  (<hi ‘thing’ + ci ‘pl’ + -n ‘human.cl’) (Murasaki 1979:85). Murasaki (1979:51) also notes that the suffix  *-hci*  often marks an impersonal subject. Moreover, it sometimes appears as the prefix  *ci*  - (Murasaki 1979:48) with the same impersonal subject function; the accompanying examples are, however, more suggestive of the resultative function.

As for Hokkaido Ainu dialects, they lack the suffix  *-hci-n*  (pl.poss) entirely but some Hokkaido dialects, especially those of Northeastern group (e.g. Shizunai), employ the suffix  *-ci*  in the iterative function (Nakagawa 2010). As to the above-mentioned resultative prefix  *ci*  -, it is found in Hokkaido dialects such as Saru. Tamura (2000:210) prefers to regard it as a special usage of 1PL.(exc) and analyzes the respective form as an unproductive perfect participle. I suggest that this 1PL.(exc)  *ci*- has probably been reanalyzed into the passive resultative, but unfortunately there are no good syntactic tests to support this view; as mentioned in §1, the coordination test, as in (18), is irrelevant for Ainu.

(18)  
cisekitay  Ø=ci-cari  wa  Ø=an

house-roof  3.s=res-scatter.about  and  3.s=exist.sg

‘The roof of the house is scattered about.’  

(T2 210); Saru

(19)  

rukoc-ihi,  Ø=ci-nina-nina  hine  sir-an

footprint-poss  3.s=res-jam-jam  and  appearance-exist.sg

‘(It seems that from olden times people were visiting each other, therefore it is jammed with footprints.’  

(T5 10); Saru
In many Hokkaido dialects, including Saru and Ishikari, the prefix *ci*- in the impersonal function (conventionally glossed as 1PL.(EXC).A) is retained in a large number of lexicalized nominalizations naming an argument, which is indicative of its old origin (20). On some stems in lexicalized nominalizations, *ci*- is interchangeable with *a(n)-*(INDF.A), as in (21).

(20)  
cep 'fish' < *ci*-e-p (1PL.(EXC).A-eat-thing/place/time.NR) lit. 'the thing people eat'  
(before vowels the final vowel of *ci*- is dropped)  
cironnup 'fox' < *ci*-ronnu-p (1PL.(EXC).A-kill.PL-thing/place/time.NR)  
lit. 'the thing people kill'  
cikoykip 'beast, game' < *ci*-koyki-p (1PL.(EXC).A-catch-thing/place/time.NR)  
lit. 'the thing people catch'  
cihoki 'trade goods' < *ci*-hok-hi (1PL.(EXC).A-buy/do business with-thing/place/time.NR)  
lit. 'the thing people buy/do business with'  
citatap 'minced raw salmon head, roe, and head cartilage' < *ci*-tata-p  
(1PL.(EXC).A-chop-thing/place/time.NR)  
lit. 'the thing people chop'  
(Tamura 1996)

(21)  
a.  
coypep 'tableware' <ci-o-ipe-p (1PL.(EXC).A-from.APPL-eat-thing/place/time.NR) lit. 'the thing people eat from'  
b.  
aoypep 'tableware' <a-o-ipe-p (INDF.A-from.APPL-eat-thing/place/time.NR) lit. 'the thing people eat from'  
(Tamura 1996)

It is obvious that the grammaticalization of *ci* involves a number of parallel developments; a preliminary grammaticalization scenario of *ci* can be summarized as follows. Some of the developments have previously been registered in other languages (I, Ia, and II), other seem to be rare and have not, to my knowledge, been registered elsewhere (Ib).

I.  
THIRD PLURAL (proclitic) > IMPERSONAL (proclitic/enclitic)  
(-hci < hi 'thing' + *ci* 'pl.' of Sakhalin) (-hci of Sakhalin, *ci*- (RES) of Hokkaido)  
a.  
IMPERSONAL (proclitic/enclitic) > IMPERSONAL PASSIVE (enclitic)  
(-hci of Sakhalin, *ci*- (RES) of Hokkaido) (ci= of Ishikari, Central Hokkaido)  
b.  
IMPERSONAL (proclitic/enclitic) > PERS-PRON, FIRST PLURAL (enclitic)  
(-hci of Sakhalin, *ci*- (RES) of Hokkaido) (ci= 1PL.EXC.A of Hokkaido) for politeness (?)  

II.  
THIRD PLURAL (proclitic) > PLURAL (proclitic)  
(-hci < hi 'thing' + *ci* 'pl.' of Sakhalin) (-ci of Hokkaido, mainly Northeastern)  
To summarize, ci- should be regarded as one more impersonal marker which originated from the third person plural marker and is now in the process of reanalysis into the passive. More research is required on the exact nature of these developments in particular dialects of Ainu.

5. Impersonal passive or personal passive?

Having excluded action nominalizations, we are finally left with two impersonal passive constructions:

a. the impersonal passive marked with a(n)= which is employed for all persons of O in Saru and other dialects of the Southwestern group and for 3sg/pl.o in Ishikari and other dialects of the Northeastern group;
b. the impersonal passive marked with ci= which is employed for 1sg/pl.o, viz. en=ci=/un=ci= in Ishikari (the resultative ci-, as in (18) and (19), is beyond the scope of the discussion).

In both constructions, there is a clear evidence for actor-demotion: the actor-phase can appear only as an oblique and cannot be expressed as an argument, which means that reanalysis into passive has already started.

A further important question is whether there is any evidence for the promotion of O to S, i.e. whether the constructions in question are still impersonal passive and have not been reanalyzed as personal passive.

5.1 A(n)= -marked construction

Based on such syntactic phenomena as the word order and honorific use of the plural verb forms, Shibatani (1985:824; 1990:57–59) made an attempt to prove that O in the a(n)= -marked construction has been reanalysed as an S and that the construction in question is passive.

With regard to the word order phenomenon, Shibatani (1990:58) claims that the patient in the a(n)= -marked construction is treated more like a subject because it occurs before the ablative nominal encoding actor and “normally the subject occurs before the ablative nominal, whereas the object follows such a nominal occurring immediately before the verb.”

(22)  
\[ e=kor \text{ hampe eper orowa an=rayke} \]
\[ 2sg=have father bear from \text{ pass=kill} \]
'Your father was killed by a bear.'  \(\text{(SK 48); Ishikari} \)

(the glossing and translation are as in Shibatani (1990:57))
Shibatani’s passive analysis (1990) has been criticized by Satoo (1995) who argues that the parameter of the relative position of the patient NP marked as an argument and the actor NP marked as an oblique may hardly be regarded as evidence for the promotion of O to S, because the actor NP is not an argument. Additionally, Satoo notes that there are examples where the oblique NP with the actor appears before the patient NP (23).

(23) toan nispa or-o wa taan kur a=Ø=kusa
that rich.man place-poss from this man INDEF.A=3.O=take.across.a.river
a ruwe?
PERF INF.EV
‘Has this man been taken across a river by that noble man?’
(Satoo 1995:7); Chitose

As to my own observations, examples of the a(n)= -marked construction containing both the overt patient NP and actor NP in any order are very rare in texts. In the case of first and second person patients, there is overt pronominal marking on verbs and independent pronouns are normally not used, recall §2.2. In the case of the third person, one may, in principal, expect two overt NPs, but, in practice, both NPs very rarely appear together. Neither (22) nor (23) is common, because of the general tendency to omit topical NPs in Ainu.13 As exemplified in (24) pirka menoko ‘beautiful woman’ introduced in the first clause becomes topical and is omitted in the a= -marked construction where it has a patient role.

(24) sine-p pirka menoko,
one-thing be.beautiful/good woman
nean pe anak kim ta Ø=an i ta
that thing TOP mountains at 3.s=exist time/thing at
nispa eturen wa sonno nispa or
rich.man be.obsessed.with and really rich.man place
wa an=Ø=omap
from INDEF.A=3.O=love
‘(There were three nurses in the hospital.) One was a beautiful woman. When that woman was in the mountains, the doctor became obsessed with (her), (she) was really loved by the doctor.’
(SK 73); Ishikari

Generally, any topical NP in Ainu is subject to left-dislocation, as in (22), or to zero-anaphora, as in (24). Thus, Shibatani’s example (22) with the inverse order of NPs, as well as my example (24) without an overt patient NP, are indicative of the fact that the patient has become more topical in discourse and that the construction in question

13. The full source text (about 10,000 words) from which Shibatani cited his example, contains no other examples of two overt NPs in the a(n)= -marked construction.
Anna Bugaeva

has acquired a patient-focusing function in addition to the actor-defocusing function. Another piece of evidence for the increase in the patient’s topicality comes from the relative order of the patient NP and temporal (adverbial) NP in (13b): normally, a temporal NP occurs before a patient NP but in the a(n)= -marked construction the order is reverse.

The second phenomenon that is regarded by Shibatani (1990: 59) as evidence of “the subject status of the patient of the passive is the honorific use of plural verb forms.” He notes that there is no independent case of the object honorific use of plural verb forms, and thus that the honorific plural verbs used in reference to the passive patient indicate that it is treated like a subject, cf. (25a) and (25b).

(25)  

a.  

\[ \text{amset-kasi} \quad a-i-o-resu \]  

\[ \text{(Itadori)} \]  

bed-top \quad \text{PASS-1SG/O[INDEF.O - A.B.]-APPL-raise}  

‘On top of (this) bed I was raised.’

b.  

\[ \text{nekonan-kur} \quad a-o-res-pa \]  

\[ \text{(Itadori)} \]  

what kind of person \quad \text{PASS-APPL-raise-PL}  

‘What kind of person is being raised.’

(the glossing and translations are as in Shibatani (1990: 59))

In fact, unless there has been undertaken a corpus-based research on the use of plural verb forms, we cannot draw immediate conclusions about the impossibility of the object honorific use of the plural verb forms because no other part of Ainu grammar points at this potential impossibility. Generally, the use of plural verbal forms in Ainu operates on an ergative basis: in the case of intransitive verbs, plural forms are used to refer to the plural subject (S) and, in the case of transitive verbs, plural verbal forms are used to refer to the plural object (O). If plural verbal forms occur in the a(n)= -marked construction, they usually refer to the plural patient, as in (26), which should probably be regarded as the object.

(26)  

\[ \text{ku=Ø=kor} \quad \text{hampe ray-ke eper an=Ø=panakte,} \]  

\[ \text{1SG.A=3.O=have father die-CAUS bear INDEF.A=3.O=punish} \]  

\[ \text{nimak-ihi ka an=Ø=tuye} \]  

\[ \text{tooth-POSS even INDEF.A=3.O=cut.SG} \]  

\[ \text{am-ihi ka an=Ø=tuy-pa wa} \]  

\[ \text{nails-POSS even INDEF.A=3.O=cut-PL and} \]  

‘They punished the bear who killed my father, they cut his \text{tooth} and they cut his \text{claws} (to disable him for rebirth).’  

(SK 57); Ishikari

A further important contribution to the study of the syntactic status of the patient in the a(n)= -marked construction was made by Satoo (1995 and 2008) who focused on such syntactic phenomena as equi-deletion and verb raising in the desiderative construction.
Desiderative meaning in Ainu is expressed by the auxiliary verb *rusuy* ‘want’ which is attached directly to the verb if its subject is coreferential with the subject of the main predicate; in this case, *rusuy* itself is unmarked for the person and number of the subject and functions as an auxiliary of the main predicate, as in (27a). In the case of the non-coreferential subject, *rusuy* is preceded by the verb *ki* ‘do sth’ (vt) and complementizer *hi*; the *hi* phrase functions as a complement of the verb *ki*, the latter is marked for the person and number of the non-coreferential subject (the one who wishes) and *rusuy* functions not as an auxiliary of the main predicate but as an auxiliary of the verb *ki* ‘do sth’, as in (27b).

(27) a. \[ ku=ype \ rusuy \\
1sg.s=eat DESID \]
‘I want to eat.’

b. \[ [sir-pirka \ hi]o ku=Ø=ki \ rusuy \\
appearance-be.good COMP 1sg.a=3.o=do DESID \]
‘I wish that the weather becomes good.’ (Satoo 2008:205); Chitose

Which of the above-mentioned strategies is to be chosen in the case of the *a(n)* -marked construction? It appears that the subject of the desiderative is marked as non-coreferential with that of the main predicate (Satoo 1995), as in (27b). This means that O in the *a(n)* -marked construction has not been reanalyzed as subject. If it were reanalyzed as subject, we would have had here *a=en=omap rusuy*, like in (27a), but what we have is the pattern (27b).

(28) \[ húci \ or-o \ wa [a=en=omap \ hi]o \\
grandmother place-poss from INDF.A=1sg.o=love COMP \]
\[ ku=Ø=ki \ rusuy pe ne korka, \\
1sg.a=3.o=do DESID NR COP but \]
\[ húci \ Ø=en=omap \ ka \ somo \ ki \\
grandmother 3.a=1sg.o=love even NEG do \]
‘I want to get the attention of grandmother but she does not devote her attention to me.’ (Satoo 2008:206); Chitose

Another piece of evidence strongly suggestive of the non-promotion of O comes from reflexive binding (Satoo 1995). The reflexive marker *si-* may be used on predicates and relational nouns but only if its referent is coreferential with the subject, as in (29a). Satoo (1995:9) has shown that in the case of the *a(n)* -construction, *si-* cannot be used, which is consistent with the claim that the patient is not the subject. Compare the grammatical (29b) with the ungrammatical (29c).

14. Recall that *hi* also has a nominalizing function which was mentioned in §2.3 and §4, see (20).
(29) a. **si-osmak un ku=inkar**
   refl.-back to 1sg.s=see
   ‘I looked back.’ lit. ‘I looked at oneself’s back.’

b. **nispa wen-kamuy utar or wa**
   rich.man bad-god pl place from
   osmak-e wa a=Ø=kik
   back-poss from indef.a=3.o=hit
   ‘The rich man was hit by evil gods from his back.’

*c. **nispa wen-kamuy utar or wa**
   rich.man bad-god pl place from
   si-osmak wa a=Ø=kik
   refl.-back from indef.a=3.o=hit
   ‘The rich man was hit by evil gods from oneself’s back.’

(Satoo 1995: 9); Chitose

In this section, I have shown that, on the one hand, there is some evidence in Ainu, viz. word order and supposedly honorific use of plural verb forms, for reanalysis to personal passive (adduced by Shibatani 1985, 1990), and on the other hand, there is some evidence against reanalysis to personal passive, viz. equi-deletion and reflexivization tests (adduced by Satoo 1995), which has resulted in two different interpretations of this construction by the respective authors, i.e. as promotional vs. non-promotional. However, in the present paper, instead of taking one side or another, I would like to emphasize that, in a sense, both viewpoints are correct because what we are dealing with here is an incomplete reanalysis of O into S.

As is well-known, “the prototype of subject represents the intersection of agent and topic” (Comrie 1989: 107), which appears to be of relevance to the sequence of particular changes in the process of reanalysis of impersonal into passive. It has been pointed out in the literature that “the process of reanalysis starts with the acquisition of topicality-related properties (e.g. linear position), followed by syntactic (‘behavioural’ in terms of Keenan) properties, and finally by coding properties pertaining, in particular, to agreement.” (Malchukov 2008: 90; cf. Givón 2001). Not surprisingly, in Ainu, subject properties which are more topic-related such as word order and honorifics (Shibatani 1985, 1990) are acquired by O prior the syntactic behavior with regard to equi-deletion and reflexivization, which are more agent-related cross-linguistically (A. Malchukov, p.c.). As for the coding properties, Ainu shows no clear signs of reanalysis of the object person markers as subject markers, and complete reanalysis of indefinite subject construction into personal passives seems to be uncommon cross-linguistically (Siewierska 2010: 91).

To summarize, the **a(n)=**-marked construction cannot be regarded as impersonal because of the possibility of the inclusion of an oblique actor phrase, but it cannot be regarded as personal passive either because reanalysis of O to S has just started: only some topic-related subject properties have been acquired by O but no single
agent-related subject property has been acquired yet; the a(n)= -marked construction should be best regarded as the impersonal passive.

5.2 Ci= -marked construction

As mentioned in §4, the Ishikari dialect employs for first person O a completely different transitive marker ci= (conventionally glossed as 1PL.(EXC)) which should be regarded as one more impersonal marker originating from the third person plural marker. The marker ci= is attached in the reverse order, viz. en=ci= (1SG.O=1PL.(EXC).A=) and un=ci= (1PL.O=1PL.(EXC).A=), and the inversion of markers may be taken as an evidence for the start of reanalysis of the impersonal construction into the passive. This kind of evidence may probably be also regarded as topicality-related, similar to word order, the difference is that here we are dealing with the change in the relative order of cross-referencing prefixes. The next question is whether there is a reliable syntactic evidence for the reanalysis of ci= into a genuine passive marker and en=/un= into subject markers. Since we entirely depend on the previous documentation of the Ishikari dialect and the existing sources are very limited, it was impossible to find good examples of the desiderative auxiliary rusuy or reflexive binding used in the ci= -marked constructions. However, there are a number of auxiliaries with the meanings of psychological states, viz. oyra ‘forget’ (aux/vt), sitoma ‘be afraid of’ (aux/vt), etoranne ‘be reluctant to do’ (aux/vt), kopan ‘hate/refuse to’ (aux/vt), nukuri ‘dislike’ (aux/vt), which behave similar to the desiderative auxiliary rusuy with regard to equi-deletion: they do not take personal markers in the case of the coreferential subject (i.e. act as auxiliaries), as in (30a,c), but require person marking and a preceding complementizer, viz. hi or kuni, in the case of the non-coreferential subject, as in (30b); in the latter case they function as full-fledged verbs, which is slightly different from rusuy.

(30) a. ku=oman etoranne korka
    1SG.S=go.SG be.reluctant but
    ‘I was reluctant to go (there).’ (SK 106); Ishikari

b. sonno [en=ci= e-u-ko-itak
    really 1SG.O=1PL.(EXC).A=about.APL-REC-TOWARD.APL-SPEAK
    kuni]O ku=Ø= e-toranne kusu
    COMP 1SG.A=3.O=about.APL-be.reluctant because
    ‘I felt reluctant that they would talk about me (in a bad way).’
    (SK 86); Ishikari

cf. c. sonno en=ci= e-u-ko-itak
    really 1SG.O=1PL.(EXC).A=about.APL-REC-TOWARD.APL-SPEAK
    e-toranne kusu
    about.APL-be.reluctant because
    intended meaning: ‘Someone felt reluctant to talk about me (in a bad way).’
    (constructed example – A.B.)
In the $en=ci=$ marked construction in (30b), the complementizer $kuni$ is used and there is personal marking on $etoranne$ ‘be reluctant to do’. Therefore its subject is not coreferential with the subject of the main verb $e-u-ko-itak$ ‘talk with each other about sth/sb’. The latter has an impersonal subject encoded with the prefix $ci=$ while $en=$ retains its function of the first person object marker, cf. my constructed example without the complementizer and its coreferential interpretation in (30c). If promotion to subject had taken place, we would expect the (30c) in place of (30b). But this is not the case, hence we have no syntactic (behavioural) evidence for the promotion of $O$ to subject in the $ci=$-marked construction. Here as well we are dealing with an impersonal passive construction which has only partially been reanalyzed as a passive, which is signalled by the inverse order of markers $en=/un=$ and $ci=$ and by equi-deletion test.

6. **A tentative grammaticalization scenario for the impersonal passive in Ainu**

In this paper, I have considered the following three types of constructions:

A. action nominalization involving the existential construction in $an$;
B. impersonal passive in $a(n)=;$
C. impersonal passive in $ci=.$

Only the latter two are directly related to the topic of impersonals under the narrow subject-based approach to impersonalization. I have discussed the data from the well-documented Saru and Chitose dialects of Ainu (Southern Hokkaido, Southwestern group) and the data from the less documented Ishikari dialect of Ainu (Central Hokkaido, Northeastern group).

The data from the Ishikari dialect is of particular interest because this dialect makes use of all the three constructions (§4), while the Saru and Chitose dialects employ only construction (B) and construction (A) for intransitives (§3). I assume that the Ishikari dialect of Ainu has preserved an earlier stage of development, which clearly points to the action nominalization involving the existential construction in $an$ as a source of the impersonal construction in $a(n)=$. A remaining question is why has the action nominalization involving the existential construction in $an$ been retained in Ishikari in the case of the second person and indefinite $O$ and replaced with the impersonal in $a(n)=;$ in the case of the third person $O$? And why has it been entirely replaced with the impersonal in $a(n)=;$ in Saru?

As mentioned in §4, in the action nominalization, the object cannot be omitted while the subject should be omitted, which, in practice, results in the retention of object marking and omission of subject marking. However, the third person has no overt
verbal marking. Therefore the only way for it to maintain the action nominalization
construction is to retain the respective full NP. The latter requirement comes into con-

clict with the general tendency of omitting topical arguments both in subject and object
functions. Thus, the \textit{a(n)= -} marked impersonal construction has emerged as an action
to solve this conflict: it perfectly allows for the omission of the object NP. In the Ishikari
dialect too, one can occasionally find examples of action nominalizations involving the
existential construction in \textit{an} in the case of the third person O, as in (31a). Note that
the construction becomes ungrammatical if the object (\textit{Rikomanpecu-kamuy}) is omit-
ted. However, a more common way in Ishikari to render the same meaning is to use the
\textit{a(n)= -} marked impersonal passive construction, as in (31b),\footnote{In (31b), the object NP \textit{Akan marimo} may freely be omitted.} see also (13); the latter
construction is employed in Saru for all persons of O, recall §3.

(31) a. [\textit{Rikomanpecu-kamuy nomi}]\textsubscript{5} \textit{Ø=an} \textit{kusu paye=as}
Rikomanpecu-god \textit{worship 3.s=exist.sg because} \textit{go.pl=1pl.(exc).s}
\textit{Since there was worshipping (of) the god of Rikomanpecu, we went (there).'}
\hspace{1cm} (SK 185); Ishikari

b. \textit{Akan marimo \textit{an=Ø=nomi} \textit{kusu}}
Akan \textit{ball-shaped.duckweed indf.a=3.o=worship because}
\textit{ku=Ø=kor kur ku=Ø=tura paye=as}
\textit{1sg.a=3.o=have person 1sg.a=3.o=follow go.pl=1pl.(exc).s}
\textit{Since they (lit. 'someone') were worshipping the Akan ball-shaped}
duckweed, I took my husband (and) we went (there).'}
\hspace{1cm} (SK 184); Ishikari

As for 2\textit{sg/pl.o} and \textit{indef.o} in Ishikari, it was possible to retain the action nominaliza-
tion construction in \textit{an} even with the tendency of omitting argument NPs, because
of their overt obligatory verbal marking, viz. \textit{e=} (2\textit{sg,o/a/s}), \textit{es=} (2\textit{pl,o/a/s}), and
\textit{i=} (\textit{indef.o}).

It is worth mentioning that some dialects of the Northwestern group, e.g. Shizunai
for 1\textit{pl.o} (Refsing 1986:223), make use of both (A) action nominalization, as in (31a),
and of (B) impersonal passive in \textit{a(n)=}, as in (31b). Note that even for Ishikari Tamura
mentions both types in the case of 2\textit{sg.o}, see Table 4. Furthermore, certain dialects of
the Northwestern group, e.g. Tokachi (Kirikae 1996:196), seem to have switched com-
pletely to the type (B) impersonal passive in \textit{a(n)=} in the case of 2\textit{sg/pl.o} but show
fluctuation in the use of types (A) and (B) for \textit{indef.o}. The state of affairs in the Sakhalin
dialects is not clear but it seems that they employ both types (A) and (B) (Murasaki
(1979:50) and Chiri (1942) 1973). All these comparative dialectal data are strongly
suggestive of the fact that construction (B) impersonal passive in \textit{a(n)=} developed

\footnote{In (31b), the object NP \textit{Akan marimo} may freely be omitted.}
from (A) action nominalization involving the existential construction in *an*, and not the other way round.

Additionally, Ishikari has developed a special impersonal passive construction in *ci=* for first person O, which is not observed elsewhere. This construction employs an older impersonal morphology, namely *ci*, which has in its turn originated in the third person plural marker (registered in Sakhalin dialects), and the respective object prefixes *en=/un= (1sg/pl.o) which are attached in the reverse order, viz. (*en=/un=*)*ci=*, pointing to a greater degree of reanalysis into the passive in comparison with that of the impersonal passive in *a(n)=*. The special treatment of first person O may probably be accounted for in terms of its high topicality in discourse, which becomes particularly relevant with the patient-focusing function of the impersonal passive, which is acquired in the process of passive reanalysis. It is hardly surprising that O which is higher on the topicality hierarchy has triggered a construction which is, in a sense, closer to the prototypical (personal) passive.

However, it was shown with a number of syntactic tests (see §5) that none of the constructions in question can be classified as the personal passive, since there is no complete promotion of O to the subject: synchronically both *a(n)=* and *ci=*
-marked constructions are impersonal passive regardless of the differences in their sources.

As mentioned in §3, the *a(n)=*-marked impersonal passive allows for actor extension via an oblique phrase, viz. ablative *or-o wa* ‘from the place of’ with animate actors or instrumental *ani* ‘by’ with inanimate actors; only the former actor phrase has been attested with the *ci=*-marked impersonal passive so far, see §4. According to Siewierska (2010: 104), “the possibility of overtly expressing an agent does not appear to have any morphosyntactic correlates but rather to be tied to …contact with languages which have such passives”, which seems to be a plausible scenario for Ainu. In fact, the same “case” distinction as in Ainu is maintained in the treatment of animate vs. inanimate actors in the Japanese passive, viz. Jap. *kara* (ablative) vs. *de* (instrumental).

### 7. Conclusions

In the present paper, I have made an attempt to describe the impersonal passive construction in Ainu from diachronic and synchronic perspectives. It appears that dialects from different groups of Hokkaido Ainu make use of partially different impersonal marking, viz. *a(n)=* in Saru (Southwestern Hokkaido group) vs. *a(n)=* for 3sg/pl.o and *ci=* for 1sg/pl.o in Ishikari (Northeastern Hokkaido group), but the respective constructions seem to follow the same grammaticalization pathway: IMPERSONAL > IMPERSONAL PASSIVE > PASSIVE, which was previously suggested by typologists (Shibatani 1985; Malchukov 2008:96).
I have argued that in spite of clearly impersonal origins, neither the \textit{a(n)}= \text{marked} nor the \textit{ci=} \text{-marked} construction can be regarded as impersonal (active) synchronically because of the possibility of the inclusion of an oblique actor phrase and that none of the constructions can be regarded as personal passive either because reanalysis of \textit{O} to \text{S} is incomplete: only some topic-related subject properties (word order etc.) have been acquired by \text{O} but no single agent-related subject property has been acquired yet (as shown by equi-deletion and reflexivization tests); both constructions should be best regarded as the impersonal passive. The sequence of particular changes in the process of reanalysis of impersonal into passive in Ainu conforms with that proposed in typological literature: the emergence of an agent phrase precedes the subjectivization of the patient (Siewierska 2010:82), topic-related subject properties are acquired by \text{O} prior to agent-related subject properties (Malchukov 2008:90).

With regard to the development of the \textit{a(n)}= \text{marked} impersonal passive construction, I suggested that it has developed from action nominalization involving the existential construction in \textit{an}, which is hardly common cross-linguistically. Based on the analysis of the behaviour of ditransitive predicates, I have shown the inconsistency of the previous “circumfixal interpretation” (Asai 1969:779; Tamura 1970a: 261) of the action nominalization in Ishikari, viz. \textit{e=}...\text{=an} (2sg.o...indef.s), \textit{es=}...\text{=an} (2pl.o...indef.s), and \textit{i=}...\text{=an} (indef.o...indef.s), as it poses a lot of questions about the transitivity value of the construction.

With regard to the development of the \textit{ci=} \text{-marked} impersonal passive construction, which has not received special attention so far, I was able to propose only a very preliminary grammaticalization scenario: THIRD PLURAL > IMPERSONAL > IMPERSONAL PASSIVE; more research is required on the exact nature of these developments in particular dialects of Ainu.

\textbf{Abbreviations}

<table>
<thead>
<tr>
<th>1/2/3</th>
<th>person</th>
<th>DEC</th>
<th>decausative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ø</td>
<td>zero-marked third person</td>
<td>DESID</td>
<td>desiderative</td>
</tr>
<tr>
<td>A</td>
<td>transitive subject or Agent</td>
<td>EXC</td>
<td>(first person plural) exclusive</td>
</tr>
<tr>
<td>A.B.</td>
<td>Anna Bugaeva</td>
<td>FIN</td>
<td>final particle</td>
</tr>
<tr>
<td>ADV</td>
<td>adverbial</td>
<td>INDF</td>
<td>indefinite</td>
</tr>
<tr>
<td>ANC</td>
<td>action nominalization construction</td>
<td>INC</td>
<td>(first person plural) inclusive</td>
</tr>
<tr>
<td>APPL</td>
<td>applicative</td>
<td>INF.EV</td>
<td>inferential evidential</td>
</tr>
<tr>
<td>ARG</td>
<td>argument</td>
<td>O</td>
<td>object</td>
</tr>
<tr>
<td>CAUS</td>
<td>causative</td>
<td>NR</td>
<td>nominalizer</td>
</tr>
<tr>
<td>CL</td>
<td>classifier</td>
<td>P</td>
<td>patient</td>
</tr>
<tr>
<td>COMP</td>
<td>complementizer</td>
<td>PASS</td>
<td>passive</td>
</tr>
<tr>
<td>COP</td>
<td>copula</td>
<td>S</td>
<td>intransitive subject</td>
</tr>
</tbody>
</table>
Sources


T1  Tamura, Suzuko. 1984.


References


Referential impersonal constructions in Mandarin*

Yi Yan & Anna Siewierska
Beijing Foreign Studies University/Lancaster University

This paper provides an overview of referential impersonal constructions in Mandarin Chinese. It is shown that while like many European languages Mandarin utilizes generalized nouns, special person forms, regular personal pronouns and a zero pronoun in the subject position to encode impersonality, unlike in many other languages the impersonal use of the 3PL form ‘tā-men’ is marginal. The range of uses associated with 3PL impersonals namely the universal, corporate, vague, inferential and specific existential (Cabredo Hofherr 2006) are rather preferentially rendered by the yǒu rén construction which is formally a combination of the existential verb yǒu, a generalized noun rén and a predicate whose agent is replaced by yǒu rén.

Keywords: referential impersonalizing constructions; impersonality; Mandarin

1. Introduction

This paper considers the properties of referential impersonal constructions in Mandarin Chinese in the light of what is currently known about the corresponding European constructions. Inspired by Siewierska (2008b), we categorize impersonal constructions into those involving referential strategies and verbal strategies. In verbal strategies a sense of impersonality is created through manipulation of argument structure, as in the case of agentless passives or reflexive or participle impersonals in Romance, Slavic or Finnic. In referential impersonal constructions, on the other hand, impersonality is conveyed by the choice of referentially impoverished expression for the subject or the agent. The most common referential strategies for expressing impersonality include the use of generalized nouns and impersonal uses of personal pronouns. It is on such expressions of impersonality that we will be concentrating.

* The authors would like to thank the anonymous reviewers for their helpful suggestions and insightful comments. Of course, we alone are responsible for any potential errors in the paper.
1.1 Types of impersonality

As discussed in the introduction to this volume and documented in full in the contribution by Malchukov & Ogawa, the notion of impersonality is a wide and disparate one with different understandings of it creating different expectation as to the scope of the constructions involved. Of the morphological, structural and semantic characterizations of impersonality that have been employed in the analysis of European languages, the first two are basically inapplicable to Sinnitic. Morphological impersonality is associated with verb forms which while normally characterized for person are either invariant for person or lack person specification altogether. Mandarin has no verbal person marking at all and thus the issue of this type of impersonality does not arise. Structural impersonality relates to the presence of a canonical subject which in the European context has typically been identified with morphological nominative case and/or thematic content. Again due to the absence of case distinctions in Sinnitic no impersonality such as that found with dative or accusative subjects in, for example, Germanic occurs. Nor are there any expletive subject pronouns to feature in say weather constructions such as the English *It is raining*. By contrast to the above, the third characterization of impersonality – the semantic – can be applied to Sinnitic. Semantic impersonality revolves around two notions, i.e. agentivity and reference. The first type of semantic impersonals lack agents, while the second do have agents be it ones which are not fully referential. Needless to say, we are interested in the latter, i.e. the referential impersonals. Given that agents are prototypically subjects, from a cognitive-functional perspective the function of referential impersonals may be seen as that of ‘hiding an agent’ (Afonso 2008: 190), ‘agent backgrounding’ (Fried 2006) or ‘agent defocusing’ (Myhill 1997).

Even among reference-based interpretations of impersonality, the type of reference that qualifies as impersonal is not uniformly understood. Some scholars, most notably Kitagawa & Lehrer (1990), make a distinction between ‘impersonality’ and ‘vagueness’ with ‘impersonality’ defined as necessarily implying anyone or everyone with the possible inclusion of speaker and addressee and ‘vagueness’ as indicating a specific group of individuals who are not identified or identifiable by the speaker and exclude the speaker and addressee. Scholars such as Cinque (1988) and Cardinaletti & Starke (1998), on the other hand, emphasize the difference between ‘generic’ and ‘impersonal’. According to them, generic constructions express law-like propositions which hold for all the members of a group and are prototypically timeless while impersonal constructions express propositions which apply to some unspecified individual or set of individuals and are associated with a specific time. In the formal semantics literature this distinction is often referred to as generic vs. arbitrary (see e.g. Cabredo Hofherr 2006 & Siewierska this volume).
In the ensuing discussion of referential impersonals in Mandarin we use the term impersonal to cover both the vague and generic and arbitrary and generic distinctions. Our presentation is organized around the linguistic forms employed in the expression of reference impersonality in Mandarin. The presentation of the referential strategies that are used to convey this notion will be in each case accompanied by a brief discussion of their morpho-syntactic features, and then followed by a consideration of the type of impersonality involved. Comparisons with relevant forms in other languages (mainly in English) will also be made where possible or necessary.

1.2 Nouns and pronouns in Mandarin

As the chief referential strategies for encoding impersonality in Mandarin involve generalized nouns and personal pronouns, a brief description of the two relevant paradigms is in order. Mandarin nouns are not marked for gender or case. Nor are they obligatorily marked for number. For instance, zhuōzi can mean either ‘a table’ or ‘tables’. A plural suffix -men can be added to NPs denoting people, but its use is neither obligatory nor frequent (Li & Thompson 1981). The general nouns that are employed to denote a loosely defined group of people in Mandarin include rén ‘person/people’, yī-ge rén ‘one-cl person, one’, rén-men ‘people, pl.’

As with nouns, Mandarin pronouns are not distinguished in terms of grammatical role. Neither is gender reflected in the spoken language. In the written language there are distinct forms for third person pronouns. All three persons have singular and plural forms. The latter are rendered by the previously mentioned suffix -men added to the singular forms. The six basic pronouns of Mandarin are shown in Table 1.

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st person</td>
<td>wǒ</td>
<td>wǒ-men</td>
</tr>
<tr>
<td>2nd person</td>
<td>nǐ</td>
<td>nǐ-men</td>
</tr>
<tr>
<td>3rd person</td>
<td>tā</td>
<td>tā-men</td>
</tr>
</tbody>
</table>

In addition to the forms in Table 1 two other personal pronouns are widely used in northern dialects of Mandarin; nín is a polite form for the second person singular and zán-men is an inclusive pronoun referring to the speaker, other people associated with the speaker, and to the addressee. By contrast, wǒ-men can be used either inclusively or exclusively and is much more commonly used than zán-men.
Indefinite pronouns that are used to refer to humans include *dà-jia* 'everybody', referring to people in general, *rénjia* 'other person/people' and *biérén* 'other person/people' and *shé*¹ 'everybody/nobody', placed before the verb and always with *dōu* 'all' or *yě* 'also'.

In what follows we will demonstrate that like many European languages, Mandarin utilizes generalized nouns, regular personal pronouns and a zero pronoun in subject position to encode impersonality. Interestingly though, in contrast to many other languages, the 3pl pronoun *tā-men* is not utilized as a major impersonalizing device. We will see that the various impersonalizing uses associated with 3pl imp (see Cabredo Hofherr 2006; Siewierska this volume) are in Mandarin preferentially conveyed by the zero pronoun construction and the *yòu rén* construction.

---

1. *shé* can be used either to mean 'everybody/nobody' or as an interrogative pronoun:

   (i)  
   *nèi-ge rén shì shé?*  
   (constructed)  
   that-cl. person cop who  
   Who is that person?

Therefore, the affinity between interrogative pronouns and indefinite pronouns witnessed in many languages (Haspelmath 1997 & Bhat 2004) is also found in Mandarin. The indefinite use of *shé* is restricted to the subject position or the pre-verb position and is typically used with the universal quantifier *dōu* (all), or its negative form *yě*

   (ii)  
   *shé* dōu xīhuān tā  
   everybody all like 3sg  
   'Everybody likes him/her.'

   (iii)  
   tā xīhuān *shé*  
   3sg like everybody  
   (intended meaning)'S/He likes everybody.'

   (iv)  
   tā xīhuān shé  
   3sg like whom  
   'Who does s/he like?'

   (v)  
   tā shé dōu xīhuān  
   3sg everybody all like  
   (Lit) S/he everybody likes  
   'S/He likes everybody.'

   (vi)  
   tā shé yě bù xīhuān  
   3sg everybody also neg like  
   (Lit) S/he everybody not likes  
   'S/He likes nobody.'
2. Referential impersonalizing strategies in Mandarin

2.1 Generalized nouns

Generalized nouns in Mandarin that are used to encode impersonality include: rén 'people not specified for number', yī-ge rén ‘one-cl person, one’, rén-men ‘people, pl’.

We will consider each of these in turn.2

2.1.1 rén

While Myhill (1997) is correct in saying that rén is a dummy form meaning “people in general”, it is by no means the only or most frequently used generalized noun in Mandarin. This is partly due to monosyllabicity of rén which runs counter to the overall preference in Mandarin for disyllabic words. So although sentences with rén alone as subject can be found in Mandarin, as illustrated by (1), rén is most often used as a morpheme in di- or multi-syllabic free forms such as yī-ge rén ‘one-cl person, one’, rén-men ‘people, pl’ biérén ‘other people, others’, rénjia ‘other people’, all of which can be used to denote impersonality in the sense of genericity and/or non-specificity of the agent.

(1) (yī-ge) rén huó-zhe jiù yào xuéxi4
(one-cl) people live-DUR if...then should learn
‘People/we live to learn.’ / ‘One lives to learn.’

3. Most examples of this paper come from Corpus of Chinese Language (CCL) complied by PeKing University (http://ccl.pku.edu.cn:8080/ccl_corpus/index.jsp?dir=xianpai) or other published works in Mandarin. The examples that are constructed by the authors are so marked.
In (1) rén appears in subject position picking as its referent the whole body of human beings. Bare rén is also used in the following three grammatical relations other than the subject. First, rén can appear as an object with a generic meaning:

(2) zuòjiā zhìyóu shēn cēngcí de lìjiě (*yī-ge rén)
writer only deep level ADV understand (one-CL person)
ren cái néng xiēchū zhēnzhèng fǎnyìng rénlèi gǎnqíng
people then can write really reflect mankind emotion
de zuòpǐn.
nom works

‘Only when a writer understands people deeply can s/he produce works that truly reflects the emotions of human beings.’

Second, rén is often used as a possessor in a possessive NP in the form rén de (a person’s) as shown in (3).

(3) wǒ xīhuān zhè-zhòng tiáozhàn (yī-ge rén) rén de
1sg like this-CL: type challenge (one-CL person) people poss
zhìhuì de zhīyè.
intelligence nom occupation

‘I like the type of occupation that challenges one’s/your intelligence.’

---

4. In Mandarin the following construction with rén in the object position is very common:

(i) lìng rén jīdòng
make people excite
‘exciting’

The construction functions an adjective (phrase) because it can either be used as a noun modifier or as a predicate:

(ii) a. lìng rén jīdòng de xiāoxī
make people excite nom news
‘exciting news’

b. zhè gè xiāoxī zhēn lìng rén jīdòng
this CL news really make people excite
‘This piece of news is really exciting.’

As a matter of fact, in Mandarin combining lìng (make) with a generic noun rén and a lexeme denoting emotions or psychological actions is a fairly productive way of forming adjectives denoting qualities of things, e.g. lìng rén jīngyà (make people surprise, surprising), lìng rén hàipà (make people afraid, frightening), lìng rén yíhuò (make people doubt, puzzling). In a sense we could say that rén, or rather lìng rén (make people) has become an adjective forming affix.
Third, the dummy 人 is often found as the object of 被 in the 被-construction which is widely considered a syntactic passive form in Mandarin, e.g.: 5

(4) suíshí kěyì bèi rén (*yī-gè rén) mài diào!

Anytime can 人 people (one-cl. person) sell out

'(Lit.) Ø can be sold at any time.’

The 被人 construction is fairly common in 被-constructions. The 人 ‘people/person’ after 被 in (4) is a non-referential or generic one. It could be deleted and the resulting construction will become something like an English short passive:

被人卖掉 = 被卖掉 ‘be sold.’

In all the examples above, the reference of 人 is the general body of human beings with the speaker and addressee included. As we shall see below 被 is different from other generalized nouns and indefinite pronouns in two aspects. Firstly 人 may function as an object. Secondly, it is necessarily interpreted as generic or kind-referring. When a definite interpretation is demanded, 人 is not allowed. Therefore 人 could be regarded as a dedicated generalized noun. By contrast, the other forms we are going to discuss are all potential impersonalizing devices in the sense that on the phrase level they are all ambiguous between a generic/impersonal reading and a definite reading.

2.1.2 yī-gè rén

As indicated in (1) and (3) dummy 人 can be replaced by another NP yī-gè rén ‘one-cl. person’ though not always. In form, yī-gè rén is an indefinite noun that corresponds to the English ‘a/one person.’ Like singular indefinites in English, yī-gè rén is open to a specific reading, as in (5) and a nonspecific potentially generic reading as in (1) and (3). 6

(5) wǒ zài jiēshàng pèngshàng-le yī-gè rén, (tà) hěn

1sg on street run into -PEF one -CL person (3sg) very

xiàng wǒ mèimei (constructed)

like 1sg sister

‘I ran into a person on the street. (She) looked like my sister.’

5. Whether 被 is a verb or a preposition is a source of controversy among Chinese linguists. The readers are referred to Huang (1999) for a review of this issue.

6. yī-gè rén could also be used as an adnominal emphatic with the meaning of ’alone’ or ’on one’s own’ as in the following example:

zhè shì tā yī-gè rén zuò de

this COP 3sg one-cl. person do PAT

‘This is done by him alone.’ or

‘He did it all by himself.’
In fact, as shown by (5) and (4), *yī-ge rén* in the object position favors a specific reading except in a negative sentence as demonstrated by (6):

(6)  
```
 wò jìntiān zài jiēshàng méi kànjiàn yī-gè  rén.
1sg today on street NEG see one-CL person
 'I saw nobody on the street today.'
```

As a possessor, *yī-ge rén* is typically nonspecific or generic in reading as in (3). As a subject, however, it is open to the generic interpretation as in (1) or a specific interpretation as in (7):

(7)  
```
yī-gè  rén shàngwǔ lái zhāo nǐ, nǐ bù zài
one-CL person morning come look for 2sg 2sg NEG in
 'A person/someone wanted to see you in the morning but you were out.'
```

Given that *yī-ge rén* may receive a specific or a generic reading, the question arises what is this determined by. Simply put, the answer lies in the nature of the context. In episodic contexts, particularly perfective ones as exemplified by (7) the reading will be specific whereas in various irrealis contexts e.g. law-like propositions, negative, deontic and conditional, the reading will be generic as shown in (8).

(8)  
```
a. yī-ge  rén zuò diān hào shì bīng bù nán,
one-CL person do little good thing INT NEG difficult
 nán de shì yī bèizi bù zuò huài shì
difficult NOM COP one life NEG do bad thing
 'It's not difficult for a person/one to do a good deed, but it's difficult not to do wrong all his life.'
```
```
b. yī-ge  rén ruò bù néng dú zhéxué zhùzuò yuánwén,
one-CL person if not can read philosophy works original
yào xiāng duì tā-men wánquán liǎnxiăo shì hén
will want to 3PL completely understand COP very
kùnmán de
difficult PAR
 'If a person/one can't read works on philosophy in the original language or version, it will be difficult for him/her to fully understand them.'
```
```
c. yī-ge  rén bù kěnéng yòngyuán jùyòu jīngzhēnglì
one-CL person NEG possible forever have competitiveness
 'One cannot remain competitive forever.'
```
```
c'. shé  dōu bù kěnéng yòngyuán jùyòu jīngzhēnglì
everybody all NEG possible forever have competitiveness
 'Nobody can remain competitive forever.'
```
```
d. yī-ge  rén yīnggāi qīngchū zījǐ de zérén
one-CL person should clear refl NOM responsibility
 'One should be very clear about one's own responsibility.'
```
d’.  

\[ \text{shé dōu yìnggāi qīngchǔ zìjí de zérén} \]

everybody all should clear \text{REFL NOM} responsibility

‘Everybody should be very clear about one’s own responsibility.’

e.  

\[ \text{yī-ge rén ruò shì bā tā míngbái de} \]

one-cl person if \text{COP ba-particle 3SG understand NOM}

\[ \text{wánquán zuòdào-le, tā yē shì shèngrén} \]

fully \text{act-PERF 3SG also COP} saint

‘If one can act fully according to what he understands (i.e. rules in life) he is also a saint.’

(8a) through (8d) illustrate what Cinque (1988) and Cardinaletti & Starke (1998) and Brehens (2005) define as law-like propositions or generic sentences, which is a typical type of irrealis situation. One can see that \text{yī-ge rén} in them is interpreted as a generic NP including the speaker and the addressee. As a matter of fact, this type of law-like propositions is typically employed by the speaker to explain his or her behavior, that is, to talk about himself/herself, as shown by the contexts of (8c) and (8d) which are not given due to the limited space. (8c) and (8d) also illustrate how negative and deontic situations give rise to the generic reading of \text{yī-ge rén}. Interestingly, the genericity can be rendered by \text{shé … dōu} as shown by (8c’) and (8d’). In (8e) which is less like a law-like proposition, the genericity of \text{yī-ge rén} still obtains largely due to the fact that the context is a conditional one.

As we have seen, in most instances, \text{yī-ge rén} constitutes a good translation of the English generic pronoun \text{one}. However in the following sentence from the British National Corpus the generic \text{one} cannot be translated as \text{yī-ge rén}.

(9)  

(Some semantic links manifest inheritance. For instance, if the network connects the node ‘student’ to the node ‘person’ with the link ‘is a’, then)

\[ \text{one can infer that the properties of “student”} \]

\[ ^*\text{yī-ge rén} \text{kèyì tuīzhī …} \]

\[ \text{Ø/rén-men/wó-men} \text{kèyì tuīzhī …} \]

are inherited from those of “person”.

To express the idea of the English sentence, the Mandarin speaker will have to choose an impersonalizing form other than \text{yī-ge rén}, either \text{rén-men} ‘people’, \text{wó-men} ‘we’ or a zero pronoun, all of which will be discussed later.\footnote{It seems that the incompatibility of the generic reading of \text{yī-ge rén} and (9) can be explained by the distinction between a generic sentence and a generic NP proposed by Brehens (2005). Brehens (2005) points out that generic NPs are kind-referring phrases which do not necessarily have to occur in generic sentences and generic sentences are law-like propositions which do not necessarily have to contain a generic noun phrase. Viewed in this framework, when \text{yī-ge rén} is employed in real generic constructions or law-like propositions.}
To sum up, *yī-ge rén* is not a dedicated general noun but only a form allowing a generic interpretation with the speaker and addressee included. The possessor's role and most irrealis contexts are the linguistic contexts that will trigger a generic reading of it. In episodic sentences, *yī-ge rén* will have a specific interpretation. The preclusion of the arbitrary meaning in non-generic context is a reflex of the form of *yī-ge rén*. Recall that being primarily an indefinite noun, it is ambiguous between an unspecific and a generic reading. Therefore its interpretation is sensitive to the context.

With respect to matters of style, *yī-ge rén*, like English *one*, is most frequently used in formal registers.

2.1.3 *rén-men*

As the plural form of *rén*, *rén-men* 'people' can be used definitely, indefinitely and non-specifically. In its non-specific use, as in its referential uses, it can occur as subject, as in (10), as direct object (11a), as prepositional object (11b) and as a possessor (11c).

10. *rén-men* fāxiàn wúlǐshù bīng zuīzhòng chéngrèn
   people    discover irrational number and finally acknowledge
   wúlǐshù jīnglì-le yī-duàn tòngkǔ
   irrational number experience-PERF one-CL:period painful
de    guòchéng
   NOM  process

'It took a painful period for people to discover irrational numbers and accept them.'

as exemplified by (8a through d) and in conditionals as in (8e), it has the generic reading. (9), on the other hand, is not a generic sentence or not even a characterizing one because it does not serve to express the typical characteristic of the subject ‘one’ (Behrens 2005). In other words, the sentence is not about the typical characteristic of its subject ‘one’ or the fact that anyone/all member of a certain group can perform something. Rather it is about the possibility of making an inference and the content of it. This interpretation is corroborated by the possibility of rendering the same idea as expressed in (9) with 'it is possible' or a short passive 'it can be inferred' in which the subject place is held by a pleonastic ‘it’ the reference of which can be seen as even more impersonal than generalized nouns or indefinite pronouns. That is, generalized nouns and indefinite pronouns portray the situation as having an agent but give no further information on its identity than its humanness whereas the construction with a pleonastic form as subject depicts the situation as completely lacking in an agent. If this interpretation is true, it is likely that in an irrealis context that is nonetheless non-generic, *yī-ge rén* cannot have the generic reading. This hypothesis is of course to be verified by future work.
Referential impersonal constructions in Mandarin

557

(11) a.  

高自春从前可不是人们不看印地语电影

Ever since last year this organization began to call on people not to watch films in Hindi.

b.  

田斯智卓尔丽莎李来中国，向人们展示这个国家

TV anchorwoman Lisa Lee came to China to show how this country is changing.

c.  

再人们的感觉中，大家对演员/女演员的

As people see it, most actors or actresses are extrovert.

It is worth noting that when 人们 appears as the direct object in (11a) with generic reference it is also the semantic agent of the predicate of the subordinate clause, i.e. ‘不看印地语电影’ ‘not (to) watch movies in Hindi’. This is different from 人 which can appear as a simple object as demonstrated in (2). Note that in (2) 人 cannot be replaced by 人们.

In (10) and (11), the referent of 人们 is clearly people in general to the exclusion of the speaker and the addressee. But 人们 could also be used in contexts where the speaker or addressee is included as in the constructed example in (12).

(12)  

(在) 这里人们很早去睡觉

People here go to bed early at night.

In (12) whether the speaker is included within the range of referents subsumed by 人们 depends on the discourse context. For example, if the speaker is a tourist guide explaining the customs of a particular locality, i.e. “this place/here”, of which he/she is not a member then 人们 excludes the speaker. If the speaker is a local, 人们 includes the speaker. But in both cases whether the addressee is included is not important for the meaning of the utterance. In the second case where the speaker is a local 人们 can be replaced by a non-anaphoric, non-deictic use of 我 ‘we’, but 人们 is even vaguer than 我 because it doesn’t pick out its referent by making the distinction between the speaker and the hearer. Now consider a situation where (12) is uttered by an owner of an apartment to a person who is going to move into that
apartment or by a neighbor to a new tenet who plays loud music at night. Then wǒ-men will clearly include the speaker but exclude the addressee though it doesn’t spell out the exclusion in form. In this case, the unspecified exclusion of the addressee as the agent is important for understanding the illocutionary meaning of the utterance as a piece of advice or as a complaint. Note this ambiguity is also what makes (12) more polite as a piece of advice or complaint than wǒ-men zhèlǐ shuǐjiào hěn zǎo ‘We go to bed early here’.

It is also possible to use rén-men when the intended referent is only the addressee. In (13a) the speaker is a reporter asking about a practice at a place in which the addressee is involved but which excludes the speaker. Siewierska (2008b) notes that in Polish the reflexive impersonal, (13b) is also possible in a similar context:

(13) a. wúxiū qìjiān rén-men gàn shénme?
lunch break during people do what?
‘What do people do during the lunch break?’

b. Co się robiło na przerwach?
what refl did:3sg on breaks
‘What did you used to do during the break?’

Unlike its English equivalent, rén-men is not restricted to informal style. It is a common way of achieving impersonality in academic writing. Our survey of the Sinica Corpus consisting of 4,892,324 words in Mandarin found 935 tokens of indefinite rén-men within which 183 cases are found in written expository writings. (http://dbo.sinica.edu.tw/ftms-bin/kiwi1/mkiwi.sh).

So far we have seen that rén-men is the most unrestricted generalized NP in terms of referential range. By only specifying the type of referent as human and plural, it remains uncommitted as to whether the speaker and the addressee are included or not. This ambiguity or vagueness lends itself to situations where explicitly indicating the inclusion of the speaker or addressee will have pragmatic-politeness consequences. Unlike the English equivalent ‘people’, rén-men is not confined to colloquial style. Its flexibility with respect to genre and referential range resembles the Swedish man-construction (Ragnarsdóttir & Strömquist 2005), though the Swedish form as a generic pronoun is syntactically constrained to subject position whereas rén-men being a generalized noun can appear in the subject, object and possessor’s role with a generic reference.

Since both the bare rén and its morphological plural form rén-men can be used impersonally, it is important to understand the distinction between them. The bare form rén is typically used to express generic reference including the speaker and addressee and it differs from other impersonal forms in that it can be interpreted generically even in object function. By contrast, rén-men ‘people’, the plural form of rén can either be interpreted as generic or as an unspecified group of people to the
exclusion of the speaker and addressee. In the latter sense rén-men is interchangeable with the impersonal use of two other forms that will be introduced later i.e. biérén ‘other(s)’ and rénjia ‘other(s)’.

There are morphological distinctions between rén and rén-men as well. Firstly, rén can be used in the sense of the human species whereas the morphological plural form rén-men cannot, as shown in (14a). Secondly rén can be used to form complex nominals whereas rén-men cannot as shown, in (14b). Thirdly, rén-men, just like any noun suffixed with -men cannot be further modified by a quantifier, as shown in (14c) and (14d). But as (14e) through (14g) show, other types of modification such as by means of classifiers, adjectives and relative clauses are allowed with rén-men.

   bird can fly but people neg can
   ‘Birds can fly but people can’t.’ cf:
   a’. The bird can fly but man can’t.

b. fāguó rén
   France people
   ‘the French, French people’

b’. *fāguó rén-men
   France people
   Intended meaning: ‘the French, French people’

c. háizi-mén hěn gāoxìng
   child-cl:people very happy
   ‘The children are very happy.’

c’ *sān-ge háizi-men
   three-cl:child-cl:people
   ‘three children’

d. sān-ge rén

d’. sān-ge *rén-men

e. sūnchuántíng dui zhè-xiē rén-men shì tuīchūqù hài
   proper name to this-cl:pl people-cl cop drop out or
   shì liuxiàlái bìng bù zhūyì
   cop stay int neg care
   ‘Sūnchuántíng didn’t care if these people would drop out or stay.’

f. shānliáng de rén-men bù nèng bù wèi zhì jīngtì
   kind nom people-cl neg can neg for it vigilant
   ‘Kind people have to be on guard against it.’

g. rèài hépíng de rén-men, qíng jiārù wǒ-men
   love peace nom people-cl please join 1pl
   ‘People who love peace, please join us.’
In addition to rén-men Mandarin also has three other forms that could be used to refer to people in general, dàjia ‘everyone’, biérén ‘other person, other person/people’ and rén-jia ‘other person/people’ but they are more like pronouns that can be used impersonally. Therefore they will be discussed in the following section to which we now turn.

2.2 General pronouns

2.2.1 Impersonal use of 1st and 2nd person pronouns

As in English and many other languages (Siewierska 2004: ch 5) in Mandarin the 1st person plural personal wò-men ‘we’ and second person singular pronoun nǐ ‘you’ can be used to denote anyone or everyone with the possible inclusion of speaker and addressee.

Like its English counterpart, wò-men can be used to refer to a loosely defined group of people, especially with the intention to involve the addressee in the situation or to create a sense of mutual participation, commitment and often responsibility (Siewierska 1984):

(15)  wò-men zài huánghé shàng zuò-le hěnduō gōngchéng
     1pl at Yellow River on do-pef many project
     yóuqí jiěfāng yīlái, diànhàn bùduàn de xiù
     in particular liberation since power station continually ADV build
     jiéguò huánghé duānlǐú, érqī duānlǐú de
     result Yellow River dry up furthermore dry up NOM
     tiānshù yuéláiyuèduò
     day number more and more

‘We did many projects on the Yellow River, especially since the liberation (1949). Power stations are being built on it one after another. As a result, the river often dries up. Moreover the time it is dried up becomes longer and longer.’

This excerpt was taken from a talk show transcript on water conservation. The speaker is the Minister of Construction. The wò-men in the first clause is a typical case of inclusive impersonal use where its referent is people in China. A reading where wò-men could be interpreted as referring just to the Ministry of Construction is excluded given the phrase “especially since the liberation” which implies that the minister was thinking about all the projects on the Yellow River and their effect on it. The following sentence makes clear that the speaker’s choice of the impersonal wò-men is motivated by an intention to create a sense of responsibility.

Semantically impersonal nǐ is similar to you in English which is not only generic but also implies special interest in the addressee (Quirk et al. 1985). Like you, nǐ is less formal and is mainly used in speech or narratives.
Referential impersonal constructions in Mandarin

(16) zhè-zhòng wèntí a, nǐ děi xiǎng hǎojiù cái néng
this-cl:type problem part 2SG have to think long then can
xiāng de chū bàn fā ne
think res out solution part

(Lit) 'This type of problem, you have to think for a long time before you find a solution.'

According to Myhill (1997:818), in English when the agent can be any person, vague you is always used in preference to the passive and the generic one and especially so in informal usage, e.g.

(17) a. Joanie, to really understand suffering, you have to experience it first-hand.

In Mandarin, however, the impersonal nǐ is probably not as preferred as the English you because it has a competitor, i.e. the zero pronoun which matches its referential flexibility while being less stylistically restricted. So in (17b) which is the Mandarin equivalent of (17a), the zero pronoun would be equally fine. More will be said about using the zero pronoun to convey impersonality in 2.4.

(17) b. yào zhènzhèng lǐjiě kùnnán (nǐ/Ø) jiù bǐxū
to really understand suffering (2SG/Ø) then have-to
qīnzi jīnglì tā
self experience 3SG

‘To really understand suffering you/Ø have to experience it first-hand.’

An important difference between the English we and you, on the one hand, and the Chinese impersonal wǒ-men and nǐ on the other is in regard to the genre they appear in. Whereas we and you are sometimes used in English proverbs as impersonal subjects, wǒ-men and nǐ are not the preferred subjects in such sentences. So although the proverbs in (18) could sometimes be translated with wǒ-men and nǐ as seen in (18a’) through (18c’), the more idiomatic way to express proverbs or rather law-like propositions is to use zero pronouns and rén as exemplified by (18a”) through (18d”) and (19):

(18) a. You may take a horse to the water, but you cannot make him drink.

a’. ?wǒ-men/ní/?rén kěyī bā mǎ qiāndào
1pl/2sg/person can ba-particle horse take
hén bìa, dānshì bù néng qiāngpò tā hē shuǐ
river side but NEG can force 3SG:N drink water

a”. qiān mǎ dào shuǐ yì, bǐ mǎ hē shuǐ nán
take horse to water easy force horse drink water difficult

(Lit) ‘It’s easy to take the horse to the water but it’s difficult to make him drink.’
b. We live to learn.

b’. wǒ-men/*ni/rén huó-zhé jiù yào xuéxí
1PL/2SG/person live-CRS if...then will learn

b”. huó dào lǎo xué dào lǎo
Ø live till old Ø learn till old

c. We never know the worth of the well till the well is dry.

c’. bù dào jǐng gàn-le, wǒ-men/*ni/rén/Ø yòngyuǎn
NEG till well dry-PERF 1PL/2SG/person never

bù zhīdào jǐng de jiàzhí
NEG know well ass value

c”. jǐng gàn Ø fǎng zhī shuǐ kēguì
well dry Ø only...until understand water valuable

(19) a. rén wú wán rén.
person no perfect person

(Lit) ‘There is no such man who is perfect.’/’Every man has his faults.’

b. hé huà rén dājiàodào Ø bǐxū tíguō jīngtī
with bad person deal with must raise vigilance

‘He who deals with bad people must be alert.’ or
‘He that sups with the devil must have a long spoon.’

2.2.2 Impersonal use of 3rd person pronouns

3rd person IMPs are constructions with a third person pronominal subject which is non-referential and human and which lacks an antecedent in the discourse (Siewierska & Papastathi, forthcoming). In Mandarin, the 3rd person singular pronoun tā is unavailable for impersonal use. The closest tā comes to being impersonal is when it is anaphoric to a non-specific antecedent, for instance, yī-ge rén in (20).

(20) yī-ge rén zhīyào bānyàn-le yì-zhōng shèhuí juéshè,
one-CL person as long as play-PERF one-CL: type social role
tā jǐu huì mànman de ànzhāo zhè-ge
then will slowly ADV according to this -CL
juéshè de yāoqíu qù zuò
role ass requirement to do

‘Once a person assumes a social role, he will gradually act according to the requirements of that role.’

It has been mentioned that cross-linguistically, 3rd person plural pronouns are the most common form of referential impersonal devices. In Mandarin, however, the 3PL form tā-men is not used impersonally or at best only in very restricted circumstances. Significantly, it is not used to render any of the five types of 3PL impersonals identified
by Cabredo Hofherr (2006) and further explicated in Siewierska and Papastathi (Forthcoming), namely the universal, corporate, vague, inferential and specific existential exemplified in (21a) through (21e):

(21)  a. In Spain, they eat late. (universal)  
b. They changed the tax laws last year. (corporate)  
c. They've found his bike in the back of a barn. (vague)  
d. They've been frying chips here. (inferential)  
e. They're knocking at the door. (specific existential)

(22) demonstrates that the Mandarin equivalent of the 3pl universal impersonal is a construction with rén-men in (22c), not the one with the 3pl pronoun, i.e. tā-mén (22b).

(22)  a. They usually clean cows in Switzerland.  
   tā-men tōngcháng qǐngjié nàiniú zài ruìshì  
b. zài ruìshì *tā-mén tōngcháng qǐngjié nāiniú  
c. zài ruìshì rén-mén tōngcháng qǐngjié nāiniú

The preferred rendering of a 3pl corporate impersonal is via a zero pronoun subject in an active construction, as in (23a) or in what is commonly referred to in Chinese linguistics as a notional passive construction or patient-verb construction (PVC) as in (23b).

(23)  a. qùnián Ø gǎi-le shuì fǎ  
   last year change-PVF tax laws.  
   ‘They changed tax laws last year’  
b. qùnián shuì fǎ gǎi-le  
   last year tax laws change-PVF  
   ‘Last year tax laws were changed.’

Li and Thompson (1981:659) also give an example where a construction with a zero pronoun in the initial subject position is used to create a sense of impersonality:

(24) Ø yòu xiū lǜ-le  
    again repair road-PVF  
    ‘(They)’re fixing the road again.’

We’d like to point out that (24) could in fact be interpreted not only as a corporate impersonal but also as an inferential one, if the speaker has not seen any workers, or even as a vague one. However, the vague and inferential readings of 3pl impersonals and also the specific existential one are all preferentially expressed in Mandarin by the yòu rén construction. This is illustrated in (25a) for the vague readings, in (25b) for the inferential and in (25c) for the specific existential.
a. They’ve found his bike in the back of a barn.

a’. você ren zai niushè houmiàn fàxiàn-le tā
de zixingché
exist person/people in barn back find-third person
poss bike

b. They’ve been frying chips here.

b’. você ren zai zhēli zhá-guò shūtiáo.
exist person/people in here fry-exp chip

c. They’re knocking at the door.

c’. você ren qiáo mén.
exist person/people knock door

The implied agent in the você ren imp has a flexible referential range which is granted by the form of the construction. The construction is not specified for number because the generalized noun ren like most Mandarin nominals does not specify number. Therefore in (25a’,b’,c’) whether one person or a group of people carried out the action in question is unclear, which is similar to the case with the 3pl imp in English and Dutch.

The você ren construction will be discussed in more detail in §2.3.

Although tā-men is not normally used non-referentially in Mandarin, Siewierska and Papatathi (2011) have observed that it does occur in the Mandarin translation of Harry Potter to render English impersonal they:

(26) a. They say there’s dragons guardin’ the high security vaults. (HP p. 49)

a’. tā-men shuō nà-xiē fāngfān zuì yánmì de jīnkuò dōu
3-pl say those security most tight nom vault all
yóu lóng bāshōu
have dragon guard

b. (‘Red card!’ said Dean furiously. ‘In soccer you get shown the red card and you’re out of the game!’ ‘But this isn’t soccer, Dean,’ Ron reminded him. Hagrid, however, was on Dean’s side.)

‘They oughta change the rules. Flint could a knocked Harry outta the air.’

(HP p.150)

b’. ….

tā-men yìnggāi gài biàn gūi zē …
3-pl should change rule

We have also come across a case in written Mandarin where tā-men could be seen as being used impersonally.

(27) ① zhè-ge huàidàn! ② Zhōuzhào lù gàndào bùān
this-cl. bastard, Proper Name feel uneasy
3 lão liú wūfēi shì gěi tā yídàn nánkān,
Old Liu, just COP to 3SGj little embarrassment

4 dàn yè bù páichú nà-ge rén duì tā hé
but also NEG rule out that-CL personi about 3SGj and

Huánáqiàn de guànxì de mìngān
Proper Namek ASS relationship ASS sensitivity

5 tā-men méi yǒu zhèngjù
3PLj NEG have evidence

6 tā-men bù kěnéng yǒu zhèngjù
3PLj NEG possible have evidence

7 tā shì yànjūshì fūzérén, duì xiàshū jīnxíng
3SGj COP research lab director to subordinate do

ywù shàng de zhídào wú kě fēiyì
profession about NOM direction nothing should bad say

‘That bastard! Zhōuzhàolù felt uneasy. Old-Líu just wanted to embarrass him.

But it could also be that this bastard sensed that something was going on

between him and Huánáqiàn. They don't have evidence. They couldn't

possibly have any evidence. He is the director of the research lab and nobody

can say anything bad about him for giving guidance to his subordinate

at work.’

This excerpt from a fiction (clauses ① to ②) is a direct thought of the protagonist,
Zhōuzhàolù. What is narrated before it is that Lao Líu, a colleague of the protagonist
said something implicating that the protagonist is giving special favor to his secrete
lover, Huánáqiàn, who is also his subordinate in the research lab. Although the two
tā-men in clauses ③ and ④ clearly include Lao Líu and we can therefore say that the
3PLj is to some degree anaphoric, it is also possible to argue that the referent of tā-men
is not fully evoked in the previous context because tā-men clearly refers to a group of
people whose identity is specified only by the joint common activity or feature they
shared, that is, those people who could have the evidence of the special relationship
between the protagonist and his lover. Crucially there is no antecedent of tā-men. By
virtue of this, tā-men in this excerpt may qualify as a case of vague existentials or joint
common activity IMPS.

2.2.3 Between noun and pronoun, dàjia, biérén and rénjia
In addition to rén-mén ‘people, pl’ Mandarin has another form which is used to refer
to people in general – dàjia. In terms of referential range dàjia is close to the English
indefinite pronouns everyone or everybody except that it is plural in number. dàjia has
both deictic/anaphoric uses as in (28a) and nonspecific/impersonal uses as in (28b)
through (28d).
As shown in (28b) through (28d), like many other generic pronouns, the non-referential use of *dàjia* is restricted to subject and possessor positions. *Dàjia* is considered a pronoun because it manifests the two properties characteristic of pronouns in Mandarin. As shown by (28a), it is typically used to refer to an entity whose identity is already (situationally) established at the time it is used. Moreover, like other pronouns in Mandarin it does not allow any modifier, be it a classifier/measure phrase, an associate phrase or a modifying phrase (Li & Thompson 1981: 133). On the other hand, *dàjia* clearly retains some nominal features, as discussed in Sugamoto (1989). In the first place, it carries a sociolinguistic connotation. When it has 2nd person reference it is more polite than the 2 plural *ni-men* ‘you’ and when it used to refer to people in general with the speaker included, it gives more emphasis on the solidarity between the speaker and the addressee. Secondly, *dà-jia* has more semantic content than the prototypical first and second plural pronouns with which it can sometimes be used interchangeably. While *wǒ-men* ‘1-pl’, and *nǐ-men* ‘2-pl’ only contain the grammatical meaning of person and number, *dàjia* is made up of two parts with lexical meaning, that is, *dà* means ‘all/ every’ and *jia* is a suffix denoting ‘a certain type or group of people’. As a matter of fact, *dàjia* was historically used as a noun in classical Chinese.

Mandarin also has two nonspecific forms that refer to people other than the speaker and/or the addressee, i.e. *biérén* ‘other person, other person/people’ and *rénjia* ‘other person/people’. Like *dàjia* ‘everyone’, *biérén* and *rénjia* contain more semantic content than prototypical pronouns. *bié* means ‘other’; *rén* means ‘person or people’ with no specification for number, and *jia* means ‘person or a group of people’. So the literal meaning of *biérén* and *rénjia* is ‘people other than the named person’, and they
are similar to the English other(s) as a reference to people in general excluding the speaker and the hearer. Stylistically, rénjia is primarily used in the colloquial language while biérén is more neutral. Due to their literal meaning, the two forms are often used to emphasize the contrast between others and the speaker:

(29) biérén/rénjia néng gànchú chéngjì lái, zán-mén jiù

Others can make achievements out 1PL INT

bù néng? neg can?

‘If others can do it, why can’t we?’

Note that in this context, reference by means of rén-men would be less appropriate because it would include the speaker. When the contrast is not emphasized, rénjia and biérén can be used interchangeably with rén-men because all the forms share the meaning of people in general:

(30) měidāng biérén/rénjia/rén-men/yǒu rén zhèměi háng

whenever others/people/exist person this way call

tā, tā jiù yào chénxià zuì jiào dūnang …

3SG 3SG then will lower mouth corner mutter

‘Whenever others/people address him in this way, he will droop the corners of his mouth and mutter…’

While discussing the typology of 3PL IMPS Siewierska & Papastathi (2011) proposed a separate group of impersonals: 3PL IMPS with the speech act verb “say” the example of which is found in (26a). Although as mentioned earlier rendering they by means of tā-men in Mandarin causes no problem in understanding among Chinese readers, the more idiomatic way to background the source of the information is through biérén/rénjia/rén-men, and yǒu rén plus the verb shuō (say). Among the 30 native speakers of Mandarin who the authors consulted with only 5 consider tā-men acceptable in this context in a non-anaphoric sense whereas all of them agree that rén-men, biérén, rénjia and yǒu rén are appropriate. There is a slight difference between these impersonal forms though. rén-men picks people in general as its referent while biérén and rénjia exclude the speaker and hearer in both its form and reference.

Below is an example of rénjia from Mandarin that resembles the 3PL IMP-say use of they:

(31) yǒu huí wǒ wèn LàoShē xiānshēng: “rénjia shuō

exist time 1SG ask proper name Mr. others say

wǒ wénzhāng xiě de bù shūnlíù,

1SG writing write RES NEG fluent

‘Once I asked Mr. Lao-shé: “Others they say my writing is not fluent…'
The writer of the excerpt is asking another famous Chinese writer LàoShé for advice. rénjiā does not refer to anyone evoked in the previous context. It is used to refer to people in general to the exclusion of the speaker and the addressee, much in the same way the English they is used.

2.3 Impersonals based on the verbal expression yǒu rén

As mentioned in §2.2.2 a preferred way of backgrounding third person agents in Mandarin is via the yǒu rén construction. yǒu can be a possessive or an existential verb as shown in the two constructed examples in (32): 8

\[
(32) \begin{align*}
\text{a. } & \text{wǒ yǒu yì-dòng fāngzì} \\
& \text{1sg have one-cl house} \\
& \text{‘I have a house.’}
\end{align*}
\]

\[
\begin{align*}
\text{b. } & \text{zhūō shàng yǒu yì-běn shū} \\
& \text{desk on exist one-cl book} \\
& \text{‘There is a book on the desk.’}
\end{align*}
\]

In the two examples above, yǒu is a verb because it exhibits the syntactic features of verbs in Mandarin. In (32a) where it denotes possession it takes a direct object; as a typical verb such as zǒu ‘walk’, it can be negated by the negator méi: méi yǒu ‘don’t/doesn’t have’, can appear in the ‘verb NEG verb’ pattern: yǒu méi yǒu, and can be modified by dōu ‘all’ as well. Unlike ordinary Mandarin verbs, however, it doesn’t take aspect markers, but neither does the copula shì or other modal verbs for that matter.

yǒu also appears frequently as the first verb in serial verb constructions or pivotal constructions 9 which are commonly found in Mandarin (Lǚ 1980/1995:63; Li & Thompson 1981; Zhu 1983/2003). In (33) the second verbs in the serial verb constructions are double underlined.

---

8. Linguists identify 5 uses of yǒu in Mandarin: (1) possessive, (2) locative-existential, (3) presentational, (4) perfective and (5) assertive. See Huang (1987) inter alia.

9. A serial verb construction is a sentence that contains two or more verb phrases or clauses juxtaposed without any marker indicating what the relationship is between them (Li & Thompson 1981:594). The following is the schema given by Li & Thompson:

\[(NP) V (NP) (NP) V (NP)\]

A pivotal construction is a special type of serial verb construction which contains a noun phrase that is both the direct object of the first verb and the subject of the second verb. Despite of the abundance of the serial verb constructions in Mandarin, there is no consensus regarding the taxonomy and treatment of them among Chinese linguists. See Li & Thompson 1981 and Zhu (1983/2003) for further discussion.
Referential impersonal constructions in Mandarin

(33) a. wǒ yǒu (‐jiàn) shí gèn nǐ shàngliàng
    1sg have (‐CL) matter with 2sg discuss
(Lit) ‘I have a matter to discuss with you.’ or
‘I have something to discuss with you.’

b. yǒu (gè) rén diū‐le yī‐bā fūzī
   exist (‐CL) person lose‐PEF one‐CL ax
(Lit) ‘There is someone who lost an ax.’ or
‘Someone lost an ax.’

Li & Thompson (1981:611) treat sentences like (33b) as a subtype of serial verb constructions – descriptive clauses. They are so named because they involve a transitive verb (in this case yǒu) whose direct object is described by a following clause i.e. (diū‐le yī‐bā fūzī, ‘lost an ax’). On the other hand, (33b) is also regarded as a presentative construction in that its function in discourse is to present or introduce an NP to be described. Note that the classifiers of the object of yǒu in (33) can be omitted and the resulting forms become yǒu shì and yǒu rén respectively. The two forms, yǒu rén in particular, are very common in Mandarin. It is possible that yǒu rén has lexicalized into a compound meaning as ‘someone’ or ‘some’ (when referring to persons).10 The correspondence between yǒu rén and the English indefinite pronoun some is shown by the second translation of the two sentences in (33b) and (35b) below. However it should be pointed out that the yǒu rén form is not a full‐fledged noun form yet because unlike a compound NP or the indefinite pronoun ‘someone’, it can never appear in object position; compare (34a) with (34b).

(34) a. tā kàn‐dào‐le *yǒu rén.
   3sg see‐PERF yǒu person
   Intended meaning: ‘He saw someone.’

b. tā kàn‐dào‐le yī‐ge rén.
   3sg see‐PERF one‐CL person
   ‘He saw someone.’

Actually in Mandarin yǒu + NP + verb is quite a productive pattern of creating serial constructions. Zhu (1983/2003) lists six types of serial constructions with yǒu as the first verb and regards yǒu rén as a variant of yǒu + (number) +‐CL+ NP with the classifier omitted and consequently a type of pivotal construction. For these two reasons, we will call the pattern ‘yǒu rén + Verb (+NP)’ yǒu rén constructions.

10. We thank one of the anonymous reviewers for pointing this out.
The following example illustrates the agent backgrounding function of \textit{yǒu rén} construction.

(35) a. ① \textit{wǒ-men xiǎoqū bùduàn yǒu rén lái kàn fángzi,} ② \textit{fángshì hái shì gāowēn bú tui a drop PAT}

(Lit) 'Our neighborhood continuously \textit{has people} coming to check out the houses (on sale). The real estate fever is still here.'

'There are always people coming to check out houses on sale in my neighborhood. It seems that the real estate fever is not lost.'

b. ① \textit{yǒu rén jiànyì zhōngyāng cáizhèng shělǐ zhì exist people suggest central finance establish treat huái zhuānxìàng zìjīn,} ② \textit{guówùyuàn shīfōu zhūnbèi Huai (river) special fund the State Council if or not prepare zuò zhè fāngmiàn gōngzuò? do this side job}

(Lit) 'There are some people who suggest establishing a special fund for the treatment of Huai River. Is the State Council prepared to take this type of action?' or

'Some people suggested establishing a special fund to treat Huai River…' or

'A special fund is proposed for the treatment of Huai River…'

Note that the \textit{yǒu rén} constructions in (35) are somewhat different from what is commonly regarded as a presentational \textit{yǒu rén} construction illustrated by (33b). While (33b) performs the function of a typical presentative construction, that is, to introduce an NP into the discourse (Li & Thompson 1981 inter alia) or to establish some entity in/into/from a domain or of presenting the existence of some entity in a domain (Yang & Pan 2001), (35) cannot be said to serve the same function. This can be seen from the following contexts (e.g. clauses ②) where no further information or description is given about the referent of \textit{rén} and the text goes on to discuss about something else, i.e. the real estate market or the State Council. Notably, (35a) cannot be translated with \textit{someone} or \textit{some people}.

What we argue is that the \textit{yǒu rén} construction as exemplified by (35) could be seen as syntactically impersonal because the existential verb \textit{yǒu} is not governed by an agent and the sentence can be said to lack a canonical subject. More importantly they are also semantically impersonal under the agent/instigator defocusing view (Siewierska 2008a) by virtue of the fact that the actual agent of the verb, i.e. \textit{lái kàn fángzi} 'come
to check out the houses on sale’ or jiànyì ‘suggest’ is under-elaborated. In other words, in the yǒu rén construction a situation that could have been presented as involving an agent carrying out some action is encoded as simply existing as a situation or state of affair and the only information given about the agent is that it is human and it carries out an action denoted by the verb that comes after rén. On these grounds we will call sentences like (35) yǒu rén imps.

The yǒu rén imp is reminiscent of the English there-existentials as discussed by Ziv (1982) and the haver-construction in European Portuguese as discussed by Afonso (2008). Both linguists have shown with examples such as (36a) from Ziv (1982: 757) and (36b) from Afonso (2008: 192) that what might be considered a pre-representative-existential sentence out of the context actually performs an agent defocusing function:

(36)  

(a. The President: There will be a total and complete cooperation with the agencies of government to get at the facts.

b. A: Gostava desse trabalho?
   ‘Did you like that job?’

   B: Aqui não havia gostos, foi por conveniência
   ‘There were no likes here, it was convenient.’

Ziv (1982) argues that the existential in (36a) is used not to present ‘total and complete cooperation’ but to allow the speaker not to mention whoever is involved in the cooperation with the government. Analogously in (36b), the haver-construction is not employed to set the scene or background, as is the norm for presentatives, but to downgrade the agent/instigator.

Returning to the yǒu rén construction, syntactically it is a hybrid combining the verbal strategy (existential verb yǒu) and the nominal strategy (a generalized noun rén) and a predicate whose agent is replaced by yǒu rén. The generalized noun is used to give only a minimal specification of the agent of the event, i.e. it is human; the existential is used to assert the existence of such an agent. The amount and nature of information provided about the agent corresponds to that given in the English agentless passive. Hence the yǒu rén construction is sometimes rendered as short passives in English as shown earlier in (35b).

Although the agent in the yǒu rén imp is typically interpreted as an unspecified person not including the speaker or the addressee, as in the examples in (25) given in §2.2.2, the construction can be used when the referent is actually known to the speaker
or addressee, a situation similar to the Spanish and Italian 3pl IMP but different from the English 3 pl IMP (Cinque 1988: 543 and Siewierska 2008b):

(37) a.  
\[ \text{gāngcái } yóu \text{ rén } \text{gěi } nǐ \text{ dǎdiànhuà, kěnéng shí } nǐ \text{ jiējie} \]
   Cop 2sg sister

   (Lit) 'Just now there was a person who called you. It might be your sister.'
   'Someone called you just now. It might be your sister.'

b.  
\[ \text{Prima hanno telefonato: mi pareva tua sorella è arrivata} \]

   earlier have-3pl telephoned; me seemed your sister

   'Someone (*They) telephoned earlier. It seemed to me that it was your sister.'

The \text{yóu rén} IMP can even be used when the implied agent is actually the speaker:

(38)  
\[ \text{yóu rén } \text{kànjiàn nǐ hé tā zài yīqī} \]

   exist 2sg and 3sg being together

   (Lit) '(There) is a person/someone who sees you with him/her.'
   'You are seen hanging out with him/her.'

(38) could be uttered by a speaker who actually saw the addressee hanging out with someone, but for tactfulness s/he chooses not to mention her/himself as the agent of seeing. Notice that the \text{yóu rén} construction enables the speaker not to rule out the possibility that s/he is included in the possible range of the implied agent of seeing because this construction only asserts the existence of such a person and gives no information about the identity of the agent other than that. In contrast, other NPs such as \text{tā-men, biérén, rénjia} in Mandarin or the English 'they' won’t be appropriate for the speaker’s purpose of avoiding mentioning him/herself without violating the truth because they explicitly exclude the speaker as the agent of the action.

Like the \text{rén-men} construction, the \text{yóu rén} IMP is not stylistically restricted. It is common both in colloquial and formal academic Mandarin. Nor is it restricted to semantic transitive events as demonstrated in (35a).

In sum, the \text{yóu rén} construction is a highly flexible impersonalizing device in terms of the referential range of the implied agent, the types of situations encoded by the verb and genre distribution. What grants it this flexibility is again its form. By being an existential it is open to all kinds of existential meanings that are related to vague reference and impersonality; by only asserting the existence of a most generalized noun \text{rén}, it is open to both individual and group interpretation, both to an unspecified group excluding the speaker and including the speaker; by achieving the impersonality meaning through only asserting the existence an agent rather than promoting the patient, it places no demand on the semantic transitivity of the event.
Yet another way employed in Mandarin to background the source of information or action is through a zero pronoun in subject position to which we will now turn.

2.4 Impersonalization via zero pronouns

Mandarin is known for its freedom in the use of zero pronouns (Li & Thompson 1981; Huang 1984 and Chen 1984, 1986 inter alia). We’ve already seen in previous sections that using a zero pronoun in subject place is a common way of backgrounding the agent in Mandarin. We call this type of zero pronouns ‘impersonal zero pronouns’ and the construction containing them zero pronoun IMP Cs. Impersonal zero pronouns are to be distinguished from zero pronouns used anaphorically as in (39) taken from Huang (1984:533):

(39) Speaker A: zhāngsān kànjiàn lǐsì le ma?
   Proper name see proper name PEF PAT
   ‘Did Zhāngsān see Lǐsì?’

Speaker B: a. tā kànjiàn tā le
   3sg see 3sg PEF
   ‘He saw him.’

b. Ø kànjiàn tā -le
   [he] see 3sg-PEF
   ‘He saw him.’

c. tā kànjiàn Ø le
   3sg see [him] PEF
   ‘He saw him.’

d. Ø kànjiàn Ø le
   [he] see [him] PEF
   ‘He saw him.’

The zero forms in (39b) through (39d) are all definite and have an overt antecedent in the discourse.

In terms of proposition, the Mandarin zero pronoun IMP Cs often express generalizations that apply to all the people in a group, that is, they often appear in law-like propositions.

(40) zhǔzhù-le zhè-ge zhúyào máodì, yǐqiè wèntí jiù yīnɡrènèrjiē -le
       grasp-PEF this-cl principal contradiction all problem then readily solve -PEF.

(Lit) ‘Once Ø grasp this principle contradiction, all problems can be readily solved.’

‘Once this principal contradiction is grasped, all problems can be readily solved.’
Like Polish (Siewierska, 2008b) and Modern Hebrew (Berman 1980: 768), Mandarin also employs zero pronouns to express deontic modality:

(41) a. Trzeba odezći
necessary to leave
'It’s necessary to leave.'

b. CARIX le hasbir lo (et ze).
must to explain to-him (ACC it)
‘He needs to have it explained’ or
‘It must be explained.’

c. zhè-zhong qíngkuàng xià yínggài jínkùài lìkāi
this-cl:condition under should immediately leave
(Lit.) ‘Under this condition, Ø should leave immediately.’
‘Under this condition, one should leave immediately.’

The ‘zero pronoun + modal verb + complement’ construction is widely used in Mandarin, especially in written expository and argumentative writings.

Zero pronouns are also frequently utilized to urge or advocate certain action, i.e. as imperatives. While imperatives typically involve zero pronouns, in English agentless passives are used more often to perform the directive function politely. By contrast, in Mandarin rules and regulations are usually expressed by means of zero pronoun IMP Cs. Compare the following two library announcements in English taken from Leech and Svartvik (1975) and in Mandarin which is from the National Library of China.

(42) a. It has been noted with concern that the stock of books in the library has been declining alarmingly. Students are asked to remind themselves of the rules for the borrowing and return of books, and to bear in mind the needs of other students. Penalties for overdue books will in the future be strictly enforced.

b. ① wài jiè shùkān zǐliào bìxū ànqǐ guíhuán
out loan book material must on time return
‘Borrowed books and materials must be returned on time.’

② yùqí guíhuán xū jiàonà yūqǐ shíyòng fèi
overdue return must pay overdue use fee.
(Lit.) ‘If Ø return overdue Ø must pay overdue fee.’
‘Books must be returned on time. Overdue fees will be charged for overdue items.’

In (42a) three passives are used to achieve impersonality and formality while Mandarin in (42b) uses a patient verb construction (PVC) in ① and a zero pronoun construction in ②. All three constructions background the agent, i.e. the authority behind the command by leaving it unexpressed but there are minute differences between the effects achieved. The English agentless passive backgrounds the agent of the directive through foregrounding the patient "students”, so the resulting construction binds
the patient/addressee explicitly. The Mandarin PVC leaves out the sense of directive altogether by specifying the desirable result, i.e. books being returned on time. The Mandarin zero pronoun IMP C simply leaves the subject of the directive unspecified without explicitly mentioning the patient/addressee.

The referential range of zero pronouns from (40) to (42) is a loosely specified group of people or generic in the sense of Cinque (1988) and Cardinaletti & Starke (1998), but it is also possible to background a more specific agent, e.g. 1st or 2nd person agent by the zero pronoun IMP C. In this respect, the zero pronoun construction functions like the English short passive and is in competition with two verbal impersonal constructions in Mandarin i.e. bèi-constructions and PVCs. The syntactic features of the relevant constructions will help decide which construction will be used.

In the following example given by Myhill (1997:822):

(43) a. Anyone know how many shots were fired?

the speaker is the policeman addressing a group of people involved in a shooting spree. It is likely that at least some of the addressees actually fired shots themselves, so the passive instead of impersonal you or they is used to avoid either directly implicating the addressees or excluding them. For the same situation, however, the passive in Mandarin i.e. the bèi-construction is not possible because the semantics of the predicate kāiqiāng ‘fire a gun’ will preclude the use of it. Neither is a PVC possible because of the presence of jǐ ‘how many’. The only impersonal way of expressing the idea like (43a) is through using a zero pronoun as in (43b):

(43) b. shé zhīdào kāi-le jǐ qiāng?
who know fire-PERF how many shot
(Lit) ‘who knows Ø fire how many shot?’
‘Who/anyone knows how many shots were fired?’

Below is an example from Report on the Work of the (Chinese) Government 1999:

(44) rúguò Ø rush into mass action, rashly put up establishments, randomly launch redundant and inferior projects, Ø will put a heavy burden on government budget, which will sooner or later lead to inflation.

(Lit) ‘If Ø rush into mass action, rashly put up establishments, launching redundant and inferior projects will put a …’
The conditional clause has a zero pronoun as its subject because the agent of “rash into mass action, put up establishments, launch redundant and inferior projects” is not spelt out. The use of the zero pronoun is motivated by two reasons. One is to achieve vagueness about the real agent of the action. The covert agent could be interpreted either as a general body of people including the speaker and addressee, analogous to English we, or any agent or organization who is in the capacity of carrying out the actions expressed by the predicate, in which case the speaker is excluded but the addressee included. The other reason for using a zero pronoun instead of the verbal impersonals such as the bèi-construction is that ‘rush into mass action (in putting up establishments)’ is a semantically intransitive clause.

Another type of situation that will call for the use of a zero pronoun is the reporting situation. In English the short passive, they, or a construction with a noun like rumor or news as the subject are used to background an incompletely specified source of information as in (45).

(45)  
a. It is reported/said that 75 people died from yesterday’s clash.  
b. Rumor/word/news has it that 75 people died from yesterday’s clash.  
c. They say 75 people died from yesterday’s clash.

Mandarin will use the following construction:

(46)  
jù Ø shuō/chēng/chuán(shuō)/bàodào zuòtiān  
according to say/claim/spread (word)/report yesterday  
chōngtū zhōng si-le 75-ge rén  
clash in die -PEF 75-CL person  
(Lit) ‘According to Ø say/claim/report in yesterday’s clash died 75 people.’  
‘It is said/reported that 75 people died in yesterday’s clash.’

It’s obvious that what Siewierska and Papatathi (2011) suggests as a speech act 3PL IMP in English (e.g. 45c.) could be rendered by the constructions like jù shuō in Mandarin. The following is an example from the translation of Harry Porter:

(47)  
a. They say he met vampires in the Black Forest.  
b. jù shuō tā zài hēi sènlin yǔdào-le xīxuēguì  
according to say 3SG:M in black forest meet PEF vampire  
(HP translated p. 43)

Syntactically, jù is a preposition meaning ‘according to’ and it is directly followed by a verb denoting speech act e.g. ‘say,’ ‘claim,’ ‘spread (word)’ and ‘report’. It is therefore reasonable to say that the agent of the speech act and the source of the information is realized by a zero pronoun. The content of the information is in the subordinate clause after the verb shuō, chēng etc. An alternative analysis of (46) is to view ‘jù+verbs
of speech act’ as an evidential marker in Mandarin. Consequently what comes after shuō becomes the main clause. What we argue is that however it is analyzed, the agent of the speech act is realized by a zero pronoun with the purpose to obliterate the exact identity of the source of information.

So far we have seen that expressing the real agent with a zero pronoun is a widely used impersonalizing strategy in Mandarin. Since the agent-backgrounding function is iconically encoded in the use of zero form, unlike the verbal strategies used in the bèi-construction, PVCs or the English short passives, the zero pronoun IMP C places no demand on the semantic transitivity of the event it may express. On the other hand, the referential range of the implied agent in the construction is also flexible: it can be generic, first or second person.

3. Concluding remarks

In this paper we have introduced the major referential impersonalization strategies in Mandarin. Table 2 intends to capture the distribution of Mandarin forms that encode various type of impersonality relative to the corresponding English forms.

We see that broadly speaking there are many points of similarity between the Mandarin way of encoding impersonality by way of nominal strategies and the means employed in English (and also other European languages). For instance, Mandarin utilizes both distinct person forms and regular personal pronouns to encode impersonality. It has a morpheme/generalized noun rén that is exclusively used to denote a general body of humans much in the same way English uses one and German uses man. Like many European languages Mandarin employs a first person plural form, and a second person form to create a sense of impersonality. Although the 3pl tā-men only has a very restricted impersonal use, two other indefinite pronouns that have 3rd person reference biérén and rēnjiā have an impersonal use that is analogous to 3pl imps in many European languages. Like Modern Hebrew, Spanish and many pro-drop languages Mandarin allows an iconic way of expressing impersonality by using a zero pronoun in the subject position.

Against the background of these broad similarities certain phenomena have emerged as possible features of the Mandarin way of expressing impersonality, though further study both within the language and across languages is needed to substantiate the posited claims. The most striking thing is the yǒu rén construction. Although the use of existentials to background the agent is also found in various other languages, including English and Portuguese, the Mandarin yǒu rén construction is unique both

---

11. This was suggested by one of the anonymous reviewers.
Table 2. Mandarin IMPCs contrasted with English forms

<table>
<thead>
<tr>
<th>Types of IMP</th>
<th>English Forms</th>
<th>Mandarin Forms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Form</td>
<td>Referential Type</td>
</tr>
<tr>
<td>Generic (±Sp./Add.)</td>
<td>we 1PL ±F</td>
<td>wǒ-men 1PL ±F</td>
</tr>
<tr>
<td>Generic</td>
<td>you 2SG –F</td>
<td>nǐ 2SG –F</td>
</tr>
<tr>
<td>Law-like propositions</td>
<td>he 3SG +F</td>
<td>zero pron., rěn gen.n ±F</td>
</tr>
<tr>
<td>Generic</td>
<td>one IND.SG +F</td>
<td>rěn gen.n ±F</td>
</tr>
<tr>
<td>Law-like propositions/characterizing</td>
<td>one IND.SG +F</td>
<td>yì-ge rěn gen.n ±F</td>
</tr>
<tr>
<td>generic –Sp./Add.</td>
<td>everybody IND.SG ±F</td>
<td>dàjia IND.PL ±F</td>
</tr>
<tr>
<td>–Sp./Add.</td>
<td>other(s) IND.3SG/PL ±F</td>
<td>biérén 3SG/PL ±F</td>
</tr>
<tr>
<td>Generic (±Sp./Add.)</td>
<td>people GEN.N –F</td>
<td>rěn-men GEN.N ±F</td>
</tr>
<tr>
<td>Generic (±Sp./Add.)</td>
<td>/ / /</td>
<td>zero pron. Pron. ±F</td>
</tr>
<tr>
<td>Universal</td>
<td>they 3PL –F</td>
<td>rěn-men GEN.N ±F</td>
</tr>
<tr>
<td>Corporate</td>
<td>they 3PL –F</td>
<td>zero pron. GEN.N ±F</td>
</tr>
<tr>
<td>Vague</td>
<td>they 3PL –F</td>
<td>yǒu rén 3PL –F</td>
</tr>
<tr>
<td>Inferential</td>
<td>someone, somebody 3SG –F</td>
<td>yǒu rén Hybrid ±F</td>
</tr>
<tr>
<td>Specific-existential</td>
<td>they 3SG –F</td>
<td>yǒu rén Hybrid ±F</td>
</tr>
<tr>
<td>Short Passive</td>
<td>verbal +F</td>
<td>jù shuō(say)/ bǎodào (report) etc. Prep.+ zero pron.+ speech act verb –F</td>
</tr>
</tbody>
</table>

(Sp.=speaker; Add.=addressee; Gen.N =generalized noun; Ind.= indefinite pronoun; F=formal)
in form and in use. In form the 有 人 construction is a hybrid way of backgrounding the agent: a combination of the existential verb 有 (a verbal strategy), a generalized noun 人 (a nominal strategy) and a predicate whose agent is replaced by 有 人. In use the 有 人 construction seems to have much wider range than existentials in European languages. It is the preferred impersonalizing strategy to denote vague, specific, inferred and speech act 3pl impersonal uses identified in the European languages. Secondly, Mandarin does not have dedicated referential IMP forms other than the generalized noun 人 which is restricted to law-like propositions. All of the other forms that can be used impersonally are open to both impersonal and definite interpretations at the phrase level. The discourse context plays an important role in deciding which meaning is the intended one and in deciding the referential range of the backgrounded agent. The lack of dedicated referential IMP forms is probably a reflex of having to rely on forms that are less grammaticalized than their European counterparts for expressing impersonality. Mandarin does not have a dedicated indefinite pronoun as the English one; its rough equivalent 一个 人 ‘one-cl. person’ is primarily an indefinite noun. The plural pronouns that can be used impersonally i.e. 二 人, 人 and 大家 demonstrate less pronominality and more noun features in that they all have rather concrete semantic content; their impersonal meaning is tainted by the literal meaning of the form and they carry sociolinguistic connotations. To what extent these features are different from other languages and how and to what extent these features are related to the typological features of Mandarin constitute important topics for further research.

References


Impersonal constructions in some Oceanic languages

Claire Moyse-Faurie
CNRS-Lacito

Kanak and the Polynesian languages exhibit a wide range of impersonal constructions which may involve: (i) the lack or frequent omission of arguments; (ii) optional impersonal construction of monovalent verbs or intransitive construction of bivalent verbs, associated with different meanings; (iii) differential agent marking such as oblique adjuncts or agents/experiencers expressed as a possessor; (iv) the use of impersonal pronouns or non-referential ‘dummy’ pronouns. These constructions are considered to be impersonal from two perspectives: (a) as constraints (impersonal verbs, dummy pronouns) and options (labile verbs) offered by a language system, and (b) as discourse strategies offered to the speaker of a language to make the agent impersonal (through its omission, by making it peripheral as an oblique adjunct or a possessor).

Keywords: Oceanic languages; labile verbs; argument omission; differential agent marking; impersonal pronouns

1. Introduction

The Kanak languages spoken in New Caledonia and the Polynesian languages belong to different subgroups of the Oceanic language family, one of the main branches of the Austronesian family. They possess several morpho-syntactic features which contribute to a rich array of impersonal constructions. From an Indo-European point of view the most unusual feature of these languages is the fairly frequent use of verbs with no overt subjects and objects. In the case of the Polynesian languages and in the Kanak languages spoken in the South of the Mainland of New Caledonia, this holds despite that fact that these languages have no verbal person marking. The Kanak languages of the Centre and the North of the Mainland...

---

1. I would like to thank the editors and an anonymous reader for their fruitful comments, as well as Anna Siewierska and Ekkehard König for their help in improving my English version.
land, as well as Iaai (Uvea, Loyalty islands), by contrast, do have obligatory pro-
nominal clitics, but only when the relevant arguments refer to a human being. The
second important morpho-syntactic characteristic of the Kanak and Polynesian
languages pertinent to this paper is that, like most Oceanic languages, they have
no syntactic passive, that is, a construction which demotes a would-be subject to
peripheral position and places another noun phrase in the subject slot. However,
the semantics generally associated with the passive voice may be obtained by the
non-expression of the agent, without additional marking of the verb, or be sig-
naled by a type of flexibility in the argument structure (lability), a lexical charac-
teristic of quite a few verbs which are licensed in different constructions without
derivational affixes.

As we shall see some of the constructions to be discussed in this paper qualify
as impersonal by virtue of lacking a referential subject, others by virtue of involving
agent backgrounding, and yet others on both counts. My interpretation of agent-
backgrounding is, however, quite broad be it mainly in line with many agent-based
approaches in that I include under agent defocusing the speakers choice to use of
an anticausative or ambitransitive verb as opposed to a transitive one (see The intro-
duction to this volume). We will begin in §2 with constructions which have only
non-referential subjects or no subjects at all. In §3 we will consider impersonality
in relation to labile verbs, so characteristic of the Kanak and Polynesian languages.
Then in §4 we will turn to the issue of differential agent marking which I also view
as a form of agent backgrounding, since the agent is expressed not as an argument
but, either as a possessor or as an oblique adjunct. Another form of agent defocus-
ing motivated by politeness factors will be presented in §5. Finally §6 will deal with
impersonal pronouns and their occurrence in relative clauses.

2. Impersonal constructions with restricted or no overt subjects

Impersonal verbs, i.e. verbs which cannot have any argument (a-valent verbs) or verbs
which are restricted in their choice of arguments are found in both Kanak and Poly-
nesian languages.

2.1 No argument at all

In Western Polynesian languages, there are a few a-valent predicates referring to mete-
orological phenomena or temporal segments, such as East Uvean 'ua ‘rain', matagi
‘wind, be windy', Pileni/Taumako malamake '(be) morning', Tongan mofuike 'earth-
quake', etc., in each case preceded by a tense-aspect marker:
East Uvean (Wallis, Nuclear-Polynesian)²

(1)  a. ’E matagi.  
    NPAST wind  
    ‘It is windy.’

Pileni/Taumako (Solomon islands, Polynesian Outlier) (Naess 1998: 161)

b. Ko malamake.  
    PFV morning  
    ‘Morning came.’

Tongan (Polynesian, Tongic subgroup) (Churchward 1953: 70)

c. Na’e mofuike.  
    PAST earthquake  
    ‘There was an earthquake’. (Lit. It earthquaked).

In Kanak languages, such a-valent predicates – mostly nominal predicates – are also attested. They may or may not be preceded by tense-aspect markers, as is shown by the following examples from Cèmuhi and from Xàràcuù.

Cèmuhi (Centre of the Mainland, New Caledonia) (Rivierre 1980: 63)

(2)  a. Utè.  
    rain  
    ‘It rains.’

b. Bwén.  
    night  
    ‘It’s dark.’

Xàràcuù (South of the Mainland, New Caledonia) (Moyse-Faurie 1995: 144)

(3)  a. Nää jômââdéé.  
    PAST.IPfv flood receding  
    ‘The flood had been receding.’

b. Wà mò.  
    PFV night  
    ‘Night is falling.’

In Kanak languages such meteorological phenomena or verbs denoting temporal segments may also occur in more explicit monovalent constructions, as in (4b), even

². Except when otherwise indicated, examples come from my own field notes.
though the subject can be tautological, i.e. already given by the predicate, as in (4c), both examples contrasting with the a-valent construction (4a):

Xárâcùù

(4) a.  *Puxârû, wâ kwiè!*
    run  PFV rain
    ‘Get running, it is raining!’

b.  *Kwiè xwa.*
    rain  fall
    ‘It rains.’

c.  *Daa wâ daa.*
    day  PFV day
    ‘It’s getting light.’ (Lit. day becomes day.)

2.2 Verbs only occurring with dummy pronouns as subjects

‘Dummy’ pronouns are not attested in Polynesian languages, since no pronominal form is required when the subject is not referential. The situation in Kanak languages is quite different. In these languages we do find impersonal pronouns (see § 6.), even though constructions without subjects are possible, as we have seen with meteorological verbs and as we will see in § 5.1. with active verbs. We also find ‘dummy’ pronouns, that is obligatorily non-referential pronominal subjects. The use of such non-referential subjects is one of the typical impersonal constructions found across languages.

In Xárâcùù, for example, two verbs only admit such a dummy 3rd singular pronominal argument in constructions expressing potential modality. These verbs are *dù* ‘be possible’, ‘be enough’ and *xwaé* ‘depend’; they are often followed by complements, either juxtaposed nominalized verbal phrases in which the agent may be expressed as a possessor as in (5c), or clausal adjuncts introduced by the complementizer *mè* (5b):

Xárâcùù

(Moyse-Faurie 1995:92)

(5) a.  *È dû.*
    *3sg possible*
    ‘It’s possible,’ ‘It’s enough.’

b.  *È dû kau mè rè mè wèrè-nâ?*
    *3sg possible INTERR COMP 3sg come COM-1sg*
    ‘Is it possible for him to come with me?’

c.  *È sii dû kêè-fârâ rè ri pèci a.*
    *3sg NEG possible NMLZ-read POSS 3PL book DEIC*
    ‘They cannot read this book.’ (Lit. it is not possible, their reading this book)

In other Kanak languages, not even a dummy pronoun is required with verbs of similar meaning but these verbs are followed by a clausal adjunct, which, as it was the case in (5b), includes a referential agent:
Impersonal constructions in some Oceanic languages

Nyelâyu (North of the Mainland, New Caledonia) (Ozanne-Rivierre 1998:47)

(6) Anooyu me hava temwa bwee-r.
possible COMP 1PL.EXCL go top-its
'We can walk on top of it.' (Lit. possible that we walk on its top)

In Polynesian languages, no complementizer is needed with such modal verbs; the clausal adjunct is just juxtaposed, be it a verbal clause (7) or a nominal clause (8) as, for example, in East Uvean with verbs such as *lagi* ‘seem' or *mātu‘aki* ‘be useful'. The resultant constructions are very similar to the corresponding English ones, but without any dummy pronoun.

East Uvean

(7) 'E lagi ne'e tōtō ia Sioli ke talanoa atu ki te
npast seem PAST forget ABS Sioli that tell DIR OBL SPEC
teuteu o te fakahāhā faiva o te Pasifika.
preparation POSS SPEC show arts POSS SPEC Pacific
'It seems Sioli forgot to tell you about the preparations for the Pacific Festival.'

(8) 'E lagi ko tonatou 'api 'āē atu.
npast seem PRED their home DEIC DIR
'It looks like it is their house.'

The constructions presented so far include lexically impersonal verbs, which cannot occur with referential subjects. Let us turn now to verbs which may occur in different types of constructions.

3. Labile verbs

We will use the terms ‘labile verbs’ to cover different types of verbs which can occur in two (or more) different constructions, without derivation, but with a change in meaning or in their orientation.

In fact, besides the intransitive impersonal constructions involving modal verbs which require a dummy 3rd singular pronoun just discussed, quite a number of intransitive verbs may occur with or without an argument. Without an argument, in an a-valent construction, they have a non-dynamic impersonal meaning, while an overt subject argument is always referential, as will be shown in §3.1.

In some Kanak languages, there are also verbs which occur both in intransitive and in transitive constructions, in correlation with a change in the semantic role of the subject (§3.2.). Such labile verbs may be considered as impersonal when they occur in an intransitive construction: their subject does not refer anymore to an agentive participant, in contrast to their subject argument in a transitive construction. Moreover, the
event in the intransitive construction is spontaneous, or conveys a passive, resultative or middle meaning. However, such labile verbs are only few in number. In most cases, derivation is necessary to allow an intransitive verb to become transitive or vice versa.

3.1 Optional impersonal construction of intransitive verbs

In many cases, the same verb may occur in impersonal/a-valent constructions, without any dummy pronoun, and in intransitive ones in combination with a referential subject. In Kanak languages, the relevant lexical items are stative verbs denoting qualities, such as Xârâcùù mëgi 'warm', müü 'cold and humid' and xùpè 'cold', on the one hand, and verbs related to temporal situations such as Xârâcùù xutuè 'be a long time' and cokwa 'be finished' or Nyelâyu toven 'finish', on the other. Examples (9a), (10a) and (11a) are instances of impersonal a-valent constructions, contrasting with (9b), (10b) and (11b), which contain a referential subject in an intransitive construction.

Xârâcùù (Moyse-Faurie 1995:71)

(9) a. Xùpè sé na amù.
cold big PAST yesterday
‘It was very cold yesterday.’

b. Nà xùpè.
1SG cold
‘I am cold.’

(10) a. Wà xutuè.
PFV last.long
‘It’s been a long time.’

b. È xutuè rô a.
3SG last.long LOC here
‘S/he has been here for a long time.’

Nyelâyu (Ozanne-Rivierre 1998: 165)

(11) a. Bwa harivan cu toven.
ASS NEG still finish
‘It is not finished yet.’

b. Hon toven timi mwa-ja.
PFV finish DEIC house-1PL.INCL
‘Our house is finished.’

Similarly, in East Uvean, stative verbs can occur without any argument (12a), or take a referential absolutive argument as in (12b):

East Uvean
I mpersonal constructions in some Oceanic languages

(12) a. *Kua vevela.*
   PFV hot
   ‘It is hot.’

b. *Kua vevela te vai.*
   PFV hot SPEC water
   ‘Water is hot.’

Note that in the Kanak language Drehu, an a-valent construction may receive either an impersonal or a personal interpretation, as in example (13), but in this later case, it can only refer to a 3rd singular unidentifiable individual.

Drehu (Lifu, Loyalty islands) (Sam 2007: 267)

(13) *Kola hnoétr.*
   PROG cold
   ‘It’s cold.’ or ‘S/he is cold.’

In Fagauvea (West Uvean), the Polynesian Outlier spoken in Uvea (Loyalty islands), the meteorological verb *ua* ‘rain’ can also occur in two types of constructions without derivation. With no argument (14a), the verb only expresses a meteorological event, while in (14b), the argument referring to a specific location is affected by the event, which is perceived as being more dynamic.

Fagauvea (Uvea, Polynesian Outlier) (A. Djoupa pers. com.)

(14) a. *E ua.*
   Npast rain
   ‘It is raining.’

b. *E ua dona potu.*
   Npast rain his place
   ‘His place is covered with rain.’
   (Lit. [it] rains [on] his place)

In Xârâcùù, the basic word order is SVO. However, existential and non-existential predicates are V-initial. They occur without any argument (15a) in an impersonal construction, or take an unmarked postposed argument (15b); these verbs may also express possession and non-possession when the postposed argument is a possessive noun phrase (15c):

Xârâcùù (Moyse-Faurie 1995: 142)

(15) a. *Va siè*
   ASS not exist
   ‘It doesn’t exist.’

b. *Wâ siè laasi.*
There is no rice left.

I don't have any basket. (Lit. not exist my basket)

Such impersonal constructions are also attested in some Polynesian languages in connection with the same kinds of verbs. When no overt argument is expressed, these verbs may only refer to a non-dynamic event without any referential indication. This is the case in East Uvean, with verbs such as ‘osi ‘be finished’, fe‘auga ‘be useful’, uta ‘be enough’, noa ‘be nothing, vain’, which can occur either in an a-valent construction (16a), or take an absolutive argument referring to the patient, with an optional oblique complement referring to the experiencer (16b):

East Uvean

(16) a. ‘Ena
    Npast nothing
    ‘It is nothing.’

b. Fakakata‘i fëia kae noa ia ki te tama.
    joke so but be nothing deic obl spec boy
    ‘[They] were kidding but this didn’t mean anything to the boy.’

The experiencer can also be expressed by a possessor in a nominalized oblique phrase (17b), but in any case as an argument.

East Uvean

(17) a. Kua ‘osi.
    Pfv finish
    ‘It is finished.’

b. Kua ‘osi te fenua nei i tuku mamata.
    Pfv finished spec country deic obl my visit
    ‘I just visited this country.’
    (Lit. the country is finished as far as my visit is concerned)

3.2 Verbs employed both transitively and in transitively

Verbs occurring in both transitive and intransitive constructions (like English melt or French brûler ‘burn’) without any formal change in the verb are usually called “labile” or “ambitransitive” verbs (cf. Creissels 2006). Depending on the construction these verbs may have different orientations. I consider the intransitive use of such labile verbs as impersonal since the subject cannot refer to an animate agent but only to an inanimate entity (see also the discussion of anticausatives in the Introduction to this volume).
Xârâcùù has a few such labile verbs, consisting of resultative vs, causative pairs such as têi 'be empty, empty', xwi 'exist, build', cokwa 'be finished, finish' nàâbu 'begin', kê 'be burned, burn', sùù 'suffer, treat', xwêê 'fall, pour', and English loanwords such as sukwa 'be sugared, sugar'. Some of these verbs are known to have labile counterparts in many other languages (Haspelmath 1993). In intransitive constructions (cf. (18a) and (19a)) the verb has a passive, resultative or middle meaning, and the subject always refers to the patient. In transitive constructions (cf. (18b) and (19b)), the verb is active, with the agent as subject, and the patient as object:

(18) a. Laasi kwètaa.
    rice salt
    ‘Rice is salted.’

    b. Nà kwètaa laasi.
       1sg salt rice
       ‘I am salting the rice.’

(19) a. Ké têi.
    basket empty
    ‘The basket is empty.’

    b. Nà têi kê.
       1sg empty basket
       ‘I am emptying the basket.’

According to Du Feu (1996:65–66), Rapanui verbs may occur without any subject, either in the case where “the subject of a verb is known from the immediate context, linguistic or pragmatic”, as in (20a), or in impersonal constructions. In such impersonal constructions, the object of a transitive predicate, normally marked by the particle i (20a), will take “the marking of the first valency element, i.e. the zero marking of the subject” as in (20b). Here again, the intransitive construction in (20b) conveys a non-dynamic or resultative meaning, contrasting with the active meaning of the transitive construction, even though the predicate undergoes no derivation.

Rapanui (Eastern Polynesian) (Du Feu 1996:65)

(20) a. He heriki i te ʻana hai mauku.
    TAM strew OBJ SPEC cave with grass
    ‘(We) used to cover the floor of the cave with tufted grass.’

    b. He heriki te ʻana hai mauku.
       TAM strew SPEC cave with grass
       ‘The cave was strewn with tufted grass.’

Besnier (2000:127) describes similar constructions in Tuvaluan: “[…] the verbs are in their transitive use dynamic verbs, but in impersonal constructions they acquire a
non-dynamic meaning”. Since in Tuvaluan, as in other Polynesian languages, “verbs of all valencies may appear without an agent when this agent is generic, unidentifiable, non-referential, or simply not worth identifying” (id.:126), the following example has two interpretations:

Tuvaluan (Polynesian, Proto-Ellicean) (Besnier 2000: 127)

(21) *Te saamala ne tuku i lalo i te moega.*
    SPEC hammer PAST place OBL under OBL SPEC bed
    ‘Someone placed the hammer under the bed.’
    Or: ‘The hammer is under the bed.’

Only the second interpretation is relevant to our understanding of impersonal constructions.

In the following Tokelauan example, no explicit arguments are expressed, and outside of a specific context, the verb *fanau* could mean either ‘give birth’ or ‘be born’, that is, the verb *fanau* may have a dynamic interpretation with a non-expressed agent, or a non-dynamic interpretation with a non-expressed patient.

Tokelauan (Nuclear Polynesian) (Hooper 2000: 167)

(22) *Na fanau ifo loa, hiki loa lava oi kave.*
    PAST give birth DIR at once lift at once INT SEQ take
    be born
    ‘As soon as [she (ABS)] gave birth, [the men (ERG)] picked [it (ABS)]
    up and took [it (ABS)].’ Or: ‘As soon as [the child (ABS)] was born, …’

Hooper (2000:167–168) explains the conditions under which such very common omissions are possible in Tokelauan as follows: “The woman who is about to give birth and the child itself have been mentioned five clauses earlier, the men, three clauses earlier. The only ‘obligatorily’ omitted NP is the ergative argument of *kave* which is subject to Equi.” (See § 5.2. below for other examples of zero-anaphora constraints.)

4. Differential agent marking

Two main types of differential agent marking are attested in Kanak and the Polynesian languages, and are relevant for impersonal constructions in the sense that they both background the role of the agent, which is no longer expressed as a plain argument. Moreover, its degree of agentivity is reduced or becomes zero.

In the first type the agent or the experiencer is still an argument morphosyntactically, or is expressed as a possessor in a noun phrase (§ 4.1.). This possibility is attested in Kanak languages for a number of verbs (§ 4.1.1.). As first described by Duranti (1994) for Samoan, the Polynesian ergative languages are known for what is now called ‘ergative case avoidance’, consisting in the agent expressed as a possessor in
a noun phrase which is the unique – absolutive – argument of the verb, instead of being an ergative argument by itself; our examples of this construction are from East Futunan (§ 4.1.2.). In the second type of differential agent marking experiencers or agents may be expressed as oblique adjuncts rather than as arguments. Thus, in Polynesian ergative languages, the agent may be marked in the oblique case in an intransitive construction instead of being expressed as an ergative argument in a transitive construction (§ 4.2.). In both of the above types of differential agent marking the semantic orientation of the verb itself does not change with the way the agent/experiencer is expressed, in contrast to what we found with labile verbs. The only difference lies in the encoding of the experiencer/agent, i.e. whether the verb applies to a more ‘patientive’ or to a more ‘agentive’ experiencer who is characterized as being more or less affected, depending on the way the relevant phrase is marked morpho-syntactically (as a plain argument, as a possessor or as an oblique adjunct).

4.1 Agent or experiencer expressed as possessor

4.1.1 In Kanak languages

With a few Xârâcùù verbs there is a choice between a genitive and a nominative experiencer, which has clear semantic effects. A nominative experiencer describes a controlled emotional state, while one marked as a possessor strongly expresses the experiencer’s deeper and uncontrolled feelings. Xârâcùù verbs manifesting this choice for the encoding of the experiencer are verbs of perceptions or emotions such as saa ‘bad, bad looking’, ‘feel bad’; kwèti ‘be tired’, ‘feel tired’; wîrî ‘disgusting’, ‘feel disgusted’; mârâ ‘be worried’, ‘feel dizzy’. The experiencer expressed as a possessor is totally affected and powerless (examples 23b and 24b), in the way described by Creissels (2007: 28) for his ‘affective covert impersonals’, while examples (23a) and (24a) are less emotional, with a more agentive experiencer expressed as a nominative argument.

Xârâcùù (Moyse-Faurie 1995: 72)

(23) a. Nâ kwèti û-sööpö rè xöu 1sg tired NMLZ-wash poss clothes
   ‘I am tired of washing clothes.’
   b. Nâ sii fè ti nuò dóbwa wà-nâ kwèti. 1sg neg go at bush because inner-1sg tired
   ‘I don’t go to the bush because I feel tired.’

(24) a. Nâ mârâ-ri kèè-paii a. 1sg be worried-TR NMLZ-sick DEIC
   ‘This disease worries me.’
    b. Wâ-pa xûuchî mârâ. inner-coll child be worried
   ‘Children are feeling dizzy.’
Similarly, the Xârâcùù verb xöru ‘good, well’ may have a non-dynamic interpretation in combination with a pronominal argument (25a); it may also mean ‘be happy’ with an experiencer expressed as a possessor (25b), and it may mean ‘be pleasant’ with an oblique adjunct experiencer as in example (25c), where it is introduced by the benefactive preposition xù, a possibility I will discuss in § 4.2.3.

Xârâcùù

(25) a. È xöru.
   3SG good
   ‘It’s good/fine’ or ‘S/he is good looking.’

   b. Wâ-nâ xöru.
      inner-1SG good
      ‘I feel happy.’

   c. Xiti sé va xâñi xöru iï xù rè nà.
      feast big ass often good always BEN ass 1SG
      ‘A big feast is always pleasant to me.’ (Moyse-Faurie 1995:82)

In Nêlêmwa, the transitive verb pweede '(re)turn' may occur with only one argument, which refers to the experiencer expressed as a possessor:

Nêlêmwa (North of the Mainland, New Caledonia) (Bril 2002: 160)

(26) Pweede hna-nanamiwo i na.
      turn NMLZ-think POSS 1SG
      ‘I changed my mind.’ (Lit. turn my thought)

4.1.2 In Polynesian languages

As was first shown by Duranti (1994) for Samoan, marking the agent as the possessor in a genitive noun phrase is another way of backgrounding the role of the agent in ergative Polynesian languages. Here again, the agent is morphologically expressed as the possessor. With intransitive verbs, such sentences express a low degree of agentivity, i.e. an involuntary process.

East Futunan

(27) a. Kua puli lana sele.
      PFV absent his knife
      ‘He lost his knife (accidentally).’ (Lit. his knife is lost)

3. The verb xöru is also used in a modal impersonal construction, with a dummy pronoun as subject and a complement clause introduced by a complementizer: è xöru mè altogether means ‘it would be better if’:

   d. È xöru mè ke wita fè!
      3SG well COMP 2SG must not go
      ‘You should not go away!’ (Lit. it is well that you must not go)
The East Futunan verb *puli* must undergo causative derivation in order to license an ergative argument denoting high, conscious agentivity:

\[ (b) \quad \text{Kua fakapuli e i a lana sele talie e le'ése} \]

\[ \begin{array}{llllll}
\text{PFV} & \text{lose} & \text{ERG} & \text{3SG his knife} & \text{because NPAST NEG} \\
\text{fia} & \text{ga'oi.} & \text{feel like work} \\
\end{array} \]

‘He lost (intentionally) his knife because he doesn’t feel like working.’

With such bivalent verbs, there is always the choice of using an ergative construction (28a), which emphasizes the role of the agent, while the possessive construction (28b) backgrounds this role and is considered to be much more polite (cf. Moyse-Faurie 2000).

**East Futunan**

(28) a. \( E \text{ feave'aki e Atelea ana fakapaku i lamatu'ā.} \)

\[ \begin{array}{llllll}
\text{NPAST peddle} & \text{ERG Atelea his doughnut on road} \\
\end{array} \]

‘Atelea is peddling his doughnuts along the road.’

b. \( E \text{ feave'aki a fakapaku a Atelea i lamatu'ā.} \)

\[ \begin{array}{llllll}
\text{NPAST peddle} & \text{ABS doughnut POSS Atelea on road} \\
\end{array} \]

‘Atelea is peddling his doughnuts along the road.’

(Lit. Atelea’s doughnuts are peddled along the road)

Of course, example (28b) has two possible interpretations: the doughnuts (*fakapaku*) may be sold by Atelea, or it is Atelea’s doughnuts that are being sold by someone not mentioned in the sentence, since the verb *feave'aki* remains potentially bivalent. But this latter interpretation is never a first choice. Another example of the agent expressed as possessor is given in (29).

**East Futunan**

(29) \( Na \text{ tao' lana puaka lasi ke ma'iloga ai lona tagata} \)

\[ \begin{array}{llllll}
\text{PAST cook his pig big so that show ANAPH his manhood} \\
\text{i le fakatasi.} & \text{OBL SPEC feast} \\
\end{array} \]

‘He baked a big pig for the feast to show what a fine man he is.’

(Lit. was baked his big pig…)

In Tuvaluan, Besnier (2000: 283) describes a similar choice for the expression of the agent, which occurs as possessive modifier in the absolutive noun phrase argument in an intransitive construction (30a) instead of the ergative pronominal argument in a transitive construction (30b).
Tuvaluan (Besnier 2000: 283)

(30) a. Ne kkati telotou niu.
   past cut spec.poss.3pl coconut tree
   ‘They cut down the coconut tree.’ (Lit. cut down their coconut tree)

b. Ne kkati nee laatou te niu.
   past cut erg 3pl spec coconut tree
   ‘They were the ones who cut down the coconut tree.’

The English translation of (30b) underlines the ‘high degree of responsibility’ of the agent, its control and authority in the event described by the clause.

In ergative Polynesian languages the expression of the agent as possessive modifier, rather than as ergative argument, is very often used to background its role, in order to manifest a socially polite behavior, making it as much as impersonal as possible!

4.2 Oblique marked agents

According to Mosel & Hovdhaugen (1992: 424–425) and as noted in Malchukov (2008: 207 and this volume), a few bivalent Samoan verbs may occur either in an ergative/transitive construction with one argument in the absolutive and one in the ergative, or in an intransitive construction with one absolutive argument and an oblique adjunct, without derivation.


(31) a. Na tapuni e le matagi le faitoto’a.
   past close erg spec wind spec door
   ‘The wind closed the door.’

b. Na tapuni i le matagi le faitoto’a.
   past close obl spec wind spec door
   ‘The wind closed the door.’

In East Futunan, similar constructions are attested, as shown in (32a–b), but their meanings are different from the Samoan ones. In (32b), the ergative argument is not mentioned anymore (no agent is expressed) and the event becomes impersonal. The noun phrase i le matagi ‘due to/because of the wind’ is just an oblique complement expressing an involuntary causation and no longer the agentive argument as it is the case in (32a).

East Futunan (Moyse-Faurie 1992: 216)

(32) a. Na tapono le matapā e le matagi.
   past close spec door erg spec wind
   ‘The wind closed the door.’

b. Na tapono le matapā i le matagi.
   past close spec door obl spec wind
   ‘[Someone] closed the door because of the wind’
Consequently, these two constructions in East Futunan are not a case of differential agent marking consisting in a choice between an ergative vs. an oblique agent, but just a case of omission of the ergative argument, one of the commonly attested strategies used to background the agent in Polynesian languages; the oblique complement in (32b) barely conveys an involuntary causation, not an agentive volition, which could only be expressed by the addition of an ergative agent.

c. \textit{Na tapono le matapā e Soane i le matagi.}
\hspace{1cm} \text{PAST close SPEC door \text{erg} Soane OBL SPEC wind}
\hspace{1cm} \text{‘Soane closed the door because of the wind’}

However, there is a way of preventing the expression of an ergative agent. It consists in a derivational process. East Futunan verb \textit{tapono} ‘close’ may be combined with the resultative prefix \textit{ma-}, giving rise to an intransitive verb, \textit{mapono} ‘be closed’. In contrast to (32b), no ergative argument can be added to (32d):

d. \textit{Na ma-pono le matapā i le matagi.}
\hspace{1cm} \text{PAST res-close SPEC door OBL SPEC wind}
\hspace{1cm} \text{‘The door has been closed because of the wind.’}

However, this intransitivizing process involving Proto-Polynesian *ma- is no longer productive in present-day Polynesian languages and is only possible with about 20 verbs. The resultative prefix \textit{ma-} expresses a sort of impersonal passive, as shown again in the East Uvean examples (33b) and (34b). Whereas no ergative argument can occur with such derived verbs, an oblique adjunct referring to an involuntary causation may be added (34b). With the non-derived verbs (33a) and (34a), by contrast, the ergative-marked argument refers to an agent highly responsible for his/her actions.

\textbf{East Uvean}

(33) a. \textit{Kua pelu te lauākau e te finematuā.}
\hspace{1cm} \text{PFV fold SPEC pandanus leaf \text{erg} SPEC old woman}
\hspace{1cm} \text{‘The woman folded the pandanus leaves.’}

b. \textit{Kua ma-pelu te fala.}
\hspace{1cm} \text{PFV res-fold SPEC natte}
\hspace{1cm} \text{‘The plait is folded.’}

(34) a. \textit{‘E huhua e te tamasi’i te pāketē.}
\hspace{1cm} \text{NPAST \text{red,pour} \text{erg} SPEC child \text{spec} bucket}
\hspace{1cm} \text{‘The child is pouring the water out of the bucket.’}
\hspace{1cm} \text{(the root *hua is no longer attested)}

b. \textit{‘E ma-hua te pākete i te tamasi’i.}
\hspace{1cm} \text{NPAST res-pour \text{spec} bucket OBL \text{spec} child}
\hspace{1cm} \text{‘The child is spilling the water of the bucket (accidentally).’}
\hspace{1cm} \text{(Lit. the bucket is spilled due to the child)}
Lexically intransitive verbs, on the other hand, have to add a transitivizing affix in order to admit an ergative argument (35b).

East Futunan

(35) a.  
\[ \text{Kua mate le afi i le matagi.} \]
\hspace{1cm} PFV die SPEC fire OBL SPEC wind

‘The fire went out because of the wind.’

b.  
\[ \text{Kua mate-i le afi e le matagi.} \]
\hspace{1cm} PFV die-TR SPEC fire ERG SPEC wind

‘The wind blew up the fire.’

The transitivized verbs behave just like the lexically transitive ones, and may occur with both an ergative argument (the agent) and an oblique adjunct referring to an involuntary causation.

4.3 Multiple choices for one and the same verb

In East Futunan, a few intransitive verbs such as ‘aoga ‘be useful’ or galo ‘forget’ may be employed in different constructions, each one involving a different morpho-syntactic encoding of the experiencer argument. In (36a), the verb galo has one argument with the experiencer coded as a possessor (the agentive/alienable possessive marker a is part of the possessive pronoun laku). This verb may also be employed in an unaccusative construction, with a ‘more affected’ experiencer in the oblique case (36b), and an optional absolutive argument referring to the patient. In (36c), the experiencer is only specified by the agentive argument of the subordinate clause, and the main verb has no argument.

East Futunan

(36) a.  
\[ \text{Na gal o laku tosi i fa leako.} \]
\hspace{1cm} PAST forget my book LOC school

‘I forgot my book at school [it is not very important].’

(Lit. my book is forgotten at school)

b.  
\[ \text{Na gal o fuli i ate au (le tosi).} \]
\hspace{1cm} PAST forget all OBL 1SG (SPEC book)

‘I totally forgot (the book) [and I feel very sorry for that].’

(Lit. is totally forgotten to me (the book))

c.  
\[ \text{Na gal o ke kau ‘aga o fai ni kaso a.} \]
\hspace{1cm} PAST forget that 1SG face COMP do NSPEC.PL necklace

‘I forgot to make some necklaces.’

(Lit. is forgotten that I was about to make some necklaces)
5. **Zero-anaphora**

It is of course very important to make a distinction between zero anaphora due to contextually given reference (§5.1.), and zero-anaphora due to syntactic constraints or pragmatic/behavior (§5.2.). Omission of arguments is a very common phenomenon in many Oceanic languages.

5.1 **Zero-anaphora due to discourse strategies**

More often than not, argument omissions are related to contextually given discourse, and do not imply any change in the semantic role of the remaining arguments. There is no ambiguity, since reference is guaranteed by the context. We are familiar with cases of omission of the object (‘They eat.’), less so with missing subjects, or omission of both the subject and the object.

Omission of the object is exemplified in (37a). As is the case with labile verbs, the lexically bivalent verb has a stative/non-dynamic meaning when it occurs in an intransitive construction. Whereas labile verbs, however, undergo a change in their orientation depending on the construction in which they occur, non-labile bivalent verbs do not. In examples (37), the verb *dapwii* ‘be greedy (for something)’ keeps the same subject and the same orientation whatever the construction, intransitive in (37a) or transitive in (37b).

Xârâcùù (Moyse-Faurie 1995:74)

(37) a. È *dapwii.*
   3sg greedy
   ‘S/he is greedy.’

b. È *dapwii* döö.
   3sg greedy earth
   ‘S/he wants more land.’ (Lit. she is greedy [for] earth)

Omission of the subject of bivalent verbs is also non-ambiguous, since the basic word order in Xârâcùù is SVO. There is no need for choosing a dummy pronoun, or an impersonal pronoun, although the use of the later is always possible, as we will see later in §6.2. Compare, however, examples without subjects like (38a) and (39a), on the one hand, and examples with explicit subjects like (38b) and (39b), on the other. In both cases, the verb has the same orientation, and is not derived. The agent may simply be omitted for pragmatic reasons, it is not syntactically peripheral. There is no change in word order. We are not dealing with any kind of passive voice.
Xârâcûù

(38)  a. Wà cheè kwâjûù-rê.
  PFV pull sheet-poss3sg
  ‘[Someone] is pulling the sheets of the sail.’

b. Apuukwâ wà cheè kwâjûù-rê.
  wind chief PFV pull sheet-poss3sg
  ‘The chief of the wind is pulling the sheets of the sail.’

(39)  a. Wà tara miï mwà ri wà kê xêêdi.
  PFV see pl house 3pl PFV ready totally
  ‘[One] can see these houses which are totally ready.’

b. Mô tara è.
  night see 3sg
  ‘Nightfall caught him.’ (Lit. night sees him)

The omission of the subject is also possible in the other Kanak languages of the South, as for example in Drubéa, where, just as in Xârâcûù, the SVO word order precludes any ambiguity. In example (40), the discontinuous tense-aspect marker té...re signals that the verb is bivalent (otherwise, only the prefix part of the TAM is required); the clause only includes one argument, the object, and the subject position remains empty.

Drubéa (Extreme-South of the Mainland, New Caledonia)

(40)  Tê kwîá-re Paul.
  tam kill-tam Paul
  ‘[Someone] killed Paul.’  (Shintani & Païta 1990:88)

Omission may involve the subject as well as the object. This corresponds to what Lambert (1998) has called “le degré-zéro du verbe fini”. Only the event is expressed, the participants can be inferred from the context.

Drubéa  (Shintani & Païta 1990:88)

(41)  Pâ tóomwéré ngî trá.
  ass cover with earth
  ‘[Someone] cover [them] with earth.’

Drehu  (Sam 2007:52,528)

(42)  a. Kola sa.
  prog cut(tr)
  ‘[Someone] is cutting [something].’

One or two arguments can be added to this Drehu sentence; in the progressive aspect, the patient will follow the verb, without any marking (42b), while the agent is introduced by the agent marker hne- (42c).
b. *Kola pane sipo hne-ng!
   PROG first ask for AGT-POSS.1SG
   (‘I am asking for silence!’)
One simply has to say:

c. \textit{Kola pane sipo!}\tabularnewline \textit{prog} first ask for\tabularnewline ‘Silence!’

In fact in everyday discourse, one prefers to speak in an impersonal way for reasons of humility, that is, without overtly mentioning anybody, even in direct speech as, again, in (45).

Drehu \hspace{1cm} (Sam 2007: 270)

(45) \textit{Ame la ewekë hnapan kola pane ketre ipië…}\tabularnewline \textit{pred spec thing first prog at first int(‘other’) humble}\tabularnewline ‘First of all [I] humble myself…’

This type of omission, involving the speaker, is also attested in Polynesian languages when one wants to be very polite and humble. An East Uvean guest arriving at his host’s, for example, will simply respond to the words of welcome by saying:

East Uvean

(46) \textit{Kua ha’u.}\tabularnewline \textit{pfv come}\tabularnewline ‘Here [I] am.’

These constructions involving the omission of arguments can be considered as impersonal, even when the omission is not due to syntactic constraints and only depends on the speaker.

6. Impersonal pronouns

Few Oceanic languages have special impersonal pronouns such as the French (\textit{on}), German (\textit{man}) or English (\textit{one}) (see Siewierska this volume). Rather they use certain personal pronouns impersonally. I will first discuss the impersonal use of what is reconstructed as the 1st plural inclusive pronoun attested in several Polynesian languages (6.1.) and then the impersonal use of the 3rd plural pronoun in Nêlêmwa (6.2.). Finally, I will consider the rather exceptional impersonal pronoun found in Xârâcûû, whose putative origin is a nominalizing prefix (6.3.).

6.1 Impersonal use of the Polynesian 1sg inclusive pronoun

In Western Polynesian languages, there is a reflex of what was originally a plural inclusive pronoun, reconstructed for Proto-Austronesian as *kita, but which has been reinterpreted in Proto-Polynesian languages as a “1st singular inclusive pronoun”
(Churchward 1953:126), after the introduction of new dual and plural inclusive forms into these languages. Reflexes of Proto-Polynesian *kita are often used in familiar speech to express the emotional involvement of the speaker, indicating “self-abasement, humility, or an appeal for help” as stated by Ulrike Mosel & Even Hovda-haugen in their Samoan grammar (1992:121), instead of the otherwise ordinary 1st singular pronoun. Similar uses are found in other Polynesian languages, such as East Futunan, East Uvean or Tongan.

East Futunan

(47)  
\[E \text{ loto 'ita'ita a kita o tupu ko leia a ne' a.}\]

npast heart angry abs 1sg.incl because deic app person

‘I am very angry because of this man.’

East Uvean

(48)  
\[E \text{'ui mai ke kita 'alu kita.}\]

npast say dir that 1sg.incl go 1sg.incl

‘I am told to go away.’

This kita pronoun may also refer to any person and is often translated into French either by ‘on’ or by ‘tu’ in its impersonal use. In his Tongan grammar, C. Maxwell Churchward lists “the main uses of te/kita as follows: (a) = the indefinite ‘one’ which means, in effect, I (or me) or you or anyone else; (b) = ‘I’ or ‘me’: in the language of politeness or humility” (1953:127).

Indeed, the use of kita is also frequent in proverbs and common sayings:

East Futunan

(49)  
\[E \text{kita moe fā'i i le kāiga na kita nofo}\]

npast 1sg.incl sleep restr obl spec estate past 1sg.incl stay
\[i ai.\]

obl anaph

‘One only sleeps in the house where one lives.’

East Uvean

(50)  
\[E \text{kita fafaga pē te kuli ke ina kai kita.}\]

npast 1sg.incl nourish restr spec dog so that 3sg eat 1sg.incl

‘You feed a dog, so that it (later) bites you.’

(51)  
\[Maumau si'i vaihū lelei kala kita kai\]

spoil emot fish.dish good neg 1sg.incl eat
\[pea kita moe ke a'u ki te 'aho!\]

and 1sg.incl sleep until spec day

‘What a pity to eat so good a dish of fish and not to be able to sleep until the morning!’
An excellent and comprehensive study of both the impersonal/non-referential and the ‘integrative’/inclusive uses of this pronoun PPn *kita is given in Lichtenberk (2005).

6.2 The impersonal use of the 3rd plural pronoun

In some Kanak languages the 3rd person plural pronoun may be used with a similar non-specific interpretation (see Siewierska this volume):

Nêlêmwa (Bril 2002: 160)

(52) *Hla pewede nanamiwo i na.
   3PL turn thought POSS 1SG
   ‘They/someone made me change my mind.’

The use of a third person plural pronouns to express similar cases of impersonal constructions is also attested in Paamese, a language of Vanuatu, in which, according to Crowley (1982: 189), verbs always occur with preverbal prefixes. An impersonal construction will consist of an ‘ambient verb’ bearing the third person plural subject prefix, but with no actor subject expressed, as in (53a) and (53b), which differ from the actor-subject construction (53c), in which both the prefix (indicating in this case the person and the number of the subject), and the subject pronoun are present:

Paamese (North Central Vanuatu) (Crowley 1982: 189–190)

(53) a. *Alonge tahos.
   3PL.REAL.feel 3SG.REAL.good
   ‘It feels good.’

b. *Amunumun Vauleli.
   3PL.REAL.drink Vauleli
   ‘There is drinking going on at Vauleli.’

c. *Kail amunumun Vauleli.
   3PL.S 3PL.REAL.drink Vauleli
   ‘They are drinking at Vauleli.’

6.3 Impersonal pronouns in Kanak languages

Besides the impersonal use of Polynesian *kita ‘1SG inclusive pronoun’ or the impersonal uses of 3rd plural pronouns found in Nêlêmwa and Paamese, true impersonal pronouns seem quite rare in Oceanic languages. However, there are some in a few Kanak languages of the South of the Mainland such as Ajië, Xârâcùù, and Tîrî. I have already presented Xârâcùù examples with no syntactic subjects (see examples (38a) and (39a) above), even though the relevant participants may often be inferred from the context. It is also possible to add an impersonal pronoun, which makes the construction slightly less impersonal than with a zero subject!
Xàràcùù (Moyse-Faurie 1995: 124,153)

(54)  Èè nàà  chùrà mè bìkòr rè pìi-kòfi.
3sg.impers past.ipfv grill and grind ipfv grain-coffee

'We used to grill and grind coffee beans.'

(55)  Dòbwananà èè sòpurù kwé, nà nù na pàâàì rè
when.past 3sg.impers cut water 1sg send past children poss
nà ti xwàrè.
1sg at river

'After the water had been cut off, I sent my children to the river.'

This impersonal pronoun èè is also used as a nominalizing prefix in relative clauses, conferring a resultative meaning to the action expressed by the verb:

Xàràcùù (Moyse-Faurie 1995: 44–45)

(56)  a.  pwì èè-chutâà
banana nmlz-cook

‘cooked banana’ (Lit. banana which is/has been cooked)

The corresponding main clause would have an active orientation:

b.  Èè chutâà pwì
3sg.impers cook banana

‘Someone is cooking banana.’

The relative clause may include an agent, which has to be expressed as a possessor:

c.  pwì èè-chutâà rè anyàà
banana nmlz-cook poss mummy

‘banana cooked by Mummy’
(Lit. banana which is the cooking result of Mummy)

This construction is reminiscent of a type of differential agent marking we described earlier, where the agent is expressed as a possessor.

7. Conclusion

The formal properties of impersonal constructions are language-specific. On the other hand, I have presented several examples of features fairly wide-spread in the world’s languages. Besides impersonal verbs, referring to meteorological or temporal phenomena, we have discussed several cases of labile verbs and of differential agent marking, which demonstrate a high flexibility in the choice of constructions for verbs without any derivation. The main characteristic of the languages discussed in this paper is their lack of constraints regarding the overt (obligatory) encoding of arguments. Many verbs do not require any overt encoding of their arguments. In Kanak languages, only a few verbs require a 3rd singular dummy pronoun when the subject is non-human.
Impersonal pronouns are rare in Kanak languages and non-existent in Polynesian languages. The impersonal use of a 1st person inclusive plural or a 3rd plural pronoun, however, is attested in both Polynesian and Kanak languages.

A distinction was drawn between impersonal constructions as a result of constraints imposed by a system (impersonal verbs; some cases of zero-anaphora; obligatorily expression of ‘dummy pronouns’) and as a result of discourse options offered to the speaker of a language (no argument expressed for contextual or politeness reasons; choice of constructions leading to different degrees of assigning a peripheral status to the agent which is marked as possessor in a noun phrase or as an oblique adjunct, instead of being a plain argument).

We have seen that one of the ways of demoting the agent is to express it as possessor instead of as an argument of its own; it may occur as possessive modifier in a subordinate clause, in a nominalized oblique phrase, in a nominalized relative clause, or as part of a noun phrase argument. This syntactic expression of the agent or the experiencer as possessor is semantically correlated with a reduction of agentivity, for the former, and more emotional involvement of the latter.

**Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABS</td>
<td>absolutive</td>
</tr>
<tr>
<td>AGT</td>
<td>agent marker</td>
</tr>
<tr>
<td>ANAPH</td>
<td>anaphoric</td>
</tr>
<tr>
<td>APP</td>
<td>apposition</td>
</tr>
<tr>
<td>ART</td>
<td>article</td>
</tr>
<tr>
<td>ASS</td>
<td>assertive</td>
</tr>
<tr>
<td>BEN</td>
<td>benefactive</td>
</tr>
<tr>
<td>COLL</td>
<td>collective article</td>
</tr>
<tr>
<td>COM</td>
<td>comitative</td>
</tr>
<tr>
<td>COMP</td>
<td>complementizer</td>
</tr>
<tr>
<td>DEIC</td>
<td>deictic</td>
</tr>
<tr>
<td>DIR</td>
<td>directional</td>
</tr>
<tr>
<td>EMOT</td>
<td>emotional article</td>
</tr>
<tr>
<td>ERG</td>
<td>ergative</td>
</tr>
<tr>
<td>EXCL</td>
<td>exclusive</td>
</tr>
<tr>
<td>IMPERS</td>
<td>impersonal</td>
</tr>
<tr>
<td>INCL</td>
<td>inclusive</td>
</tr>
<tr>
<td>INT</td>
<td>intensifier</td>
</tr>
<tr>
<td>INTERR</td>
<td>interrogative</td>
</tr>
<tr>
<td>IPFV</td>
<td>imperfective</td>
</tr>
<tr>
<td>LOC</td>
<td>locative</td>
</tr>
<tr>
<td>NEG</td>
<td>negation</td>
</tr>
<tr>
<td>NMLZ</td>
<td>nominalizer</td>
</tr>
<tr>
<td>NPAST</td>
<td>non past</td>
</tr>
<tr>
<td>NSPEC</td>
<td>non specific article</td>
</tr>
<tr>
<td>OBL</td>
<td>oblique</td>
</tr>
<tr>
<td>OBJ</td>
<td>object marker</td>
</tr>
<tr>
<td>PFV</td>
<td>perfective</td>
</tr>
<tr>
<td>PL</td>
<td>plural</td>
</tr>
<tr>
<td>POSS</td>
<td>possessive</td>
</tr>
<tr>
<td>PRED</td>
<td>predicative</td>
</tr>
<tr>
<td>PROG</td>
<td>progressive</td>
</tr>
<tr>
<td>RED</td>
<td>reduplicate</td>
</tr>
<tr>
<td>REAL</td>
<td>realis</td>
</tr>
<tr>
<td>RES</td>
<td>resultative</td>
</tr>
<tr>
<td>RESTR</td>
<td>restrictive</td>
</tr>
<tr>
<td>SG</td>
<td>singular</td>
</tr>
<tr>
<td>SEQ</td>
<td>sequential</td>
</tr>
<tr>
<td>SPEC</td>
<td>specific article</td>
</tr>
<tr>
<td>TR</td>
<td>transitive</td>
</tr>
</tbody>
</table>
Impersonal constructions in some Oceanic languages

References

Impersonal constructions in Umpithamu and the Lamalamic languages

Jean-Christophe Verstraete
University of Leuven

In the four Pama-Nyungan languages Umpithamu, Morrobolam, Mbarrumbathama and Rimanggudinhma there is a core set of impersonals centred around experiencer object constructions. They describe involuntary physical processes, and are formally characterized by lack of nominative pronominal cross-reference, and optional absence of ergative agent nominals. In addition, systematic lack of nominative cross-reference is found in constructions with inanimate agents in all four languages, and in experienced action constructions in Umpithamu, in both cases with ergatively-marked nominals. It is argued that nominative cross-reference is the basic criterion for subject status, with ergative marking merely indicating agent status. Given the lack of any specific valency-changing morphology, impersonals with ergatively-marked nominals are functional equivalents of a voice mechanism, with agents demoted from subject status. This process has developed furthest in Umpithamu where the experienced action construction is systematically available as an alternative construal for a subset of transitive clauses.

Keywords: impersonal; experiencer object; inanimate agent; passive; Umpithamu; Lamalamic

1. Introduction

Impersonal constructions are relatively well-attested in a number of non-Pama-Nyungan families of northern and northwestern Australia (see, for instance, Walsh 1987; Evans 2004; Singer 2006), but they seem to be much rarer in the large Pama-Nyungan family that covers three-quarters of the continent. In this study, I will analyse

1. My greatest debt is to the Umpithamu, Morrobolam and Mbarrumbathama speakers who so patiently taught me their languages, Florrie Bassani, Bobby Stewart, Daisy Stewart and Joan Liddy. I am also greatly indebted to Bruce Rigsby for introducing me to the Lamalama people, and for generously sharing his earlier recordings and his vast knowledge of Princess Charlotte Bay languages. Fieldwork on Umpithamu, Morrobolam and Mbarrumbathama was
a set of impersonal constructions – defined here as lacking a canonical subject (see Siewierska 2008) – in four languages from the Paman subgroup of Pama-Nyungan, from Cape York Peninsula in northeastern Australia. I will use data from the Middle Paman language Umpithamu, and from its southwestern neighbours Morrobolam, Mibarrumbathama and Rimanggudinhma, all Lamalamic languages. The analysis will focus on constructions like the ones illustrated in (1) and (2), from Morrobolam and Umpithamu, respectively. These can be contrasted with the standard transitive constructions illustrated in (3) and (4).

(1) Morrobolam 
``
awar ‘war’
garra-m=na ‘head pull’

I have a headache.’ [Lit. ‘(It) pulls me at the head’] (own fieldnotes, elicitation)
``

(2) Umpithamu 
``
warrki-ku=ampunguna ‘follow’
pot=1plinc.gen

‘We will be chased.’ [Lit. ‘(It) will chase us.’] (own fieldnotes, text)
``

(3) Morrobolam 
``
maga-n=ya-rhan ‘look’

‘I was looking at you all.’ (own fieldnotes, elicitation)
``

(4) Umpithamu 
``
ama-mpal ‘person’
wama-n=ina-ingku ‘grab’
apii ‘here’

The men grabbed him here.’ (own fieldnotes, text)
``

sponsored over the years by the Fund for Scientific Research-Flanders, which funded two stays at the University of Melbourne, by the Australian Department of Communication, IT and the Arts (DCITA), and by a grant from the Hans Rausing Endangered Languages Documentation Programme (MDP0133). Many thanks to Nick Evans and the Department of Linguistics in Melbourne for hosting my stays, and to Clair Hill for organizing the DCITA grant. I am also indebted to Bruce Sommer and the staff of the Fryer Library at the University of Queensland for providing access to the Sommer collection, and for helping out in many practical ways. Work on this paper was carried out within the project “Optional ergative marking and the architecture of case systems”, funded by the Research Council of the University of Leuven (OT/08/011). Part of this material was presented at the University of Melbourne and the 7th Conference of the Association for Linguistic Typology in Paris. I am grateful to Scott DeLancey, Nick Evans, Stefanie Fauconnier, Cliff Goddard, Clair Hill, Bill McGregor, Åshild Naess, Peter Petré and An Van linden for comments on these and other occasions, and to the editors and reviewers of this volume for very useful comments on an earlier draft. I am, of course, solely responsible for any problems or inaccuracies.
What the constructions in (1) and (2) have in common is that they contain a standard transitive verb that normally has two arguments, but they lack any indication of a subject argument. In both languages, a subject argument would show up as a nominative-marked pronominal enclitic on the verb, as in (3) and (4), which can cross-reference an ergatively-marked nominal, as in (4). The verb in (1) and (2) does not show any indication of valency-changing morphology, which is structurally absent in both languages. In terms of the analysis in Malchukov (2008), therefore, these constructions could be classified as ‘trans impersonal’, combining a basic transitive structure with indications of impersonal status.

In this study, I will argue that the core of impersonal constructions in Umpithamu and in the Lamalamic languages is centred around a set of experiencer object constructions describing involuntary physical processes. In §2, I will provide a small sketch of some relevant features in the grammar of these languages, focusing on the basic criteria for subject status. In §3, I will analyse the core of impersonal constructions in Umpithamu and the Lamalamic languages, showing that there are two basic types of experiencer object constructions: there is the basic impersonal type as found in (1) and (2) above, and there is a related type with a demoted agent argument. In §4, I will show that there are a number of extensions from this core, towards constructions with inanimate agents in all four languages, and towards an “experienced action” construction in Umpithamu, which provides an alternative construal for a subset of transitive constructions. In the final section, I will suggest that impersonal constructions in these languages are functionally not unlike a voice mechanism, especially in Umpithamu, with the cross-reference system taking up functions of argument demotion that in other languages are more typically associated with valency-changing morphology.

2. Subject status in Umpithamu and the Lamalamic languages

The languages studied in this paper are all Paman languages (as defined by Hale 1964, 1966, 1976a) from the east coast of Cape York Peninsula in northeastern Australia. Morrobolam, Mbarrumbathama and Rimanggudinhma are three closely related languages of the Lamalamic subgroup, which is characterized by a number of salient innovations, like the development of fricative phonemes and phonemic voice contrasts, and the development of enclitic agreement marking (see, for instance, Laycock 1969; Sommer 1976; Rigsby 1997). The genetic status of Umpithamu is not certain, but there is comparative and lexicostatistic evidence to suggest that it is a Middle Paman language (as used in Hale 1976b), probably most closely related to Ayapathu (see Evans 2005 for some published material). None of the languages has been described in detail. The data used in this study come from my own fieldwork.
(mainly on Umpithamu and Morrobolam, and to a lesser extent on Mbarrumbathama),
from recordings made by Bruce Rigsby (all four languages), and from two sketches
of Morrobolam (Ogilvie 1994; Sommer 1998), two sketches of Rimanggudinhma
(Godman 1993; Sommer 1999b), and a sketch of Mbarrumbathama (Sommer 1999a).
Further information on the languages and the people associated with them can be
found in Rigsby (1992, 1997), Rigsby & Chase (1998), and Thomson (1934). In this
section, I will provide a brief outline of some relevant features in the grammar of the
four languages, as well as the criteria that can be used to determine subject status. The
basic grammatical structure of the languages is very similar, even for Umpithamu,
which belongs to a different subgroup but shows strong morphosyntactic convergence
towards the Lamalamic model through intensive contact (see Rigsby 1997).

Nominals can take a number of cases, and core arguments are organized on an
ergative-absolutive basis: intransitive subjects and objects remain unmarked, while
transitive subjects can take an ergative marker. In all four languages, ergative mark-
ing is of the so-called ‘optional ergative’ type (see McGregor 2006, 2010): the ergative
marker is distributionally restricted to transitive subjects, but it is not used with all
transitive subjects. In Umpithamu and Morrobolam, its actual distribution is deter-
mined by principles of animacy and information structure (see Verstraete 2010 on
Umpithamu). Inanimate transitive subjects typically take ergative marking – obligatory
in Umpithamu – as shown in (5) and (6).\(^2\)

(5) Umpithamu

\[
\text{aykirri-mpal umpa-n=ilu-ungku yuku} \\
\text{wind-erg break-pst=3sg.nom-3sg.acc tree} \\
\text{‘The wind knocked down the tree.’ (own fieldnotes, elicitation)}
\]

(6) Morrobolam

\[
\text{aθa-l a lam la-n=na} \\
\text{fire-erg hand burn-pst-1sg.acc} \\
\text{‘The fire burned me on the hand.’ (own fieldnotes, elicitation)}
\]

Animate transitive subjects only take ergative marking when they are in focus, as illus-
trated in the contrast between (7a) and (7b) below. In (7a), the transitive subject \text{woy pu}
‘ghost’ is in focus relative to a presupposition in the preceding context (it is the ghosts
who are responsible for inserting foreign material into people’s bodies), and takes erga-
tive case. In (7b), by contrast, the transitive subject \text{amitha athuna} ‘my mother’ is not

\(^2\) The term ‘transitive subject’ is used here in a general typological sense, to introduce the
basic case system of the language. The discussion in the following sections will show that erga-
tive marking actually signals an agent rather than a subject, and that the construction in (6) is
subjectless in a more precise language-specific use of the term.
in focus (this clause simply starts one phase of the description of a river crossing), and
does not take ergative case.

(7) Umpithamu

a. [Context: You (interlocutor) have something in your body]
   
   \textit{woypu-mpal} \textit{ayngki-n=ina}

   \textit{ghost-ERG} \textit{throw-PST=3PL.NOM}

   ‘It’s the ghosts who threw it.’ (own fieldnotes, text)

b. [Context: description of a difficult river crossing]

   \textit{amitha athuna yu} \textit{kurrun kali-n=iluwa}

   \textit{mother 1SG.GEN gear carry-PST=3SG.NOM}

   ‘My mother carried the gear.’ (own fieldnotes, text)

The data available for Mbarrumbathama and Rimanggudinhma suggest that the basic
options for ergative marking are similar: ergative case is systematically found with
inanimates, and can be left out for animates, as shown by the contrast between (8a) and
(8b). There is not enough textual material for these languages, however, to determine
whether the actual distribution for animates follows the same principle of information
structure found in Umpithamu.

(8) Rimanggudinhma

a. \textit{uta-w ki-n=ni-la}

   \textit{dog-ERG follow-PST=1SG.ACC-3SG.NOM}

   ‘The dog followed me.’ \hspace{1cm} (Godman 1993:76)

b. \textit{kolpatha-igurr mbi-y=lo boy}

   \textit{man-bad throw-FUT=3SG.NOM stone}

   ‘The old man will throw the stone.’ \hspace{1cm} (Godman 1993:76)

In addition to nominal case, a second system of marking that is available in all four
languages consists of pronominal forms encliticized to the verb, illustrated in all of the
examples above. This system is organized on a nominative-accusative basis, with nom-
inative forms for transitive and intransitive subjects, and accusative forms for objects.
The pronouns can establish reference by themselves, or they can cross-reference nomi-
inals, as in most of the examples above – note that third-person forms are not zero in
any of the four languages. I prefer to use the term cross-reference for these languages,
because it is not as systematic as a full-fledged agreement system would be: even if we
leave aside the constructions discussed here, which never take subject cross-reference
(see further below), nominals are not systematically cross-referenced. In Umpithamu,
subject nominals are cross-referenced in about 2/3 of the cases, while object nominals
are cross-referenced in about 1/3 of the cases (see further in Verstraete & De Cock
2008 for more details). There is not enough textual material to have reliable figures for
the other three languages, but there too the general tendency seems to be that object
cross-reference is less systematic than subject cross-reference.

In addition to the default position of pronouns encliticized to the verb, all four
languages also have another position available for pronouns, viz. the clause-initial
position, illustrated in (9)–(11) below. For Morrobolam and Umpithamu, where
enough textual material is available, the choice between encliticized and clause-ini-
tial position can be shown to be determined by information structure, with clause-
initial pronouns occurring in contexts of focus and contrastive topic (see further in
Verstraete 2010).

(9) Morrobolam
    [Context: preceded by ‘I listened to you.’]
    ola-nan ngayi-n
    3SG.NOM-2SG.ACC listen-PST
    ‘He listened to you.’ (own fieldnotes, elicitation)

(10) Umpithamu
    [Context: People telling interlocutor to take it easy]
    antyampa kali-ku
    1PLEXC.NOM carry-POT
    ‘We will carry it (for you).’ (own fieldnotes, text)

(11) Mbarrumbathama
    yaw-ngunh li-y
    1SG.NOM-3SG.ACC eat-FUT
    ‘I will eat it.’ (Sommer 1999a: 56)

Within the system sketched so far, there seem to be two basic criteria for subject
status: the availability of nominative cross-reference, and the availability of ergative
case. Given the optionality of both features in the languages, the actual criterion is of
course not whether a particular nominal has an ergative marker or nominative cross-
reference, but whether it could have these. This distinction will become crucial when
we investigate subjectless constructions in the following section. None of the languages
studied in this paper has any form of valency-changing morphology on the verb, to
form passives, antipassives, middles and the like. At first sight, this may seem to imply
that the criteria for subject status could just as well be regarded as criteria for agent
status (in the broad sense of Van Valin’s (eg 2005) macro-roles), given that there is
no mechanism for re-arranging roles. As we will see in the following sections, how-
ever, the existence of impersonal constructions, and a number of related construc-
tion types, shows that this type of construction can take over from verbal marking
to encode changes of valency. One of the consequences will be that only nominative
cross-reference can be regarded as a criterion for subject status, and that ergative marking for nominals is a criterion for agent status rather than subject status.

3. The core of impersonal constructions: Experiencer objects

In this section, I will discuss the ‘core’ type of impersonal construction, which in the four languages studied here is an experiencer object construction describing sudden, involuntary physical processes. The next section will then discuss a number of exten-
tions, towards transitive constructions with inanimate agents in all four languages, and
towards a more general type of experiencer construction in Umpithamu.

The structures in (12)–(14) below illustrate the relevant construction in
Umpithamu, Morrobolam, and Rimanggudinhma. All these structures have a tran-
sitive verb that normally takes two arguments, like ‘pull’ in (13) or ‘spear’ in (12)
and (14), but they lack any trace of subject marking. There is never a nominative
pronoun encliticized to the verb, and there is no nominal with ergative case. In this
sense, they are a good example of Malchukov’s (2008) category of transimpersonal
constructions, which use transitive verbs with object marking but without any trace of
subject marking.

(12) Umpithamu

   manu watyu-ngka=athuna
   neck spear-PRS=1SG.GEN
   ‘I have hiccups.’ [Lit. ‘(It) spears me in the neck.’] (own fieldnotes, elicitation)

(13) Morrobolam

   avar garrar=m=na
   head pull-PRS=1SG.ACC
   ‘I have a headache.’ [Lit. ‘(It) pulls me at the head’] (own fieldnotes, elicitation)

(14) Rimanggudinhma

   noka rhi-m=thum
   neck spear-PRS=1SG.DAT
   ‘I have hiccups.’ [Lit. ‘(It) spears me in the neck.’] (Godman 1993)

Semantically, the constructions all describe involuntary physical processes that are
punctual or take an abrupt start – other physical processes more typically use a stan-
dard intransitive construal in the four languages, as illustrated in (15)–(17) below for
Umpithamu, Morrobolam and Rimanggudinhma.

(15) Umpithamu

   waympi ethanema-ngka=ayuwa
   back ache-PRS=1SG.NOM
   ‘My back is sore.’ [Lit. ‘I am sore in the back’] (own fieldnotes, text)
Given its association with involuntary physical processes, the impersonal construction as a whole can be analysed as an experiencer construction with the experiencer role mapped onto the non-subject argument (compare Evans 2004 for parallel cases of experiencer object constructions in other Australian languages). This semantic specialization in experiencers may also explain why in at least two languages the non-subject argument takes an oblique case, a genitive in Umpithamu and a dative in Rimanggudinhma (see further in Verstraete 2010 on Umpithamu). Note that the genitive in the Umpithamu structure in (12) cannot be analysed as marking a possessive relation with kayku ‘head’, as this is an inalienably possessed noun that does not take genitive-marked possessors. In other construction types, this type of body part would be cross-referenced with nominative or accusative pronouns in an external possession relation. This is the case with the nominative 1st person pronoun in the intransitive structure in (15), which functions as an external possessor for the body part waympi ‘back’, or with the accusative in the transitive structure in (18), which functions as an external possessor for the body part apa’ala ‘lower leg’. In this sense, the body parts in (12)–(18) can be contrasted with alienably possessed nouns, like uuuku in (19), which simply take a genitive possessor.

(18)  Umpithamu

apa’ala  atha-n=iluwa-athungku
lower.leg  bite-pst=3sg.nom-1sg.acc
‘It bit me in the leg.’ (own fieldnotes, text)

(19)  Umpithamu

uuuku  antyangana  yula-n=ilu
language  1plexc.gen  make-pst=3sg.nom
‘He spoke our language.’ (own fieldnotes, text)

In all four languages, there is a second subtype of experiencer object construction, with the same basic meaning of an abrupt involuntary physical process, but with a nominal encoding the stimulus role, which is always marked with ergative case. This type is illustrated in (20)–(21) below for Umpithamu and Morrobolam. As with the
other type, there is no nominative pronoun to cross-reference the ergatively-marked element. Incidentally, the case pattern in (20) also provides independent evidence that the body-part nominal in experiencer object constructions cannot be regarded as the stimulus role, but is in an external possession relation with the experiencer. The use of ergative case shows that the stimulus in these structures is the abstract term denoting the process (like *aatyarra* 'cramp'), rather than the body part that is affected by it (*apaāla* 'lower leg'), which remains unmarked.

(20) Umpithamu

\[\text{aatyarra-mpal atha-n=athuna apaāla} \]
\[\text{cramp-ERG bite-PST=1SG GEN lower.leg} \]
\[\text{‘I had a cramp in my lower leg.’ [Lit. ‘Cramp bit me in the lower leg.’]} \]
\[\text{(own fieldnotes, text)} \]

(21) Morrobolam

\[\text{marrunyen-h-u tha-m=na} \]
\[\text{sneeze-ERG bite-PRS=1SG ACC} \]
\[\text{‘I am sneezing.’ [Lit. ‘Sneeze bites me.’] (own fieldnotes, elicitation)} \]

To sum up, the basic type of impersonal construction in Umpithamu and the Lamalamic languages is a transimpersonal construction encoding sudden or punctual involuntary physical processes. The verb is a standard transitive action verb, but the construction as a whole has experiential meaning, projecting an experiencer role, and in the second subtype also a stimulus role. The experiencer is marked with a pronominal enclitic in an accusative or an oblique case, and optionally an absolutive nominal denoting the specific body part that is affected, in an external possession relation. The construction need not make reference to any stimulus: it never has a nominative pronoun, but the second subtype has a nominal referring to the stimulus, which is obligatorily marked with ergative case. Table 1 below summarizes the basic characteristics of the two subtypes of this construction type.

<table>
<thead>
<tr>
<th>Action</th>
<th>Transitive verb (<em>‘pull’, ‘bite’, ‘spear’</em>)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undergoer/</td>
<td>Pronominal Accusative or oblique (dative, genitive)</td>
</tr>
<tr>
<td>Experiencer</td>
<td>Nominal Absulative (body part, in external possession relation with the referent of the pronoun)</td>
</tr>
<tr>
<td>Agent/Stimulus</td>
<td>Pronominal No nominative pronoun</td>
</tr>
<tr>
<td></td>
<td>Nominal Subtype 1: No nominal</td>
</tr>
<tr>
<td></td>
<td>Subtype 2: Ergative-marked nominal (abstract process)</td>
</tr>
</tbody>
</table>

Table 1. Experiencer object constructions
Looking at these constructions from the perspective of impersonal status, it is only the first subtype, without a nominal referring to the stimulus, that can be regarded as unquestionably impersonal. The structures in (12)–(14) lack any trace of a subject, both syntactically and semantically. There is not even an obvious referent for such a subject argument, unless one were to take the abstract stimulus nominals of the second subtype (like aatyarra 'cramp' in (20) or marrunyenh 'sneeze' in (21)) as the model. The second subtype has a candidate for subject status, namely the ergatively-marked stimulus nominal. Given that it is never cross-referenced with a nominative pronoun, however, it is hardly a standard type of subject in the languages studied here. In the following sections, I will discuss a number of extensions of the experiencer object construction, which together with the data presented in this section suggest that even the second subtype can equally be regarded as impersonal, with the ergatively-marked nominal as a demoted agent.

4. Extensions: Inanimate agents and experienced action constructions

The experiencer object constructions discussed in the previous section can be regarded as the core of impersonal constructions in the languages studied here, because they are the only type that can consistently use a transitive verb without any trace of a subject, lexical or pronominal. If we take the lack of nominative cross-reference in the second subtype as a weakened version of impersonal status, there are also a number of extensions that are relevant, and that can shed further light on impersonal constructions in these languages. In §4.1, I will discuss transitive constructions with inanimate agents and human undergoers, which systematically lack nominative cross-reference in all four of the languages discussed here. In §4.2, I will discuss experienced action constructions, a construction type that is only found in Umpithamu, and that could be regarded as a further extension of experiencer object constructions, carrying over its defining morphosyntactic features to a basically non-experiential domain.

4.1 Inanimate agents

As already mentioned, ergative marking is optional in all of the languages discussed here, driven by a principle of information structure for animate transitive subjects. For inanimates, by contrast, ergative marking is typical, and in Umpithamu it is obligatory. In itself, this is not very remarkable: it is in line with well-known hierarchies like Silverstein’s (1976), and it can be motivated functionally by the fact that inanimates do not typically instigate actions, and therefore require explicit marking when occurring in this role. In the languages studied here, however, transitive constructions with inanimate agents have an additional feature that distinguishes them from their
standard counterparts. In constructions with inanimate agents and human undergoers, the agents systematically lack cross-reference with a nominative pronoun, as illustrated in examples (22)–(25).

(22) Umpithamu

\[
\text{ngoki-mpal} \quad \text{ungka-n=anyangana} \\
\text{water-ERG} \quad \text{wet-PST=1PLEXC GEN}
\]

'The water made us wet.' (own fieldnotes, text)

(23) Morrobolam

\[
oxar-aw \quad \text{the-m=ungan} \\
\text{mud-ERG} \quad \text{bog-PRS=3SG OBL}
\]

'Mud is bogging him.'

(Sommer 1998: 22)

(24) Mbarrumbathama

\[
?alanh-a \quad \text{ka-n=na} \\
\text{rain-ERG} \quad \text{wet-PST=1SG ACC}
\]

'The rain wet me.'

(Sommer 1999a: 19)

(25) Rimanggudinhma

\[
arta-w \quad \text{ti-n=ni} \\
\text{sun-ERG} \quad \text{burn-PST=1SG ACC}
\]

'I got burned by the sun.' [Lit. 'The sun burned me.]

(Sommer 1999b: 13)

For constructions that have both inanimate agents and inanimate undergoers, there is not enough data on all languages, but at least for Umpithamu there is evidence that these constructions take standard patterns of cross-reference, i.e. they can have a nominative pronoun for the inanimate subject, and an accusative one for the inanimate object, as shown in (26).

(26) Umpithamu

\[
\text{aykirri-mpal} \quad \text{umpa-n=ulu-ungku} \quad \text{yuku} \\
\text{wind-ERG} \quad \text{break-PST=3SG NOM-3SG ACC} \quad \text{tree}
\]

'The wind knocked down the tree.' (own fieldnotes, elicitation)

From a morphosyntactic perspective, the constructions in (22)–(25) are identical to the second subtype of the experiencer object constructions discussed in the previous section: there is an ergatively-marked nominal, but never nominative cross-reference on the verb. In Umpithamu, this identity even extends to the use of an oblique case marker for the non-subject function. What distinguishes these constructions from experiencer object constructions is the semantic contribution of the verb. In structures
like (22)–(25), the semantics of the verb (‘wet’, ‘bog’, ‘burn’) is directly relevant to the final interpretation: the construction describes an action of wetting, bogging or burning, even if experiential semantics still plays a role in the semantics, as also reflected in the use of an oblique marker for the undergoer (see further in Verstraete 2010). In the experiencer object construction, by contrast, the semantics of the verb is subsumed in a fully experiential interpretation, with ‘spear’, ‘bite’ etc contributing to a metaphor of physical experience. Verstraete (2011) provides further details on the semantic analysis of these constructions, and their link with experiential semantics.

4.2 Experienced action constructions

In Umpithamu, there is one further construction type that is similar to the ‘core’ impersonal type of experiencer object constructions. The other three languages show no evidence of this type, which I call an ‘experienced action construction’. Basically, this type is available as an alternative construal for any transitive clause with an animate agent and a human undergoer. Thus, for instance, the structures in (27a) and (27b) are available as an alternative for their standard transitive counterparts in (27c) and (27d).

(27) Umpithamu

a. ama-mpal yitha-n=inangana
   person-erg leave-pst=3pl.gen
   ‘The man left them.’ (own fieldnotes, text)

b. yitha-n=anyangana Bamaga
   leave-pst=1plexc.gen Bamaga
   ‘(They) left us at Bamaga.’ (own fieldnotes, text)

c. ama-mpal yitha-n=iluwa-inangku
   person-erg leave-pst=3sg.nom-3pl.acc
   ‘The man left them.’ (constructed example)

d. yitha-n=ina-anyangku Bamaga
   leave-pst=3pl.nom-1plexc.acc Bamaga
   ‘They left us at Bamaga.’ (constructed example)

Morphosyntactically, the structures in (27a) and (27b) are identical to the experiencer object constructions and the inanimate-agent constructions discussed in the previous sections. They have a genitive-marked non-subject argument, they lack nominative cross-reference for agents, and they can even lack nominal agents altogether, as in the entirely impersonal structure in (27b). The reason why I treat them in a distinct section is that unlike the other two types they are only found in Umpithamu, and their use is a matter of choice, providing an alternative construal for standard transitive clauses. Moreover, their semantics is distinct from experiencer object constructions.
First, there is the interpretation of the absent agent in constructions like (27b). In experiencer object constructions without a nominal subject like (12)–(14) above, it is hard to imagine any referent for the absent agent, except perhaps for the referent of the abstract process term found in (20)–(21). In experienced action constructions, by contrast, one can often find a potential referent for the absent agent, typically one that has been mentioned in the preceding context (for instance white officials in (27b)). And even if there is no such actual referent, it does have a schematic shape, in the sense that it will typically be human and animate, unlike in experiencer object constructions. A second semantic difference from experiencer object constructions is that, as with inanimate-subject constructions, the specific semantics contributed by the verb continues to play a role in the final interpretation, and is not subsumed in a purely experiential interpretation. Instead, the construction has an interpretation that combines action semantics with experiential semantics. Verstraete (2011) provides further details on the semantics of this construction type, which – in a nutshell – is modelled on the experiencer object construction in that it allows speakers to construe the action denoted by the verb as ‘overcoming’ people beyond their control, rather than simply affecting them.

5. Interpretation: Subject criteria, impersonal status and voice

In this section, I will discuss the more general typological status of the construction types examined in this paper. If we sum up the data discussed in the preceding sections, impersonal constructions in Umpithamu and in Lamalamic languages have the following characteristics. They use a transitive verb, with accusative or oblique object cross-reference, but without nominative subject cross-reference, and they may have an ergatively-marked nominal. From a semantic perspective, these constructions are best described in terms of the ‘core’ type of experiencer object constructions, either with entirely experiential semantics (as in §3), or with mixed experience/action semantics (as in §4). Stipulating the presence of an experiencer role captures both the semantics and the morphosyntax of these constructions best. The details of the semantic analysis go beyond the scope of this paper, but there is a good range of evidence for this analysis, as argued in Verstraete (2010).

From a more general functional perspective, however, there is a bit more that can be said about these constructions. In the languages described here, which do not have any specific valency-changing morphology on the verb, they are also the closest thing there is to a mechanism of voice. I will argue that the systematic absence of nominative cross-reference in a construction is the strongest criterion for impersonal status in these languages, removing the agent argument from the set of core arguments
Jean-Christophe Verstraete

normally associated with a transitive verb and therefore also making nominal reference to this argument optional. There are four sources of evidence for this: the general function of nominative cross-reference in the languages studied here, the formal make-up of the impersonal constructions, the types of configurations for which they are used, and more general typological developments observed for experiencer object constructions.

In the languages studied here, the basic function of nominative cross-reference is to establish a predicative relation with a subject argument. There is interesting independent evidence for this from the domain of non-verbal predicates, where the presence or absence of a nominative pronoun is the basic distinction between predicative and attributive uses, as shown in the Umphathamu structures in (28) below. With nominative cross-reference, as in (28a), the adjective *yawul* 'big' becomes a clausal predicate that predicates a property of *yaanti* 'beach'. Without nominative cross-reference, however, as in (28b), *yawul* 'big' is merely an attribute of the nominal *yaanti* 'beach', and together they form a noun phrase.

(28) Umphathamu

a. *yaanti* yawul=iluwa  
   beach big=3sg.nom  
   'The beach is big.' (own fieldnotes, text)

b. *yaanti* yawul  
   beach big  
   'A/the big beach.' (constructed example)

In this perspective, the systematic absence of nominative cross-reference in the constructions studied here can be interpreted as the absence of a predicate-subject relation. Even if there is an ergatively-marked nominal, as in the subtypes illustrated below in (29) and (30), this can no longer be regarded as the subject of the verb if it is not possible to have nominative cross-reference. In this sense, ergative marking in these constructions can be regarded as a marker of agent status rather than of subject status (see McGregor 1997 for a similar approach to split ergative marking): ergatively-marked nominals become subjects when they can be cross-referenced with a nominative, but remain agents demoted from a subject role in constructions where they cannot be. In this sense, the presence of an ergatively-marked nominal in structures like (29) and (30) is not a counter-indication for impersonal status, because the systematic lack of nominative cross-reference ensures that the ergative nominals are not subjects.

(29) Morrobolam

*iθam-i* tha-m=na  
cramp-erg bite-prs=1sg.acc  
'I have a cramp.' [Lit. 'Cramp is biting me.'] (own fieldnotes, elicitation)
Under this interpretation, the impersonal constructions studied here are functionally not unlike passives, with an optional agent nominal that is not a subject. Obviously, they are still distinct from typical passives (see Siewierska 2005) in a number of ways. Most importantly, unlike with typical passives the impersonal construction is not available as a genuine alternative construal for transitive clauses in three of the four languages: it is the only construal available for describing involuntary physical experiences and inanimates affecting humans. Still, the fact that Umpithamu has gone somewhat further along this path – with the experienced action construction described in §4.2 as an alternative construal for transitive clauses with human undergoers – at least strongly suggests that the potential is there to develop into a genuine voice system.

This interpretation is actually supported by three further types of evidence. First, there is the formal make-up of the constructions. Syntactically, the characteristics of the impersonal constructions studied here can be interpreted as signs of reduced transitivity. Even if the verb does not register a change in valency, the construction as a whole shows reduced transitivity, due to the systematic absence of nominative cross-reference for agents (as opposed to potential absence in ordinary transitive constructions), and in some languages also the use of oblique case for undergoers (as opposed to accusative case in ordinary transitive constructions). Further evidence for the non-argument status of ergative-marked agents is the fact that they can remain absent, at least for the experiencer object constructions discussed in §3, and for the experienced action constructions discussed in §4.2. In some views of the passive (e.g. Dixon & Aikhenvald 1997), the formal absence of valency-changing morphology on the verb to go with these case features would be problematic. However, there are good cases for passives without specific verbal morphology in the literature, for instance as reported in Arka & Kosmas (2005), and Lüpke (2005: 289–304), which show convincingly that verbal marking is a typical but not a criterial feature.

Second, if we look at the contexts that require impersonal constructions in the four languages, they are precisely the types of configurations that are more generally problematic in terms of subject assignment, and that typically lead to alternative construal in languages with valency-changing morphology. Both the experiencer

3. I am grateful to Anna Siewierska for helping me to refine this part of the analysis, and for pointing me towards crucial references.
object construction discussed in §3 and the inanimate-agent construction discussed in §4.1 are configurations where an inanimate agent-like element affects a human undergoer-like element. This type of configuration goes against well-known hierarchies of animacy and empathy, and is typologically precisely the situation that will require valency-changing mechanisms when they are available. In this perspective, it is not suprising that as soon as the human undergoer is replaced by an inanimate one, as illustrated in (26) above and in (31) below, the construction takes standard nominative cross-reference again.

(31) Umpithamu

\[
\text{muumpam-mpal ngaympi-n=ilu-ungku wungku}
\]

\begin{flushright}
stone-\text{ERG hit-PST=3SG.NOM-3SG.ACC house}
\end{flushright}

‘The stone hit the house.’ (own fieldnotes, elicitation)

A final piece of evidence to support a functional link with passive constructions is found in the work of Malchukov (2008), who reports on a number of cases where the non-subject argument of experiencer object constructions is taking on a number of behavioural subject properties even if its coding remains that of a non-subject. Thus, for instance, the Papuan (Trans-New Guinea) language Amele has an experiencer object construction not unlike that in Umpithamu, with an experiencer coded as an object, and with a dummy subject marker, as illustrated in (32a) below. Even though the experiencer is not coded as a subject, it does show behavioural subject properties, in the sense that it can affect the marking of switch-reference in a preceding clause, as illustrated in (32b) below (see Malchukov 2008:87–88). The first clause has a first-person subject, and even though the second clause has the dummy 3rd person as its morphological subject and the first person as its morphological object, switch-reference marking in the first clause is still same-subject, which suggests that at least for this parameter the experiencer object behaves as a subject.

(32) Amele

\[
\text{Amele}
\]

---

4. Given the association with animacy hierarchies, an interpretation as an inverse rather than a passive should not be excluded, as suggested by Anna Siewierska (p.c.). However, there are a number of arguments that favour a passive analogy over an inverse one. First, the construction shows signs of reduced valency, which points towards a passive rather than an inverse. Second, even if we disregard the syntactic point, the inverse analysis really only works for the inanimate-agent type discussed in §4.1. It does not work for experienced action constructions (discussed in §4.2), which do not necessarily involve reversals of typical animacy relations. And it does not work very well for the core experiencer type (discussed in §3), because these structures often lack any reference to an agent (subtype 1), or if they have a formal agent, its actual referent remains vague and abstract (subtype 2). In this sense, the passive looks like a more promising candidate for a functional analogy in the domain of voice.
None of the four languages studied here have the rich interclausal morphology that would allow us to test whether the non-subject argument in the impersonal constructions is acquiring any behavioural subject properties. Moreover, a close semantic analysis of the constructions shows that traces of experiential semantics are found throughout the set of structures studied here (as argued in Verstraete 2011), which are not necessarily explained very well by simply cataloguing them under passives. So for the time being, the more conservative conclusion is that these are impersonal constructions best analysed in terms of their relation to experiencer object structures, but that they show at least a good functional similarity with voice mechanisms in other languages. This has developed furthest in Umpithamu, which uses the construction not only in experiencer object constructions and constructions with inanimate agents, but also as an alternative construal for constructions with animate agents. The fact that the system seems to have moved furthest in Umpithamu is probably not a coincidence, as there is good evidence for a relatively recent restructuring of this language towards the head-marking pattern typical of Lamalamic languages (Rigsby 1997). Within the regional context, the system of cross-reference with verbal enclitics is a clear Lamalamic innovation, as all three Lamalamic languages show the same pattern. In Umpithamu, intensive contact with Lamalamic caused a shake-up of the morphosyntax, leading not only to a system of cross-reference that allows clear marking of impersonal status, but perhaps also to the expansion of the Lamalamic core of experiencer object constructions towards experienced action constructions that infringe upon the domain of standard transitives.

Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACC</td>
<td>accusative</td>
</tr>
<tr>
<td>DAT</td>
<td>dative</td>
</tr>
<tr>
<td>ERG</td>
<td>ergative</td>
</tr>
<tr>
<td>EXC</td>
<td>exclusive</td>
</tr>
<tr>
<td>FUT</td>
<td>future</td>
</tr>
<tr>
<td>GEN</td>
<td>genitive</td>
</tr>
<tr>
<td>INC</td>
<td>inclusive</td>
</tr>
<tr>
<td>NOM</td>
<td>nominative</td>
</tr>
<tr>
<td>OBL</td>
<td>oblique</td>
</tr>
<tr>
<td>POT</td>
<td>potential</td>
</tr>
<tr>
<td>PRS</td>
<td>present</td>
</tr>
<tr>
<td>PST</td>
<td>past</td>
</tr>
<tr>
<td>SS</td>
<td>same subject</td>
</tr>
<tr>
<td>SIMUL</td>
<td>simultaneous</td>
</tr>
</tbody>
</table>
References


Impersonal constructions in Umpithamu and the Lamalamic languages

De Cock, B. 611
Delbrück, B. 101, 235, 241
Denny, J.P. 492
Diewald, G. 199
Dimmendaal, G.J. 47
Divjak, D. 4
Dixon, R.M.W. 33, 176, 520, 621
Dombrowski, A. 361
Donaldson, B.C.
Donohue, M. 5, 27, 46, 48, 170, 285
Dowty, D. 23
Drapeau, L. 6, 11, 489–491, 508, 511
Dryer, M. 25, 36, 38, 45–46, 497–498
Du Feu, V. 589
Duncan, L. 82
Duranti, A. 590, 592

E
Egerland, V. 62, 66, 79, 84–85
Elizarenkova, T. Ja. 239, 243
Ellis, C.D. 496, 498
Erelt, M. 401, 407
Erelt, T. 401, 407
Eriksen, P.K. 395, 398, 433
Ernout, A. 177
Evans, N. 607–609, 614
Eythórsson, T. 193

F
Faarlund, J.T. 25, 75, 151, 155–157
Fedin, M.S. 403
Feltenius, L. 182–183
Fleischman, S. 327
Flobert, P. 182, 246–247
Forsberg, H. 73, 78, 395
Fortescue, M. 401
Frajzyngier, Z. 29, 36–37, 92, 118, 189, 192, 311–312, 315, 385
Franks, S. 81
Fraser, N.M. 358
Fried, M. 548
Friedmann, N. 325

G
Gadžieva, N.Z. 431
Gair, J.W. 37
Ganenkov, D. 49
Gasparini, A. 313
Gawronska, B. 114
Geldner, K.F. 239, 243
Geniusiené, E. 230, 517
Gerola, B. 179
Giacalone Ramat, A. 6, 9, 57, 61, 68, 70, 79–80, 189
Givón, T. 22, 27, 30–32, 39, 43, 45, 49, 323, 325, 333, 346, 524, 527, 538
Goddard, I. 492, 608
Godman, I. 610–611, 613–614
Gonda, J. 239
Görski, R. 118, 376–377, 381, 384
Grassmann, H. 242
Greenberg, J. 5, 9, 45, 275, 280, 282
Gregores, E. 25
Grimshaw, J. 3, 25

H
Haas, M. 5–8, 29–30, 127, 461
Haider, H. 114
Hakanen, A. 398, 401, 407, 416–417, 433
Hakulinen, A. 428
Hale, K. 609
Harris, A.C. 176
Harshaw, B. 326
Hattori, S. 520
Haverling, G. 171
Haudry, J. 241
Heine, B. 45, 232, 335, 523, 533
Helasvuoto, M.-L. 49, 72, 400
Helimskij, E., also:
Helimskij, E. 406, 419–420
Herman, J. 1, 171–173, 182
Hetzron, R. 29, 128, 136
Hewitt, G. 34, 58
Hinchliffe, I. 75
Hoeckstra, T. 97
Holmberg, A. 3, 7, 11, 13, 28–29, 57, 77–79
Holmberg, H. 3, 7, 11, 13, 28–29, 57, 77–79
Holmes, P. 75
Holton, G. 5, 46, 184
Holvoet, A. 72
de Hoop, H. 34
Hooper, R. 590
Hopkins, B 286
Hopper, P. 23, 81, 98, 264, 327
van Hout, A. 99
Huang, C.-T. J. 553, 568, 573
Hudson, J.A. 340

I
Idström, A. 411
Jleva, T.M. 403, 424
Itkonen, E. 31, 422, 426, 428, 430

J
Jacobson, S. 469
Janda, L.A. 4
Janhunen, J. 396–397, 399, 432
Jaxontov, S.J. 373–374
Jelinek, E. 494
Jensen, T.J. 58, 65, 74, 84
Jeze, E. 202
Jisa, H. 60, 84, 342, 345–346
Johansson, S. 62
Jónsson, J.G. 120

K
Kallas, K. 372
Kanerva, J. 31, 285
Kawasha, B. 265
Kazenin, K.I. 192
Keenan, E.L. 3, 7, 19–24, 30, 36, 38, 44–46, 50, 510, 538
Keller, F. 99
Kibort, A. 3–4, 6–7, 10, 37, 45, 85, 190, 357, 375, 381, 386
Kibrik, A. 33
Kimball, G.D. 34, 47
Kindaichi, K. 530
Kirikae, H. 523, 541
Kirschaubm, E.-G. 100–101
Kirsner, R.S. 194–195
Kitagawa, C. 548
Kittila, S. 34, 398, 433
Klainman, M.H. 491, 508
Klamer, M. 43–44
Klumpp, G. 422
Koch, P. 5, 128, 135, 138–139, 154
Kolehmainen, L. 395, 398, 415–416, 433
Koljaděnkov, M.N. 402, 407, 422, 426
van der Auwera, J. 20, 40, 50, 70
van der Wal, J. 285
Van Hell, J.G., also van Hell, J. 74, 339
Van Valin, R.D. Jr. 494, 612
Velankar, H.D. 242–243
Verhoeven, E. 15
Verhoeven, L. 324, 331, 339
Viguié, A. 60, 84, 345
Vilkuna, M. 49, 72, 400, 416, 428
Vincent, N. 128, 155–156
Vogel, P. 189, 194

W
Walsh, M. 607
Ward, G. 129–130, 360
Watkins, L.J. 25
Waugh, L.R. 84
Wedekind, C. 315
Wedekind, K. 315
Weerman, F. 63, 84
Wehr, B. 206, 220
Weiner, E. 232
Weyhe, E. 85
Wichmann, S. 5, 32, 170, 285
Wierzbicka, A. 398
Wheeler, M.W. 62, 71, 84
Wolfart, H.C. 491, 496–498, 510–511, 513
Wolvengrey, A. 497
Wright, J.C. 242

X
Xrakovskij, V.S. 36–37

Y
Yang, S. 570

Z
Zamboni, A. 171, 176–177
Zamora, A. 339
Zhu, D. 568–569
Ziff, P. 249
Zifonun, G. 84
Zimmerling, A. 3, 45
Ziv, Y. 325, 571
Zubair, C. 23
Zúñiga, F. 491
Language index

A
Abaza 72
Abkhaz 34, 58, 70, 72
Afroasiatic (also Afroasiatic), 6, 69, 307–323
Ainu 5–6, 11, 13, 46, 517–544
Albanian 70, 72
Algonquian languages 46, 489, 491–492, 496–497
Amharic 310–311
Arabic 27, 69, 171, 312, 319
Atlantic languages 285, 287
Austronesian 527, 581, 600

B
Babungo 28, 46, 69
Balto-Finnic languages 395
Bantu languages 22, 31–32, 265, 285–287
Basque 58, 60, 70, 72
Beja 315
Bengali 35
Berber 309, 312, 314–315, 318–319, 321
Breton 70, 72

C
Catalan 60, 62, 70–71, 84–85, 139, 142, 158, 161–162
Celtic 70, 86, 179
Cmu 583
Central Alaskan Yupik (CAY) 8, 11, 47, 459–486
Chadic 309–314
Chichewa 31
Chinese, see Mandarin Chinese 67, 547–580
Cree 489–490, 497–498, 510, 514
Cushitic 309–310, 312–317, 321

D
Danish 58, 65, 70, 72, 74–75, 84–85, 118, 142, 149–154, 156, 158
Dhivehi 37
Drehu 587, 598–600
Druba 598

E
East Futunan 45, 591–596, 601
East Uvean 582–583, 585–586, 588, 595, 600–601
Eskimoan 11, 459
Estonian 59, 72, 397–398, 401, 410, 416–418, 426, 428, 430–432, 527
Even 26

F
Fagauvea 587
Faroese 433
Finnic 59, 72, 79, 86, 395, 397–400, 410, 426, 547
Finno-Ugric 399–400

G
Galician 71
Gawwada 314, 316–318, 321
Georgian 72
Greek 60, 70–71, 77, 82–83, 85–86, 102, 114, 247, 295, 330
Guarani 25

H
Hebrew 6, 10, 39, 69, 77–78, 178, 323–351, 574, 577
Hungarian 64, 70–71, 85–86, 330, 365, 391, 396–397,
Language index

399–401, 405, 408, 415, 422–423, 427, 430–432

I
Iaai 582
Indo-Iranian 239, 244, 247
Ingrian 397–398

J
Japanese 19, 26–27, 39, 69, 423, 517–518, 529, 532, 542
Jóola-Banjal 6, 9, 285–304

K
Kabardian 72
Kabyle 312, 314–316, 318–321
Kamas 397, 399, 406, 421, 430–432
Kambera 43
Kanak 12, 581–587, 590–591, 598, 602–604
Karelian 397–398
Kartvelian 72
Ket 5–6, 11, 13, 47, 439–457
Kiowa 25
Koasati 34, 43, 47
Komi 72, 395, 397–400, 403, 407, 411, 420, 424, 426, 430–432

L
Lamalamic 607–612, 619, 621
Latin, see also Latin 6, 8, 11, 169–189, 230, 239, 246, 461
Late Latin 6, 8, 111, 116, 169–189, 196, 199, 202
Lele 69, 312
Lezgian 34, 40–41, 49, 58–59
Lithuanian 58, 70, 72, 79

M
Maa 5–6, 9, 12–13, 42, 257–284, 410, 476
Maasai, see Maa 5, 6, 9, 12, 42, 259–283, 410, 476
Maltese 70, 72
Manam 26
Mandarin, also Mandarin Chinese 547
Mari 72, 395, 397, 399–400, 402, 407, 411, 418, 424, 426, 430–432
Mbarrumbathama 607–612, 617, 621
Mexican Spanish 133, 139
Modern Hebrew, see Hebrew 6, 10, 323–353, 574, 577
Mordvin 72, 397, 399–400, 402, 407, 411, 413, 418, 422–424, 426, 430–432
Morrobolam 607–610, 612–615, 617, 620

N
Nenets 397, 399, 405, 408–409, 413–415, 425, 427, 430–432
Nganasan 397, 406, 425, 430–432
Nilotic 257, 282
Nkore-Kiga 69
Norwegian 62, 70, 72, 75, 108–110, 118, 141–142, 156–157, 433
Norwegian, Nynorsk 74, 75
Nyeláyu 585–586

O
Ob-Ugric 297, 399, 400, 406, 408–409, 420, 424
Old Church Slavonic 239, 247–248
Old English 149–150
Old Italian 195–197, 200, 202, 204, 206, 209, 214
Old Norse 151, 155–158
Old Russian 102, 106, 111
Oromo 31, 41, 310–312

P
Paamese 69, 602
Pama-Nyungan 12
Pileni 582–583
Polish 6, 10, 29, 37, 45, 58, 60, 65, 71, 76–77, 79, 85, 92, 94, 118–119, 189–193, 206, 218, 250, 330, 357–394, 558, 574
Polynesian 595, 601–602
Portuguese 2, 63, 70–71, 77, 139, 142–145, 148, 158–161, 164, 571, 577
Proto Indo European 247

R
Rapanui 589
Rimanggudinhma 607–611, 613–614, 617
Romance 182
Rumanian 70

S
Samoaan 33, 40–41, 45, 590, 592, 594, 601
Samoyed, Samoyedic
Sanskrit 229–245
Sardinian 71, 149
Semitic 309, 310, 385
Selkup 397, 399, 406, 413, 415, 425, 430–432
Serbo-Croatian 71, 76, 85
Sesotho 31–32, 39–40
Sidamo 313–314
Sinitic, also Sinnitic 548
Swahili 31–32, 40
Swedish 62, 66, 70, 72, 74–75, 84, 108, 142, 345–347, 351, 428, 433, 558

T
Taba 38
Tewa 37–38
Tlapanec 32, 35, 41
Tokelauan 590, 599
Tongan 582–583, 601
Ts’umakko 313–314, 316
Tukang Besi 27
Turkana 47
Turkish 71, 80, 171, 432
Tuvaluan 69, 589–590, 593–594

U
Udmurt 72, 397, 399, 402–403, 413, 419, 424, 426, 430–433
Umpithamu 12–13, 607–623

V
Vedic Sanskrit, see also Sanskrit 229–245
Veps 72, 397–398, 422
Volgaic 72, 79, 397, 399
Votic 397

W
Wandala 311, 314
West African 9, 285–286, 292
Wolof 292, 295

X
Xârâcùù 583–584, 586–587, 589, 591–592, 597–600, 602–603
Subject index

A
accusative-subject construction 8
actant H 30, 133–34, 139n7, 145
action nominalization 517, 525, 530–31, 540–43
active system (active alignment) 169, 171n2, 176
agent
demoted agent 20, 197, 518, 607, 609, 616, 620
inanimate agent 34, 38, 41, 47, 264, 273, 276, 495, 501, 512–13, 607, 609, 613, 616–17, 622–23
agreement
default agreement 25, 32, 48, 78n11, 174, 177, 196n9, 198, 318, 358, 360, 380, 386, 509, 512, 514
loss of 19, 30, 31–40, 42, 135. See also default agreement
topic agreement 32
Aktionsart 100–101, 103–4, 112, 509

Animacy Hierarchy 37, 260, 489–91, 494. 497, 507, 514
anticausative 9, 12, 35, 38, 44, 49n16, 171n3, 172–73, 177n8, 180, 181, 184–85, 192, 196, 229, 231–32, 234, 236, 238, 240–49, 251 303, 308, 377, 582, 588
anti-impersonal 4
arbitrary reading 48n14, 285, 301
argument omission 498, 503, 519n2, 535, 540–41, 581, 590, 595, 597–98, 600
atmospheric verb 20, 26, 290, 309–11, 314, 317, 319–21. See also meteo-verb, meteorological predicate, weather verb
auxiliary selection 91, 95–100, 113

B
Burzio’s generalization 3, 93, 114, 122

C
cline of depersonalization 325, 344–45, 350
cline of impersonalization 10, 323
cognitive construction 26, 27, 38, 40, 426–27
detotalization 137–38
detransitivization 259, 264, 265, 445–46, 448, 456–57, 459, 528
Diathetic Null Hypothesis 114
differential agent marking 12, 581–82, 590–96, 603
differential subject marking 12, 35
discourse genre 324, 338, 340, 350
discourse referentiality 43, 45, 274
dummy subject 3, 20–27, 38, 325–6, 396, 426, 622. See also expletive subject
dummy verb 38, 44n11

e
ellipsis 28, 47–49, 58, 329
episodic (reading) 57, 59, 62, 63–64, 66–67, 70–71, 79, 80, 84, 85, 327, 332, 334, 350, 554, 556
ergative language 4, 12, 22, 33, 35, 40, 41, 49n17, 590–91
Subject index

grammaticalization path 5, 11, 50, 232, 517, 519, 542. See also cline of depersonalization, cline of impersonalization

I

impersonal construction

impersonal reflexive 59

impersonal to passive

reanalysis, see under reanalysis

primary impersonal verb 459–61, 470, 479, 486

of transitive verbs, see transimpersonal constructions


function-based approach 4, 7, 9, 12, 13

hybrid approach 258

prototype approach 21–24, 36, 38, 40

structural approach 1–2, 5, 7, 9, 10, 11–13, 114, 191n2, 206, 257–59, 323–24, 333, 390, 529, 531, 548

subject-based approach 2, 3, 20, 21, 190, 518, 526, 540

impersonalization of passive constructions, see evolution from passive to impersonal impersonals

A-impersonals 4, 6, 7, 13, 17, 19, 44–45, 51


of necessity 11, 459, 461, 467, 479, 480, 486

man-impersonals, see man-constructions, MAN-IMPS

man-impersonal passive 479, 486

man-impersonalized 232, 250–51

man-impersonal pronoun 28, 282, 2468, 581, 584, 597, 600–603, 604. See also indefinite pronoun, expletive pronoun, MAN-IMPS

man-impersonal reflexive 59


Global Repulsion

Hypothesis 133, 137–39, 151, 158

lexical 11, 489–90, 498–99, 507, 512, 514

primary impersonal verb 459–61, 470, 479, 486

of transitive verbs, see transimpersonal constructions


function-based approach 4, 7, 9, 12, 13

hybrid approach 258

prototype approach 21–24, 36, 38, 40

structural approach 1–2, 5, 7, 9, 10, 11–13, 114, 191n2, 206, 257–59, 323–24, 333, 390, 529, 531, 548

subject-based approach 2, 3, 20, 21, 190, 518, 526, 540

impersonalization of passive constructions, see evolution from passive to impersonal impersonals

A-impersonals 4, 6, 7, 13, 17, 19, 44–45, 51


of necessity 11, 459, 461, 467, 479, 480, 486

man-impersonals, see man-constructions, MAN-IMPS

man-impersonal passive 259, 336

R-impersonals 4, 6–8, 12–13, 44–45, 51, 57–61, 68–70, 85–86, 330

T-impersonals 4, 6–7, 13, 44, 51, 128

topographic impersonals 501–2

inanimate subject, see under subject

incorporation 117, 119, 160–61, 403, 413, 440, 446, 448,
inversion (word order)
involuntary process
involuntary action
indefinite pronouns
locative arguments (location
universal
corporate
vague
indefinite subject, see under
subject. See also
man- constructions
inverse construction
inversion (word order
inversion)
involution passive
involuntary action
involuntary agent, see
non-volitional subjects
(agents), impersonals:
A-impersonals
involuntary process
labile (verb)
language development
locative arguments (location
arguments)
locative inversion
man- constructions
MAN-IMPS
meteorological predicate
meteo-verb (meteo-
construction, 
meteo-predicate)
multifactorial definition of
subject
subject
natural forces
non-volitional subjects
non-volitional knowledge
oblique actor phrase
omission of arguments
passive
passive
prominence
raising verb, see under verbs
reanalysis
non-promotional
passive
passive of intransitive
passive participle
personal passive
personal presentational
construction
copular presentational
non-formulaic
copular
transitive
modality marker
pronouns
indefinite
involuntary
uninverted
specific existential
universal
vague
See also
–
9,
–
37
–
563
n
452
92,
578
49, 557
3, 290,
79
81, 191, 287
40
149, 150, 151, 292, 333, 332,
535, 539–40
34, 37–38
44, 260–61, 281, 491,
496–97, 511, 513, 622
5, 29, 30, 31,
38–39, 40n9, 42, 129–32,
149, 150, 151, 292, 333, 332,
535, 539–40
12, 581, 582–97
324, 344, 347
144, 147,
149–50, 152, 161–62, 164,
362, 365–66, 388
5, 31, 39–40,
129–32, 149, 151, 292
4, 7–8, 28,
37n7, 57–71, 73–86, 191,
307, 338, 345–46, 347n9,
351, 558, 577, 600. See also
MAN-IMPS
4, 7–8, 28, 57,
60–71, 73, 77, 79–81,
83–86
(state, event, verb,
construction) 2, 10, 257,
269, 290–91, 292, 304,
307–21, 395–433, 499–501,
506, 507, 512, 582–84, 587,
603
(agents) 24, 33–35, 37–38,
42, 45, 47, 50, 98. See also
inanimate subject under
subject, impersonals:
A-impersonals
A-impersonals
knowledge 229, 249
46,
49n17, 383, 517, 534, 538,
543, 581–82, 591, 594–96,
604
38, 42,
248–49, 451, 540, 541, 581,
590, 595, 597–600
96,
107n12, 109–10, 342
1–3, 5,
7–11, 29, 36–37, 41, 44,
46, 49, 72n6, 82n14, 83,
91–122, 170, 178, 185,
190–92, 194–95, 229,
250–51, 293, 336, 344,
357–58, 362, 375, 377–84,
387–88, 390, 396, 459,
461, 480, 508, 511–12,
517–43, 595
non-promotional
passive 189
passive of intransitive
191–92, 509
passive participle 93, 96, 111,
118, 247, 372
personal passive 9, 12, 46,
110, 383–84, 390, 508,
511, 518–19, 532, 534, 538,
542–43
politeness 533, 558, 582,
601, 604
postverbal subject 5, 30,
128, 130, 134–35, 137–39,
144–45, 154–55, 158–59,
163, 204, 266, 315
construction
4–5, 7,
8, 20, 24, 29, 37–38, 39,
41–43, 109, 127–64, 285,
292, 304, 499, 505–7,
568n8, 570
135, 138, 140–43, 148–58, 163,
315, 320
formulic 8, 127–64, 336,
348
personal
presentational
127–28, 137
non-formulic 129, 132–33,
336
personal presentational
127, 129
presentational focus
construction 19, 285, 304
presentational
transitive 141, 158–62, 163
pro-drop 3, 7–8, 28, 42, 57, 61,
77–79, 81, 83, 86, 191, 301,
324–25, 520, 577
192, 333
from impersonal suffix to
modality marker
from impersonal to personal
passive 9, 46, 519
from O to S 12, 517, 538, 543
from object to subject, see
reanalysis from O to S

Subject index 639
from passive to anticasative 9, 229, 244
from passive to impersonal 9, 194. See also evolution from passive to impersonal from transimpersonal to intransitive 11
referential impersonalizing construction 547
resultative participle 113, 365, 373–78, 384, 385, 37, 390
S
semantic map 4, 7, 9, 19–21, 23–24, 38, 40–44, 50–51
sentence-focus construction 29, 132, 137–38, 292, 316, 490, 502–7. See also presentational construction
split ergativity 95, 115, 121, 620
subject
canonical subject 2, 23, 38, 44, 48, 144n9, 191, 285, 325, 548, 570. See also prototypical subject experiencer subject 514
expletive subject 2, 176, 151, 157, 191, 193, 286, 300, 347, 548
facultative subject 409, 416, 433
generic subject 38–40, 45, 47, 49, 50, 311, 324, 328, 329, 341–42, 345
inanimate subject 11, 12, 32–33, 39, 41, 172n5, 489–90, 495n12, 505n17, 514, 617, 619
indefinite subject 6, 27–29, 30, 36–37, 42, 45–46, 439, 519, 538
non-canonical subject 3, 5, 9–10, 47, 50, 133–39, 258, 307–9, 369
non-topical subject 24, 29–32, 39, 317
non-volitional 24, 33–38, 42, 45, 47, 50
oblique subject 20, 47, 122, 325
partitive subject 416
patientive subjects 47
prototypical subject 19, 22–23, 51, 264. See also canonical subject
acquisition of subject properties 22n1, 30, 39–40, 128, 139, 147, 151, 158, 159, 160, 162–64, 197, 259, 264–68, 517, 527, 538–39, 543, 623
coding 19, 21–24, 30, 36, 38–41, 44–45, 49, 51, 169–70, 175–76, 260n4, 308–9, 316–21, 494–98, 534, 538, 622
functional 19–20, 24, 50
subjecehood 23, 57, 100, 133–34, 137, 139, 145, 148, 160, 189, 194, 198–99, 307–9, 489, 519
criteria for 134, 145n9, 148, 609–10, 612, 619–23
there-existentials 129–30, 138, 149, 571
with modal operators 323, 328, 334, 337
with plural predicates 328
T
thctic construction 499, 505
thctic languages
thctic-V1 languages 8, 127, 129, 134, 139, 143–44, 146, 154, 162–64
thctic-XV languages 127, 129, 134, 137, 139, 142–43, 146, 162–64
thcticity 10, 110, 307, 308, 316–21
third-person plural impersonals 28, 56, 60, 329. See also 3pl-IMPS 3pl-IMPS 7, 8, 57, 60–76, 79–83, 85–86
transimpersonal constructions 5–7, 11–13, 28, 44–50, 439, 453, 457, 459, 461, 475n18, 480, 490, 512–14, 609, 613, 615
transitivity parameters 23, 91, 98
U
underspecification 4, 10, 91, 105–7, 357, 387, 390
unspecified agent 9, 259, 307, 514
V
verbs
Animate Intransitive 493, 502, 505, 510n21, 511
avalent verbs 396, 400–409, 419, 421, 423, 427, 499, 504, 512
Subject index

verb-focus constructions 31, 44
verbless constructions 10, 295, 307
voice, see passive, reflexive

W
weather verb 1, 24, 4110, 400–401, 407, 471–72, 499. See also meteo-verb, meteorological predicate, atmospheric verb

Z
zero-anaphora 590, 597–600, 604
zero argument 499, 506
zero lexeme 3
zero pronoun 12, 547, 550, 555, 561, 563, 573–77
zero subject 26, 28, 602
zero valence 121, 395–96