

# Linguistic Typology

---

## at the Crossroads



ISSN 2785-0943

Volume 1 – issue 1 – 2021

Issue DOI: <https://doi.org/10.6092/issn.2785-0943/v1-n1-2021>

This journal provides immediate and free open access. There is no embargo on the journal's publications. Submission and acceptance dates, along with publication dates, are made available on the PDF format for each paper. The authors of published articles remain the copyright holders and grant third parties the right to use, reproduce, and share the article according to the [Creative Commons Attribution 4.0 International](#) license agreement. The reviewing process is double-blind. Ethical policies and indexing information are publicly available on the journal website:

<https://typologyatcrossroads.unibo.it>

### Editors

*Nicola Grandi* (University of Bologna, Editor in chief)

*Caterina Mauri* (University of Bologna, Editor in chief)

*Francesca Di Garbo* (University of Aix-Marseille)

*Andrea Sansò* (University of Insubria)

### Publisher

*Department of Classical Philology and Italian Studies* (University of Bologna)

*Department of Modern Languages, Literatures and Cultures* (University of Bologna)

The journal is hosted and maintained by [AlmaDL](#)



# Linguistic Typology

---

## at the Crossroads



### Editorial board

*Mira Ariel* (Tel Aviv University)  
*Sonia Cristofaro* (Sorbonne Université)  
*Chiara Gianollo* (University of Bologna)  
*Matti Miestamo* (University of Helsinki)  
*Marianne Mithun* (University of California Santa Barbara)

### Scientific Board

*Giorgio Francesco Arcodia* (Università Ca' Foscari, Venice)  
*Peter Arkadiev* (Johannes-Gutenberg University of Mainz)  
*Gilles Authier* (École Pratique des Hautes Études, Paris)  
*Luisa Brucale* (University of Palermo)  
*Holger Diessel* (University of Jena)  
*Eitan Grossman* (The Hebrew University of Jerusalem)  
*Corinna Handschuh* (Universität Regensburg)  
*Guglielmo Inglese* (University of Turin)  
*Elisabetta Magni* (University of Bologna)  
*Francesca Masini* (University of Bologna)  
*Susanne Maria Michaelis* (MPI EVA – Leipzig)  
*Emanuele Miola* (University of Bologna)  
*Anna Riccio* (University of Foggia)  
*Eva van Lier* (University of Amsterdam)

### Responsible Editor

*Caterina Mauri*, University of Bologna

Department of Modern Languages, Literatures and Cultures, via Cartoleria 5, 40124 Bologna. Email: [caterina.mauri@unibo.it](mailto:caterina.mauri@unibo.it)

### Production editors

*Silvia Ballarè* (University of Bologna)  
*Alessandra Barotto* (University of Insubria)  
*Simone Mattioli* (University of Bologna)  
*Eleonora Zucchini* (University of Bologna)

### Assistant production editors

*Antonio Bianco* (University of Pavia)  
*Valentina Di Falco* (University of Bologna)  
*Sara Gemelli* (University of Pavia)  
*Maria Cristina Lo Baido* (University of Cagliari)  
*Nicola Perugini* (University of Bologna)  
*Antonia Russo* (University of Bergamo)



# Linguistic Typology at the Crossroads



## CONTENTS

### Linguistic Typology at the Crossroads - Editorial paper

*Francesca Di Garbo, Nicola Grandi, Caterina Mauri, Andrea Sansò* ----- 1-5

### Hypothetical manner constructions in world-wide perspective

*Jesús Francisco Olguín Martínez*----- 6-46

### Expressing equality and similarity in English, Italian, and Ladin: Interlingual contrastive features and micro-variation

*Martina Irsara*----- 47-93

### The comparative cycle in crosslinguistic perspective

*Agnes Jäger* ----- 124-178

### Comparative of Inferiority: Marking and Aspects of Use

*Valeria Modina*----- 179-222

### Small events. Verbal diminutives in the languages of the world

*Jenny Audring, Sterre Leufkens, Eva van Lier* ----- 223-256

### Diachronic evolution of Russian standard markers *kako* and *aky*

*Irina Kobozeva, Natalia Serdobolskaya* ----- 257-287

### Comparative disparity: comparative constructions in the Walser German islands of Italy

*Livio Gaeta* ----- 288-332



## Linguistic Typology at the Crossroads

CATERINA MAURI<sup>2</sup>, NICOLA GRANDI<sup>2</sup>, FRANCESCA DI GARBO<sup>1</sup>, ANDREA SANSÒ<sup>3</sup>

<sup>1</sup>UNIVERSITY OF HELSINKI, <sup>2</sup>UNIVERSITY OF BOLOGNA, <sup>3</sup>UNIVERSITY OF INSUBRIA

- |  |   |
|--|---|
| 1 <i>Two roads diverged in a yellow wood,<br/>And sorry I could not travel both<br/>And be one traveler, long I stood<br/>And looked down one as far as I could<br/>To where it bent in the undergrowth;</i>         | 11 <i>And both that morning equally lay<br/>In leaves no step had trodden black.<br/>Oh, I kept the first for another day!<br/>Yet knowing how way leads on to way,<br/>I doubted if I should ever come back.</i> |
| 6 <i>Then took the other, as just as fair,<br/>And having perhaps the better claim,<br/>Because it was grassy and wanted wear;<br/>Though as for that the passing there<br/>Had worn them really about the same,</i> | 16 <i>I shall be telling this with a sigh<br/>Somewhere ages and ages hence:<br/>Two roads diverged in a wood, and I—<br/>I took the one less traveled by,<br/>And that has made all the difference.</i>          |

**Robert Frost, 'The road not taken'**

Welcome to *Linguistic Typology at the Crossroads* (LTC), a new journal dedicated to the crossroads where linguistic typology meets its neighboring fields.

A crossroads is not only and necessarily a place where a choice must be made: this is the point of view of the traveler, wondering which way to go. If we take a bird's-eye view and observe the crossroads from above, we see much more than choices. A Crossroads is the place where different directions, and different travelers, meet or follow each other. Thanks to the crossroads, it is possible to change and exchange, and the very concept of 'step forward' opens itself to diagonals and curves. The crossroads is where future is imagined and innovation occurs, thanks to the reciprocal influence of intersecting perspectives.

Crossroads means meeting, exchanging, converging, choosing, diverging, changing, and possibly making the difference.

This journal aims to take the point of view of the crossroads, capturing the moment when linguistic typology intersects other fields, changing and exchanging methods, theories and data, in the belief that a closer look at the crossroads may reveal converging paths and new directions to go.

Aiming to find, describe and explain linguistic diversity, linguistic typology naturally crosses the study of the emergence and evolution of the world's grammars (Bybee 2015) and the observation of language-internal variation (Croft 2000, 2007). Likewise, explanations for language universals and cross-linguistic tendencies are typically found in discourse use, paths of language change and/or cognitive mechanisms (Keller 1994, Hawkins 2004 and 2014, Haspelmath 2021, among others). This means that the methodological and theoretical tools of linguistic typology are by definition at least partially shared with neighboring fields, in a dynamic intersection that can be itself described as an ever-changing crossroads.

*Linguistic Typology at the Crossroads* aims to give space above all, but not exclusively, to studies exploring the crossroads at which linguistic typology meets areas of linguistics, such as language documentation, language change, sociolinguistics, psycholinguistics, neurolinguistics, language technology, educational linguistics, corpus-based analysis of speech and discourse.

Thanks to the advanced infrastructure provided by ABIS-AlmaDL at the University of Bologna and to the support of the Department of *Classical Philology and Italian Studies* (FICLIT) and the Department of *Modern Languages, Literatures and Cultures* (LILEC), the journal meets the highest international standards for scientific publications. *Linguistic Typology at the Crossroads* is indeed the first peer-reviewed journal in the field of linguistic typology that provides immediate and free open access to all publications, with no embargo and no publication fees (Diamond OA).

The journal publishes two issues per year. Papers accepted for publication are selected solely on the basis of scientific quality and scholarly standing, after undergoing a double-blind peer-review. Members of the Editorial and Scientific Committee have been invited based on their scientific profile and their expertise in different areas and approaches to linguistic typology.

The authors of published articles remain the copyright holders and grant third parties the right to use, reproduce, and share the article according to the Creative Commons Attribution 4.0 International license agreement. Authors are welcome to post pre-submission versions, the original submitted version of the manuscript (preprint) and the final post-refereeing draft (postprint) on other platforms (personal

website, a collaborative wiki, departmental website, social media websites, institutional repository or non-commercial subject-based repositories).

---

This first issue is thematic and focuses on **comparative constructions across languages**. It gathers selected contributions from the TypoBO Workshop,<sup>1</sup> organized for the Summer 2020 and canceled due to the COVID pandemic.

Comparative constructions are a set of grammatical strategies that the languages of the world use to compare two or more items in order to highlight both differences and similarities among them (cf. Dixon 2008, Stolz 2013, Treis 2018). We can recognize different types of comparative constructions depending on the kind of relationship existing between the two (or more) items (Fuchs 2014, Treis 2018). The first relevant distinction to be made is between quantitative comparison and qualitative comparison (cf. Treis 2018: iii): quantitative comparison can be further divided into comparison of inequality (superiority, e.g. *taller than*, *tallest*, or inferiority, e.g. *less tall than*, *least tall*) or equality (e.g. *as tall as*), while qualitative comparison can be further distinguished into similarity (e.g. *like a horse*) and simulation (e.g. *as if he were a horse*).

In this issue, the contributions by Ivani and Gaeta take into account comparative constructions of inequality, in Tibeto-Burman languages and across German minorities of Italy, respectively. The comparison of inferiority is instead studied by Modina, both from a cross-linguistic point of view and with a focus on Russian, and by Audring & Leufkens & van Lier, who provide a comprehensive typology of verbal diminutives. Equality and similarity relations are the object of Irsara's paper, who examines Ladin data in a comparative perspective with Italian and English, while simulation (or hypothetical manner) is investigated in a typological language-sample by Martínez.

The languages of the world exhibit several different formal strategies to express these functions (e.g. Ultan 1972, Andersen 1983, Stassen 1985, Cuzzolin & Lehmann 2004, Dixon 2008, Stolz 2013, Treis & Vanhove 2017, Treis 2018). However, we can generally identify some elements that are cross-linguistically recurrent within a comparative construction (cf. Stolz 2013: 9 and Treis 2018: ii). The comparee and the

---

<sup>1</sup> Organizers: Alessandra Barotto, Nicola Grandi, Simone Mattiola, Caterina Mauri.

<https://eventi.unibo.it/typobo-2020>

standard are the items being compared to each other, respectively, the *primum comparationis* (comparee) and the *secundum comparationis* (standard). The parameter or quality is the property on which comparee and standard are compared (*tertium comparationis*). The degree (or parameter) marker explicates the kind of relationship that comparee and standard have with respect to the parameter/quality (e.g. *more, less, as ... as* in English). Finally, the tie or standard marker is the grammatical function or relation that connects comparee, standard, and quality (e.g. *than* in English).

Standard markers of comparison seem to develop out of a restricted set of recurrent sources (cf. Heine & Kuteva 2002), such as verbal forms meaning ‘exceed/defeat/surpass’ (Cantonese *kwo* ‘surpass’ > *kwo* ‘than’) and interrogative adverbs (Hungarian *mint* ‘how’ > *mint* ‘than’). In this issue, the paper by Jäger traces the diachronic development of comparison constructions crosslinguistically, highlighting a recurrent pattern of change with respect to standard markers, i.e. what she calls the comparative cycle. A diachronic perspective is also adopted by Kobozeva & Serdobolskaya, who follow the evolution of Russian standard markers *kako* and *aky*.

Papers in this first issue combine studies based on typological samples (cf. papers by Audring & Leufkens & van Lier, Martínez, and Jäger) with studies focusing on particular linguistic areas or families (cf. Gaeta on German minorities in Italy, Ivani on Tibeto-Burman, Irsara on Ladin, Modina and Kobozeva & Serdobolskaya on Russian), integrating the analysis of descriptive grammars with the analysis of corpus data, to actually understand the discourse use of comparatives (cf. Modina on Russian, Irsara on Ladin). The typological approach thus fruitfully stands at the intersection with a sociolinguistic and areal perspective (especially in Ivani’s, Irsara’s and Gaeta’s contributions), with diachronic methodology (cf. papers by Jäger, Kobozeva & Serdobolskaya, and Gaeta) and with corpus-based research (cf. contributions by Modina, Irsara, Kobozeva & Serdobolskaya).

Linguistic typology is at the crossroads. See you there!

## References

- Andersen, Paul Kent. 1983. *Word order typology and comparative constructions*. Amsterdam / Philadelphia: John Benjamins.

- Bybee, Joan L. 2015. *Language change*. Cambridge / New York: Cambridge University Press.
- Croft, William. 2000. *Explaining language change: an evolutionary approach*. Harlow: Pearson education.
- Croft, William. 2007. The origins of grammar in the verbalization of experience. *Cognitive Linguistics* 18(3). 339-382.
- Cuzzolin, Pierluigi & Lehmann, Christian. 2004. Comparison and gradation. In Geert Booij & Christian Lehmann & Joachim Mugdan & Stavros Skopeteas (eds.), *Morphologie. Ein internationales Handbuch zur Flexion und Wortbildung* 17.2, 1857–1882. Berlin / New York: Mouton de Gruyter.
- Dixon, R.M.W. 2008. Comparative constructions: A cross-linguistic typology. *Studies in Language* 32(4). 787–817.
- Fuchs, Catherine. 2014. *La comparaison et son expression en français*. Paris: Ophrys.
- Haspelmath, Martin. 2021. Explaining grammatical coding asymmetries: Form–frequency correspondences and predictability. *Journal of Linguistics* 57(3). 605–633.
- Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.
- Hawkins, John A. 2014. *Cross-linguistic variation and efficiency*. New York: Oxford University Press.
- Heine, Bernd & Kuteva, Tania. 2002. *World lexicon of grammaticalization*. Cambridge / New York: Cambridge University Press.
- Keller, Rudi. 1994. *On language change: the invisible hand in language*. London: Routledge.
- Stassen, Leon. 1985. *Comparison and universal grammar*. Oxford: Basil Blackwell.
- Stolz, Thomas 2013. *Competing comparative constructions in Europe*. Berlin: Akademie-Verlag.
- Treis, Yvonne 2018. Comparative constructions: an introduction. *Linguistic Discovery* 16(1). i–xxvi.
- Treis, Yvonne & Vanhove, Martine (eds). 2017. *Similative and equative constructions. A cross-linguistic perspective*. Amsterdam / Philadelphia: John Benjamins.
- Ullian, Russell. 1972. Some features of basic comparative constructions. *Working Papers on Language Universals (Stanford)* 9. 117–162.



# Hypothetical manner constructions in world-wide perspective

JESÚS FRANCISCO OLGUÍN MARTÍNEZ<sup>1</sup>

<sup>1</sup>UNIVERSITY OF CALIFORNIA, SANTA BARBARA

Submitted: 17/11/2020    Revised version: 23/03/2021

Accepted: 25/03/2021    Published: 31/08/2021

## Abstract

Similatives (e.g. *she swims like a fish*) have been the focus of a number of investigations (e.g. Treis & Vanhove 2017). However, hypothetical manner constructions (e.g. *She treats me as if I were a stranger*) have received little attention cross-linguistically. Therefore, our typological knowledge of this type of comparative clause is still in its infancy. This paper offers an analysis of the cross-linguistic variation in the expression of hypothetical manner constructions in a sample of 61 languages. Among the most common strategies found are simulative ‘like’ markers and free adverbial conjunctions. Also discussed are other rare strategies, which seem to show clear areal patterns. In particular, some languages from Mesoamerica use correlative words, some Australian languages use counterfactual mood markers and some African languages employ head nouns meaning ‘thing’. This paper also explores whether hypothetical manner constructions show formal resemblances to other constructions.

**Keywords:** Comparative clauses; adverbial clauses; hypothetical manner constructions; similatives; real manner constructions.

## 1. Introduction

Qualitative comparative constructions “do not express a quantitative gradation on a particular parameter, but bring together the two terms of the comparison on the basis

of similarity or likeness” (Fuchs 2014: 133).<sup>1</sup> Qualitative comparative constructions are divided into similarity (e.g. *she swims like a fish*) and hypothetical manner (e.g. *She treats me as if I were a stranger*) (Treis 2018: iii). The linguistic typological literature has especially been concerned with similatives (e.g. Treis & Vanhove 2017) while hypothetical manner constructions have received little attention cross-linguistically (but see Hetterle 2015: 195). To the best of my knowledge, this construction has been explored mostly in individual languages, such as Pesh (Chamoreau 2017: 331-332), Zaar (Caron 2017: 183) and North Saami (Ylikoski 2017: 275) and in European languages (e.g. Kortmann 1997: 284). Therefore, our typological knowledge of this type of comparative construction is still in its infancy. Martowicz (2011: 144) mentions that this type of semantic relation has not as yet received serious linguistic interest, let alone special attention in any cross-linguistic study. In a similar fashion, Hetterle (2015: 195) points out that hypothetical manner constructions are unexplored territory in that no previous studies have addressed the precise semantic and discourse functions of this type of comparative clause.

The aims of this paper are two-fold. First, this research offers an analysis of the cross-linguistic variation in the expression of hypothetical manner in a sample of 61 languages. Kortmann’s (1997) study of free adverbial conjunctions in the languages of Europe shows that hypothetical manner constructions tend to be formed by phrasal clause-linking devices, such as Spanish *como si* ‘as if’, German *als ob* ‘as if’, Portuguese *como se* ‘as if’ and Icelandic *likt og* ‘as if’. However, it is not entirely clear whether languages from other areas of the world also use free adverbial conjunctions to encode hypothetical manner constructions. The present research should make it clear that there may be more to the story, in that languages may use other clause-linking devices to express hypothetical manner. Some languages from Mesoamerica use correlative words, as in (1), some Australian languages use counterfactual mood markers, as in (2) and some African languages employ head nouns meaning ‘thing’, as in (3).

(1) Silacayoapan Mixtec (Oto-Manguean/Mixtecan; Shields 1988: 431-432)

<i>tá</i>	<i>tuhūn</i>	<i>ndáā</i>	<i>nā</i>	<i>xyoko,</i>	<i>xá</i>	<i>ndáā</i>	<i>nā</i>
if	word	appear:PL	3PL.SBJ	San.Andrés	thus	appear.PL	3PL.SBJ

<sup>1</sup> Translated from the French original: “... une comparaison qualitative qui n’opère pas de gradation quantitative sur un paramètre, mais rapproche les deux termes de la comparaison sur la base d’une similarité ou d’une analogie.”

*ndahví.*

poor

‘As if they were from San Andrés, those poor people look (lit. about like the people from San Andrés appear, so appear those people).’

- (2) Kayardild (Tangkic; Evans 1995: 378)

*jani-jani-ja niwan-ju, yakuri-ya buru-tha thaa-tha marak.*

search-RDP-ACT 3SG-PROP fish-LOC take-ACT return-ACT CF

‘They searched for him, as if they were going out to get fish.’

- (3) Dogul Dom (Dogon; Heath 2016: 269)

*è-wé gòŋ ñǎ: bè ñà:l yáŋ, yǎ: yê:-n.*

child-PL thing meal 3PL.SBJ eat-PFV.NEG.PL like tears weep-IPFV.3PL.SBJ

‘As if the children have not eaten, they are crying.’

Given that these strategies seem to be only attested in particular areas forming areal clusters, it is proposed that the most obvious explanation seems to be language contact. This is because: (1) the languages are spoken in the same region, (2) they are not genetically related and (3) the probability of chance resemblance is low, given the rarity of the strategies. Interestingly, the forms of the strategies are not the same. Accordingly, speakers seem to have replicated these clause-linking strategies with native material. This is known as pattern replication. In this scenario, only the patterns of the other language are replicated, i.e. the organization, distribution and mapping of grammatical or semantic meaning, while the form itself is not borrowed (Weinreich 1964: 39; Heath 1978; Sakel 2007: 15; Heine & Kuteva 2008). Therefore, this research aims at contributing to theories of language contact in that it can help us to better understand how a particular grammatical pattern may have spread to different neighboring languages not genetically related (e.g. the different possible directions from which a particular development could have been stimulated; Mithun 2012).

Second, this paper explores whether hypothetical manner constructions show formal resemblances to other constructions. Some work has shown that hypothetical manner constructions exhibit formal and functional resemblances to similatives (e.g. *The man swims like a fish*; Fortescue 2010: 131; Chamoreau 2017: 331-332). Another construction that is also very similar to hypothetical manner constructions is that of

real manner adverbial clauses (e.g. *do as I told you*; Darmon 2017: 372-373). In achieving this goal, the guiding questions are: do hypothetical manner constructions tend to resemble similative or real manner constructions formally? How can hypothetical manner constructions be classified according to whether they resemble similatives and/or real manner constructions? Are any of these systems frequent in particular areas of the world? If hypothetical manner, similative and real manner constructions are expressed by the same marker in a particular language, how do the different meanings arise (e.g. context, specific TAM values)? Put it another way, if 'X' language employs the same marker to express hypothetical manner, similative and real manner, how do speakers differentiate these meanings? Is context the only factor that can distinguish them? Or do specific TAM markers aid in the disambiguation process? Some work has shown that TAM markers may differentiate one adverbial meaning from others (Hetterle 2015). A case in point comes from Lango (Nilotic/Eastern Sudanic). This language expresses *when*-relations and *after*-relations by means of the device *àmê*. Interestingly, the 'when' interpretation arises in combination with the progressive aspect and the 'after' interpretation arises in combination with perfective aspect (Noonan 1992: 243-246).

The structure of the paper is as follows: Section 2 provides some remarks on hypothetical manner constructions. Section 3 addresses methodological questions relating to the language sample of the present study and the collection and analysis of the data. Section 4 discusses the range of strategies used to express hypothetical manner in the language of the sample, highlighting the role of language contact in the spread of some strategies. Section 5 explores the formal and functional resemblances that hypothetical manner constructions show to other constructions, in particular to similatives and real manner constructions. Section 6 summarizes the main points of the study as a whole and presents its overall conclusions.

## 2. Hypothetical manner: some remarks

A large number of unrelated languages scattered throughout the world share a complex sentence construction that portrays an imagined ('do X as if it was caused by Y'), or counterfactual ('do X as if Y were true') situation (Dixon 2009: 35; Hetterle 2015: 54; Darmon 2017: 372-373). Because of the lack of typological studies, there is as yet no consensus on the proper terminology for referring to this construction.

Kortmann (1997: 195) employs the term “comparative clause” to refer to hypothetical manner clauses. He mentions that ‘comparative clauses’ encoded by morphemes whose meaning is close to ‘as if’ express a situation that is typically hypothetical. Hengeveld (1998: 355) employs the term “unreal circumstance clause”. He notes that this construction is introduced by a counter-factive conjunction characterizing the situation as not real. Vanhove (2017: 206), in her description of similative, equative and comparative constructions in Beja (North-Cushitic), uses the term “pretence clauses”. Roulon-Doko (2017: 226) calls this construction “modus essendi”. Heath (2014, 2016), in various grammatical descriptions of Dogon languages, uses the term ‘counterfactual manner adverbial clauses’. Finally, Treis (2017: 125) employs the term “hypothetical similarity clause”. In this paper, Dixon’s terminology (i.e. hypothetical manner clauses) has been adopted in that it seems to be the most accessible term to refer to this construction. The term ‘simulative’ may be confused with the term ‘similative’ because of their phonological similarity. The term “comparative clause” used by Kortmann (1997) is also ambiguous in that there are different types of comparative constructions (e.g. inequality, superlatives).

Hypothetical manner constructions may be considered adverbial clauses or complement clauses. In order to flesh out this claim, however, it is necessary to elaborate somewhat on what is meant by adverbial clauses and complement clauses. Hypothetical manner constructions may be adverbial clauses, that is, non-argument clauses that relate to the predicate or the entire proposition expressed by another clause (i.e. the main clause) (Schmidtke-Bode & Diessel to appear: 2). This is nicely illustrated in the West Coast Bajau example in (4), in which the non-argument clause introduced by *masam* ‘as if’ fulfils a semantic and syntactic role in another unit. Accordingly, the dependent clause spells out part of the setting of the main-clause situation.

(4) West Coast Bajau (Austronesian/Sama-Bajaw; Miller 2007: 418)

*be-sinar-sinar no emas e masamkeadaan kampung e tunu.*  
DISTR-shine-RDP FOC gold DEM as.if condition village DEM burn  
‘The gold shimmered as if the village were burning.’

Hypothetical manner constructions may also be considered complement clauses, that is, the predicate of one clause “entails reference to another proposition or state of affairs expressed in a second clause” (Cristofaro 2003: 95). They would function as a

syntactic argument of a higher clause (Dixon 2006: 15). The range of semantic classes of complement-taking predicates in this construction is rather limited in that only some verbs may appear in this environment. First, hypothetical manner constructions in which the verb of the main clause means ‘to act’ or ‘to pretend’ are known in the literature as MISTAKEN IDENTITY constructions (see Spronck 2015; Spronck & Vuillemet 2019). The mistaken identity involves a reversal of polarity by expressing some ideas inconsistent with behavior of a particular participant (Qian 2016: 220), as is shown in the Donno So example in (5).

(5) Donno So (Dogon; Heath 2016: 269)

*ù = η      bëndé-dê-η      gìnè      kán-jê-Ø.*  
 2SG = ACC hit-IPFV-LOG.SBJ as.if      act-IPFV-3SG.SBJ  
 ‘He acts as if he’s going to hit you.’

Second, hypothetical manner constructions in which the verb of the main clause means ‘to look’ or ‘to seem’, as in (6), are known in the literature as EPISTEMIC-JUDGEMENT PREDICATES (Schmidtke-Bode 2014: 44) and they belong to the domain of propositional modality (Palmer 2001: 8). This stems from the fact that speakers express their judgments about the factual status of the proposition (Palmer 2001: 8). This type of hypothetical manner construction is a subject complement clause. Schmidtke-Bode (2014: 44) mentions that “the experiencer, or holder, of the propositional attitude is normally the speaker, and the proposition whose truth is evaluated is coded as a complementation pattern in a main clause.”

(6) Ojibwe (Algonquian; Valentine 2009: 214)

*dibishkoo      miznaakide-g      izhinaagwad-w.*  
 as.if      be.printed-CNJ look-IND.OBJ  
 ‘It looked as if there were printing on it.’

This study only takes into account adverbial hypothetical manner clauses. However, based on a close inspection of the languages of the sample, it has been found that hypothetical manner constructions have usually developed by extension from the adverbial domain to the complementation domain. This theoretical fact has not gone unnoticed. López-Couso & Méndez-Naya (2015: 193) show that this development is not restricted to English and other Indo-European languages, such as Spanish, Dutch

and German, but can also be found in other languages (e.g. Caucasian languages). What this seems to indicate is that this connection cannot be considered a language specific phenomenon, but rather a development common in many languages not genetically related. López-Couso & Méndez-Naya (2015: 196) mention that this development is a case of secondary grammaticalization, that is, it refers to “increased grammaticalization of already grammatical items in specific contexts” (Hopper & Traugott 2008: 175). With this in mind, the hierarchy put forward in (7) aims at capturing this tendency.

- (7) Adverbial clause > complement clause (‘to look’ or ‘to seem’) > complement clause (‘to act’ or ‘to behave’) > in subordinate clause<sup>2</sup>

There are two main theoretical observations to be gleaned from the hierarchy in (7). First, if adverbial hypothetical manner constructions and mistaken identity constructions (‘to act’ or ‘to behave’) are encoded by the same linking device in ‘X’ language, epistemic-judgement predicates (‘to look’ or ‘to seem’) will also tend to be encoded in the same way. Second, one further development attested in the languages of the sample of the present study is that of insubordination, that is, once adverbial hypothetical manner constructions develop into complement clauses, they may develop into in subordinate clauses, i.e. “the conventionalized independent use of a formally subordinate clause” (Evans 2007: 377). This development is only attested in a few languages of the sample, mainly Indo-European. For instance, in the Spanish example in (8), the *como si* ‘as if’ in subordinate construction has the illocutionary force of an exclamation, that is, the *como si* ‘as if’ has come to serve another function (e.g. incredulity, disbelief, repulsion, disgust), but at the same time the construction retains its irreality value. Given that only a few languages have in subordinate clauses functioning in this way, further studies will enable us to explore in more detail the functions served by this type of in subordinate construction.

- (8) Spanish (Indo-European/Romance)  
*como si tuvieras suficiente dinero!*  
as if have.SBJV enough money  
‘As if you had a lot of money!’

---

<sup>2</sup> In English the ‘as if’ conjunction can be used colloquially on its own (e.g. “Miss me, honey?” *As if. Where’s my crab cake?*; Brinton 2014: 95). This usage has not been attested in the languages of the sample.

Before leaving the present section, mention should be made of another domain relevant to the study of hypothetical manner constructions. Hypothetical manner constructions may be encoded by TAM markers appropriate to this context in the languages of the sample, such as irrealis, subjunctive markers and counterfactual mood markers. For instance, Chafe (1995) mentions that Caddo (Caddoan) has a realis/irrealis distinction encoded within pronominal prefixes on verbs. He notes that irrealis pronouns are used in several contexts, such as yes/no questions, prohibitions, obligations, conditions and hypothetical manner constructions. The fact that hypothetical manner constructions appear with TAM markers that harmonize with their meaning is the expected scenario in that, as explained by Darmon (2017: 372-373), hypothetical manner constructions portray an imagined ('do X as if it was caused by Y'), or counterfactual ('do X as if Y were true') situation. However, there are languages in which hypothetical manner constructions are encoded by past tense marking, perfective marking and/or completive marking. This is only attested in a few languages of the sample. In Korean, the hypothetical manner clause has to appear with the past tense marker *-at*. In a similar fashion, in Uyghur, the hypothetical manner construction occurs with the past tense marker *-d*.

(9) Korean (Koreanic)<sup>3</sup>

*ki-nin matɕʰi njʌn toŋan maktɕi anin kat-tɕʰarʌm mʌk-ʌt-ta.*  
 3SG.SBJ-TOP as.if year during eat NEG thing-as.if eat-PST-DECL  
 'He ate as if he had not eaten in years.'

(10) Uyghur (Turkic)<sup>4</sup>

*u xuddi hëchqachan tamaq ye-mi-gen-dek yë-d-i.*  
 3SG.SBJ like never food eat-NEG-PTCP-SIM eat-PST-3SG  
 'S/he ate as if s/he had never eaten before.'

Given that hypothetical manner constructions portray an imagined ('do X as if it was caused by Y'), or counterfactual ('do X as if Y were true') situation, it seems reasonable to explore whether other constructions expressing imagined or counterfactual situations also occur with the same TAM markers. A possible candidate to this analysis is counterfactual conditionals, a complex sentence construction in which the relation

<sup>3</sup> Example provided by Jiyoung Jang (personal communication).

<sup>4</sup> Example provided by Michael Fiddler (personal communication).



between the two clauses is that of an imagined situation that did not happen ('if it had not been for him, we could have got lost') (Olguín-Martínez & Lester to appear). Before proceeding to the analysis of this domain, it is important to bear in mind that many of the sources taken into account in the present study do not contain a detailed analysis of counterfactual conditionals. Therefore, this pioneering research can make only a modest contribution to the understanding of this domain.

Of the 61 languages of the sample, 16 languages provide a detailed analysis of counterfactual conditionals. As is shown in Table 1, in most languages, the *as if*-clause and the *if*-clause of a counterfactual conditional construction tend to appear with different TAM markers. Note that only in two languages, the *as if*-clause and the *if*-clause appear with the same TAM markers (i.e. Korean and Finnish). Accordingly, these results seem to indicate that although hypothetical manner constructions and counterfactual conditional constructions express imagined situations, they tend to appear with different TAM markers. However, caution needs to be exercised with these results given that they are based on 16 languages. Furthermore, these languages are for the most part African languages and Eurasian languages. Future studies will have to find out whether these results hold in a larger sample.

### **3. Sample and methodology**

Since this is primarily an explorative study that seeks to characterize a type of construction that has been traditionally neglected, the method for language sampling employed here is the Genus-Macro-area method proposed by Miestamo et al. (2016). This is a variety sample method, which aims at capturing as much variety as possible with respect to the expression of the phenomena under investigation and to reveal even the rarest strategies (Miestamo et al. 2016: 234). In this method, the primary genealogical stratification is made at the genus level, and the primary areal stratification at the level of macro-areas. In particular, a bottom-up variant of the method has been employed in this research. In what follows, the structure and motivations behind the selection of the languages of the current sample are introduced.

An ideal sample would include one and only one language from each genus of the classification of the world's genera in Dryer (2013). However, for some genera (e.g. Alacalufan, Camsá, Tacame, Shabo, Yurimangí), it has not been possible to find any source that provides a description of hypothetical manner constructions.

Language	<i>As if</i> -clause TAM marking	<i>If</i> -clause TAM marking
Alto Perené	Irrealis and conditional mood (Mihas 2015: 285)	Irrealis and counterfactual mood (Mihas 2015: 263)
Armenian	Present tense (Dum-Tragut 2009: 440)	Past tense and irrealis (Dum-Tragut 2009: 263)
Ben Tey	Progressive (Heath 2015: 253)	Past tense (Heath 2015: 258)
Boko	Progressive (McCallum Jones 1998: 263)	Past perfect (McCallum Jones 1998: 269)
Cuwabo	Past perfect (Guérois 2015: 350)	Counterfactual mood (Guérois 2015: 410)
Finnish	Conditional mood (Sulkala & Karjalainen 1992: 53)	Conditional mood (Sulkala & Karjalainen 1992: 315)
Goemai	Consecutive marker (Hellwig 2009: 333)	Past tense (Hellwig 2011: 463)
Ingush	Simultaneous converb (Nichols 2011: 574)	Irrealis (Nichols 2011: 305)
Iraqw	Past tense (Mous 1992: 167)	Past infinitive tense and perfective (Mous 1992: 329)
Ket	Non-past tense (Nefedov 2015: 201)	Past tense and irrealis (Nefedov 2015: 187)
Korean	Past tense (Jiyoung Jang, personal communication)	Past tense (Chang 1996: 159)
Lezgian	Aoristic participle (Haspelmath 1993: 247)	Past tense and aoristic converb (Haspelmath 1993: 396)
Lumun	Incompletive (Smits 2017: 669)	Completive (Smits 2017: 390)
Supyire	Perfect or potential mood (Carlson 1994: 570)	Counterfactual mood (Carlson 1994: 576)
Tundra Nenets	Dubitative (Nikolaeva 2014: 104)	Perfective aspect (Nikolaeva 2014: 374)
Udihe	Present participle (Nikolaeva & Tolskaya 2001: 748)	Past tense and irrealis (Nikolaeva & Tolskaya 2001: 750)

**Table 1:** TAM marking of the *as if*-clause and the *if*-clause in the languages of the sample.

As can be observed in Table 2, Eurasia has a stronger representation in the languages of the sample. This is because many genera of this macro-area contain languages with good grammatical descriptions of hypothetical manner constructions. Note that many languages from Australia and South America could not be taken into

Macro-area	Number of genera	Number of genera in the sample	Coverage
Africa	77	12	15.58%
Australia	43	3	6.97%
Eurasia	82	21	25.60%
North America	95	11	11.57%
Papunesia	136	10	7.35%
South America	110	4	3.63%
Total	543	61	-----

**Table 2:** Genera covered in the sample.

consideration. This is due to the fact that various grammars provide detailed descriptions of hypothetical manner constructions. However, they do not explain the encoding of similatives and real manner constructions. Table 3 shows a complete list of the languages taken into account for each macro-area.

Macro-area	Sample languages	Sum
Africa	Beja, Ben Tey, Boko, Cuwabo, Gbaya, Goemai, Iraqw, Kusaal, Lumun, Supyire, Tadaksahak, Yulu	12
Australia	Arabana, Kayardild, Warrongo	3
Eurasia	Armenian, Biak, Chinese, English, Finnish, Georgian, Greek, Ingush, Karbi, Ket, Korean, Latvian, Lezgian, Mongolian, North Saami, Spanish, Tundra Nenets, Turkish, Udihe, Yakkha, Yukaghir (Kolyma),	21
North America	Crow, Barbareño Chumash, Francisco León Zoque, Magdalena Peñasco Mixtec, Pech, Sahaptin, Sochiapan Chinantec, Huasteca Nahuatl, Warihio, Xicotepec Totonac, Yuchi	11
Papunesia	Komnzo, Makasae, Mali, Manambu, Marind, Moskona, Urim, Samoan, West Coast Bajau, Yimas,	10
South America	Alto Perené, Cavineña, Chamacoco, Piapoco,	4
		61

**Table 3:** Languages of the sample per macro-area.

The uneven distribution of comprehensive descriptions of hypothetical manner constructions and the limitations mentioned above cause what Bakker (2011: 106) calls a bibliographical bias. In this regard, Schmidtke-Bode (2014: 49) notes that the sampling procedure for complex sentence constructions is complicated because of the

of lack of comprehensive reference materials. This gives rise to areal biases that cannot be controlled for.<sup>5</sup>

#### 4. Range of strategies

Across languages, the semantic relation between the adverbial clause and its associated main clause may be indicated by various strategies (Hetterle 2015: 106). In this section, the focus is on individual items that serve this function. Section 4.1 first concentrates on the most common clause-linking strategies in the languages of the sample, viz. simulative ‘like’ markers and free adverbial conjunctions. Section 4.2 then proceeds to explaining the less common strategies attested in the languages of the sample, to which more time is devoted in that some of them seem show clear areal patterns. In this regard, some languages from Mesoamerica use correlative words, some Australian languages use counterfactual mood markers and some African languages employ head nouns meaning ‘thing’.

##### 4.1. Most common strategies

Across the languages of the sample, simulative ‘like’ markers are more common in the expression of hypothetical manner, as in the Cuwabo example in (11). Among the languages of the sample, 39 seem to show this scenario in that they employ the simulative ‘like’ marker to express hypothetical manner. This clause-linking device tends to introduce clauses whose internal structure shows no evidence of dependent status, dependent verb forms or a combination of both. Accordingly, they operate in clauses that appear with the same properties of main clauses. This finding echoes Hetterle (2015: 173) who shows that ‘as if’ constructions exhibit the lowest degree of downgrading in comparison to other types of adverbial clauses (e.g. purpose, cause).

(11) Cuwabo (Niger-Congo/Bantoid; Guérois 2015: 350)

<i>nyúwó</i>	<i>mu-ní-óná</i>	<i>nínga</i>	<i>ddi-a-kweńt-ílé</i>	<i>iíyí</i>
2PL.SBJ	2PL.SBJ-IPFV-see	like	1SG.SBJ-PST-copulate-PFV	while

---

<sup>5</sup> The sample taken into account in the present study is first and foremost a variety sample. Accordingly, the areal bias is not directly harmful for the general aims of variety sampling (Miestamo et al. 2016: 251).

*ka-ddi-a-kweñt-île.*

NEG-1SG.SBJ-copulate-PFV

‘You see me as if I had had sex, whereas I had not.’

Free adverbial conjunctions (Kortmann 2001: 842) are morphemes which mark adverbial clauses for their semantic relationship to the main clause. Of the languages of the sample, 22 have free adverbial conjunctions encoding hypothetical manner constructions, as in the Comaltepec example in (12), where the adverbial clause is introduced by the adverbial conjunction *la<sup>h</sup>huaʔ* ‘as if’. Free adverbial conjunctions also tend to introduce clauses whose internal structure shows no evidence of dependent status. Therefore, simulative ‘like’ markers and free adverbial conjunctions tend to operate in clauses that appear with the same properties as main clauses.

(12) Comaltepec Chinantec (Oto-Manguean/Chinantecan; Anderson 1989: 50)

*ʔi<sup>l</sup> hiú<sup>l</sup> na<sup>l</sup>-b zé la<sup>h</sup>huaʔ ŋó<sup>l</sup> hnä<sup>l</sup>.*

REL child that-AFF go as.if go 1SG.SBJ

‘That child is going as if I were going.’

Two general remarks on free adverbial conjunctions are in order here. First, in some languages of the sample, the free adverbial conjunction seems to have been derived from a verb meaning ‘to say’. In Lezgian (Nakh-Daghestanian/Lezgian), *guja na luhudi* ‘as if’ is a fixed expression that is synchronically best regarded as a free adverbial conjunction. Etymologically, *na luhudi* means ‘you would say’ (*na* is the ergative case of *wun* ‘you’ and *luhudi* is the archaic future of *luhun* ‘say’) (Haspelmath 1993: 247). Another example comes from Georgian (Kartvelian). In this language, the free adverbial conjunction *titkos* ‘as if, as though’ derives from *tu ttkva* ‘if it was/is said’ or *tu ttkva* ‘if it transpires that it is said’ (Hewitt 1995: 589). An interesting example is found in Supyire (Niger-Congo/Senufo). In this language, hypothetical manner may be expressed by a construction, which literally means ‘you would say’ (cf. French *on dirait*). Carlson (1994: 570) notes that this clause is beginning to function as a phrasal connective meaning ‘as if’ because it is not possible to pronounce it with pauses. This finding echoes Heine & Kuteva (2002: 268), in which verbs meaning ‘to say’ may grammaticalize into devices expressing the meaning ‘as if’. Although not explicitly mentioned by them, the examples shown by Heine & Kuteva (2002: 268) seem to indicate that this is common in Mande languages, such as Koranko (Mande/Western

Mande), e.g. *kó* ‘say’, > *íko* (‘you say’) ‘as if’ and Vai (Mande/Western Mande), e.g. *ro* ‘say’ > *i:ro* (‘you say’) ‘as’, ‘as if’.

Second, there are languages in which a similative ‘like’ marker and a free adverbial conjunction may occur at the same time in a clause, as in (13), (14), and (15). Interestingly, in these languages, it is the similative ‘like’ strategy that has become optional. Cross-linguistically, various types of adverbial clauses may appear at the same time with two clause-linking devices expressing the specific semantic relation in question. In this scenario, one of the markers is usually dropped (Hetterle 2015: 108; Schmidtke-Bode & Diessel to appear: 15). One possible hypothesis is that hypothetical manner constructions in these languages appeared first marked by similative ‘like’ markers. After that, speakers gradually developed a more specialized way of expressing hypothetical manner in order to differentiate similatives from hypothetical manner constructions. Once the two strategies converged in the same construction, the similative ‘like’ marker gradually became optional. Schmidtke-Bode & Diessel (to appear: 15) mention that in the recent typological and psycholinguistic literature, such patterns have attracted increasing attention under the label of redundancy management in grammar.

(13) Boko (Mande/Eastern Mande; McCallum Jones 1998: 263)

<i>má</i>	<i>kã</i>	<i>zu</i>	<i>gbẽ</i>	<i>pĩ</i>	<i>sàé</i>	<i>láńdõ</i>	<i>málé</i>
1SG.SBJ.FUT	arrow	shoo	rock	that	beside	as.if	1SG.SBJ.PROG
<i>pɔ</i>	<i>bã</i>	<i>wà.</i>					
animal	fire	like					

‘I will shoot an arrow beside that rock as if I am firing at an animal.’

(14) Makasae (Timor-Alor-Pantar/Makasae-Fataluku-Oirata; Huber 2008: 116)

<i>gi</i>	<i>nagar</i>	<i>seu</i>	<i>meti</i>	<i>wa’a</i>	<i>lor</i>	<i>hani.</i>
3SG.SBJ	as.if	meat	sea	REL	swim	like

‘He swims as if he were a fish.’

(15) Piapoco (Arawakan/Inland Northern Arawakan; Klumpp 2019: 332)

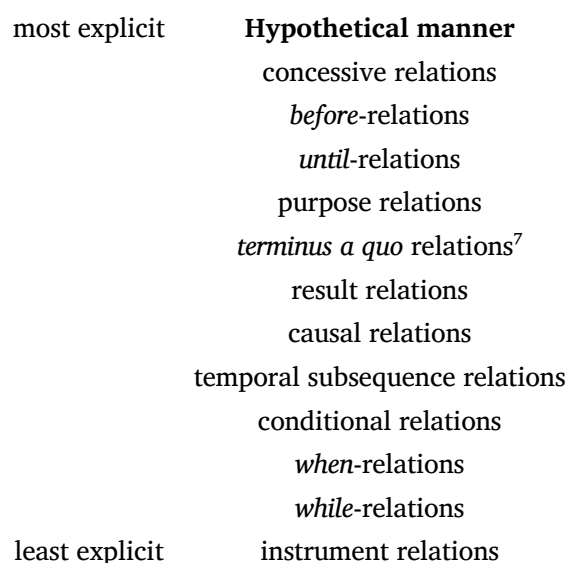
<i>báawa-ca</i>	<i>na-icá-ca</i>	<i>wía,</i>	<i>càide</i>	<i>iyúwa</i>	<i>wa-dé-ca</i>	<i>nacaicaalí</i>
bad-DECL	3PL-see-DECL	1PL.OBJ	thus	like	1PL-attain-DECL	as.if

*wa-màni-ca báawa-iri.*<sup>6</sup>

1PL-do-DECL bad-M

‘They look badly upon us (i.e. they hate us), as if we had done (something) bad to them.’

The mono/polyfunctionality of the strategies shown above seems to be another domain relevant to the study of hypothetical manner constructions. By polyfunctionality is meant the range of meanings within the domain of adverbial subordination that a particular temporal clause-linking strategy can have (Kortmann 1997: 89; Hetterle 2015: 202). Hetterle (2015: 214) shows that hypothetical manner and concession are the two relations most commonly expressed by monofunctional markers, as is illustrated in Figure 1. However, it is important to mention that her analysis is based on only one example found in her sample. Accordingly, this seems to suggest that there may be more to the story. One question that arises at this point is: does the form of the clause-linking strategy encoding



**Figure 1:** The explicitness scale of adverbial relations (Hetterle 2015: 218).

<sup>6</sup> One reviewer mentions that the example (15) could be analyzed as a correlative construction. However, it is important to stress that in correlative constructions, both clauses must appear with a clause-linking device. In (15), the first clause does not appear with a clause-linking device.

<sup>7</sup> Clauses that express *terminus a quo* refer to a semantic relation in which the situation of the dependent clause indicates a starting point or starting period of time in the (relative) past from which the situation in the main clause has been true (Kortmann 1997: 85).

hypothetical manner constructions play a role in the degree of mono/polyfunctionality? That is, do similative ‘like’ markers used to express hypothetical manner and free adverbial conjunctions develop different degrees of mono/polyfunctionality?

One interesting observation on the languages of the sample is that while free adverbial conjunctions tend to be monofunctional, similative ‘like’ markers used in the expression of hypothetical manner tend to be polyfunctional in that they are also used to express other adverbial semantic relations. Accordingly, this suggests that the mono/polyfunctionality of clause-linking devices encoding hypothetical manner constructions will vary depending on their form.

#### ***4.2. Less common strategies***

After having explored the most common strategies that languages may use to express hypothetical manner, I can now proceed to explaining some rare strategies attested in the languages of the sample. Interestingly, these strategies form particular areal clusters. Since these strategies are cross-linguistically rare and are only found in languages not genetically related spoken in the same area, diffusion through language contact is most likely to have taken place.

##### ***4.2.1 Correlative words***

From a cross-linguistic perspective, languages may use pairs of correlative words to connect clauses together into complex sentences (see Haspelmath 2004 for an overview of correlative coordinators). Perhaps the best known case is that of comparative correlatives, such as *the more money you have, the more you want to travel* (see Culicover & Jackendoff 1999) and immediate temporal subsequence (e.g. *No sooner had I left home than the phone rang*), to name but a few.

In the languages of the sample, Silacayoapan Mixtec and Huasteca Nahuatl express hypothetical manner by means of a correlative construction. In this correlative construction, the first connective is a conditional marker which can be optionally followed by a lexical item meaning ‘word’. Note that the main clause must appear with a linker meaning ‘thus’. Interestingly, in both languages, the verbs of both clauses must be the same. In the Silacayoapan Mixtec example in (16), the verb of the dependent clause *ndáā* ‘to appear’ must occur in the main clause. In a similar fashion,



in the Huasteca Nahuatl example in (17), the two verbs must be the same in both clauses. Furthermore, the first connective is a conditional marker optionally followed by a lexical item meaning ‘word’ and the main clause must appear with a linker meaning ‘thus’. The most obvious explanation to this parallelism seems to be language contact. This is because: (1) the languages are spoken in the same region, (2) they are not genetically related and (3) the probability of chance resemblance is low given the rarity of the strategies.

(16) Silacayoapan Mixtec (Oto-Manguenan/Mixtecan; Shields 1988: 431-432)

*tá tuhūn ndáā nā xyoko, xá ndáā nā*  
if word appear:PL 3PL.SBJ San.Andrés thus appear.PL 3PL.SBJ  
*ndahví.*  
poor

‘As if they were from San Andrés, those poor people look (lit. about like the people from San Andrés appear, so appear those people).’

(17) Huasteca Nahuatl (Uto-Aztecan/Aztecan)

*tla tlahtol mayana-h,<sup>8</sup> yekah mayana-h ki-kua-yaya baka naka-tl,*  
if word be.hungry.PL thus be.hungry.PL 3SG.OBJ-eat-IPFV COW meat-ABS  
‘As if they were hungry, they ate the cow meat.’

This pattern is also attested in other Mixtec languages. In the Alacatzala Mixtec example in (18), the verb of dependent clause *ndóō* ‘to sit’ must appear in the main clause. In this correlative construction, the first connective is a conditional marker which can be optionally followed by a lexical item meaning ‘word’. Note that the main clause must appear with a linker meaning ‘thus’. It is important to stress that other Nahuatl varieties do not express hypothetical manner in this way. For instance, Tetelcingo Nahuatl expresses this semantic type by means of the phrasal linker *kiem tlō* ‘as if’, composed of the similative marker *kiem* ‘like’ and the conditional marker *tlō* ‘if’ (Tuggy 1979: 129). Accordingly, this seems to indicate that Mixtec languages served as the source. The fact that several Mixtec adverbial clause-linking strategies may have spread to Huasteca Nahuatl is an interesting finding in that it has been proposed for the most part that Nahuatl served a prominent role in the formation of Mesoamerica as a linguistic area (Brown 2011: 201). This stems from the fact that

---

<sup>8</sup> The Huasteca Nahuatl example comes from the fieldwork of the author of the paper.

this language served as a widely used lingua franca. However, it is important to stress that this does not necessarily mean that Nahuatl did not copy linguistic traits from other Mesoamerican languages (Brown 2011: 201). Speakers of Mixtec languages and Nahuatl languages have been in contact through intermarriage, alliances and warfare at least since the colonial period (Sousa & Terraciano 2003: 353), which has resulted in a complex network of interactions and bilingualism (Terraciano 1990: 142). Therefore, it seems reasonable to assume that Nahuatl may have copied various patterns from Mixtec languages and vice versa.

(18) Alacatzala Mixtec (Oto-Manguenan/Mixtecan; Zylstra 1991: 149)

*tá ndóō ñĩĩ kĩti sãá ndóō nā.*

If sit.CONT.PL one animal thus sit.CONT.PL 3PL.SBJ

‘They live as if they were animals (lit. as animal live, so they live).’

Before leaving the present section, it is important to mention that in spite of the fact that neighboring languages have not copied the same Mixtec pattern for expressing hypothetical manner, they seem to have copied some properties for expressing other meanings closely related to the hypothetical manner meaning. Chiapas Zoque has a construction expressing real manner in which the verbs of both clauses must be repeated as in (19), where *kips* ‘think’ is repeated in the second clause. Bear in mind that this is a feature similar to one attested in Mixtec hypothetical manner constructions. Similatives in Francisco León Zoque show an interesting scenario. As can be seen in the example in (20), the second clause must appear with the linker *jetse* ‘therefore’; a feature similar to one attested in Mixtec hypothetical manner constructions. What this seems to indicate is that some of the properties of a construction may be copied from one language to another to express a similar meaning. This can set the stage for further processes of development, which may be internally motivated.

(19) Chiapas Zoque (Mixe-Zoque; Faarlund 2012: 172)

*uj-t te’ = se = ti ñu ø-kips-ke’t-u*

1SG.SBJ-ERG DET = SIM = LIM PROG 1SG.SBJ-think-REP-PROG

*m-kips-u-se.*

2SG.SBJ-think-PROG = SIM

‘I think the same way as you think.’

(20) Francisco León Zoque (Mixe-Zoque; Bartholomew & Engel 1987: 358)

*como ncastillo nø ज्या'pøtyøju-se jetse ajnøpya.*  
like castle PROG.AUX fire.PST.TERM-like therefore sound.CONT

'It sounds like a castle is on fire (lit. like a castle is on fire, therefore it sounds).'

#### 4.2.2 Counterfactual mood markers

Hypothetical manner constructions encoded by counterfactual mood markers are only found in Australia in the languages of the sample. This is attested in Warrongo, a Pama-Nyungan language, as in (21), and Kayardild, a Tangkic language, whose genetic affiliation lies with the non-Pama-Nyungan languages of Arnhem Land (Evans 1995: 239), as in (22).

(21) Warrongo (Pama-Nyungan/Northern Pama-Nyungan; Tsunoda 2011: 683)

*nyola yaji-garra-n jilbay-Ø = gaji.*  
3SG.NOM laugh-ITER-NON.FUT knowing-NOM = CF

'He is laughing as if he knew (i.e. understood Warrongo).'

(22) Kayardild (Tangkic; Evans 1995: 378)

*jani-jani-ja niwan-ju, yakuri-ya buru-tha thaa-tha marak.*  
search-RDP-ACT 3SG-PROPI fish-LOC take-ACT return-ACT CF

'They searched for him, as if they were going out to get fish.'

What is interesting to note is that these languages show a striking parallelism in a type of pattern that is quite unusual cross-linguistically. Therefore, this pattern cannot be explained by chance. Another important aspect to bear in mind is that these languages are not genetically related. Accordingly, the fact that both express hypothetical manner by means of counterfactual mood markers cannot be due to common inheritance. The most likely explanation is language contact because the languages are spoken in the same geographical region. In this regard, Evans (1995: 239) notes that there is evidence for sustained contact with Pama-Nyungan languages now spoken on the northern fringe of the Central Australian desert.

One important aspect that further supports the idea that this construction may have spread through language contact comes from the fact the markers in both languages

have almost the same range of functions. The counterfactual marker *maraka* in Kayardild expresses not only hypothetical manner, but also other meanings. First, it indicates the course of action which should have been taken, but was not, as in (23). Second, it may refer to events that could have happened but did not, as in (24). Third, it may be used to express mistaken identity or belief, that is, it indicates that, at the time of the situation, someone either held a false belief about the identity or characteristics of the relevant entity, or acts as if they had such a belief, as in (25). Fourth, this marker may be used to express similitive meanings, as in (26).

(23) Kayardild (Tangkic; Evans 1995: 378)

*nyingka maraka raba-nangku dathin-ku dulk-u.*  
 2SG.NOM CF tread-NEG.POT that-PROP place-PROP  
 ‘You should not have set foot in that place.’

(24) Kayardild (Tangkic; Evans 1995: 378)

*maraka yuuma-thu barruntha-y.*  
 CF drown-POT yesterday-LOC  
 ‘He could have drowned yesterday (but did not).’

(25) Kayardild (Tangkic; Evans 1995: 379)

*kurri-ja manharr-iy maraka dangka-karran-ji birra niwan-ji.*  
 see-ACT torch-LOC CF man-GEN-LOC too his-LOC  
 ‘They saw a bark torch, and thought it was the man’s, that it too was his.’

(26) Kayardild (Tangkic; Evans 1995: 381)

*kaban-d maraca kamarr.*  
 stargazer-NOM CF stone-NOM  
 ‘The stargazer (fish) is like a stonefish.’

In a similar fashion, the Warrongo counterfactual mood marker =*gaji* expresses not only hypothetical manner, but also other meanings similar to those expressed by the Kayardild marker *maraka*. First, it may be used for expressing epistemic-judgements, that is, the speaker expresses his judgments about the factual status of the proposition, as in (27). Second, it may also be used to express mistaken identity or belief, that is, it indicates that, at the time of the clause, someone either held a false belief about the

identity or characteristics of the relevant entity or acts as if they had such a belief, as in (28).

(27) Warrongo (Pama-Nyungan/Northern Pama-Nyungan; Tsunoda 2011: 586)

*gibagiba-Ø = gaji jombi-Ø yino.*  
mushroom-NOM-CF penis-NOM 2SG.GEN

‘It looks as if your penis were a mushroom.’

(28) Warrongo (Pama-Nyungan/Northern Pama-Nyungan; Tsunoda 2011: 677)

*yarro-Ø ngalnga = gaji jojarra-Ø.*  
this-NOM CF = CF urine = NOM

‘I thought it was urine, but in fact it was not (it was semen).’

This pattern is also attested in other Pama-Nyungan languages, such as Bidyara, Gungabula and Wunambal. Tsunoda (2011: 984) mentions that this marker is a cognate attested in various Pama-Nyungan languages that can be reconstructed. Accordingly, this seems to indicate that Pama-Nyungan languages served as the source, that is, Kayardild seems to have replicated this pattern from Pama-Nyungan languages by using native material.

#### 4.2.3 Nouns meaning ‘thing’

In the sample used for the present study, hypothetical manner constructions encoded by head nouns meaning ‘thing’ appear as an African singularity. As can be seen in the Dogul Dom example in (29), hypothetical manner is expressed by means of the noun *gòŋ* ‘thing’ plus the similative *yáŋ* ‘like’. This is also attested in other Dogon languages, such as Donno So, as in (30). Jeffrey Heath (personal communication) informs me that the noun meaning ‘thing’ is not an argument (subject or object) within the predicate of the ‘as if’ clause. Accordingly, this construction should be understood as ‘like the thing (situation) in which the children have not eaten’ in (29), and as ‘like the thing (situation) in which I had hit him’ in (30).<sup>9</sup>

---

<sup>9</sup> As correctly pointed by one reviewer, the examples in (29) and (30) are similar to those shown in §4.1, in which a similative ‘like’ marker and a free adverbial conjunction occur at the same time in a clause.

(29) Dogul Dom (Dogon; Heath 2016: 269)

è-wé      *gòŋ*    *nǎ:*      *bè*            *nà:-l*            *yáŋ*,    *yǎ:*  
 child-PL    thing    meal    3PL.SBJ    eat-PFV.NEG.PL    like    tears  
*yê:-ñ.*  
 weep-IPFV.3PL.SBJ

‘As if they the children have not eaten, they are crying.’

(30) Donno So (Dogon; Heath 2014: 269)

Î            *yǎ:*      *yà:-dê-Ø*,            *kidê*    *wò = ř*      *mí*            *bënd-é-Ø*  
 child    tears    weep-IPFV-3SG.SBJ    thing    3SG = ACC    1SG.SBJ    hit-PFV-3SG.SBJ  
*ginè.*  
 like

‘The child is crying, as if I had hit him.’

Interestingly, a similar pattern is also attested in another African language of the sample. In Gbaya, hypothetical manner is expressed by a simulative ‘like’ marker accompanied by a noun meaning ‘thing’, as in (31). Dogon languages and Gbaya are not genetically related and the probability of chance resemblance is low given the rarity of the strategies. Although it is very tempting to propose that language contact may have played a role, Jeffrey Heath (personal communication) informs me that this scenario is highly unlikely in that Dogon languages have not been contact with this language. Therefore, care should be taken when interpreting these results.

(31) Gbaya (Niger-Congo/Gbaya-Manza-Ngbaka; Roulon-Doko 2017: 227)

*zà*      *gòná*            *gásá zó*    *hé*    *mè*    *né*    *Gbàmbòndó*    *měi*            *gá.*  
 3SG.SBJ    carve.out.PFV    big    grass SIM    thingbe    Gbàmbòndó    over.there SIM

‘He prepared a large plot of lands as if it were Gbambòndó (the village’s largest hunting territory) over there.’

## 5. Hypothetical manner and formal resemblances to other constructions

This paper also explores whether hypothetical manner clauses show formal resemblances to other types of constructions. It is well-known that comparative and superlative constructions are more similar to each other than to other types of

comparative clauses (Ulan 1972). Equative and similative constructions are more similar to each other than to other types of comparative clauses (Haspelmath & Buchholz 1998: 278). With this in mind, the question is: are there any structural similarities between hypothetical manner clauses and other types of comparative constructions? Various language-specific investigations have shown that hypothetical manner clauses show formal and functional resemblances to similatives. In this regard, Chamoreau (2017: 331-332) notes that hypothetical manner expressions and similatives are related in various languages. These concepts are distinct but connected, as hypothetical manner means “to imitate, pretend, aspire to the appearance of something” and similarity means “to give the same appearance as something” (Chamoreau 2017: 331-332). In a similar fashion, Fortescue (2010: 131) mentions that most languages have expressions related to those they use as similative markers that express hypothetical manner relations. However, as he acknowledges, no typological study has explored this aspect in more detail.

Another construction that is also very similar to hypothetical manner is that of real manner clauses (e.g. *do as I told you*). Although real manner clauses are not a subtype of comparative construction, they show formal resemblances to hypothetical manner constructions in many languages. Darmon (2017: 372-373) notes that while hypothetical manner portrays an imagined (‘do X as if it was caused by Y’) or counterfactual (‘do X as if Y were true’) situation, real manner adverbial clauses depict an action or state identical to that of the main clause. Hetterle (2015: 54) mentions that both hypothetical manner and real manner constructions answer the questions ‘how?’, but they differ from one another in that real manner describes the character of a situation comparing it to a real situation and hypothetical manner compares a situation to a hypothetical or counterfactual situation.

Here it is proposed that hypothetical manner constructions can be classified into three main types according to whether they are encoded in the same way as similatives and/or real manner clauses. In what follows, this classification is discussed in more detail.

### ***5.1. Hypothetical manner, real manner and similatives marked in the same way***

Hypothetical manner, real manner and similative meanings may all be expressed by the same clause-linking device, as is illustrated in the Chamacoco examples in (32), (33), and (34). Note that *itso* is used to express a similative meaning in (32), while it

is employed to express hypothetical manner in (33) and real manner in (34). This is the most common system in that 30 languages in the sample have this type. Note that all constructions in the languages showing this pattern are always marked by a similative ‘like’ marker. This indicates that from a diachronic perspective hypothetical manner constructions and real manner constructions seem to have developed from similatives. Although this pattern is attested in all macro-areas, in the languages of the sample, it seems to be more frequent in Papunesia.

(32) Chamacoco (Zamucoan)<sup>10</sup>

*o-ho naraje oti-ch shi itso awi-t.*  
 3PL-drink orange juice-M.SG only like water-M.SG  
 ‘They drink orange juice like water.’

## (33) Chamacoco (Zamucoan)

*ese ir oti-ch nosh = āha ōr ish-u-wo shi itso uje o-ch-ūrhu*  
 that 3SG liquid-M.SG spill = PREP 3PL dress-M.SG only like SUB PL-3-wash  
*l-asu-wo = ho wir erze wino.*  
 REFL-dress-M.PL-PREP DET.PL that.PL wine  
 ‘Its liquid spills from their dresses, as if they had washed their dresses with wine.’

## (34) Chamacoco (Zamucoan)

*uhu itso uje y-ikitkēryēr.*  
 2SG.do like SUB 1PL-talk REC.F  
 ‘You do the way we talk to each other’

One question that may arise at this point is: if hypothetical manner, real manner, and similative constructions are realized by the same clause-linking device in a particular language, how are the various meanings differentiated?

In almost all languages showing this system, contextual factors seem to be the only factor disambiguating the different meanings. In Pesh, a Chibchan language spoken in Honduras, the similative clitic =*kán* appears in hypothetical and real manner constructions. In this scenario, the distinction is only given by context (Chamoreau 2017: 331-332). In a similar fashion, Luca Ciucci (personal communication) informs me that, in Chamacoco, hypothetical manner, real manner and similative

<sup>10</sup> Examples provided by Luca Ciucci (Personal communication).



constructions are expressed by the simulative marker *itso* ‘like’. He mentions that the only way one can distinguish them is based on the context.

For only a small number of languages, scattered pieces of information are available regarding this disambiguation process. Therefore, this pioneering research can make only a modest contribution to the understanding of this domain. In some languages for which this sort of information is available, hypothetical manner constructions are marked by a simulative ‘like’ marker plus a TAM marker that aids in disambiguation. A case in point comes from Alto Perene. In this language, hypothetical manner, real manner and simulative constructions appear with the simulative marker *ki-* ‘like’. Hypothetical manner constructions occur with the simulative marker *ki-* plus the irrealis marker *-ia* and the conditional clitic *=rika*, which allows speakers of this language to disambiguate this adverbial meaning from the others, viz. simulative and real manner meanings.<sup>11</sup>

(35) Alto Perene (Arawakan/Pre-Andine Arawakan; Mihas 2015: 285)

<i>a = ny-i = ri</i>	<i>nihaa-tsapya-ki</i>	<i>kisaa-tsantsana-ite</i>	<i>katari</i>
1PL = see-REAL = 3SG.OBJ.M	river-bank-LOC	be.black-wide-AUG	duck
<i>i = ki-t-ak-a</i>	<i>i = mitsaink-ia = rika</i>	<i>y = ovayeri-t-ia-ranki.</i>	
3PL = SIM-EP-PFV-REAL	3PL = be.in.line-IRR = COND	3PL = fight-EP-IRR-ADV	

‘We see black ducks on the river banks as if they were all warriors standing in lines.’

Another example comes from Karbi. In this language, hypothetical manner, real manner and simulative constructions are encoded by *asón* ‘like’. Hypothetical manner constructions appear with the irrealis marker *-jí*, as in (36), to disambiguate this meaning from the simulative and real manner meaning. Interestingly, hypothetical manner constructions can also appear with *thàngbāk* ‘as if’, considered another constructional property used to disambiguate hypothetical manner from simulative and real manner. This seems to be an instance of compositional encoding; i.e., specific

---

<sup>11</sup> Mithun (1995: 384) explains that the notion IRREALIS portrays a state of affairs as purely within the realm of thought, knowable only through imagination. A source of potential confusion in any discussion on irrealis is that it has been applied to different concepts and constructions in languages from many areas of the world. It is therefore important to clarify what is meant when using this term. In this paper, irrealis is considered a specific marker (rather than notional descriptions of non-encoded meanings of constructions) in the form of verbal affixes and clausal enclitics (Brooks 2018: 4).

constructional properties of a clause combine to dictate a particular adverbial reading (Hetterle 2015: 106) and also to disambiguate various meanings from one another.

(36) Karbi (Sino-Tibetan/Kuki-Chin; Konnerth 2014: 409)

<i>mh</i>	<i>è-li = ke</i>	<i>ke-rè-èt</i>	<i>a-tum = ke thàngbāk = si</i>
pause	1PL.INCL-HON = TOP	NMLZ-be.alive-all	POSS-PL = TOP as.if = FOC
<i>ke-lè-dūn-tām</i>		<i>thèk-jí</i>	<i>asón</i>
NMLZ-reach-join-impossible		know.how.IRR	like
<i>nang = pinkhát-táp</i>		<i>nang = pinkhát-phrú.</i>	
1SG.NON.SBJ = advise-here.and.there		1SG.NON.SBJ-advise-here.and.there	

‘Since we are alive, how can you give me so many pieces of advice as if we could reach (the place where my wife has gone after she died).’

## 5.2. Hypothetical manner and similitive constructions marked alike; Real manner encoded differently

The second most common system in this study is that of languages in which hypothetical manner and similitive constructions are marked in the same way, while real manner constructions are encoded differently. In Tadaksahak, hypothetical manner and similitive constructions occur with the similitive marker *injin* ‘like’, as in (37) and (38), respectively. Note that real manner clauses appear with the head noun *ammək* ‘manner’, as in (39). Of the languages of the sample, 18 languages seem to have this system. These languages are scattered in the different areas of the world showing no effects of areal grouping.

(37) Tadaksahak (Songhay; Christiansen-Bolli 2010: 268)

<i>ee-dag</i>	<i>aro</i>	<i>senda</i>	<i>injin</i>	<i>ni = yyikkəl-a</i>	<i>ni = dd = a</i>	<i>suubu</i>	<i>ka.</i>
SG-place	DET	DEM	as.if	2SG = lift-3SG	2SG-put-3SG	hay	LOC

‘This matter is as if you take it (fire) up and put it to the hay.’

(38) Tadaksahak (Songhay; Christiansen-Bolli 2010: 268)

<i>feeji</i>	<i>ar(o)</i>	<i>ooda</i>	<i>injin</i>	<i>ar = wani.</i>
sheep	DET	DEM	like	2SG = of

‘This sheep is like mine.’

(39) Tadaksahak (Songhay; Christiansen-Bolli 2010: 284)

*əmmək*    *aro*    *ənda ar = tə-d-a...*  
manner    DET    with 1SG-FUT-do-3SG  
'The manner in which I do it...'

Another example comes from Crow. In this language, hypothetical manner and simulative constructions are encoded by the simulative marker *kummah* 'like', as in (40) and (41), respectively. Real manner clauses are marked by a headless relative construction in which a head noun meaning 'manner' has been omitted, but it is understood from context, as in (42). Note that in this example the construction still keeps the relativizer *am-*.

(40) Crow (Siouan/Core Siouan; Graczyk 2007: 350)

*"kuss-dee-ssaa-la-h"*    *he-m*    *kummah*    *"naa-la-h"*    *he-lahth*    *dee-laa.*  
GL-go-NEG-PL-IMP    say-DS    as.if    go-PL-IMP    say-even.if    go-SS  
'"Do not go", he said, and as if he had said "go", they went.'

(41) Crow (Siouan/Core Siouan; Graczyk 2007: 350)

*kummah*    *issaxpuatahchewishke*    *ahkaash-dak...*  
like    sheep    many-COND  
'They were like sheep...'

(42) Crow (Siouan/Core Siouan; Graczyk 2007: 255)

*biaxaake*    *am-ma-lasitt-uua*    *ko*    *kala-koot-uu-k.*  
ducks    REL-1PL.SBJ-happy-PL    PRO    PREF-like.that-PL-DECL  
'(The way) that we ducks are happy, it is like that.'

One interesting correlation shown by languages having this type of system is as follows: while hypothetical manner and simulative constructions are encoded by a simulative 'like' marker, real manner constructions tend to be formed by a relative clause appearing with a head noun meaning 'manner' or a relative clause from which a noun meaning 'manner' has been elided, as in the Crow example shown above. The fact that manner adverbial clauses are encoded in this way is not surprising. Thompson et al. (2007: 245) point out that adverbial clauses expressing time (e.g. *We'll go when Tom gets here*), location (e.g. *I'll meet you where the statue used to be*) and

manner (e.g. *She spoke as he had taught her to*) can commonly be paraphrased, in many languages, with a relative clause that appears with a generic head noun that is semantically empty, such as ‘time’ (e.g. *We’ll go at the time at which Tom gets here*), ‘place’ (e.g. *I’ll meet you at the place at which the statue used to be*) and ‘way/manner’ (e.g. *She spoke in the way in which he had taught her to*), respectively.

### ***5.3. Hypothetical manner, real manner and similatives each marked by a different strategy***

Hypothetical manner, real manner and similative constructions may each be formally distinguishable from one another in that they are encoded by a different marker. Therefore, in this type of system there does not seem to be a diachronic connection among these constructions. Note that this is the third, and least common pattern in the present research. Of the languages of the sample, 11 languages seem to have this sort of system. Instances of this type are found in all macro-areas, but they seem to show a clear areal cluster in Eurasia. Languages vary with respect to the strategies they employ to express this system. The examples discussed below do not exhaust the whole range of ways in which languages formally distinguish this type of system, but they should serve for discussion purposes. With that proviso, let us briefly discuss a couple of languages showing this system.

In Iraqw, hypothetical manner is expressed by the free subordinating conjunction *barékwa’o* ‘as if’, as in (43), similatives appear with *at* ‘like’, as in (44) and real manner is realized by means of a relative clause occurring with the head noun *adó* ‘manner’, as in (45).

(43) Iraqw (Afro-Asiatic/Southern Cushitic; Mous 1992: 329)

*hamí án qaat dí-r afá loohi, barékwa’o a-na*  
 now 1SG 3SG.M.PRS.lie place-F mouth way as.if 1SG-PST  
*gwáá’.*  
 1SG.die

‘Now I will lie at the side of the path, as if I have died.’

(44) Iraqw (Afro-Asiatic/Southern Cushitic; Mous 1992: 280)

*a at see’aay.*  
 COP like dog

‘He is like a dog.’

(45) Iraqw (Afro-Asiatic/Southern Cushitic; Mous 1992: 280)

*adór ís dawé ngi-r hlakat-i, an ahlaw-ká.*  
manner-F 3SG elephants OBJ-INSTR hunt-3SG.M 1SG 1SG.can-NEG  
'I cannot hunt elephants the way he does.'

The second example is from Tundra Nenets. This language has 15 inflectionally formed non-indicative moods which express various epistemic, deontic and evidential meanings (Nikolaeva 2014: 85). One of these moods is that of the reputative mood which is formed by means of the marker *-mána* and is used to express hypothetical manner or, as stated by Nikolaeva (2014: 85), it is employed to express "irrealis comparison", as in (46). Real manner constructions are formed by means of the postposition *p'iruw°na* 'how', as in (47), and similatives are realized by means of the similative marker *-rəxa*, as in (48).

(46) Tundra Nenets (Uralic/Samoyedic; Nikolaeva 2014: 104)

*wada-xəqnata s'íta xa-ma-m, nyi-w°n'a=w°h nəmtor°-q yet°h*  
word-3SG.LOC SG.ACC call-IPFV.AN-ACC NEG-REPUT-DUB listen-CONN NEG DP  
*tolaŋku.*  
read  
'He is reading as if he does not hear that is being called.'

(47) Tundra Nenets (Uralic/Samoyedic; Nikolaeva 2014: 372)

*t'irt'a-q məl°nə-wa-h p'iruw°naləx°nə°-n'ih.*  
bird-GEN.PL chirp-IPFV.AN-GEN how talk-1SBJ.DU  
'We talk the ways birds chirp.'

(48) Tundra Nenets (Uralic/Samoyedic; Nikolaeva 2014: 35)

*...numki°-q tu-rəxa-q.*  
star.PL.GEN fire-SIM-PL  
'...like the lights of the stars.'

Before I leave the present section, mention should be made of one system that seems to be rare cross-linguistically in that it is only found in two languages of the sample. This pattern is concerned with those instances in which similative and real manner constructions appear with the same clause-linking strategy, while hypothetical

manner constructions occur with a different one. The first example comes from Mali. In this language, similitive and real manner constructions are formed with the similitive marker *klaŋ* ‘like’, as in (49) and (50), respectively. Hypothetical manner is realized by means of the free subordinating conjunction *gisnia* ‘as if’, which seems to be a contraction of a part of a larger expression used to introduce sensory experiences *ngia tu gia snēŋ ia* ‘you would think (say) that’. Recall that this is in line with the observation that in various languages free adverbial conjunctions may have been derived from a verb meaning ‘to say’.

(49) Mali (Baining; Stebbins 2009: 377)

*ki            tneŋ            klaŋ    ka.*  
 3SG.F        dodge.PRS    like    3SG  
 ‘She dodges like him.’

(50) Mali (Baining; Stebbins 2009: 377)

*a = musnēŋ    ngē        muēŋ            vēt        gu = auj                    klaŋ*  
 SPEC = idea    3SG        arrive.NON.PRS    at        1SG.POSS = grandmother like  
*da = ithik    ia        “ngo    da        vandingu    vlek    ngu        vang”.*  
 EMPH = DEM    REL        1SG        and        DES        1SG    want    1SG.run.NON.PRS  
 ‘An idea came to my grandmother the way in which she thought it, “I want to try and run away.’

(51) Mali (Baining; Stebbins 2009: 377)

*da    koki    ka        tet        gisnia    kule        ka        pe        mēni    aut*  
 and    just    3SG.M    go.PRS    as.if    stay.PRS    3SG.M    there    on        1PL.POSS  
*gling-igēl.*  
 place-EXC.SG  
 ‘He had just gone as if he was staying at our place.’

The second example is found in Warrongo. In this language, similitive and real manner constructions appear with *yamanyon* ‘similar’, as is shown in the examples in (52) and (53), respectively. This item seems to have acquired the status of an enclitic in that it receives stress, so it should not be considered a suffix (Tsunoda 2011: 671). Etymologically, *yamanyon* ‘similar’ contains the demonstrative member of adverbs *yama* ‘in such a way’. Note that the etymology of *-nyon* is unknown (Tsunoda 2011:

671). Hypothetical manner is not expressed by *yamanyon* ‘similar’. Rather, the language has developed a different formal way of expressing this meaning. As can be seen in (54), hypothetical manner constructions are realized by the counterfactual clitic =*gaji*. Etymologically, this counterfactual clitic comes from the adverb(ial) of modality *gaji* ‘maybe, might’ and ‘you try!’ (Tsunoda 2011: 679).

(52) Warrongo (Pama-Nyungan/Northern Pama-Nyungan; Tsunoda 2011: 673)

*ngaya = yamanyon nyola gawa-l.*

1SG.NOM = like 3SG.NOM call.out-NON.FUT

‘He is calling out like me.’

(53) Warrongo (Pama-Nyungan/Northern Pama-Nyungan; Tsunoda 2011: 683)

*yinda yani-ya yangga-gali-ya ngaya = yamanyon.*

2SG.NOM go-IMP search.for-ANT-IMP 1SG.NOM = like

‘Go and look for it the way I do it.’

(54) Warrongo (Pama-Nyungan/Northern Pama-Nyungan; Tsunoda 2011: 683)

*nyola yaji-garra-n jilbay-Ø = gaji.*

3SG.NOM laugh-ITER-NON.FUT knowing-NOM = CF

‘He is laughing as if he knew (i.e. understood Warrongo).’

The present study has faced some challenges and is not without its limitations. In this regard, determining the classification of particular types of systems has been one of the most time-consuming parts of the analysis. A case in point comes from Cuwabo. In this language, hypothetical manner, real manner and simulative constructions occur with *nínga* ‘like’. However, real manner may also be expressed by means of the head noun *mikálélo* ‘way’ or the Portuguese loanword *mánéera* ‘way’ (Guérois 2015: 484). Given that real manner is more frequently expressed by means of *nínga* ‘like’ (Guérois 2015: 484), the present study has classified Cuwabo as a language in which hypothetical manner, real manner and simulative meanings are expressed by the same clause-linking device. These examples do not exhaust the whole range of problematic cases that have been encountered during the analysis of the data. However, they suffice to provide the reader with an idea of some of the difficulties that have arisen

during the analysis. Nonetheless, such problematic cases are rather few and do not detract from the validity of the overall conclusions.

(55) Cuwabo (Niger-Congo/Bantoid; Guérois 2015: 350)

*nyúwó mu-ní-óná nínga ddi-a-kweñt-îlé iiyí*  
 2PL.SBJ 2PL.SBJ-IPFV-see like 1SG.SBJ-PST-copulate-PFV while  
*ka-ddi-a-kweñt-île.*  
 NEG-1SG.SBJ-copulate-PFV

‘You see me as if I had had sex, whereas I had not.’

(56) Cuwabo (Niger-Congo/Bantoid; Guérois 2015: 341)

*ńttítti Rosa ni-luw-ey-ilé nínga árigóra.*  
 hair Rosa NC-plait-NEUT-PFV like ring  
 ‘Rosas’s hair is like rings.’

(57) Cuwabo (Niger-Congo/Bantoid; Guérois 2015: 483)

*ddi-ní-fúná ddi-kál-é nínga e-á-ligí = ímí*  
 1SG-IPFV-want 1SG-be-SBJ like NC-PST.IPFV-be-HAB = 1SG  
*va-tákúlú = vènyu.*  
 NC-house = 2PL.POSS

‘I want to be the way I used to be in your house.’

(58) Cuwabo (Niger-Congo/Bantoid; Guérois 2015: 484)

*mikálélo dhi-á-lí = iye...*  
 way NC-PST.IPFV-be = 3SG.SBJ

‘The way she was...’

## 6. Final remarks

This paper has set out to describe hypothetical manner constructions in a sample of 61 languages. It has been demonstrated that similitive ‘like’ markers and free adverbial conjunctions are more common in the expression of hypothetical manner than other types of strategies (e.g. counterfactual markers, etc.). These devices tend to operate in clauses that appear with the same properties as main clauses. One



interesting observation in the languages of the sample is that while free adverbial conjunctions tend to be monofunctional, similative ‘like’ markers used in the expression of hypothetical manner tend to be polyfunctional in that they are also used to express other adverbial semantic relations. It has also been shown that some rare strategies are only attested in particular areas of the world. In particular, some languages from Mesoamerica use correlative words, some Australian languages use counterfactual mood markers and some African languages employ head nouns meaning ‘thing’. Interestingly, the forms of the strategies are not the same. Given that these strategies are cross-linguistically rare and are only found in languages not genetically related spoken in the same area, diffusion through language contact is most likely to have taken place.

It has been proposed that hypothetical manner constructions can be classified into three main types according to whether they are encoded in the same way as similatives and/or real manner clauses: (1) hypothetical manner, real manner and similatives marked in the same way; (2) hypothetical manner and similative constructions marked alike, real manner encoded differently; and (3) hypothetical manner, real manner and similatives marked by different strategies. It has been demonstrated that in the majority of the languages, hypothetical manner, real manner and similative meanings are all expressed by a similative ‘like’ marker. Contextual factors are the most common factor used to disambiguate the different meanings of this type of system. Scattered pieces of information seem to suggest that TAM values may also aid in such a disambiguation process. However, at the current stage of our typological knowledge, much more work needs to be done in this area.

There are a number of aspects relevant to the study of hypothetical manner constructions that this study could not address. Accordingly, they remain to be investigated by future studies, and in what follows some potentially fruitful areas are mentioned. First, the diachronic origin of clause-linking devices seems a promising area. As was shown in this paper, in some languages of the sample, the free adverbial conjunction seems to have been derived from a verb meaning ‘to say’. Second, another candidate for larger-scale future investigations is the number of clause-linking devices that may appear in the construction. In various languages of the sample, the complex sentence construction may appear with two clause-linking devices at the same time (e.g. similative marker and free adverbial conjunction). Interestingly, for the most part, one of the devices is always optional. It remains an open task to explore the range of factors that lead to this optionality. Third, for some large genera, this study

could only take into account one language (e.g. Oceanic). Therefore, the next step is to explore the typology of the expression of hypothetical manner within particular large genera. This will enable us to explore internal diversity and try to come up with more fine-grained typological generalizations. Fourth, in most languages of the sample, the adverbial clause tends to appear post-posed to the main clause. The motivations for the positioning of hypothetical manner clauses are an unexplored territory and open for future research (cf. Hetterle 2015: 127). Fifth, hypothetical manner constructions usually develop from the adverbial domain to the complementation domain. After that, they tend to develop into insubordinate constructions. It remains an open task to explore whether this holds in a larger sample. Furthermore, Huddleston & Pullum (2002: 1152) mention that insubordinate ‘as if’ constructions usually develop an exclamatory function. This holds for various Indo-European languages. However, it is not clear whether other languages with insubordinate ‘as if’ construction develop this function. This also remains an unexplored territory and open for future research.

## Abbreviations

1 = 1 <sup>st</sup> person	ERG = ergative	PFV = perfective
2 = 2 <sup>nd</sup> person	EXC = excised	PL = plural
3 = 3 <sup>rd</sup> person	F = feminine	POSS = possessive
ABS = absolutive	FOC = focus	POT = potential
ACC = accusative	FUT = future	PREF = unglossable prefix
ACT = actor	GEN = genitive	PREP = preposition
ADV = adverbial	GL = goal	PRO = emph./contr. proform
AFF = affirmative	HAB = habitual	PROG = progressive
AN = action nominal	HON = honorific	PROP = proprietive
ANT = anterior	IMP = imperative	PRS = present
AUG = augment	INCL = inclusive	PST = past
AUX = auxiliary	IND = indicative	PTCP = participle
CF = counterfactual	INSTR = instrumental	RDP = reduplication
CNJ = conjunct	IPFV = imperfective	REAL = realis
COND = conditional	IRR = irrealis	REC = reciprocal
CONNEG = connegative	ITER = iterative	REFL = reflexive
CONT = continuous	LIM = limitative	REL = relativizer
COP = copula	LOC = locative	REP = repetitive
DECL = declarative	LOG = logophoric	REPUT = reputative
DEM = demonstrative	M = masculine	SBJ = subject

DES = desiderative	NC = noun classifier	SBJV = subjunctive
DET = determiner	NEG = negative	SG = singular
DISTR = distributional	NEUT = neutral	SIM = similative
DP = discourse particle	NMLZ = nominalizer	SPEC = specifier
DS = different subject	NOM = nominative	SS = same subject
DU = dual	NON.FUT = non future	SUB = subordinator
DUB = dubitative	NON.PRS = non present	TERM = terminative
EMPH = emphatic	NON.SBJ = non subject	TOP = topic
EP = epenthesis	OBJ = object	

## References

- Anderson, Judi L. 1989. *Comaltepec Chinantec syntax*. Dallas: The Summer Institute of Linguistics and the University of Texas at Arlington.
- Bakker, Dik. 2011. Language sampling. In Jae Jung Song (ed.), *The Oxford Handbook of Linguistic Typology*, 100-129. Oxford: Oxford University Press.
- Bartholomew, Doris A. & Engel, Ralph. 1987. Gramática Zoque. In Ralph Engel & Mary Engel (eds.), *Diccionario Zoque de Francisco León*, 329-416. Mexico: Instituto Lingüístico de Verano.
- Brinton, Laurel J. 2014. The extremes of insubordination: Exclamatory as if! *Journal of English Linguistics* 42(2). 93-113.
- Brooks, Joseph. 2018. *Realis and irrealis: Chini verb morphology, clause chaining, and discourse*. Ph.D. Dissertation, University of California Santa Barbara.
- Brown, Cecil. 2011. The role of Nahuatl in the formation of Mesoamerica as a linguistic area. *Language Dynamics and Change* 1(2). 171-204.
- Carlson, Robert. 1994. *A grammar of Supyire*. Berlin/New York: Mouton de Gruyter.
- Caron, Bernard. 2017. Comparison, similarity and simulation in Zaar, a Chadic language of Nigeria. In Yvonne Treis & Martine Vanhove (eds.), *Similative and equative constructions: A cross-linguistic perspective*, 167-188. Amsterdam/Philadelphia: John Benjamins.
- Chafe, Wallace. 1995. The realis-irrealis distinction in Caddo, the Northern Iroquoian languages, and English. In Joan Bybee & Suzanne Fleischman (eds.), *Modality in grammar and discourse*, 349-365. Amsterdam/Philadelphia: John Benjamins.
- Chamoreau, Claudine. 2017. Multifunctionality and polysemy of the similative marker =kan in Pesh. In Yvonne Treis & Martine Vanhove (eds.), *Similative and*

- equative constructions: A cross-linguistic perspective*, 321-337. Amsterdam/Philadelphia: John Benjamins.
- Chang, Suk-Jin. 1996. *Korean*. Amsterdam/Philadelphia: John Benjamins.
- Christiansen-Bolli, Regula. 2010. *A grammar of Tadakshak: A northern Songhay language of Mali*. Leiden: Rijksuniversiteit te Leiden.
- Cristofaro, Sonia. 2003. *Subordination*. Oxford: Oxford University Press.
- Culicover, Peter & Jackendoff, Ray. 1999. The view from the periphery: The English comparative correlative, *Linguistic Inquiry* 30(4). 543-571.
- Darmon, Chloé. 2017. The morpheme *-(a)ŋa* in Xamtanga: Functions and grammaticalization targets. In Yvonne Treis & Martine Vanhove (eds.), *Similitive and equative constructions: A cross-linguistic perspective*, 359-385. Amsterdam/Philadelphia: John Benjamins.
- Dixon, R. M. W. 2006. Complement clauses and complementation strategies in typological perspective. In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *Complementation: A cross-linguistic Typology*, 1-48. Oxford: Oxford University Press.
- Dixon, R. M. W. 2009. The semantics of clause linking in typological perspective. In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *The semantics of clause linking: A cross-linguistic typology*, 1-55. Oxford: Oxford University Press.
- Dryer, Matthew S. 2013. Genealogical language list. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at: <https://wals.info/languoid/genealogy>, Accessed on 2019-06-18).
- Dum-Tragut, Jasmine. 2009. *Armenian*. Amsterdam/Philadelphia: John Benjamins.
- Evans, Nicholas. 1995. *A grammar of Kayardild: With historical-comparative notes on Tangkic*. Berlin/New York: Mouton de Gruyter.
- Evans, Nicholas. 2007. Insubordination and its uses. In Irina Nikolaeva (ed.), *Finiteness: Theoretical and empirical foundations*, 366-431. Oxford: Oxford University Press.
- Faarlund, Jan Terje. 2012. *A Grammar of Chiapas Zoque*. Oxford: Oxford University Press.
- Fortescue, Michael. 2010. Similitude: A conceptual category. *Acta Linguistica Hafniensia* 42(2). 117-142.
- Fuchs, Catherine 2014. *La comparaison et son expression en français*. Paris: Ophrys.
- Guérois, Rozenn. 2015. *A grammar of Cuwabo. Mozambique, Bantu P34*. PhD Dissertation, Lumière Lyon.

- Graczyk, Randolph. 2007. *A grammar of Crow: Apsáaloke Aliláau*. Lincoln: University of Nebraska Press.
- Haspelmath, Martin. 1993. *A grammar of Lezgian*. Berlin/New York: Mouton de Gruyter.
- Haspelmath, Martin. 2004. Coordinating constructions: An overview. In Martin Haspelmath (ed.), *Coordinating constructions*, 3-39. Amsterdam/Philadelphia: John Benjamins.
- Haspelmath, Martin, & Buchholz, Oda. 1998. Equative and similative constructions in the languages of Europe. In Johan Van der Auwera (ed.), *Adverbial constructions in the languages of Europe*, 277-334. Berlin/New York: Mouton de Gruyter.
- Heath, Jeffrey. 1978. *Linguistic diffusion in Arnhem Land*. Canberra: Australian Institute of Aboriginal Studies.
- Heath, Jeffrey. 2014. *A grammar of Donno So (Kamma So), Dogon language family, Mali*. Ms.
- Heath, Jeffrey. 2015. *A grammar of Ben Tey (Dogon of Beni)*. Ann Arbor: University of Michigan.
- Heath, Jeffrey. 2016. *A grammar of Dogul Dom (Dogon language family, Mali)*. Ann Arbor: University of Michigan.
- Heine, Bernd, & Kuteva, Tania. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- Heine, Bernd & Kuteva, Tania. 2008. Constraints on contact-induced linguistic change. *Journal of Language Contact* 2. 57-90.
- Hellwig, Birgit. 2009. The semantics of clause-linking in Goemai. In R. M. W. Dixon & Alexandra Aikhenvald (eds.), *The semantics of clause-linking: A cross-linguistic typology*, 318-335. Oxford: Oxford University Press.
- Hellwig, Birgit. 2011. *A grammar of Goemai*. Berlin/New York: Mouton de Gruyter.
- Hengeveld, Kees. 1998. Adverbial clauses in the languages of Europe. In Johan van der Auwera (ed.), *Adverbial constructions in the languages of Europe*, 335-420. Berlin/New York: Mouton de Gruyter.
- Hetterle, Katja. 2015. *Adverbial clauses in cross-linguistic perspective*. Berlin/Boston: Mouton de Gruyter.
- Hewitt, B. George. 1995. *Georgian: A structural reference grammar*. Amsterdam/Philadelphia: John Benjamins.
- Hopper, Paul J. & Traugott, Elizabeth C. 2008. *Grammaticalization* (2nd edition). Cambridge: Cambridge University Press.

- Huber, Juliette. 2008. *First steps towards a grammar of Makasae: A language of East Timor*. Munich: Lincom.
- Huddleston, Rodney & Pullum, Geoffrey K. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
- Klumpp, Deloris A. 2019. *A grammar of Piapoco*. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.
- Konnerth, Linda. 2014. *A grammar of Karbi*. PhD Dissertation, University of Oregon.
- Kortmann, Bernd. 1997. *Adverbial subordination: A typology and history of adverbial subordinators based on European languages*. Berlin/New York: Mouton de Gruyter.
- Kortmann, Bernd. 2001. Adverbial conjunctions. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher, & Wolfgang Raible (eds.), *Language typology and language universals: An international handbook*, 842-854. Berlin/New York: Mouton de Gruyter.
- López-Couso, María-José, & Méndez-Naya, Belén. 2015. Secondary grammaticalization in clause combining: From adverbial subordination to complementation in English. *Language Sciences* 47. 188-198.
- Martowicz, Anna. 2011. *The origin and functioning of circumstantial clause linkers: A cross-linguistic study*. PhD Dissertation, University of Edinburgh.
- McCallum Jones, Ross. 1998. *The Boko/Busa language cluster*. Munich: Lincom.
- Miestamo, Matti & Bakker, Dik & Arppe, Antti. 2016. Sampling for variety. *Linguistic Typology* 20(2). 233-296.
- Mihas, Elena. 2015. *A grammar of Alto Perené (Arawak)*. Berlin/Boston: Mouton de Gruyter.
- Miller, Mark T. 2007. *A grammar of West Coast Bajau*. PhD Dissertation, University of Texas at Arlington.
- Mithun, Marianne. 1995. On the relativity of irrealis. In Joan Bybee & Suzanne Fleischman (eds.), *Modality in grammar and discourse*, 367-388. Amsterdam/Philadelphia: John Benjamins.
- Mithun, Marianne. 2012. Morphologies in contact: Form, meaning, and use in the grammar of reference. In Thomas Stolz, Martine Vanhove, Hitomi Otsuka, & Aina Urdze (eds.), *Morphologies in Contact*, 15-36. Berlin: Akademie Verlag.
- Mous, Maarten. 1992. *A grammar of Iraqw*. Leiden: Research School of Asian, African and Amerindian Studies, University of Leiden.
- Nefedov, Andrey. 2015. *Clause linkage in Ket*. Utrecht: LOT Dissertation Series.
- Nichols, Johanna. 2011. *Ingush grammar*. Berkeley: University of California Press.

- Nikolaeva, Irina & Tolskaya, Maria. 2001. *A grammar of Udihe*. Berlin/New York: Mouton de Gruyter.
- Nikolaeva, Irina. 2014. *A grammar of Tundra Nenets*. Berlin/Boston: Mouton de Gruyter.
- Noonan, Michael. 1992. *A grammar of Lango*. Berlin/New York: Mouton de Gruyter.
- Olguín Martínez, Jesús & Lester, Nicholas A. To appear. A quantitative analysis of counterfactual conditionals in typological perspective. *Italian Journal of Linguistics*.
- Palmer, Frank R. 2001. *Mood and modality* (2nd edition). Cambridge: Cambridge University Press.
- Qian, Yong. 2016. *A typology of counterfactual clauses*. PhD Dissertation, City University of Hong Kong.
- Roulon-Doko, Paulette. 2017. Expressing comparison in Gbaya, a Ubangian language of the Central African Republic. In Yvonne Treis & Martine Vanhove (eds.), *Similitive and equative constructions: A cross-linguistic perspective*, 213-237. Amsterdam/Philadelphia: John Benjamins.
- Sakel, Jeanette. 2007. Types of loan: Matter and pattern. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 15-29. Berlin/New York: Mouton de Gruyter.
- Schmidtke-Bode, Karsten. 2014. *Complement clauses and complementation systems: A cross-linguistic study of grammatical organization*. PhD Dissertation, Friedrich-Schiller-Universität Jena.
- Schmidtke-Bode, Karsten, & Diessel, Holger. To appear. The typology of non-argument clauses. In Manfred Krifka & Mathias Schenner (eds.), *The Oxford Handbook of Embedding*. Oxford: Oxford University Press.
- Shields, Jäna K. 1988. A syntactic sketch of Silacayoapan Mixtec. In C. Henry Bradley & Barbara E. Hollenbach (eds.), *Studies in the syntax of Mixtecan languages* 1, 305-449. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.
- Smits, Helena Johanna. 2017. *A grammar of Lumun, a Kordofanian language of Sudan*. Utrecht: LOT Dissertation Series.
- Sousa, Lisa & Terraciano, Kevin. 2003. The original conquest of Oaxaca: Nahuatl and Mixtec accounts of the Spanish conquest. *Ethnohistory* 50(2). 349-400.
- Spronck, Stef. 2015. Refracting views: How to construct complex perspective in reported speech and thought in Ungarinyin. *STUF – Language Typology and Universals* 68(2). 165-185.

- Spronck, Stef & Vuillermet, Marine. 2019. Mistaken-belief constructions across two continents. Presentation given at the 13th Conference of the Association for Linguistic Typology, University of Pavia 4-6 September.
- Stebbins, Tonya N. 2009. The semantics of clause linking in Mali. In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *The semantics of clause linking: A cross-linguistic typology*, 356-379. Oxford: Oxford University Press.
- Sulkala, Helena & Karjalainen, Merja. 1992. *Finnish*. London: Routledge.
- Terraciano, Kevin. 1990. Mixtec and Nahuatl in Colonial Oaxaca. *UCLA Historical Journal* 10. 122-146.
- Thompson, Sandra & Longacre, Robert & Hwang, Shin. 2007. Adverbial clauses. In Timothy Shopen (ed.), *Language typology and syntactic description volume II: Complex constructions*, 237- 300. Cambridge: Cambridge University Press.
- Treis, Yvonne. 2017. Similative morphemes as purpose clause markers in Ethiopia and beyond. In Yvonne Treis & Martine Vanhove (eds.), *Similative and equative constructions: A cross-linguistic perspective*, 91-142. Amsterdam/Philadelphia: John Benjamins.
- Treis, Yvonne 2018. Comparative constructions: An introduction. *Linguistic Discovery* 16. i-xxvi.
- Treis, Yvonne & Vanhove, Martine (eds.) 2017. *Similative and equative constructions: A crosslinguistic perspective*. Amsterdam/Philadelphia: John Benjamins.
- Tsunoda, Tasaku. 2011. *A grammar of Warrongo*. Berlin/Boston: Mouton de Gruyter.
- Tuggy, David H. 1979. Tetelcingo Náhuatl. In Ronald W. Langacker (ed.), *Studies in Uto-Aztecan Grammar 2: Modern Aztec Grammatical Sketches*, 1-140. Dallas: The Summer Institute of Linguistics and the University of Texas at Arlington.
- Ultan, Russell 1972. Some features of basic comparative constructions. *Working Papers on Language Universals (Stanford)* 9. 117-162.
- Valentine, J. Randolph. 2009. The semantics of clause linking in Ojibwe. In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *The semantics of clause linking: A cross-linguistic typology*, 193-217. Oxford: Oxford University Press.
- Vanhove, Martine. 2017. Similative, equative, and comparative constructions in Beja (North-Cushitic). In Yvonne Treis & Martine Vanhove (eds.), *Similative and equative constructions: A cross-linguistic perspective*, 189-211. Amsterdam/Philadelphia: John Benjamins.
- Weinreich, Uriel. 1964. *Languages in contact*. London/The Hague/Paris: Mouton.



- Ylikoski, Jussi. 2017. Similarity, equality and the like in North Saami. In Yvonne Treis & Martine Vanhove (eds.), *Similative and equative constructions: A cross-linguistic perspective*, 259-290. Amsterdam/Philadelphia: John Benjamins.
- Zylstra, Carol F. 1991. A syntactic sketch of Alacatlazala Mixtec. In C. Henry Bradley & Barbara E. Hollenbach (eds.), *Studies in the Syntax of Mixtecan Languages* 3, 1-177. Dallas: The Summer Institute of Linguistics and the University of Texas at Arlington.

**CONTACT**

[olguinmartinez@ucsb.edu](mailto:olguinmartinez@ucsb.edu)

# Expressing equality and similarity in English, Italian, and Ladin: Interlingual contrastive features and micro-variation

MARTINA IRSARA<sup>1</sup>

<sup>1</sup>FREE UNIVERSITY OF BOZEN-BOLZANO

Submitted: 12/11/2020 Revised version: 25/04/2021

Accepted: 26/04/2021 Published: 31/08/2021

## Abstract

Rather exceptionally among European languages, English has two standard markers, namely *as* and *like*, which make a formal distinction between equative and similative constructions. It is well known that clauses with *as* and phrasal adjuncts with *like* tend to be carefully distinguished in British formal usage. The present article uses English as a hallmark for the identification of relevant semantic distinctions within the field of equality and similarity comparison in the Ladin variety spoken in the South-Tyrolean valley of Badia, which forms part of the Rhaeto-Romance territory of Italy, and which is still under-researched in many domains. This article intends to contribute to the current discussion on quantitative and qualitative comparison by providing novel information on Ladin, which has not been an object of investigation with respect to comparative constructions. The comparative system of Ladin is illustrated from a cross-linguistic perspective, drawing comparisons with English and Italian, with which it has much in common, but from which it also differs in a number of respects. The article also shows how the Ladin system varies at a micro-level within one valley. While the vernacular in the lower part of the valley has markers that distinguish between clausal and phrasal complements, the linguistic variety in the upper part of the valley makes no such distinction. Despite highlighting cross-linguistic differences, this study serves as a further confirmation of typological tendencies.

**Keywords:** comparatives; equatives; similatives; Ladin; Italian; English *as* and *like*.

## 1. Introduction

Equative and similative constructions have been studied over a number of years from a range of perspectives and in large samples of languages. EQUATIVES express an

equivalent degree of a gradable property (*as* parameter *as* standard), while SIMILATIVES express similarity of quality or manner (parameter *like* standard). In addition to the more straightforward clauses of equality and similarity of the type of (i) *Mary is as pretty as Anne* and (ii) *Anne swims like a fish*, there are various constructions that have been viewed as somewhat related. Treis (2018) adds SIMULATIVE CONSTRUCTIONS (pretended similarity of the type of *as if*) to her analysis, while Haspelmath & Buchholz (1998) add ACCORD CLAUSES (*as we all know*) and ROLE PHRASES (*as your mother* = role). Quirk et al. (1985) distinguish between: (i) COMPARATIVE CLAUSES OF EQUIVALENCE (... *as healthy as you*), (ii) ADJUNCT CLAUSES OF SIMILARITY (*it was as I imagined*), and (iii) ADJUNCT CLAUSES OF COMPARISON (verb followed by *as if*, *as though*, *like*). Because of their formal and semantic similarity, this article explores both the more straightforward clauses of equality and similarity and a number of related constructions that employ the same markers. The article generally uses the terminology adopted by Haspelmath & Buchholz (1998) and Haspelmath (2017), naming the lexical and functional constituents of equative constructions as shown in Table 1.

Constituents	Terms
<i>My son is</i>	COMPAREE
<i>as</i>	DEGREE (OR PARAMETER) MARKER
<i>clever</i>	PARAMETER
<i>as</i>	STANDARD MARKER
<i>my husband</i>	STANDARD

**Table 1:** Terminology.

The present study is motivated by the wish to contribute to the current debate on marking patterns of equative and similative constructions in languages across Europe. Although most European languages have been described extensively, the South-Tyrolean Ladin language remains widely under-represented in language samples. The aim of this article is therefore to further investigate the equatives and similatives of Ladin and to place them in a broader context, providing novel information and increasing the sample size of languages analysed for constructions of this type. Ladin is investigated in comparison with English and with its genetically related Italian language. Mainly standard British English is considered, with occasional references

to American. Descriptions of constructions in standard Italian are supplemented with references to northern Italian dialects and to early Italian vernaculars.

Since Ladin shows pronounced areal differences, characteristic features as found in more than one part of Val Badia are included in the present report. The valley of Badia is one of the Ladin-speaking valleys that surround the Sella group of mountains in the Italian Dolomites, which are home to about 31,000 speakers of the language that is typically termed *Dolomitic* or *Sella Ladin* (*Sellaladinisch*). The latest population census in 2011 recorded that the five municipalities of Val Badia had a great majority of Ladin speakers, with percentages ranging from 90% to 98% (ASTAT 2019). The varieties spoken in Val Badia are commonly indicated as BADIOT (or LADIN DE ALTA VAL, in the upper part of the valley), LADIN DE MESA VAL (in this article *Lmv*, in the middle part of the valley and nowadays recognised as the standard form in Val Badia) and MAROU (in the municipality of Mareo-Marebbe-Enneberg, in the lower part of the valley). Because of the many features shared by *Badiot* (*Ladin de alta val*) and *Ladin de mesa val* concerning structures of comparison, this article will treat them together, unless stated otherwise. If the varieties in the upper and middle part of the valley are treated together, e.g. in the examples provided, they will be indicated as LADIN (a term that will therefore include the Ladin varieties of the upper and middle parts of the valley).<sup>1</sup>

The article therefore focuses on variation between a number of different languages and within one language as spoken in a restricted area, adopting perspectives from contrastive linguistics and dialectology or micro-variation. The contrastive analysis is linked to a number of typological findings in the field of examination. Equative and simulative constructions in the languages under consideration are investigated using corpora such as the *British National Corpus* (BNC), *the English Web* (enTenTen15) and *the Italian Web* (itTenTen16 and itTenTen10). Ladin examples are obtained from the corpora *Tratament Automatic dl Lingaz Ladin: corpus lad* (TALL), and *Corpus dl Ladin Leterar* (CLL). While the TALL corpus includes texts of various genres, CLL is made up of literary texts only and is smaller, but it is based on better controlled data sources and provides a clearer automatic differentiation of varieties. Further Ladin data are drawn from the digitalised information bulletin of the municipality of Mareo, from dictionaries, and grammar books. The linguistic evidence provided by the Ladin databases is verified and complemented orally by speakers of the language, including

---

<sup>1</sup> The use of the term *Ladin* to refer to the varieties spoken in the upper and middle parts of Val Badia is therefore a simplification adopted in this article for ease of presentation.

the author of this article. Corpora are never fully representative when the goal of the analysis is a fine-grained description of the existence of certain linguistic structures in a minority language with several varieties. Digital corpora and introspective empirical evidence therefore complement each other. Relevant data are discussed with reference to pertinent studies, some key ones of which are briefly summarised in §2 below.

## 2. Literature background

Various scholars have addressed the topic of comparison in earlier works, the most influential of which have received a comprehensive review in Stolz (2013). An extensive list of existing sources has been identified and appraised more recently by Treis (2018). The present article would do little justice to the full range of studies on the topic and will therefore limit itself to summarising the works on equative and similitative constructions from which it draws most.

Haspelmath & Buchholz (1998) belong to the pioneers who engaged in an extensive analysis of similitive and equative constructions in a high number of European languages, including English, Italian, and Friulian, which also belongs to Rhaeto-Romance and shares a number of linguistic features with the Ladin varieties considered in this article. Rhaeto-Romance is namely taken to include Ladin and Friulian in Italy and Romansh in Switzerland. With their typological study of equatives and similitives, Haspelmath & Buchholz (1998) contributed to a volume on adverbial constructions, with which they claimed to be directly concerned, since an act of comparing typically involves concepts of degree and manner. The authors identified a bundle of features of equatives and similitives across European languages that can be taken as one of the features of *Standard Average European* (SAE). English, Italian, and Friulian were classified as belonging to the core SAE languages as regards the expression of equality and similarity. Haspelmath (2017) extended his analysis to include more languages in the world, whose equative constructions can be divided into six major types. The primary type that occurs in English and in most European languages is an equative construction that includes both an equative degree marker and an equative standard marker (*as tall as*). Haspelmath (2017) concluded with three broad cross-linguistic generalisations, the first of which is that languages normally do not have a degree marker that is not accompanied by a standard marker. Secondly, languages generally have object-verb order if the parameter follows the standard.

Thirdly, if the standard comes after the parameter, then the standard marker occurs before the standard.

Basic equative constructions with a specific standard of comparison were found to be very diverse in Henkelmann's (2006) world language sample, but constructions with equative markers predominated. The European language included in his sample, namely German, was shown to attest the use of relative-based equative constructions in accordance with SAE languages. Case-transparent German examples were provided to show that standard markers are actually subordinators that introduce typically reduced clauses.

Treis & Vanhove (2017) collected a range of theoretical and analytical perspectives on equative and similitive constructions across languages, including a cognitive-typological study on *like*-concepts by Schulze (2017). Schulze (2017: 36) highlights that *like*-expressions are derived metonymically or metaphorically and that "they represent fossilised patterns of cognitive processes conventionalised over times". Schulze (2017) argues that *like*-concepts are closely related to motion-concepts and image-concepts. In the same volume, König (2017) also emphasises that the process of assessing similarity is a key cognitive activity. König (2017) maintains that deictic expressions play a significant role in the formation of equatives and similitives, arguing that the exophoric or gestural use of demonstratives of manner, quality, and degree lies at the basis of comparative constructions. Latin used the expressions *sic*, *talis*, and *tantus* for the three different domains of manner, quality, and degree, while German *so* can be employed for all three dimensions. Unlike German, English has abandoned the exophoric use of *so*, using composite constructions in a gestural context instead (e.g. *like this* or *like that*). While German *so* functions as parameter or degree marker in equative comparatives, English "uses its reinforced form (*as* < *eal swa*) as comparative marker in affirmative sentences, though retaining the original basic manner deictic *so* in negations (*Fred is not so tall as George*)" (König 2017: 156).

Belletti (1991) considers the use of the contemporary Italian standard markers *quanto*, *come*, and *che* in combination with the degree markers *tanto* and *altrettanto*, their syntactic position, the pronominal forms following the markers, and elliptic structures. Belletti (2010) analyses comparatives in early Italian texts from the 13<sup>th</sup> and beginning of the 14<sup>th</sup> century. While the comparative marker *che* in comparatives of inequivalence was used in more syntactic contexts than in modern standard Italian, no major differences were detected in comparative structures of equivalence.

The divide between *che* and *di* in comparatives of inequivalence was also addressed by Stolz (2013) in his analysis of competing comparative constructions in European languages. Stolz (2013) included Badiot in his sample of 44 Romance languages, but he provided no information on equatives or similatives in this Ladin variety. The comparative system of a Ladin variety was addressed for the first time in a systematic way by Irsara (2001, 2012, 2013), who described characteristics of the Val Badia system from a cross-linguistic perspective, including English and Italian in her analysis, and integrating the descriptive account with a discussion of potential teaching implications.

The present article moves on to addressing equative degree markers in English, Italian, and Ladin in §3. Equative and similative standard markers in these three languages will be discussed in §4, while §5 will uncover characteristic uses of the standard markers in Marou as compared to English. Although several constructions will be considered throughout the article, the main focus will be on equative and similative constructions with an adjectival parameter.

### **3. The degree markers *as* (English), *tanto* (Italian), and *tan* (Ladin) with adjectival parameters**

Equative constructions have been shown to be complex and diverse across languages, which often use degree markers as well as standard markers in their equatives, forming constructions that belong to the second of Haspelmath's (2017: 14) primary types, defined as "an equative construction that contains an ordinary predicative property-word as parameter plus differentiated comparee and standard, with both an equative degree-marker ('equally') and an equative standard-marker". This type is found in English, Italian, and Ladin, which use analytic degree markers that are always accompanied by a standard marker, in accordance with Haspelmath's (2017: 25) first generalisation that "no language has only a degree-marker, leaving the standard unmarked". Table 2 shows the degree (or parameter) markers that can be used in equative but not similative constructions in the languages and varieties considered.

A relationship of equivalence is typically expressed in English by a correlative use of the parameter marker *as* in combination with the same standard marker. Example (1) illustrates the construction 'as adjective as noun phrase'.

Varieties	Degree (parameter) markers
English	<i>as / (so)</i>
Italian	<i>tanto / altrettanto / così</i>
Badiot	<i>tan</i>
Ladin de mesa val (Lmv)	<i>tan</i>
Marou	<i>tan</i>

Table 2: Degree (parameter) markers.

(1) English (enTenTen15)

*I doubt there's a woman out there who hasn't dreamed of a spouse who looks like Brad Pitt, is **as** intelligent **as** Bill Gates, **as** kind and generous **as** Paul Newman, and **as** built **as** Arnold Schwarzenegger.*

Negated equatives are occasionally formed in English with the parameter marker *so*, which gives an archaic feeling to the constructions, according to Haspelmath & Buchholz (1998). The combination of the degree markers *as* and *so* with the standard marker *as* is exemplified in (2) and (3).

(2) English (enTenTen15)

*His injury is not **as** serious **as** that of teammate James O'Donoghue.*

(3) English (enTenTen15)

*The fiscal position of Portugal was certainly not good, but not **so** serious **as** that of Grece.*

Parameters of comparison are most typically accompanied by the degree marker *tanto* in Italian and *tan* in Ladin, as illustrated in (4) and (5).<sup>2</sup>

<sup>2</sup> The glosses used in the article follow broadly the Leipzig Glossing rules and are provided in the Appendix.



## (4) Italian (itTenTen16)

*I computer arriveranno ad essere tanto intelligenti*  
 the computers arrive.FUT to be [as much].DEGM intelligent  
*quanto gli umani.*  
 [how much].STDM the humans  
 ‘Computers will become as intelligent as humans.’

## (5) Badiot (TALL)

*L’ çiamurc, presciapüch tan gran che na*  
 the chamois, approximately [as much].DEGM big that.STDM a  
*çiora, po pesè trënta y inçe plö chili.*  
 goat, can weigh thirty and also more kilograms  
 ‘The chamois, more or less as big as a goat, can weigh thirty and even more kilograms.’

A further equative degree marker in Italian is *altrettanto*, exemplified in (6). However, *altrettanto... che* is considered bookish in contemporary Italian.

## (6) Italian (itTenTen16)

*La nostra Befana... era altrettanto simpatica*  
 the our Befana... was [just as much].DEGM nice  
*che Babbo Natale.*  
 that.STDM Father Christmas  
 ‘Our Befana (old hag) was just as nice as Santa Claus.’

As pointed out by Belletti (2010) and Pelo (2012) in their investigations into comparatives in Old Italian, the marker *(co)sì* was frequently used in combination with *come*. As (7) and (8) illustrate, *così* still occurs as a degree marker in contemporary Italian, often in negative clauses.

## (7) Italian (itTenTen16)

*... figure luminose che sono così lunghe come galassie*  
 ... figures luminous that are [like this].DEGM long like.STDM galaxies  
*intere.*  
 whole

‘... bright figures that are as long as whole galaxies.’

(8) Italian (itTenTen16)

*Jaime non è così intelligente come David, ma è più  
Jamie not is [like this].DEGM intelligente like.STDM David, but is more  
diligente.  
diligent*

‘Jaime is not as intelligent as David, but he is more diligent.’

In their attempt to identify typical features of SAE languages with respect to equative and similative constructions, Haspelmath & Buchholz (1998) concluded that parameter or degree markers are characteristically demonstrative-based in European languages. Although no deictic value is evident in the English marker *as*, its Old-English root *swá* was a demonstrative-based element. Nowadays, English *so* is not used deictically in an exophoric (or gestural) manner to refer to an extralinguistic situation, in which case adnominal demonstratives and other expressions of similarity are used, e.g. *this* + adjective, *like this* (+ gesture). König (2017) argued that the loss of the exophoric function of a basic demonstrative in English is untypical and that the original meaning of similarity of *so* is still noticeable in sentential anaphoric contexts of the type of *She said so*. The original meaning of similarity of *so* is also visible in expressions such as *Sarah is tall, and so is Jane* and in replies of the type of *So am I*. Italian *tanto* and Ladin *tan* are also demonstrative-based (like Latin *tam*) and still have a deictic function in certain contexts. The Italian utterances in (9) and (10) would most naturally be accompanied by a pointing gesture indicating height, width, and length. The Ladin questions in (11) were uttered by a saleswoman selling cheese behind a counter. By placing her knife in different positions on a wheel of cheese, the woman asked the customer how much she should cut off the larger piece, indicating potential slices.

(9) Italian (Zingarelli)

*È alto tanto e largo tanto.  
is high [this much] and wide [this much]  
‘It is this high and this wide.’*

## (10) Italian (itTenTen16)

*Aveva una pipa lunga tanto.*  
 had.3SG a pipe long [this much]  
 ‘He had a pipe as long as this.’

## (11) Badiot

*Tan nen os-te pa? Tan? Tan?*  
 [how much] PRTV want-you.SG PTCL? [this much]? [this much]?  
 ‘How much of it do you want? This much? This much?’

Ladin *tan* appears in a variety of syntactic functions, among which that of an interrogative pronoun, as can be noticed in (11) above and (12-13) below, in which case Italian uses *quanto* and not the formally similar *tanto*.

## (12) Badiot

*Tan nen os-te pa?*  
 [how much] PRTV want-you.SG PTCL?  
 ‘How much of it do you want?’

## (13) Italian

*Quanto ne vuoi?*  
 [how much] PRTV want.2SG?  
 ‘How much of it do you want?’

Hence, Ladin *tan* can be used as a quantitative pronominal word, displaying a characteristic feature of European languages. Haspelmath & Buchholz (1998: 299) maintained in fact that “an additional feature characterizing Standard Average European languages is the use of quantitative pronominal words when quantities are compared.” Quantities are compared explicitly in (14) and (15), whereas *tan* premodifies an adjective of size in (16) and an adjectival parameter that denotes a quality in (17).

## (14) Badiot

*Nos un tan de patüic ch’ os.*  
 we have [as much] of.PRTV stuff that.STDM you.PL

‘We have as much stuff as you.’

(15) Badiot

*Nos nen un tan ch’ os.*<sup>3</sup>  
we PRTV have [as much] that.STDM you.PL  
‘We have as much of it as you.’

(16) Badiot (TALL)

*L’ cíamurc, presciapüch tan gran che na cíora,*  
the chamois approximately [as much].DEGM big that.STDM a goat,  
*po pesè trènta y incé plö chili.*  
can weigh thirty and also more kilograms  
‘The chamois, more or less as big as a goat, can weigh thirty and even more kilograms.’

(17) Badiot (TALL)

*Les ères... é tan bunes che i ëi.*  
the women... are [as much].DEGM good that.STDM the men  
‘Women... are as good as men.’

Ladin *tan* can be preceded by a multiplicative numeral, as (18) and (19) illustrate. It can therefore be argued in Rett’s (2020: 182) words that Ladin demonstrates “the ability of an equative to be modified by a factor modifier like *twice*.” Examples (18) and (19) illustrate the construction called *equative-pro-COI construction* by Stolz (2013), who defines it as “the use of the syntactic format of the equative construction for the purpose of expressing the meaning of a typical COI construction,” where COI stands for *comparison of inequivalence*.

(18) Badiot (TALL)

*Al é dui iadi tan gran che le pice Martin!*  
he is two times [as much].DEGM big that.STDM the little Martin!  
‘He is twice as big as little Martin.’

---

<sup>3</sup> Pronominal *tan* is declined according to number and gender, i.e. *tan* SG.M, *tanta* SG.F, *tanc* PL.M, *tantes* PL.F.

## (19) Lmv (TALL)

*Le sach dô ester almanco trëi iadi tan gran*  
 the sack must.IPFV be at least three times [as much].DEGM big  
*co chël che al ti â sciuré ía a chë*  
 how.STDM that.DEM that.COMP he her.DAT had thrown thither to that  
*vedla.*  
 old.F

‘The bag must have been at least three times as big as the one that he had thrown to that old woman.’

Stolz (2013: 122) argues that the equative-pro-COI construction “prominently involves English”, whose speakers most naturally adopt “the frame [(comparee is X times) as quality as (standard)].” However, Stolz (2013) also points out that a proper COI construction of the type of [(comparee is X times) more quality (or quality-er) than (standard)] is also found in this context.<sup>4</sup> Ladin also uses both the *equative-pro-COI construction* and a *proper COI construction* with multiplicative numerals, as illustrated respectively in (18)-(19) above and in (20) below. Example (21) shows that the equative degree marker *tan* and the degree marker of nonequivalence *plö / plü* can be used synonymously in this context.<sup>5</sup>

## (20) Lmv (TALL)

*Chësta nëi é trëi iadi plü točia co la naturala.*  
 this snow is three times more thick how.STDM the natural  
 ‘This snow is three times thicker than natural snow.’

<sup>4</sup> Similarly, German also employs both the *equative-pro-COI construction* and a *proper COI construction* with multiplicative numerals. A corpus search in the *German Web 2013* (deTenTen13) revealed that the construction ‘three times as big as’ (2473 hits) is more frequent than the construction ‘three times bigger than’ (473 hits) and that the construction ‘three times as high as’ (2742 hits) is more frequent than the construction ‘three times higher than’ (1146 hits). The search included different spelling conventions and numerals (search carried out on 2021.04.23).

<sup>5</sup> The same is pointed out for French by Stolz (2013), who provided the example in (i).

## (i) French (Stolz 2013: 28-29)

*La France est deux fois aussi / plus grande que l’ Autriche.*  
 the France is two times as.DEGM / more big that.STDM the Austria  
 ‘France is twice as big as / bigger than Austria.’

(21) Badiot

*La Francia é dui iadi tan / plö grana ch' l'*  
 the France is two times [as much].DEGM / more big that.STDM the  
*Austria.*  
 Austria  
 'France is twice as big as / bigger than Austria.'

Italian typically employs a *proper COI construction* with multiplicative numerals, as a search in the corpus *Italian Web 2016* confirmed (see Table 3).

<b>proper COI construction</b>	<b>hits</b>	<b>equative-pro-COI construction</b>	<b>hits</b>
<i>tre volte maggiore/più grande</i> three times bigger/ more big 'three times bigger'	715	<i>tre volte tanto grande</i> three times [as much].DEGM big 'three times as big'	1
<i>tre volte più alto</i> three times more high 'three times higher'	159	<i>tre volte tanto alto</i> three times [as much].DEGM high 'three times as high'	0
numeral + <i>volte maggiore/più grande</i> numeral + times bigger/ more big 'numeral + times bigger'	3707	numeral + <i>volte tanto grande</i> numeral + times [as much].DEGM big 'numeral + times as big'	0

Table 3: Corpus search results for Italian (itTenTen16).

### 3.1. Position of the degree markers

The parameter typically precedes the standard of comparison in SAE languages, including English, Italian, and Ladin. Haspelmath (2017: 26) noted in this regard that “if the standard follows the parameter, then the standard-marker generally precedes the standard”. However, Haspelmath & Buchholz (1998) admitted that speculating on the position of the standard marker is easier than predicting the order of the parameter marker. While the standard marker is clearly the head of the standard, it is not so obvious whether the parameter marker is “the head or the dependent of the parameter” (Haspelmath & Buchholz 1998: 289). Example (22) illustrates the pre-adjectival position of the degree marker *as* in English.

(22) English (enTenTen15)

*A woman can be as intelligent as a man or even more.*

It is shown in (23) and (24) that Italian *tanto* can precede or follow the adjective. The same holds true for Ladin *tan*, which usually precedes the adjective, as in (25) and (26), but which can occasionally occur in a post-adjectival position, as illustrated in (27), which exemplifies a generic equative where the image of the world is evoked to indicate something that has existed for a long time.

## (23) Italian (itTenTen16)

*I computer arriveranno ad essere tanto intelligenti*  
 the computers arrive.FUT to be [as much].DEGM intelligent  
*quanto gli umani.*  
 [how much].STDM the humans  
 ‘Computers will become as intelligent as humans.’

## (24) Italian (itTenTen16)

*La donna è intelligente tanto quanto un uomo e a*  
 the woman is intelligent [as much] [how much] a man and at  
*volte di più.*  
 times more  
 ‘A woman is as intelligent as a man and sometimes more.’

## (25) Badiot (TALL)

*Iö sun pa bel tan sciché che tö.*  
 I am PTCL already [as much].DEGM smart that.STDM you.SG  
 ‘I am already as smart as you.’

## (26) Marou (TALL)

*Tö es tan vedla co iu.*  
 you.SG are [as much].DEGM old.F how.STDM I

## (27) Marou (TALL)

*Na verité... Ara é vedla tan co le monn.*  
 a truth... It is old [as much].DEGM how.STDM the world  
 ‘A truth... It is as old as the world / as old as the hills.’

### 3.2. Omission of the degree markers

Quirk et al. (1985) observed that the English degree (parameter) marker *as* is sometimes absent in informal and in literary contexts if a copular verb is present or implied, as in examples (28) and (29).

(28) English (BNC)

*The weapon was old **as** the world and deadly **as** poison.*

(29) English (enTenTen15)

*... his only daughter, comely **as** a summer cloud, clever **as** a cone spider, has fingers so lively she can spin straw into gold.*

Quirk et al. (1985) noted that the single *as* is close to *like* when it is followed by a noun phrase, as examples (30) and (31) illustrate. Haspelmath & Buchholz (1998: 309) specified that in SAE languages the equative parameter marker is optional and often missing in generic equatives, “whose standard does not have specific reference but refers to a class generically”. Examples (30) and (31) illustrate how the image of ivory can be evoked to describe something that is very pale. While in (30) and (31) the standard is an entity that is seen as having a certain property and is used idiomatically, in (32) the standard has specific reference and the typical correlative construction with the double *as* can be observed.

(30) English (BNC)

*Sometimes he would catch Benedicta looking at him, her lovely face pale **as** ivory...*

(31) English (BNC)

*His face was pale **like** carved ivory, his chest still.*

(32) English (BNC)

*Valerie Cass now looked **as** pale **as** her daughter.*

As in English, the Italian degree marker *tanto* can be omitted in the *tanto... quanto* construction, as illustrated in (33), whereas Ladin *tan* always needs to be expressed



in a *tan... che / co* construction of the type of (34) and can be omitted only with similitive standard markers, discussed in §4.<sup>6</sup>

(33) Italian (itTenTen16)

*Le donne non sono intelligenti quanto gli uomini.*  
 the women not are intelligent [how much].STDM the men  
 ‘Women are not (as) intelligent as men.’

(34) Badiot

\* *Iö sun sciché che tö.*  
 I am intelligent that.STDM you.SG  
 ‘I am (as) intelligent as you.’

#### 4. Equative and similitive standard markers

Many European languages characteristically use equative constructions that are based on a demonstrative expression that means ‘so’ (as noted in §3 above) and a relative expression that means ‘how’ and that is often used as an interrogative, as pointed out by Haspelmath (2017: 11-12), who explains that “the demonstrative functions as a degree-marker, and the relative as a standard-marker”. Haspelmath (2017) stresses that demonstrative-relative (correlative) constructions are typical of the European linguistic area, including Germanic and Romance territories. Table 4 enlists the equative and similitive standard markers that occur in the languages considered in the present article. Equative and similitive standard markers are not separated in the table, arguing in line with Haspelmath (2017) that equative and similitive constructions are often expressed in similar ways.

Varieties	Standard markers
English	<i>as / like</i>
Italian	<i>quanto / come / che</i>
Badiot	<i>che / sciöche / coche</i>
Ladin de mesa val (Lmv)	<i>co / sciöche / coche</i>
Marou	<i>co / desco / desche / coche</i>

**Table 4:** Equative and similitive standard markers.

<sup>6</sup> Similarly, Italian *altrettanto* cannot be omitted in the construction *altrettanto... che* (cf. example (6)).

#### 4.1. Equative and similative standard markers in English

The English standard marker *as* typically correlates with the same degree marker. Although it is possible for the standard marker *as* to be unaccompanied by a degree marker, as was discussed in §3 above, this is not particularly common in the databases considered for the present analysis. While (35) exemplifies the typical correlative construction ‘*as* adjective *as* noun-phrase’, (36) shows that it is possible for the standard marker *as* to occur without a preceding degree marker, which happens in particular in generic equatives and in literary or informal contexts, where (36) seems to have arisen, considering the rather blunt choice of words. The single standard marker *as* is usually followed by a generic lexical standard, so that the definite article in (37) is rather untypical.

(35) English (BNC)

*First time I went to London, I was about as old as Emily.*

(36) English (enTenTen15)

*We may be old as the hills, but hopefully we aren’t dumb as a box of rocks.*

(37) English (enTenTen15)

*His eyes were big as the dish-plates.*

Haspelmath & Buchholz (1998) found that the equative standard marker is identical to the similative marker in the majority of SAE languages. English has therefore been described as rather exceptional among European languages, with its two standard markers *as* and *like*, which distinguish formally between equative and similative constructions. The closeness of the single standard marker *as* to the standard *like* was pointed out in §3.2 and is further illustrated in (38) and (39), in which the structures ‘*as* adjective *as* noun phrase’ and ‘adjective *like* noun phrase’ express a semantically similar concept. In both (38) and (39) the size of a building is compared to the size of a castle. However, (38) and (39) might be argued to illustrate quantitative and qualitative comparison respectively.

(38) English (enTenTen15)

*Eight Chimneys is **as big as** a castle, but dustier and darker than Samantha imagines a castle would be.*

(39) English (enTenTen15)

*It is big **like** a castle and has many underground rooms.*

The English similitive marker *like* and the following standard of comparison often form a simple phrase that functions as an adverbial of manner, as illustrated in (40) and (41), which refer to manners of singing and swimming.

(40) English (enTenTen15)

*She sings **like** a bird.*

(41) English (enTenTen15)

*Now he swims **like** a fish.*

#### **4.2. Equative and similitive standard markers in Italian**

It was observed in §3 that the Italian degree marker *tanto* is characteristically accompanied by the standard marker *quanto*, which often occurs without the degree marker. While the standard marker *quanto* is used in combination with *tanto* in (42), it is used on its own in (43).

(42) Italian (itTenTen16)

*L' attività quotidiana può essere **tanto** efficace*  
 the activity daily can be [as much].DEGM effective  
***quanto** gli antidepressivi.*  
 [how much].STDM the antidepressants  
 'Daily activity can be as effective as antidepressants.'

(43) Italian (itTenTen16)

*La psicoterapia cognitivo comportamentale è efficace*  
 the psychotherapy cognitive behavioural is effective

*quanto*            *gli antidepressivi*.  
 [how much].STDM the antidepressants  
 ‘Cognitive-behavioural psychotherapy is (as) effective as antidepressants.’

Belletti (1991) maintains that only the standard marker *quanto* can precede a nominal standard if the adjectival parameter is preceded by *tanto*, whereas *come* cannot be used in this case. Cerruti (2009) acknowledges the standard linguistic norm according to which the degree marker *tanto* correlates with *quanto* + VP, NP, AdjP, AdvP, PrepP or with *come* + NP, but he emphasises that the structure ‘*tanto* adjective *come* adjective’ is attested in his corpus of spoken language as well. Example (44) shows a correlative use of *tanto* and *come* (+ nominal), which is therefore attested in the data examined in the present article, while the structure ‘*tanto* adjective *come* adjective’ is rare.<sup>7</sup>

---

<sup>7</sup> A typical equative construction with two adjectives in Italian is ‘*tanto* adjective *quanto* adjective’, as illustrated in (i).

- (i) Italian (itTenTen16)  
*Nina è tanto bella quanto fortunata.*  
 Nina is [as much].DEGM beautiful [how much].STDM lucky  
 ‘Nina is as beautiful as lucky.’

Ladin ‘*tan* adjective’ can correlate with both ‘*che / co* adjective’ and ‘*sciöche* adjective’, as illustrated in (ii-iii), which also show that the adverbs *avisa* ‘just’ and *ince* ‘also’ typically occur in equative and similitive constructions of this type.

- (ii) Lmv  
*Nina é (avisa) tan bela co (ince) fortunada.*  
 Nina is (just) [as much].DEGM beautiful how.STDM (also) lucky  
 ‘Nina is (just) as beautiful as (also) lucky.’

- (iii) Lmv  
*Nina é (tan) bela sciöche (ince) fortunada.*  
 Nina is ([as much].DEGM) beautiful like.STDM (also) lucky  
 ‘Nina is (as) beautiful like (also) lucky.’

German also frequently uses the emphatic element *genau* to strengthen the parameter marker semantically, as shown in (iv).

- (iv) German (deTenTen13)  
*Der Brief ist genau-so schön wie traurig.*  
 the letter is just-as beautiful as sad  
 ‘The letter is just as beautiful as sad.’

## (44) Italian (itTenTen10)

*La giudico tanto eterna come l' acqua e l'*  
 it.ACC.F judge.1SG [as much].DEGM *eternal* like.STDM the water and the  
 aria.  
 air  
 'I consider it as eternal as water and air.'

When the optional degree marker *tanto* is absent, both *quanto* and *come* are equally acceptable in specific and generic equatives, confirming the statement by Haspelmath & Buchholz (1998) that quantitative and qualitative equative markers in Italian are not always as clearly distinguished and distributed as in other languages. Examples (45-48) illustrate how *quanto* and *come* can both be used in specific and generic equatives, although it might still be argued that *quanto* and *come* convey a slightly different semantic meaning.

## (45) Italian (itTenTen16)

*Tu sei forte quanto me.*  
 you.SG are strong [how much].STDM *me*  
 'You are (as) strong as me.'

## (46) Italian (itTenTen16)

*Roberto era sano e forte come me.*  
 Roberto was healthy and strong like.STDM *me*  
 'Roberto was healthy and strong like me.'

## (47) Italian (itTenTen16)

*Era forte quanto un uomo.*  
 was.3SG strong [how much].STDM a man  
 'He was (as) strong as a man.'

## (48) Italian (itTenTen16)

*Il padre le insegna a essere forte come un uomo.*  
 the father her.DAT teaches to be strong like.STDM a man  
 'Her father teaches her to be strong like a man.'

Haspelmath (2017: 13) points out that European languages often “use a ‘how’ word to express similarity of manner”, such as *come* in Italian. Treis (2018: 22) explains

that “the main similative construction involves an intransitive stative verb ‘be like (this)’ or a transitive verb ‘do like (this)’”, which is exemplified in (49) and (50), where *come* is used to express qualitative similarity.

(49) Italian (itTenTen16)

*Libera era come suo padre.*  
Libera was like.STDM her father  
‘Libera was like her father.’

(50) Italian (itTenTen16)

*Canta come un gatto schiacciato, ma canta.*  
sings like.STDM a cat squashed, but sings  
‘He sings like a squashed cat, but he sings.’

A further standard marker in Italian constructions of comparison is *che*, which introduces the standard if the parameter is preceded by *altrettanto*, as was noted in §3 and further exemplified in (51). While Belletti (1991) rejects the use of *quanto* and *come* in correlation with *altrettanto* as unacceptable, the correlative structure ‘*altrettanto* adjective *quanto* / *come* NP’ is not categorically absent in the databases examined, as shown in (52) and (53).<sup>8</sup>

(51) Italian (itTenTen16)

... *pensa che le donne trobriandesi siano... e altrettanto*  
... thinks that the women Trobriand be.SBJV... and [just as much].DEGM  
*libere che gli uomini.*  
free that.STDM the men  
‘... he thinks that Trobriand women... are... and just as free as men.’

(52) Italian (itTenTen16)

... *oli vegetali... possono diventare altrettanto importanti*  
... oil vegetarian.PL... can become [just as much].DEGM important  
*quanto il petrolio e il carbone.*  
[how much].STDM the petrol and the coal  
‘... vegetable oils... can become just as important as petrol and coal.’

---

<sup>8</sup> When the parameter is an AP, Italian can also use *altrettanto* without a degree marker, as in *È bella ma altrettanto noiosa*. ‘She is beautiful but equally boring.’

## (53) Italian (itTenTen16)

Se l' essere umano potesse essere **altrettanto** saggio  
 if the being human could be [just as much].DEGM wise  
**come** un uccellino!  
 like.STDM a [little bird]!  
 'If the human being could be just as wise as a little bird!'

## 4.3. Equative and similative standard markers in Ladin

While the Italian standard marker *che* is used exclusively in correlation with *altrettanto*, the standard markers *che* and *co* are widely used in combination with the degree marker *tan* in Badiot, Ladin de mesa val, and Marou, as illustrated in (54-56). While *ch(e)* /kə/ is typical of Badiot, *co* is characteristic of Ladin de mesa val and Marou.<sup>9</sup>

## (54) Badiot (TALL)

Les ëres... é **tan** bunes **che** i ëi.  
 the women... are [as much].DEGM good that.STDM the men  
 'Women... are as good as men.'

## (55) Lmv (TALL)

Chël gran tachin... chël **che** é **tan** gran **co** iö?  
 that big turkey... that.DEM that.COMP is [as much].DEGM big how.STDM I?  
 'That big turkey... the one that is as big as I (am)?'

## (56) Marou (TALL)

Scë düc foss **tan** scichês **co** os...  
 if all were [as much].DEGM clever how.STDM you.PL  
 'If everybody was as clever as you...'

<sup>9</sup> The modern-spelling versions of the corpus CLL are adopted in this article for ease of comprehension. If the modern transcription of the markers does not correspond to the original form, this is pointed out in order to give a more precise picture. Transcribers of older texts often opt for the nowadays written standard form *co*, which, however, does not necessarily correspond to the original form, e.g. *qu*' in Badiot.

Badiot *che* fulfils several functions in different contexts, in some of which it corresponds to Italian *che* while in others it differs. Badiot *che* occurs in comparatives of nonequivalence independently of the syntactic context, as illustrated in (57) and (58), in which *che* is followed by a noun phrase and by a personal pronoun. In these contexts, standard modern Italian would prefer the preposition *di*, whereas *che* would be less appropriate (Dardano & Trifone, 1985; Serianni, 1989). However, due to substratum influence, the comparative marker of inequivalence *che* is overextended in northern Italian dialects, e.g. in Piedmontese (Cerruti 2009).<sup>10</sup>

(57) Badiot (TALL)

*Al se vëgn in mënt che l' scür d-la nõt sides plö*  
it us.DAT comes in mind that the darkness of-the night be more  
*sterch che la lüm d-l dé nü.*  
strong that.STDM the light of-the day new  
'It occurs to us that the darkness of the night is stronger than the light of the  
new day.'

(58) Badiot (TALL)

*Ai ne s' à nia spordü de manco che*  
they NEG.PTCL themselves.REFL have not frightened of less that.STDM  
*nos.*  
we = us  
'They were not less frightened than us.'

Both Ladin and Italian *che* can function as a complementiser in relative clauses and in subordinate clauses in general, as the Ladin sentences (59-61) illustrate.

(59) Badiot

*I á lit l' test che la studënta á scrit.*  
I have read the text that the student has written  
'I have read the text that the student has written.'

---

<sup>10</sup> Nowadays, certain uses of *de* in comparatives of nonequivalence in Badiot are arguably an influence of Italian, independently of the fact that Latin *de*, *quam*, and *quod* were long-time competitors in the older stages of various Italo-Romance varieties, as argued by Stolz (2013).



(60) Badiot

*I á baié cun la studënta **che** vëgn da-l' America.*

I have spoken with the student that comes from-the America.

'I have spoken to the student who comes from America.'

(61) Badiot

*Ël pënsa **che** ël sais döt.*

he thinks that he knows.SBJV everything

'He thinks that he knows everything.'

Unlike Italian *che*, the Badiot standard marker *che* /ke/ corresponds to the human interrogative pronoun, whereas Italian uses *chi* in interrogatives with a human reference, as in (62) and (63).

(62) Badiot

***Che** as-te pa odü?*

who have-you.SG PTCL seen

'Who have you seen?'

(63) Italian

***Chi** hai visto?*

who have.2SG seen

'Who have you seen?'

The complementiser *che* follows *wh*-elements in subordinate interrogatives in Badiot, as in some other dialects that are spoken in northern Italy, e.g. in central Trentino (Benincà & Poletto 1997; Casalicchio & Cordin 2020). This is exemplified in (64) and (65).

(64) Badiot

*Ël me á cunté **cí** **che** ël á odü.*

he me.DAT has told what that.COMP he has seen

'He has told me what he has seen.'

(65) Badiot

*Ĕl me á cunté che che ěl á odü.*  
he me.DAT has told who that.COMP he has seen  
'He has told me who he has seen.'

Marou *co* functions as a standard marker in equative and simulative structures, but it is also used as a relative pronoun, independently of whether the preceding noun phrase is a human or non-human entity, as illustrated in (66) and (67). However, *co* is used in nominative contexts, whereas *che* is employed in non-nominative contexts. This case distinction in Marou, which is not present in Badiot, is exemplified in (66)-68).<sup>11</sup>

(66) Marou (CLL, 1969)

*Le sorëdl co lomina nosta tera, é sö alalt a-l ci.*  
the sun that.NOM illuminates our earth, is up high at-the sky  
'The sun that illuminates our earth is high up in the sky.'

(67) Marou (CLL, 1967)

*... n pere co á laoré por te, co é sën belo*  
... a father that.NOM has worked for you, that.NOM is now already

---

<sup>11</sup> While no nominative vs. non-nominative distinction is marked by the relative pronoun *che* in Badiot and in Italian, French has retained a case distinction like Marou. The examples (i) and (ii) illustrate that *qui* is used in a nominative context and *que* is non-nominative.

(i) French (frTenTen17)

*... comme le soleil qui éclaire notre terre.*  
... like the sun that.NOM illuminates our earth  
'... like the sun that illuminates our earth.'

(ii) French (frTenTen17)

*L' air que je respire, l' eau que je bois, le soleil que je vois,*  
the air that.ACC I breath, the water that.ACC I drink, the sun that.ACC I see,  
*le sang qui coule dans mes veines...*  
the blood that.NOM flows in my veines...  
'The air that I breath, the water that I drink, the sun that I see, the blood that flows in my veins...'

*vedl...*

old...

‘... a father who has worked for you, who is now already old...’

(68) Marou (TALL)

*La ega de vita che iu bëri mo pai-i enstès!*

the water of life that.ACC I drink me.DAT pay-I myself

‘The spirit that I drink, I pay myself!’

Like English *how* and Italian *come*, Marou *co* is also used in manner interrogatives and subordinate clauses, as illustrated in (69) and (70).

(69) Marou (CLL, 2003)

*Co podun-se pa nos sëi cares co é tões intenziuns?*

how can-we PTCL we know which that are your intentions

‘How can we know which ones (that) are your intentions?’

(70) Marou (google)

*Le Mareo ne sa nia der co reagì a-la situaziun.*

the Mareo NEG.PTCL knows not very how react to-the situation

‘Mareo team does not really know how to react to the situation.’

The degree marker *tan* is frequently used in combination with the standard marker *sciöche* in Badiot and in Ladin de mesa val, as shown in (71) and (72), which also illustrate that the coordinated construction ‘*tan* adjective *sciöche*’ can be followed by a nominal standard with specific or generic reference. While the standard of comparison in (71) is constituted by the walls of a specific room, (72) refers to the typical smoothness of oil.

(71) Ladin (TALL)

*Al ê tan scür sciöche i mürs de chë ciamera.*

it was [as much].DEGM dark like.STDM the walls of that room

‘It was as dark as the walls of that room.’

(72) Ladin (CLL, 1987)

... mer... ch' ê datrai **tan** chât **sciöche** n öre,  
... sea... that was sometimes [as much].DEGM still like.STDM an oil,  
zënza ones intravaiades, zënza faldes.

without waves twisted, without creases

'... the sea... which was sometimes as still as oil, without twisted waves, without creases.'

*Sciöche* also occurs without *tan*, unlike the markers *che* and *co* (cf. §3). The standard of comparison can again be specific or generic. Two specific people constitute the standard of comparison in (73), while the similitive marker *sciöche* introduces a generic standard in (74) and (75), which involve the static verb 'be' and the dynamic verb 'tremble'.

(73) Ladin (CLL, 1952)

Chësc fô vistí apresciapüch **sciöch'** i atri dui sü  
this was dressed approximately like.STDM the other two his  
*fredesc.*

brothers

'This one was dressed approximately like the other two of his brothers.'

(74) Ladin (CLL, 1858)

L' Orco é gran **sciöche**<sup>12</sup> na munt!  
the ogre is big like.STDM a mountain

'The ogre is big like a mountain!'

(75) Ladin (CLL, 1964)

Al tremorâ **sciöche**<sup>13</sup> na fëia.  
he trembled like.STDM a leaf

'He trembled like a leaf.'

A further marker that is occasionally found in Badiot, Ladin de mesa val, and Marou is *coche* (composed of *co* + *che*), which is widespread in the nearby Ladin valley of

---

<sup>12</sup> Originally *sciöch'*.

<sup>13</sup> Originally *sciöch'*.

Gherdëina, but not used systematically in comparative structures in Val Badia, where it often seems to be an individual idiosyncrasy. However, Gasser (2000) equates *coche* with *sciöche* in his grammar book, where he explains that the standard of comparison can be preceded by *sciöche* or *coche* and exemplifies the latter with (76), where *coche* is followed by an adverbial phrase, and with (77), where *coche* functions as a conjunction and introduces a clause.<sup>14</sup>

(76) Ladin (Gasser 2000: 43)

*Sö-l lüch de Somavila laôr-i cíamò coche plü*  
 on-the farm of Somavila work-they still [how that].STDM more  
*dadî.*  
 formerly  
 ‘On the farm of Somavila, they still work like years ago.’

(77) Ladin (Gasser 2000: 194)

*Coche i scolars d-la cuarta tlassa cíanta, insciö cíanta bëgn*  
 [how that] the pupils of-the fourth class sing, so sing PTCL  
*inće chi d-la terza.*  
 also those of-the third  
 ‘Just as the fourth-grade pupils sing, so do the third-graders.’

Similarly, *coche* occurs alongside *sciöche* in Gallmann et al. (2013) in the example reproduced in (78).

(78) Ladin (Gallmann et al. 2013: 76)

*Cun i ciavëis lunć cíara Rita fora sciöche / coche na stria.*  
 with the hair long look Rita out like / [how that].STDM a witch  
 ‘With long hair, Rita looks like a witch.’

The form *coche* appears in the bilingual dictionary of Ladin de mesa val and Italian by Moling et al. (2016). There is no entry for *coche* in Mischí’s (2015) bilingual dictionary of Ladin de mesa val and German, but the book includes the terms *cochemai*

<sup>14</sup> It might be hypothesised that *coche* used to be widespread and employed systematically in Val Badia, where it was ousted by the co-existing *sciöche* later in history, but there is not enough evidence for this hypothesis to be proved at the present moment.

(composed of *co* + *che* + *mai*) and *cochessî* (composed of *co* + *che* + *sî*), which can both have an adjectival and adverbial function. *Cochemai* and *cochessî* convey a meaning along the lines of ‘sui generis, in whatever way’, as (79) and (80) illustrate.

(79) Ladin (CLL, 1949)

*Dlunch á-i odü iö jënt dër coche maí*<sup>15</sup>.  
everywhere have-I seen I people very [how that] never  
‘Everywhere have I seen people very much sui generis.’

(80) Ladin (CLL, 1959)

*La jënt... sona y cianta...ladin todësch talian cochessî*<sup>16</sup>.  
the people... play and sing... Ladin German Italian [how that be]  
‘People... play and sing... Ladin, German, Italian, in whatever way.’

While Badiot and Ladin de mesa val use *sciöche*, Marou has the two standard markers *desco* and *desche*.<sup>17</sup> Like *sciöche*, the standard markers *desco* and *desche* can be used with or without the degree marker *tan* and with a specific or generic standard, as shown in (81-84), and no difference between equatives and similatives could be detected.

(81) Marou (CLL, 1968)

*... checio lüront y tan sterch desco fü*.  
... red glowing and [as much].DEGM strong like.STDM fire  
‘... in a glowing red and as strong as fire.’

(82) Marou (CLL, 1930)

*I ne sun mino surt no desco os dui, ch’ i sëis*  
I NEG.PTCL am NEG.PTCL deaf not like.STDM you.PL two, that you.PL are

---

<sup>15</sup> Originally *coch’ mai*.

<sup>16</sup> Originally *co ch’si*.

<sup>17</sup> Ladin *sciöch*’ is only attested twice in the Marou texts included in the corpus of literary texts and both instances are from the same author. Additionally, three instances of *sciöche* are found in the transcribed versions, which are not totally faithful to the original in this case. The original forms *sciöco*, *söcco*, and *shö*, *ch’* should arguably be interpreted as *ensciö co*, which is also attested once.

*surc desco n tapo.*

deaf like.STDM a log

‘I am not deaf like you two, who are deaf like a log.’

(83) Marou (TALL)

*Y empò è-l sté atlò tan bel desche nos ne*

and still is-it been here [as much].DEGM nice as.STDM we NEG.PTCL

*l' ân te nosta vita mai ciamó albü.*

it had in our life never yet had

‘And still, it has been as nice here as it never was in our life before.’

(84) Marou (TALL)

*Co-le tomp s' aüs-on a chël desche le cian*

with-the time one.REFL get-used-one to that as.STDM the dog

*s' aüsa co-i pöresc.*

itself.REFL gets-used with-the lice

‘With time one gets used to that as the dog gets used to lice.’

Characteristic uses of the two markers *desco* and *desche* are discussed in §5 by drawing comparisons with English *as* and *like*, after addressing the case marking of pronominal standards in §4.4.

#### 4.4. Pronominal standard markers and case

Haspelmath & Buchholz (1998: 306) observed that “in Standard Average European, the standard marker is usually derived from a subordinator and therefore does not govern (or ‘assign’) the case of the standard, which is identical to that of the comparee.” It is often maintained that the English complement *as* has a nominative in formal style and an accusative in informal style if a verb can be added to which the pronoun is the subject. The traditional view suggests using the nominative pronominal form in this case, arguing that the pronoun is the subject of an elliptical clause. Biber et al. (1999) found that nominative and accusative forms were divided fairly equally in fiction, whereas only accusative forms were attested in their corpus of conversation. Haspelmath & Buchholz (1998: 308-309) confirmed that “in SAE languages there is a clear tendency for case-transparent standard markers to be turned into case-determining prepositions”, arguing that this tendency reflects “the gradual

loss of the connection between the ‘underlying’ relative clause and the phrasal standard and can be understood as a type of grammaticalization.”

Hence, while the standard and the comparee are often in the same case in European languages, it is also not uncommon for the standard and comparee to be in a different case, as in Italian, where the direct object pronoun is used in phrasal constructions independently of the case of the comparee. Vice versa, the standard marker does not assign the case of the pronominal standard in Ladin or Marou, as illustrated in (85-90), where the nominative forms *iö*, *iu* ‘I’ and *tö* ‘you’ are used.

(85) Badiot (TALL)

*Degügn n' é tan furbi che iö.*  
nobody NEG.PTCL is [as much].DEGM clever.PL that.STDM I.NOM  
‘Nobody is as clever as I.’

(86) Badiot (TALL)

*Iö sun pa bel tan sciché che tö.*  
I am PTCL already [as much].DEGM smart that.STDM you.NOM.SG  
‘I am already as smart as you.’

(87) Lmv (Mischí 2015: 389)

*Ëra é avisa tan eleganta co tö.*  
she is just [as much].DEGM elegant how.STDM you.NOM.SG  
‘She is just as elegant as you.’

(88) Badiot (CLL, 1878)

*... porvé de me mantigní en pesc y contènta sciöche*  
... try to me maintain in peace and satisfied like.STDM  
*tö.*<sup>18</sup>  
you.NOM.SG  
‘... try to keep myself peaceful and satisfied like you.’

(89) Marou (TALL)

*... su desco iu.*  
... lonely like.STDM I.NOM  
‘... lonely like me.’

---

<sup>18</sup> Originally *in pesc... sceoucche tou*



(90) Marou (TALL)

*Ciar-i ince iu fora desco tö?*  
 look-I also I out like.STDM you.NOM.SG  
 ‘Do I also look like you?’

In (91), the pronominal standard takes the nominative form *iu* ‘I’ even with a direct-object comparee.

(91) Marou (TALL)

*Veronica adora pö en ël madü desco... desco iu por*  
 Veronica needs PTCL a man mature like.STDM...like.STDM I.NOM for  
*tó en ejempio.*  
 take an example  
 ‘Veronica needs a mature man like... like me, to give an example.’

Despite the typical use of nominative pronominal forms in Ladin comparative constructions, the objective pronominal forms are not excluded nowadays, and this might be argued to be an influence of Italian, which exerts a considerable influence on younger speakers in particular. Cross-linguistic influence might explain the use of the accusative form *me* ‘me’ in (92), while a stylistic choice might have been made in (93), where two instances of the nominative form *tö* ‘you’ are found at first and then the accusative form *te* ‘you’, immediately after the prepositional phrase *por te* ‘for you’.

(92) Badiot (google)

*E-le... ladins sciöche me? Iö sun badiot.*  
 is-it... Ladins like.STDM me? I am Badiot.  
 ‘Are there Ladins like me? I am Badiot.’

(93) Marou (google)

*Chel co vir desco tö, vir l’ amur. Chel*  
 that.DEM that.COMP lives like.STDM you.NOM.SG lives the love. that.DEM  
*co ponsa y fej desco tö y dá la vita...*  
 that.COMP thinks and does like.STDM you.NOM.SG and gives the life..  
*Fá ch’ i viri por te y desco te.*  
 make that I live for you. ACC.SG and like.STDM you.ACC.SG

‘The one who lives like you, lives love. The one who thinks and behaves like you and gives his or her life... Let me live for you and like you.’

### 5. Marou *desco* and *desche* vs. English *like* and *as*

The preceding sections have shown that standard markers of comparison differ in the varieties that are spoken in Val Badia. The difference concerns in particular the use of the two standard markers *desco* and *desche* in Marou where Badiot and Ladin de mesa val use the same marker *sciöche* (and Italian *come*). A search of the corpus of literary texts for *desco*, *desche*, and *desch*’ in the subsection of the Marou variety returned the numbers reported in Table 5.<sup>19</sup>

Marou standard markers	Numbers of occurrence	% of the total number of words
<i>desco</i>	159	0.089%
<i>desch</i>	48	0.027%
<i>desche</i>	23	0.013%

Table 5: Marou *desco*, *desch*, *desche* in the literary corpus CLL (search carried out on 2020.11.11).

A precise explanation of the distribution of *desco* and *desche* was missing in dictionaries and grammar books, when Irsara (2001) undertook an investigation into their contexts of use, which was followed up by Irsara (2012). The existence of the two markers was acknowledged by speakers of Marou, who provided resolute judgements concerning the acceptability of *desco* and *desch(e)* in certain contexts, while being unaware of their systematic distribution. A corpus analysis of the right context of the markers immediately confuted the tentative hypothesis that the use of

<sup>19</sup> The Romansh varieties spoken in Switzerland have the forms *sco* and *scu*. A CLL search of the literary texts in their original spelling returned one instance of *skö*, as (i) illustrates.

(i) Ladin (CLL, 1910)  
 ... *ne*            *türa*    *plü*            *salč*        *skö*            *n* *sajok*.  
 ... NEG.PTCL    throws    any-more    jumps    like.STDM    a    grosshopper  
 ‘... does not jump any more like a grasshopper.’

*desco* and *desche* might be dictated by the initial sound of the following expression.<sup>20</sup> Both *desco* and *desche* are followed by vowel and consonant sounds, while the abbreviated form *desch'* is followed by a vowel in all its 23 occurrences.

A comparison of Marou *desco* and *desche* with English *like* and *as* called for a syntactic analysis. Despite the rather limited application of this rule, it was the tradition of prescriptive opposition between *like* and *as* in English that led to an investigation into the syntactic contexts of use of *desco* and *desche* in Marou, which produced the following result: *desch(e)* introduces a clause with an inflected verb, whereas *desco* normally precedes a phrase, which can be postmodified by a relative clause. *Desche* and *desch'* are freely substitutable in front of a vowel, while *desco* is not abbreviated. Table 6 shows the results that were obtained from an analysis of the right contexts of the forms *desco*, *desch'*, and *desche* in the literary texts contained in the Marou subsection of the corpus CLL. In the great majority of cases, *desco* is followed by a phrase, while *desch(e)* is followed by clause, in accordance with the generalisation made above.

Particles	+ finite clause	+ phrase
<i>desco</i>	15 <sup>a</sup>	144 <sup>b</sup>
<i>desch'</i>	48	/
<i>desche</i>	22	1

<sup>a</sup> Of these 15 exceptional instances of *desco* + clause, 9 are from the same author.

<sup>b</sup> One instance is arguably not relevant because *desco* is part of a title with no right context.

**Table 6:** Right context of Marou *desco* and *desch(e)* in the literary corpus CLL  
(search carried out on 2020.11.11)

A syntactic analysis of *desco* and *desch(e)* in the ten latest information bulletins of the municipality of Mareo yielded the results reported in Table 7, which adds validity to the findings in Table 6, in spite of the different registers of the texts considered in the two tables.

<sup>20</sup> This idea was initially proposed by Videsott & Plangg (1998) in their Marou dictionary, where the entry reproduced in (i) is found.

- (i) Marou (Videsott & Plangg 1998: 121)  
*Desco* adv., vor Vokal *desche*  
 ‘*Desco* adverb, in front of a vowel *desche*’

Particles	+ finite clause	+ phrase
<i>desco</i>	3	285
<i>desch'</i>	14	/
<i>desche</i>	26	/

**Table 7:** Right context of Marou *desco* and *desch(e)* in *Le Saltà* (numbers 21-30).

It will be illustrated in the next sections that Marou *desco* and *desch(e)* are used in various constructions that are somewhat related to the more straightforward clauses of equality and similarity, which will also explain the more extensive use of *desco* in *Le Saltà* than in the corpus CLL. While in CLL *desco* makes up 69% of the total number of occurrences of the three forms, in *Le Saltà* 88%.<sup>21</sup>

### 5.1. Equative and similitive clauses in English and Marou

<sup>21</sup> A detailed analysis of the deictic or anaphoric demonstratives *ensciö* (Marou), *così* (Italian), and *so* (English) lies beyond the scope of this article. However, their use as affirmative particles is worth pointing out. The examples (i-ii) illustrate the use of *ensciö* (Marou) and *così* (Italian) in the Lord's Prayer.

(i) Italian (itTenTen16)

*Sia fatta la tua volontà come in Cielo così in terra.*  
 be done the your will like in sky so in earth  
 'Your will be done on earth as it is in heaven.'

(ii) Marou

*Töa orenté sides fata, desco a-l ci ensciö sö-la tera.*  
 your will be done, like at-the sky so on-the earth  
 'Your will be done on earth as it is in heaven.'

König (2017) emphasizes that several languages have affirmative particles that are based on these anaphoric manner deictics, including English (*yeah swa* > *yes*) and Italian (*ëccu(m) sìc* > *così, sì*). This also holds true for Marou, where the affirmative particle *sciö* is used in a number of contexts, as illustrated in (iii).

(iii) Marou (TALL)

*Gustl ciügna de sciö.*  
 Gustl nods of yes  
 'Gustl nods.'

Equatives and similatives can have phrasal and clausal standards of comparison. Haspelmath & Buchholz (1998: 303) argued that equative clauses draw more complex comparisons and defined them as “comparisons where the verb is part of the standard, i.e., where the two situations compared differ not only in their participants, but also in their verbal core.” In (94), the compared situations differ in their participants, in (95) in their verbal core, and in (96) in both their participants and verbal core, but the standard marker *as* is followed by a clause with an inflected verb in all of them, so that *desch(e)* would be used in Marou.

(94) English (enTenTen15)

*You can walk as well as **they** can.*

(95) English (enTenTen15)

*He **moved** as beautifully as he **looked**.*

(96) English (enTenTen15)

***She respected** her father's decision as much as **he cared** for her happiness.*

While the correlative construction with the double *as* marker is used in (94-96) above, *like* is used in (97), which was given by Haspelmath & Buchholz (1998) as an example of a similative clause or manner adjunct, which specifies in what way something is done.

(97) English (Haspelmath & Buchholz 1998: 280)

*She writes **like** her brother talks.*

Haspelmath & Buchholz (1998: 320) noticed that SAE languages or “languages whose equative and similative standard marker is based on the relative pronoun ‘how’, similative clauses are generally marked just by this word or by the same marker as non-clausal similative phrases.” The relative pronoun ‘how’ in Marou is *co*, which can occur in the form *coche* in similative clauses, as in (98), although the latter type of clauses is generally uncommon, as pointed out by Haspelmath & Buchholz (1998).

(98) Marou (CLL, 1952)

*Al é t-la natöra, al é vëi, al é dërt, che önn mëss morí,  
it is in-the nature, it is true, it is right, that one must die,*

*coche*                    *n popul*                    *se*                    *perd.*

[how that].STDM    a population    itself.REFL loses

‘It is in the nature of things, it is true, it is right, that one must die, like a population gets lost.’

However, the generalization made by Haspelmath & Buchholz (1998) that similitive clauses normally use the same marker that is also employed in non-clausal similitive phrases does not hold true for Marou, which distinguishes between phrasal *desco* and clausal *desch(e)*, as illustrated in (99) and (100).

(99) Marou (TALL)

... *spo bradlâ-l*                    *desco*                    *en*                    *picio môt.*

... then    cried-he    like.STDM    a                    little boy

‘... then he cried like a little boy.’

(100) Marou (TALL)

*Al*                    *se*                    *lasciâ jí,*                    *desche*                    *en*                    *picio môt*                    *se*                    *lascia*

he    himself.REFL    let                    go,    as.STDM                    a                    little boy    himself.REFL    lets

*jí dô*                    *en*                    *toc*                    *d-la*                    *oma.*

go    after    a                    smack    of-the    mother

‘He let himself go, as a little boy lets himself go after a smack from his mother.’

While speakers do not often have occasion to compare manners or ways of doing something using similitive clauses, the formally and semantically similar accord and simulative clauses are much more common.

## 5.2. Accord clauses, simulative clauses, and role phrases in English and Marou

Accord or illocutionary clauses are typically expressed in English by adverbial finite clauses introduced by *as*, which are normally realised in a distinct tone unit, and can occur in different positions within a sentence. (101) exemplifies the use of an accord clause or parenthetical disjunct in a central position.

(101) English (BNC)

*The results are, as you might have guessed, violently unpredictable.*

Accord finite clauses are introduced by *desch(e)* in Marou. The accord clause appears as a peripheral comment at the end of the sentence in (102), whereas it occurs medially in (103) and initially in (104), with word order implications in the latter case because of the verb-second characteristics of Marou.

(102) Marou (TALL)

*Sparagné n' é-l nia gnü..., desche diüc sa.*  
 save NEG.PTCL is-it not come.., as.STDM all know  
 'Money has not been saved..., as everybody knows.'

(103) Marou (*Le Saltà 23*)

*Na banca da-l tomp n' á nia da fá con scioldi,*  
 a bank of-the time NEG.PTCL has nothing to do with money,  
*mo ara á da fá – desche l' ennom dij – con tomp.*  
 but it has to do – as.STDM the name says – with time.  
 'A time bank has nothing to do with money, but it has something to do, as the name suggests, with time.'

(104) Marou (*Le Saltà 25*)

*Desche trec se recordará, é sides le 2015 co le 2016*  
 as.STDM many REFL remember.FUT is both the 2015 how the 2016  
*stés agn dër megri de nëi.*  
 been years very thin of snow  
 'As many will remember, both 2015 and 2016 were years with very little snow.'

While English *as* also introduces elliptical clauses of the type of (105), elliptical clauses are introduced by *desco* in Marou, as illustrated in (106).

(105) English (BNC)

*As mentioned above, there are openings in both the private and public sectors.*

(106) Marou (*Le Saltà 22*)

*Desco belo anunzié dessora, ...*  
 like.STDM already announced above, ...  
 'As already announced above, ...'

Finally, English *as* is frequently found in combination with *if* and *though* in simulative hypothetical constructions of the type of (107-110), where *like* can be used in American English and in informal style, as illustrated in (111) and (112).

(107) English (enTenTen15)

*He leapt and jumped and spun around **as if** he was crazy.*

(108) English (enTenTen15)

*Jan always drove **as though** he was crazy.*

(109) English (enTenTen15)

*But Minka rolled on her back **as if** crazy.*

(110) English (enTenTen15)

*Saying ‘I have seen! I have seen!’ and dancing in ecstatic love **as though** mad.*

(111) English (enTenTen15)

*... small shrieking creatures that were dancing around **like** they were crazy.*

(112) English (enTenTen15)

*He was drinking vanilla and dancing **like** crazy.*

Both *desch(e)* and *desco* are used in simulative constructions, in accordance with their clausal and phrasal contexts of use. While *desch(e)* introduces hypothetical clauses, *desco* is used with elliptical or verbless constructions, as illustrated in (113) and (114).

(113) Marou (TALL)

*Matio... vën ete, al fej **desch’** al jorass.*

Matio... comes in, he makes as he flied.SBJV

‘Matio... enters, he acts as if he was flying.’

(114) Marou (TALL)

*Le Mazot ea post ia **desco** mort.*

the Mazot was laid thither like dead

‘Mazot lay there as if dead.’



Like accord and simulative clauses, role phrases are expressed in a manner similar to equatives and similatives in several languages, despite being semantically different, since they “express the role or function in which a participant appears” (Haspelmath & Buchholz 1998: 321). Haspelmath & Buchholz (1998: 322) found that “in most European languages, role phrases are marked by the same particle that is used as standard marker in equality and similarity constructions.” While the role marker is *as* in English, Marou uses *desco* to refer to someone’s role or job, which is preceded by no article. This function of *desco* partly explains its high number of occurrences in the municipal information bulletin *Le Saltà*, which publishes reports on people’s achievements and their roles, as in (115), and the results of public job applications, as in (116).

(115) Marou (*Le Saltà* 21)

... protagonist d-la scora, empröma **desco** maester, spo  
 ... protagonist of-the school, at first like teacher, then  
**desco** diretur.

like director

‘... a school protagonist, at first as a teacher, then as a director.’

(116) Marou (*Le Saltà* 23)

La sign.ra... é gnüda tuta sö... **desco** cöga por la scolina.  
 the Ms... is been taken up... like cook for the kindergarten  
 ‘Ms... has been hired... as a kindergarten cook.’

While resemblance to a mother is expressed by the disjunct in (117), the referent’s actual role is indicated in (118).

(117) English (enTenTen15)

**Like** a mother, she has food security close to her heart.

(118) English (enTenTen15)

**As** a mother, she wants the best for her kids.

The distinction resemblance vs. role is marked in English by the different particles *like* and *as*, whereas in Marou the difference lies in the presence or absence of the indefinite article, as illustrated in (119) and (120).<sup>22</sup>

(119) Marou (google)

*Chel-bel-dio é por nos desco n pere y desco na oma.*  
that-nice-god is for us like a father and like a mother  
'God is for us like a father and like a mother.'

(120) Marou (TALL)

*Töa fomena n' è da nia! Y desco oma ès-era ciamó*  
your wife NEG.PTCL is of nothing! And like mother is-she even  
*da manco.*  
of less  
'Your wife is worth nothing! And as a mother even less.'

Haspelmath & Buchholz (1998) maintain that role markers are the oldest and most grammaticalised markers, such as English *as*, which is older than the similative marker *like*.<sup>23</sup> Similarly, it seems reasonable to hypothesise that *desche* is innovative in Marou, while the role marker *desco* is more grammaticalised and older. An even older form appears to be *sco*, which was arguably widespread in earlier times, and is found in a Ladin version of the Lord's Prayer that appears in Bacher's early grammar book dating back to 1833. Craffonara (1995) argues that *sco* must have sounded archaic in most of Val Badia at that time already, because extant texts from the previous century usually had *[(ij)šökə]*, with different spelling patterns, in certain varieties of Val Badia, and *[dèško]* in Marou.<sup>24</sup>

---

<sup>22</sup> The same holds true for Italian, where the distinction resemblance vs. role is also marked by the presence and absence of the indefinite article (*come una madre* 'like a mother' vs. *come madre* 'as a mother').

<sup>23</sup> Haspelmath & Buchholz (1998) emphasise that this is also the case with German, which uses the older marker *als* as a role marker and not the more innovative standard marker of equality and similarity *wie*.

<sup>24</sup> *Desche* is also used in the Ladin variety spoken in the Dolomitic valley of Fascia, where it is the only extant form, whereas *desco* does not occur.

## 6. Conclusion

This article has investigated a number of marking patterns of equative and similative constructions in English, Italian, and Ladin as spoken in the South-Tyrolean valley of Badia. The analysis started from the extensively investigated English language and moved on to the well-documented Italian language, before verifying observations made for these major languages in Ladin varieties spoken in different areas of Val Badia. The article has presented a descriptive account of selected characteristics, adopting a cross-linguistic point of view, and relating the identified linguistic features to typological generalisations made in the literature on the topic under discussion. The article aimed to set the analysis of the Ladin data in particular in a wider context of qualitative research by drawing parallels with major languages and finding differences between them. The analysis was intended to make a small-scale contribution to the ongoing debate on equatives and similatives in European languages. In particular, the focus of the investigation was on the use of degree and standard markers in the more straightforward equative and similative constructions and on the use of these same markers in various related constructions, such as accord clauses, simulative clauses, and role phrases, which were analysed in particular in the second part of the article, where the Marou variety was focused upon.

In line with other European languages, English, Italian, and Ladin all use analytic degree or parameter markers, which never occur without an accompanying standard marker. Both the Italian and Ladin degree markers *tanto* and *tan* were confirmed to be demonstrative based and to have retained a deictic value in certain contexts, unlike the English degree markers *as* and *so*, which were originally also demonstrative based, but which have broadly lost their exophoric deictic value. Unlike Italian *tanto*, Ladin *tan* was shown to occur as an interrogative, thus corresponding to Italian *quanto* and English *how much*. The use of Ladin *tan* in equative-pro-COI constructions was subsequently illustrated, demonstrating that *tan* can be preceded by multiplicative numerals like English *as*, which can be preceded by factor modifiers like *twice*. The position of the Italian and Ladin markers *tanto* and *tan* was shown to be less obvious than the position of the English degree marker *as*, which is typically found in pre-adjectival position. While Ladin *tan* usually precedes the adjectival parameter in specific equatives, it is often found in post-adjectival position in generic equatives. While generic equatives sometimes lack the degree marker *as* in English, leaving the

second *as* on its own, Ladin *tan* can be omitted with certain standard markers but not with the standard marker *che / co*.

After examining the degree markers in English, Italian, and Ladin, the article went on to explore equative and similitive standard markers, which precede the standard of comparison. While the English standard markers *as* and *like* are normally taken to describe quantity and quality respectively, quantitative and qualitative standard markers often occur in similar contexts. While Italian *quanto* and *come* are equally acceptable in equative constructions if unaccompanied by a degree marker, it is *come* that is typically used to express similarity of manner in the contexts ‘to be like’ or ‘to do like (this)’. Italian uses the general subordinator *che* as a standard marker in coordination with the degree marker *altrettanto*, whereas Ladin Badiot *che* is used with *tan*, which can also be accompanied by similitive markers. The Ladin complementiser *che* has no animacy restrictions, whereas the interrogative pronoun *che* refers to human referents and corresponds to Italian *chi* and English *who*. Unlike Badiot, Marou employs the *how*-word *co* as an equative standard marker. Marou *co* also occurs as a nominative relative pronoun, while *che* is the form that is used in a non-nominative context. Case-marking issues were also addressed in relation to pronominal standards in Badiot and Marou, where the standard markers do normally not govern the case of the standard of comparison, although a tendency to use pronominal standards in their oblique case forms could also be detected and might be interpreted as cross-linguistic influence from Italian, in spite of this tendency being generally found by Haspelmath & Buchholz (1998) in languages across Europe.

Besides the different equative standard markers *che* and *co* in Badiot and Marou, a notable difference between these two varieties was detected in the use of the two standard markers *desco* and *desch(e)* in Marou where Badiot uses *sciöch(e)*. After considering the difference between clausal *as* and phrasal *like* as traditionally prescribed by norms of formal English, it was found that Marou *desch(e)* introduces a clause with an explicitly given finite verb, while *desco* is typically followed by a phrase, which can be postmodified by a relative clause. *Desch(e)* therefore occurs in equative and similitive clauses, in accord clauses, and in simulative clauses, whereas it is *desco* that is used in role phrases, where English employs its older form *as* and not *like*. It was finally maintained that the role marker *desco* is arguably the older form, in line with Haspelmath & Buchholz’s (1998) typological findings in European languages.

## Acknowledgements

The author wishes to thank Paul Videsott for reading and commenting on this article in its pre-publication format. Gratitude is extended to the anonymous reviewers and editors for their constructive suggestions.

## Abbreviations

2 = 2 <sup>nd</sup> person	F = feminine	PRTV = partitive
3 = 3 <sup>rd</sup> person	FUT = future	PTCL = particle
ACC = accusative	IPFV = imperfect	REFL = reflexive
COMP = complementiser	M = masculine	SBJV = subjunctive mood
DAT = dative	NEG = negative	SG = singular
DEGM = degree marker	NOM = nominative	STDM = standard marker
DEM = demonstrative	PL = plural	

## References

- ASTAT – Provincial institute of statistics. 2019. *Südtirol in Zahlen. Alto Adige in cifre*. ([https://astat.provinz.bz.it/downloads/Siz\\_2019\(2\).pdf](https://astat.provinz.bz.it/downloads/Siz_2019(2).pdf)) (Accessed 2020.09.28.)
- Belletti, Adriana. 1991. Le frasi comparative. In Lorenzo Renzi & Giampaolo Salvi (eds.), *Grande grammatica italiana di consultazione*, vol. 2, 832-853. Bologna: Il Mulino.
- Belletti, Adriana. 2010. La comparazione. In Giampaolo Salvi & Lorenzo Renzi (eds.), *Grammatica dell'italiano antico*, vol. 2, 1135-1143. Bologna: Il Mulino.
- Benincà, Paola & Poletto, Cecilia. 1997. *Quaderni di Lavoro dell'ASIS (Atlante Sintattico Italia Settentrionale): Strutture interrogative dell'Italia settentrionale*. Padua: Consiglio Nazionale delle Ricerche.
- Biber, Douglas & Johansson, Stig & Leech, Geoffrey & Conrad, Susan & Finegan, Edward. 1999. *Grammar of spoken and written English*. Essex: Longman.
- Dardano, Maurizio & Trifone, Pietro. 1985. *La lingua italiana: Morfologia, sintassi, fonologia, formazione delle parole, lessico, nozioni di linguistica e sociolinguistica*. Bologna: Zanichelli.
- Casalicchio, Jan & Cordin, Patrizia. 2020. *Grammar of Central Trentino: A Romance dialect from north-east Italy*. Leiden / Boston: Brill.

- Cerruti, Massimo. 2009. *Strutture dell'italiano regionale: Morfosintassi di una varietà diatopica in prospettiva sociolinguistica*. Frankfurt am Main: Peter Lang.
- Craffonara, Lois. 1995. "Nikolaus Bacher (Micurá De Rü): Versuch Einer Deütsch-ladinischen Schriftsprache, 1833". *Ladinia: Sföi Culturâl dai Ladins dles Dolomites* 19. San Martin de Tor: Istitut Ladin "Micura De Rü".
- Gasser, Tone. 2000. *Gramatica ladina por les scoles*. Balsan: Istitut Pedagogich Ladin.
- Gallmann, Peter & Siller-Runggaldier, Heidi & Sitta, Horst. 2013. *Sprachen im Vergleich: Deutsch-Ladinisch-Italienisch: Der einfache Satz*. Bozen: Innovation und Beratung im Ladinischen Bildungs- und Kulturressort.
- Haspelmath, Martin & Buchholz, Oda. 1998. Equative and similative constructions in the languages of Europe. In Johan van der Auwera (ed.), *Adverbial constructions in the languages of Europe*, 277-334. Berlin: Mouton de Gruyter.
- Haspelmath, Martin. 2017. Equative constructions in world-wide perspective. In Yvonne Treis & Martine Vanhove (eds.), *Similative and equative constructions: A cross-linguistic perspective*, 9-32. Amsterdam / Philadelphia: John Benjamins.
- Henkelmann, Peter. 2006. "Constructions of equative comparison". *Sprachtypologie und Universalienforschung (STUF)* 59(4). 370-398.
- Irsara, Martina. 2001. *The syntax of comparative structures in English, Italian and Ladin*. BA thesis, University of Padua.
- Irsara, Martina. 2012. "Comparative constructions in English, Italian and Ladin: Cross-linguistic observations and didactic considerations". *Quaderni Patavini di Linguistica* 2010 26. 39-82.
- Irsara, Martina. 2013. Awareness across languages: From English *as* and *like* to Ladin. In Doris Kofler & Hans Karl Peterlini & Gerda Videsott (eds.), *Brückenbau(e)r: Festschrift für / Miscellanea per / Publicaziun en onour de / Article collection in honour of Siegfried Baur*, 288-294. Meran: Edizioni alpha beta Verlag.
- König, Ekkehard. 2017. The deictic identification of similarity. In Yvonne Treis & Martine Vanhove (eds.), *Similative and equative constructions: A cross-linguistic perspective*, 143-164. Amsterdam / Philadelphia: John Benjamins.
- Mischí, Giovanni. 2015. *Vocabolar Ladin Val Badia-Deutsch*. (<https://www.dizionar-ladin.it>) (Accessed 2020.10.26).
- Moling, Sara & Frenademez, Ulrike & Ruggeri, Xenia & Valentin, Marlies. 2016. *Dizionar ladin Val Badia – talian: Dizionario italiano – ladino Val Badia*. San Martin de Tor: Istitut Ladin Micurá de Rü.

- Pelo, Adriana. 2012. Le proposizioni comparative. In Maurizio Dardano (ed.), *Sintassi dell'italiano antico: La prosa del Duecento e del Trecento*, 441-465. Rome: Carocci editore.
- Quirk, Randolph & Greenbaum, Sidney & Leech, Geoffrey & Svartvik, Jan. 1985. *A comprehensive grammar of the English language*. London: Longman.
- Rett, Jessica. 2020. Separate but equal: A typology of equative constructions. In Peter Hallman (ed.), *Interactions of Degree and Quantification*, 163-204. Leiden / Boston: Brill.
- Schulze, Wolfgang. 2017. Toward a cognitive typology of like-expressions. In Yvonne Treis & Martine Vanhove (eds.), *Similative and equative constructions: A cross-linguistic perspective*, 33-77. Amsterdam / Philadelphia: John Benjamins.
- Serianni, Luca (1989). *Grammatica italiana: Italiano comune e lingua letteraria*. Torino: UTET.
- Stolz, Thomas. 2013. *Competing comparative constructions in Europe*. Berlin: Akademie Verlag.
- Treis, Yvonne & Vanhove, Martine (eds.). 2017. *Similative and equative constructions: A cross-linguistic perspective*. Amsterdam / Philadelphia: John Benjamins.
- Treis, Yvonne. 2018. "Comparative constructions: An Introduction". *Linguistic Discovery* 16(1). i-xxvi.
- Videsott, Paul & Plangg, Guntram A. 1998. *Ennebergisches Wörterbuch: Vocabolar Mareo*. Innsbruck: Universitätsverlag Wagner.
- Zingarelli, Nicola. 2011. *Lo Zingarelli: Vocabolario Della Lingua Italiana*. 12th edn. Bologna: Zanichelli.

## Corpora

*British National Corpus* (BNC)

[https://app.sketchengine.eu/#dashboard?corpname=preloaded%2Fbnc2\\_tt21](https://app.sketchengine.eu/#dashboard?corpname=preloaded%2Fbnc2_tt21)

*Corpus dl ladin leterar* (CLL)

<http://vll.ladintal.it/applications/textanalysis/search.jsp>

[https://app.sketchengine.eu/#concordance?corpname=preloaded%2Fitwac3.no\\_dups.fulldocs.7&tab=basic&structs=s&operations=%5B%5D](https://app.sketchengine.eu/#concordance?corpname=preloaded%2Fitwac3.no_dups.fulldocs.7&tab=basic&structs=s&operations=%5B%5D)

*Italian Web 2010* (itTenTen10)

<https://app.sketchengine.eu/#dashboard?corpname=preloaded%2Fittenten>

*Italian Web 2016* (itTenTen16)

[https://app.sketchengine.eu/#dashboard?corpname=preloaded%2Fitten16\\_2](https://app.sketchengine.eu/#dashboard?corpname=preloaded%2Fitten16_2)

*Le Saltà*

<https://www.comun.mareo.bz.it/system/web/zeitung.aspx?sprache=7&detailonr=225171055-1042&menuonr=219899036> (Accessed 2020.11.03)

*Tratament Automatic dl Lingaz Ladin: Corpus lad* (TALL)

<http://corpuslad.ladintal.it/applications/textanalysis/search.jsp>

*English Web 2015* (enTenTen15)

[https://app.sketchengine.eu/#dashboard?corpname=preloaded%2Fententen15\\_t31](https://app.sketchengine.eu/#dashboard?corpname=preloaded%2Fententen15_t31)

*French Web 2017* (frTenTen17)

[https://app.sketchengine.eu/#concordance?corpname=preloaded%2Ffrtnten17\\_f12](https://app.sketchengine.eu/#concordance?corpname=preloaded%2Ffrtnten17_f12)

*German Web 2013* (deTenTen13)

[https://app.sketchengine.eu/#dashboard?corpname=preloaded%2Fdetenten13\\_rft3](https://app.sketchengine.eu/#dashboard?corpname=preloaded%2Fdetenten13_rft3)

*ItWAC*

<https://app.sketchengine.eu/#concordance?corpname=preloaded%2Fitwac3.nodups.fulldocs.7&tab=basic&structs=s&operations=%5B%5D>

*Italian Web 2010* (itTenTen10)

<https://app.sketchengine.eu/#dashboard?corpname=preloaded%2Fitten>

*Italian Web 2016* (itTenTen16)

[https://app.sketchengine.eu/#dashboard?corpname=preloaded%2Fitten16\\_2](https://app.sketchengine.eu/#dashboard?corpname=preloaded%2Fitten16_2)

*Le Saltà*

<https://www.comun.mareo.bz.it/system/web/zeitung.aspx?sprache=7&detailonr=225171055-1042&menuonr=219899036> (Accessed 2020.11.03)

*Tratament Automatic dl Lingaz Ladin: Corpus lad* (TALL)

<http://corpuslad.ladintal.it/applications/textanalysis/search.jsp>

## CONTACT

[martina.irsara@unibz.it](mailto:martina.irsara@unibz.it)



**This section previously contained an article that has been retracted by the author. See Editor's notes that follow for further details.**

**2021-10-11 – Editor's Note**

In agreement with the Author, the Editors have temporarily retracted the article “Comparative constructions in Suansu and the languages of northeastern India”. After publication of the article, well-founded concerns were brought to the attention of the Editors regarding the reliability of some of the cited published sources and, consequently, regarding the representation of some of the data presented. The Editors requested the author to correct these data accordingly. An amended version of the article will be published in a subsequent issue of the journal after a new round of revision. The temporarily retracted article must not be quoted.

**2022-04-14 – Editor's Note**

A revised version of this paper has been published in vol. 2, no. 1 (2022). It is available here: <https://doi.org/10.6092/issn.2785-0943/14379>

# The comparative cycle in crosslinguistic perspective

AGNES JÄGER<sup>1</sup>

<sup>1</sup>FRIEDRICH SCHILLER UNIVERSITY JENA/DEPARTMENT OF GERMAN LINGUISTICS

Submitted: 29/11/2020 Revised version: 06/05/2021

Accepted: 17/03/2021 Published: 31/08/2021

## Abstract

This paper traces the diachronic development of comparison constructions crosslinguistically, highlighting a recurrent pattern of change with respect to standard markers: the comparative cycle. Using diachronic corpus data as well as data from descriptive grammars and handbooks, it is demonstrated that comparison particles and other standard markers in many languages undergo a syntactic-semantic distributional shift from marking equality to marking inequality. More specifically, we witness a stepwise and recurrent – i.e. cyclical – shift of standard markers from similatives to equatives and to comparatives. The comparative cycle is compared to other instances of cyclical change and linked to linguistic economy and the markedness hierarchy of comparison constructions.

**Keywords:** comparative cycle; diachronic typology; similative; equative; comparative.

## 1. The comparative cycle: Aims and scope of the paper

Diachronic typological investigations may deepen our understanding of principles and regularities of language change by uncovering systematic, recurrent patterns of change that are not limited to a single language but constitute very basic, potentially universal patterns of language change. One noteworthy case in point is the stepwise and recurrent – i.e. cyclical – shift of linguistic expressions in comparison constructions, notably standard markers (e.g. comparison particles) from comparisons of equality/similarity to those of inequality/dissimilarity, more specifically from similatives to equatives and to comparatives that is aptly referred to as the comparative cycle (Jäger 2010; 2018; a terminology taken up e.g. by Reinartz et al. 2016). While previous literature only very occasionally and in passing mentions an

incidental similarity of two languages in this respect (cf. Zeilfelder 2001 for Sanskrit and German, Heine & Kuteva 2002 for Chinese and German), this paper systematically traces the comparative cycle in the history of about 20 languages and varieties on the basis of corpus data as well as data from descriptive grammars and handbooks in a diachronic typological perspective. Occasional developments in the opposite direction are also discussed. The comparative cycle is compared to other instances of cyclical change (Jespersen's cycle, subject-agreement cycle etc.) and causes for this development are explored.<sup>1</sup>

After giving an overview of the central types of comparison constructions and introducing the markedness hierarchy of comparisons as well as a typology of standard marker systems in section 2, the crosslinguistic diachronic phenomenon captured in the comparative cycle is investigated in various languages in section 3. Section 4 puts the empirical findings in the language-change theoretical context comparing the comparative cycle to other instances of cyclical language change and discussing a number of potential causes for this change, arguing for an explanation based on linguistic economy and markedness.

## 2. Comparison constructions – a taxonomy

Semantically, comparison constructions can roughly be characterised as expressing linguistically the equality/similarity or inequality/dissimilarity of two entities generally referred to as comparee and standard. This (dis-)similarity may or may not relate to degrees of a specific gradable property, the so-called *tertium comparationis* or parameter. The best-researched type of comparisons is the one in which dissimilarity between two entities in relation to degrees of a specific property is expressed, the comparative construction or simply comparative, as illustrated with English in (1)(a). Comparatives have been at the centre of attention of linguistic research because crosslinguistically they show the most grammaticalized and most specific markers. In English and many other languages, it is only in this type of comparison construction that the adjective or adverb that constitutes the parameter (*faster* in (1)(a)) may bear a specific inflectional suffix marking the comparison while there is no equivalent inflectional suffix marking the parameter in other types of comparison.<sup>2</sup> In other

---

<sup>1</sup> This paper builds on and extends parts of ch. 7 of Jäger (2018), making the results available to an English-speaking audience.

<sup>2</sup> Similarly, a language using comparison case may also only show case-marking on the standard in comparatives and not in other comparisons.

words, comparatives are the most marked type of comparisons. This can be expressed by using the two rough semantic features dissimilarity and degree for specifically relating to degrees of a gradable property. Comparatives can be characterized by [+ dissimilarity] and [+ degree].<sup>3</sup>

- |     |  |             |
|-----|--|-------------|
| (1) | a. <i>Anne walks faster than Mary.</i> | comparative |
|     | b. <i>Anne walks as fast as Mary.</i>  | equative    |
|     | c. <i>Anne walks like Mary.</i>        | similative  |

Somewhat less marked are comparisons expressing similarity in relation to degrees of a gradable property, i.e. ones that can be characterized by [- dissimilarity] and [+ degree], as illustrated in (1)(b). I will follow the usual terminology employed in the typological literature here and refer to this type of comparison as equatives.<sup>4</sup> While in English, the adjective/adverb constituting the parameter is never marked for this type of comparison by an inflectional suffix, note that it is marked in a less grammaticalized way by the free morpheme *as* in front, a type of expression that I will refer to as the correlate in the following (in the literature it is also referred to as the parameter marker).<sup>5</sup>

The least marked type of comparisons, illustrated in (1)(c), are those referring to similarity, i.e. [- dissimilarity], in a way that is not specifically restricted to degrees of a property, thus [- degree], but holds in a more unspecific way that may include a complex number of gradable or ungradable properties, referring for instance to what one might call manner.<sup>6</sup> In this type of comparison, there is accordingly typically no linguistic expression representing a parameter.<sup>7</sup> In line with the usual typological

---

<sup>3</sup> For a more detailed discussion of the markedness relations (incl. dissimilarity being marked as opposed to similarity) see sect. 4.3 below.

<sup>4</sup> Alternatively, one may call them degree equatives, using equatives as a cover term for all comparisons characterized by [- dissimilarity], cf. Jäger (2018; 2019); Hohaus (2015); see also Thurmair (2001: “Gradvergleiche” ‘degree comparisons’).

<sup>5</sup> Parameter marking by a free morpheme also occurs with certain adjectives/adverbs in comparatives in English in the form of *more* in front of the parameter. Note, however, that the most grammaticalized form, an inflectional suffix, only ever appears in comparatives in English.

<sup>6</sup> Note, however, that it is not limited to manner. Thus even the mere truth/validity of two propositions may be stated to be similar/the same, consider for example *Peter is a farmer like/as his father was* - a use that comes very close semantically to mere coordination, which is why there is a well-established crosslinguistic grammaticalization path from standard markers to coordinating conjunctions.

<sup>7</sup> If a parameter is expressed, one is not referring to specific degrees of this property. For instance in comparisons such as *He is tall like a bear* (sometimes misleadingly referred to as ‘generic equatives’) one does not refer to specific degrees of height in terms of concrete measure (in contrast to *as tall as*)

terminology (e.g. Haspelmath & Buchholz 1998), I will refer to these comparisons as similatives.<sup>8</sup> These comparisons, being the least-marked type and therefore also displaying the least specific linguistic markers crosslinguistically, have attracted the least attention by linguists so far. Yet, they play a central role in language change as will become evident below. The features of the three main types of comparison distinguished here and the resulting markedness hierarchy are summarized in (2).<sup>9</sup>

(2) Markedness hierarchy of comparison constructions:

similatives	<	equatives	<	comparatives
[- dissimilarity, - degree]		[- dissimilarity, + degree]		[+ dissimilarity, + degree]

Note that in the English examples in (1), a different standard marker (in English a comparison particle, in other languages possibly also a case marker or equivalent functional preposition)<sup>10</sup> is used for each of the three types of comparison, viz. *than*,

---

but to being tall in the way that bears are tall compared to salient other animals for instance. Note that *He is tall like a bear* may also formally be identified as a similative rather than an equative by virtue of the availability of the standard marker *like* in English (or *comme* in French etc.) which is ungrammatical in equatives, cf. \**as tall like*.

<sup>8</sup> Heine & Kuteva (2002) use the term *simile*. Alternatively, one may call them non-degree equatives (cf. Jäger 2018, 2019) or property equatives (cf. Hohaus 2015) contrasting with degree equatives. Other terms used in the literature include 'pure comparisons' (Zifonun et al. 1997: "reine Vergleiche"), 'open comparisons' (Thurmair 2001: "offene Vergleiche", with subtypes: "Modalvergleiche"/modal comparisons and "Faktizitätsvergleiche"/facticity comparisons) or 'similarity comparatives' (Alrenga 2007).

<sup>9</sup> Note that a factorial typology of the two features predicts a fourth type of comparisons whose degree of markedness would also lie inbetween that of comparatives and similatives as defined here, viz. comparisons characterized by [+ dissimilarity] and [- degree]. To my knowledge, this fourth type (which could be termed non-degree comparatives) has not been explicitly distinguished or addressed in the literature on comparisons. Comparisons including expressions meaning 'different(ly)' or '(an)other' constitute such cases, as they express dissimilarity without restricting it to specific degrees of a gradable property, but also referring for instance more generally to manner etc. (thus for instance German *Anna läuft anders als Maria* 'Anne walks differently from/in another way than Mary' is in fact the also [- degree], but comparative, i.e. [+ dissimilarity] equivalent of *Anna läuft so wie Maria* 'Anne walks like Mary'), see also Jäger (2018: 35, fn. 35; 368, fn. 330). In this paper, however, I will concentrate on the three types similatives, equatives and comparatives, leaving a more detailed discussion of this fourth type of comparison ('non-degree comparatives') to future research.

<sup>10</sup> The main types of comparative constructions in the languages of the world (cf. Stassen 1985, 2005) include languages with (i) a comparative particle, (ii) a verbal comparative construction ('exceed' construction), (iii) a 'conjunctive comparative construction' (coordination of two clauses) or (iv)

*as* and *like*, respectively.<sup>11</sup> As stated in Table 1, crosslinguistically, but also over the course of the diachronic development of one language, this may differ according to whether both features, [ $\pm$  dissimilarity] and [ $\pm$  degree] are expressed, or just one, leaving the other feature underspecified, as in the case of Spanish, which only expresses [ $\pm$  dissimilarity] by different standard markers using *como* in similatives and equatives but *que* in comparatives (type IIa in Table 1), or French, which only expresses [ $\pm$  degree] using *que* in comparatives and equatives but *comme* in similatives (type IIb in Table 1, see also sect. 3.7 on Romance languages). Some languages even use one uniform standard marker in all three types of comparison, leaving both features unexpressed by the standard marker. This is the case in Hungarian, which uses the comparison particle *mint* in similatives, equatives and comparatives alike (type I in Table 1, see also sect. 3.9 on Hungarian).

Language type	similatives	equatives	comparatives
	[- dissimilarity]		[+ dissimilarity]
	[- degree]	[+ degree]	
Type I: 1 standard marker, e.g. Hungarian	<i>mint</i>		
Type IIa: 2 standard markers, e.g. Spanish	<i>como</i>		<i>que</i>
Type IIb: 2 standard markers, e.g. French	<i>comme</i>	<i>que</i>	
Type III: 3 standard markers, e.g. English	<i>like (/ as)</i>	<i>as</i>	<i>than</i>

Table 1: Typology of standard marker systems.

---

comparative case of the standard (or case-equivalent functional preposition) with separative, allative or locative semantics.

<sup>11</sup> In similatives, *as* is also used, especially with clausal standards, in which *like* is not acceptable in the standard language but is also used in colloquial English.

The hierarchy given in (2) is also reflected in the diachronic development of comparisons in various languages as discussed in section 3, and will accordingly also play a role in explaining the changes observed crosslinguistically and captured in the comparative cycle in section 4. The selection of languages covered is largely contingent upon the availability of data. It is highly likely that many more examples of the comparative cycle in various languages will become apparent as more diachronic typological data become available. As mentioned above, the discussion in section 3 is partly based on diachronic corpus data and partly on descriptive grammars and other handbooks. For the reasons discussed above, many handbooks do not explicitly treat similatives or clearly differentiate them in their discussion of comparison constructions. Thus, in many cases we can only demonstrate that the development in the respective language follows the direction of the comparative cycle from comparisons of equality/similarity to those of inequality/dissimilarity, notably from equatives to comparatives leaving a detailed investigation of the individual steps of the development including similatives to future research. However, where there is sufficient information on all three types of comparison we can trace in detail the individual steps of the comparative cycle, which concur with the markedness hierarchy of comparison constructions.

### 3. The comparative cycle crosslinguistically

#### 3.1. Germanic languages

##### 3.1.1. German

One language in which the comparative cycle may be observed very clearly and even with several rounds of the cycle during the attested history of the language is German, cf. Jäger (2018). The stepwise syntactic-semantic shift of markers of comparison - in German as in many other languages the standard marker (comparison particle) - from similatives to equatives to comparatives that constitutes the comparative cycle, first occurred with the comparison particle *also/als*. While in classical Old High German (OHG), *thanne* ‘than’ constitutes the most common standard marker in comparatives, cf. (3), and *so* ‘as/like’ both in similatives and equatives, cf. (4) and (5), a strengthened form of the latter, *also*, cf. (6), built by univerbation with the originally adverbial

element *al* ‘all, fully’, starts to occur in similatives in (Late) OHG.<sup>12</sup> In similatives, no adjective/adverb expressing the parameter occurs between the superordinate clause originally containing the adverbial *al* on the one hand and the standard marker *so* on the other hand. Accordingly, similatives but not equatives constitute a possible bridging context for the relevant univerbation and grammaticalization of new standard markers formed according to this pattern. Univerbation of the comparison particle with a frequently adjacent expression in the preceding clause constitutes a crosslinguistically common grammaticalization path for new standard markers and occurs with different types of expressions that precede the original standard marker.<sup>13</sup>

(3) OHG (Tatian 70, 17)

<i>Eno</i>	<i>ni</i>	<i>birut</i>	<i>ir</i>	<i>furirun</i>	<i>thane</i>
Q	NEG	be:PRS;2PL	you:NOM;PL	far:CPD;NOM;PL	than
<i>sie</i>		<i>sín</i>			
they: NOM		be:SBJV;PRS;3PL			

‘Aren’t you worth more than they are?’

<sup>12</sup> Throughout the paper, comparison particles in the examples are glossed with the semantically corresponding particle in the metalanguage English for the respective type of comparison, even if the same particle is used for several types of comparison in the object language. Thus the same lexical item, e.g. French *que* etc., is glossed for instance as ‘than’ in a comparative, but as ‘as’ in an equative.

<sup>13</sup> Grammaticalization of new comparison particles by ‘strengthening’ (reanalysis of matrix-clause internal elements and adjacent comparison particle as a new comparison particle) frequently takes place with the following kinds of elements (cf. Jäger 2018: 370f.):

- (i) correlate (parameter marker): e.g. OHG *soso*, ENHG/NHG *als wie*, Latin *sicut*, *tamquam*, Gothic *svasve*, Old English *swa swa*, Middle English *so as*, Swedish *såsom*, Dutch *zoals*, French *ainsi comme/ainsi que*, *autant que*, Jiddish *azoy vi*, Romani *kade sar*
- (ii) item with identity semantics (‘same’, ‘equal’ etc.): e.g. OHG (*so*) *selb so*, *sama so* > MHG *same/(al)sam*, ENHG *gleichwie*, Dutch *gelijk*, English *like*, Norwegian *like*, Swedish *lika*, *som*, Danish *lige så*, *som*, Icelandic *eins og*
- (iii) intensifier (‘fully’, ‘completely’, ‘exactly’ etc.): e.g. OHG/MHG *also* > *als*, Dutch *als*, Old English *eallswā* > *as*, Provençal *tot aissi/atressi*
- (iv) noun (possibly within PP; ‘(in) the way/manner/look/degree’ etc.): e.g. ENHG *inmassen*, *gestalt*, (*ce*)*gleicherweis*, Italian/French/Spanish/Portuguese/Romanian *com(o)/com(m)e/cum* < Latin *quomodo* (< *quo modo* ‘in which way’), Irish *cosúil* (< *chomh/comh* ‘as’ + *samhail* ‘picture/appearance’).



- (4) OHG (Tatian 40, 3)

*inti gibit imo só manag so her bitharf.*  
 and give:PRS;3SG he:DAT as much as he:NOM need:PRS;3SG  
 ‘and gives him as much as he needs’

- (5) OHG (Tatian 44, 16)

*thaz só sí só sín meistar*  
 that so be:SBJV;PRS;3SG as his:NOM master:NOM  
 ‘that he is like his master’

- (6) OHG (Notker Ps. 35, 7)

*Din reht trûhten ist also bérga.*  
 your:NOM justice:NOM Lord:NOM be:PRS;3SG as mountain:NOM;PL  
 ‘Your justice, Lord, is like the mountains.’

In Middle High German (MHG) *also* (> *alse* > *als*) constitutes the main pattern already in similatives, as in (7), while *so* continues to be the main pattern in equatives, such as (8), and *dann(e)* (< *thanne*) in comparatives, cf. (9). However, during this period, *also* already starts to occur occasionally in equatives, cf. (10), and very exceptionally even in comparatives, cf. (11).

- (7) MHG (TrHL 10r,21f.)

*diu tvost uns also diu uil gvote*  
 you do:PRS;2SG we:DAT like the:NOM;F very good:NOM;F  
*muotir.*  
 mother:NOM  
 ‘You do unto us like the very good mother.’

- (8) MHG (Phys 151r, 18f.)

*unt izzit danne so lange so got wil*  
 and eat:PRS;3SG then as long as God:NOM want:PRS;3SG  
 ‘and then eats as long as God wants’

- (9) MHG (Phys 133r, 4-6)

*Trehtin, diniu wort diu sint*  
 Lord your: NOM;PL word:NOM;PL they:NOM;PL be:PRS;3PL

*suozzere*            *in*    *minem*    *munde.*    ***danne*** *daz*    *honich*  
 sweet:CPD;NOM;PL    in    my:DAT;M    mouth:DAT    than    the:NOM;N    honey:NOM  
*unt*    *der*                    *flade*  
 and    the:NOM;M            cake:NOM  
 ‘Lord, your words, they are sweeter in my mouth than honey and cake’

(10) MHG (Lil 8, 15-16)

*Disen*                    *zuiuel*                    *muzen*                    *wir*                    *hauen.*                    *alse*  
 this:ACC;M                    doubt:ACC                    must:PRS;1PL    we:NOM                    have:INF                    as  
*lange*    ***alse***                    *de*                    *sumer*                    *dis*                    *leuenes*  
 long    as                    the:NOM                    summer:NOM    this:GEN                    life:GEN  
*weret.*  
 last:PRS;3SG  
 ‘This doubt we must have as long as the summer of this life lasts.’

(11) MHG (SalH 097, 05-08)

*daz*    *dv*                    *nie*    *von*                    *dinge*                    *in-wordes*                    *svzer*  
 that    you:NOM    never    by                    thing:DAT    NEG-AUX:PST;2SG    sweet:CPD  
*geminit.*    ***alse***                    *von*    *gode.*  
 love:PTCP    than                    by    God:DAT  
 ‘...that you were never loved more dearly by anything than by God.’

During 15th century Early New High German (ENHG), *als* becomes the main standard marker in equatives, too. In the 16th century, also its use in comparatives, as in (13), increases, which, however, still show *dann/denn* as the main standard marker employed, as illustrated in (12). Only since 17th century New High German (NHG) does *als* also constitute the main pattern in comparatives, superseding *dann/denn*.

(12) ENHG (JBang 17r, 5f.)

*Da*    *nun*                    *die*                                    *Sachssen*                    *sahen/*                    *das* *der*  
 since    now                    the: NOM;PL                    Saxon:NOM;PL    see:PST;3PL    that    the:GEN;PL  
*Thueringer*                    *Acker*                    *besser*    *war*                    ***dann***    *jhrer*  
 Thuringian:GEN;PL    field:NOM                    good:CPD    be:PST;3SG    than    theirs:NOM;M  
 ‘Since the Saxons now saw that the field of the Thuringians was better than theirs’

## (13) ENHG (JMath 44r, 22-24)

*Denn er ist grewlicher vnd heßlicher/ als jrgend*  
 because he:NOM be:PRS;3SG. ghastly:CPD and ugly:CPD than any  
*der aller geringsten oder ergsten vnd*  
 the:GEN;PL very low:SPD;GEN;PL or wicked:SPD,GEN;PL and  
*Gottlosesten einer zugerichtet.*  
 ungodly:SPD;GEN;PL one:NOM;SG;M injure:PTCP  
 ‘because he is injured in a ghastlier and uglier way than any of the lowest, most  
 wicked and ungodly ones’

The same kind of shift as observed for *als(o)* also occurs with the standard marker *wie*: this comparison particle, which was grammaticalized from the interrogative/relative adverb ‘how’, is first very occasionally used in similatives in MHG, cf. (14), in which *also* constitutes the main pattern, as discussed above. In similatives, such as (15), *wie* becomes the main pattern superceding *als(o)* in 16th century ENHG, during which period it also starts to occur in equatives, such as (16), for the first time, which however most frequently still contain *als* at that time, as described above. During 17th century ENHG, the first attestations of *wie* in comparatives occur as illustrated in (17). In equatives, *wie* becomes the main pattern only in 19th century NHG and since that time is also increasingly used in comparatives, in which it represents the main pattern in most present-day High German dialects, as illustrated in (18), whereas the standard language has preserved *als*.

## (14) MHG (Walter 48, 7 (after Paul 2007))

*swie si sint, sô wil ich sîn*  
 how/as they:NOM be:PRS;3PL so want:PRS;1SG I:NOM be:INF  
 ‘However/as they are, so do I want do be’

## (15) ENHG (WRal 2, 7f)

*es zergehet vnd schmelzet nicht von der*  
 it:NOM dissolve:PRS;3SG and melt:PRS;3SG NEG by the:DAT  
*Sonnen/ wie das Hartz vnd Pech auß*  
 sun:DAT like the:NOM;N resin:NOM and pitch:NOM from

Norwegen

Norway

‘It does not dissolve and melt from the sun like the resin and pitch from Norway’

(16) ENHG (JMath 51v, 16-20)

*Darumb sie auch also schmeulich vnd Gotteslesterlich/*

therefore they:NOM also as disgracefully and blasphemously

*wie die Mahometisten vom Abendmal des*

as the:NOM;PL muslim:NOM;PL of.the:DAT supper:DAT the:GEN

*HERRN [...] gedencken vnd reden.*

Lord:GEN think:PRS;3PL and talk:PRS;3PL

‘Therefore, they think and talk as disgracefully and blasphemously as the muslims about the Lord’s supper’

(17) ENHG (H.U. Krafft, Reisen 248, Lit. Verein (after DWB 29: 1483f.))

*mer daran verbrechen wie gutt machen*

much:CPD. there.at break:INF than good make:INF

‘destroy it rather than making amends’

(18) NHG (Central Hessian dialect (after Jäger 2018: 316))

*Dr Thomas ess grieser wej sei*

the:NOM;M Thomas:NOM be:PRS;3SG tall:CPD than his:NOM

*Brourer.*

brother:NOM

‘Thomas is taller than his brother.’

The development in German can be summarized as given in Table 2,<sup>14</sup> highlighting the repeated stepwise shift of standard markers from comparisons of

---

<sup>14</sup> The same development as in High German, yet at a much slower pace, can also be observed in Low German, cf. Jäger & Walkden (2021: 316). The original comparative particle *than* represents the main pattern in Old and Middle Low German and is preserved in some Low German dialects in the form of *denn* even today. On the other hand, the standard marker *also*, which – as in OHG – first occurred in similatives in Old Low German became the main comparison particle used in equatives in Middle Low German. In most Low German varieties it has subsequently also become the main particle used in comparatives so that a uniform comparison particle *as* (< *also*) in all types of comparison is typical of most Modern Low German dialects. Only recently, *wie* and its Low German equivalent *wu/wo* (‘how’)

equality/similarity to those of inequality/dissimilarity, more specifically from similatives to equatives to comparatives that is typical of the comparative cycle.<sup>15</sup> While German represents a very clear instance of the comparative cycle, evidence for the same kind of change can be found in many related and unrelated languages.

	similative	equative	comparative
	[- dissimilarity]		[+ dissimilarity]
	[- degree]	[+ degree]	
OHG		<i>so</i>	<i>danne</i>
MHG	<i>also</i>	<i>so</i>	<i>dann/denn</i>
ENHG 15th cent.		<i>als</i>	<i>denn</i>
ENHG 16th cent.	<i>wie</i>	<i>als</i>	<i>denn</i>
NHG 17/18th cent.	<i>wie</i>		<i>als</i>
NHG 19th cent., Mod. Standard		<i>wie</i>	<i>als</i>
Dialects/Colloquial German		<i>wie</i>	

Table 2: The comparative cycle in German (after Jäger 2018: 364).

### 3.1.2. English

The diachronic development of comparison particles in English shows several parallels to that in German although, with *as* in equatives and partly in similatives and *than* in comparatives in present-day English, the language has, on the whole, preserved a pattern corresponding to that observed for MHG and 15th century ENHG. While in Old English similatives and equatives the standard marker *swa*, cognate of OHG *so*, was prevalent as in OHG, cf. (19), in Late Old English in the same way as in Late OHG a new strengthened standard marker arose in similatives from univerbation

---

are used in similatives and equatives, occasionally even already in comparatives in some Low German dialects, again repeating the shift observed for *also/as*.

<sup>15</sup> Another comparison particle that never constituted the main pattern, however, in historical German and therefore is not included in Table 2, but also survives in some present-day dialects is *als wie*. It was grammaticalized on the basis of the correlate (parameter marker) *als* and the frequently adjacent comparison particle *wie* (= pattern (i) in fn. 13) and first occurs in similatives in the 17th century. It is then extended to equatives and since the 18th century also occurs in comparatives, cf. Jäger (2018: 255-259). (In contrast to the statement by Dücker (1961: 216) *als wie* does not constitute an intermediate stage in the development from *als* to *wie*, but only occurs after *wie* has been firmly established as a comparison particle.) Accordingly, there is in fact evidence for three rounds of the comparative cycle in the history of German – with *als(o)*, *wie* and *als wie*.

with an item meaning ‘all, fully’, viz. *eall swa* > *alswa/ also/alse/as* (CHEL II: 357; Jäger & Walkden 2021: 317-322), cf. (20). Since then, this comparison particle has mostly been restricted to comparisons of equality, while *þonne* > *than* has been the prevalent standard marker in comparatives since Old English (Jäger & Walkden 2021: 325-327), cf. (21), for Modern Standard English see also (1) above.

(19) Old English (cobede, Bede\_3:16.226.26.2325 (after Jäger & Walkden 2021: 322))

<i>swa</i>	<i>swiðe</i>	<i>swa</i>	<i>þa</i>	<i>neowan</i>	<i>Cristnan</i>
as	much	as	the:NOM;PL	new:NOM;PL	Christian:NOM;PL
<i>þa</i>	<i>get</i>	<i>hit</i>	<i>neoman</i>	<i>meahton</i>	
then	yet	it:ACC	take:INF	may:PST;PL	

‘as much as the new Christians were capable of it’

(20) Old English (cowulf, WHom\_3:7.52 (after Jäger & Walkden 2021: 318))

<i>And</i>	<i>witodlice</i>	<i>ealswa</i>	<i>flod</i>	<i>com</i>	<i>hwilum</i>	<i>ær</i>	
and	truly	as	flood:NOM	come:PST;3SG	at.time	before	
<i>for</i>	<i>synnum,</i>	<i>swa</i>	<i>cymð</i>	<i>eac</i>	<i>for</i>	<i>synnum</i>	<i>fyr</i>
for	sin:DAT;PL	so	come:PRS;3SG	also	for	sin:DAT;PL	fire:NOM

‘And truly, just as the flood came before (to punish us) for our sins, now the fire is coming (to punish us) for our sins.’

(21) Old English (cobeowul, 16.465.391 (after Jäger & Walkden 2021: 326))

<i>se</i>	<i>wæs</i>	<i>betera</i>	<i>ðonne</i>	<i>ic</i>
this:NOM	be:PST;3SG	good:CPD	than	I:NOM

‘He was better than me’

Despite of this basic continuity, there has also been some change concerning standard markers in English comparisons. On the basis of strengthening with a regularly adjacent expression meaning ‘same’, viz. *gelice* > *like*, another common grammaticalization path for new standard markers arising in similatives (cf. (ii) in fn.13), the standard marker Old English *gelice swa* (Jäger & Walkden 2021: 319f.), Middle English *lyk as / like as* (CHEL II. 358), arose as illustrated in (22) and (23). It became more frequent during the 15th and 16th century. Eventually, the second part of this combination began to be dropped giving way to simple *like* in similatives since the early 16th century (CHEL III: 316), see (24). Note that this standard marker, which

originated in similatives, is still restricted to this type of comparison today, in which *as*, however, also still occurs.

(22) Old English (coverhom, HomS\_40.3 (after Jäger & Walkden 2021: 320))

<i>Emne</i>	<i>hit</i>	<i>bið</i>	<i>gelice swa</i>	<i>man</i>	<i>mid wætere,</i>
even	it:NOM	be:PRS;3SG	same as	man:NOM	with water:DAT
<i>þone</i>	<i>weallendan</i>		<i>wylm</i>	<i>agiote</i>	
the:ACC	surging:ACC		flame:ACC	quench:PRS;SBJV;SG	

‘It is just like putting out a surging flame with water’

(23) Middle English ([HC] Vicary 69 (after CHEL III: 316))

<i>the</i>	<i>lyuer [...]</i>	<i>should</i>	<i>be</i>	<i>plycable</i>	<i>to</i>	<i>the</i>	<i>stomacke,</i>
the	liver	should	be:INF	pliant	to	the	stomach
<i>like</i>		<i>as</i>	<i>a</i>	<i>hande</i>	<i>dothe</i>	<i>to</i>	<i>an apple</i>
like/same		as	a	hand	do:PRS;3SG	to	an apple

‘The liver should be pliant to the stomach like a hand is to an apple’

(24) Middle English (1530 Berners Arth. Lyt. Bryt. 520 (after CHEL III: 316))

<i>Ye</i>	<i>have</i>	<i>said</i>	<i>lyke</i>	<i>a noble</i>	<i>lady ought</i>	<i>to</i>	<i>say</i>
you	have:PRS;2SG	say:PTCP	like	a noble	lady ought	to	say:INF

‘You have spoken as a noble lady should speak’

Besides this rise of a new standard marker in similatives, there is crucially also evidence for a distributional shift in line with the comparative cycle in diachronic and dialectal English data. According to Small (1924: 43), since early Middle English up to Modern dialectal varieties, *as* has repeatedly ‘threatened’ to take over the function of a comparative particle. This is of course reminiscent of the development of *als* in German. The Historical Thesaurus of English (s.v. *as*)<sup>16</sup> states that *as* was used as a standard marker in comparatives in English from the mid 15th to the mid 17th century, and in archaic use even until the beginning of the 19th century. The OED (s.v. *as*, B.I.5)<sup>17</sup> lists examples for *as* in comparatives from 1300 up to the 20th century, cf. (25) and (26).<sup>18</sup>

<sup>16</sup> <http://historicalthesaurus.arts.gla.ac.uk/category-selection/?qsearch=as>.

<sup>17</sup> <http://www.oed.com/view/Entry/11307?rskey=jq3qkq&result=6&isAdvanced=false#eid>.

<sup>18</sup> Most of these contain a negation. On negated comparisons as potential bridging contexts see section 4.3.

- (25) Middle English (c 1300 St. Edward Elder (Laud) l. 38 (after OED s.v. as))

*Fellere*                      þing    n-is                      non    **ase**    wumman    ȝware  
wicked:CPD                    thing    NEG-be:PRS;3SG    none    than    woman    when  
*heo*                              wole                              to    vuele    wende.  
she:NOM                        will:SBJV;PRS;3SG    to    evil    turn:INF  
‘There is nothing more wicked than a woman if she turns to evil’

- (26) Middle English (a 1425 J. Lelamour tr. Macer Herbal f. 67v (after OED s.v. as))

*Also*    *this*            *erbe*    *haviþ*            *mo*            *vertues*    **as**            *endyue*  
also    this            herb    have:PRS;3SG    many:CPD    virtue:PL    than    endive  
*hæþe*.  
have:PRS;3SG  
‘Also, this herb has more virtues than endive has.’

- (27) Scottish English (Brian Holton: The Mossflow)

*A wad-na*                    *think*                    *you*    ‘d            *be*            *onie*            *better*  
I    would-NEG            think:INF            you    would    be:INF    any    good:CPD  
**as**            *them!*  
than    they  
‘I wouldn’t think you were any better than them.’

While this use of *as* in comparatives has never become the prevalent pattern in English, it is preserved in regional varieties such as Scottish English, cf. (27), Irish English, Yorkshire English and certain American-English varieties (Small 1929: 22; OED s.v. *as*, B.I.5).<sup>19</sup>

### 3.1.3. Dutch

In Dutch, a distributional shift of standard markers (comparison particles) according to the comparative cycle can also be observed. Like OHG and Old English, Old Dutch (Old Low Franconian) mostly used *than* in comparatives and *so* in similatives and equatives, in which, occasionally the strengthened form *also* occurs instead (cf. Jäger & Walkden 2021: 300-311, 316). Phonologically reduced to *als*, it became the main particle used in similatives and equatives in Middle Dutch, while in comparatives the

<sup>19</sup> <http://www.oed.com/view/Entry/11307?rskey=jq3qkq&result=6&isAdvanced=false#eid>.



typical comparison particle was still *dan* (cf. Postma 2006: 3). During the 14th and 15th century, however, *als* begins to appear in comparatives, replacing *dan* since the second half of the 16th century (van der Horst 2008: 728, Postma 2006: 20 with reference to Klooster 2001: 352), cf. (28). As a result, *als* was widely used as a uniform comparison particle in equatives and comparatives in Dutch during the second half of the 16th and especially the 17th century.

(28) Dutch (early 17th c., Bredero (after van der Horst 2008: 1272))

<i>Een</i>	<i>Vrouw</i>	<i>brengh</i>	<i>meer</i>	<i>te</i>	<i>weegh</i> ,	<b><i>als</i></b>	<i>dysent</i>
one	woman	bring:PRS;3SG	more	to	way	than	thousand
<i>mannen</i>	<i>souwen</i>						
man:PL	would:3PL						

‘One woman achieves more than a thousand men.’

However, since the 18th century, this change was partly turned back – according to van der Horst (2008: 1442) and Hubers & de Hoop (2013: 90) due to normative pressure by prescriptive grammarians demanding the use of *dan* in comparatives (e.g. Balthazar Huydecoper at the beginning of the 18th century). As a consequence, *dan* is used in comparatives in Standard Dutch until today, whereas *als* is generally used in comparisons expressing similarity, both similatives and equatives. According to SAND (2005: 13), *dan* is the only acceptable comparison particle in Standard Dutch comparatives; according to ANS,<sup>20</sup> however, *als* is acceptable instead of *dan* in the spoken standard language, too, but in the written language, *dan* is mostly used. In fact, as this paper is written, over 20 years after the latest edition, a new, revised edition of ANS is being prepared in which *als* is now treated as equally acceptable in Standard Dutch comparatives.<sup>21</sup> For many native speakers, however, *als* is generally not acceptable in comparatives and *dan* represents a kind of shibboleth for good Dutch. Hubers & de Hoop’s (2013) investigation of the corpus of spoken Dutch shows that in comparatives indeed mostly *dan* is used. However, in the southern Netherlands (Limburg, Noord-Brabant, Zeeland) *als* appears almost as often as *dan* in comparatives. Besides, sociolinguistic factors play a role insofar as speakers with a lower educational background use *als* slightly more often than *dan* in comparatives, while those with a higher educational background almost only use *dan*. Hubers & de

<sup>20</sup> <http://ans.ruhosting.nl/e-ans/index.html>

<sup>21</sup> Cf. <https://taalunie.org/actueel/120/hoe-de-e-ans-grondig-wordt-herwerkt-en-verbeterd>

Hoop conclude that the dominance of *dan* in comparatives is due to normative pressure against *als* taught at school and that, without this pressure, *als* would be the usual comparative particle in Dutch today. This impression is corroborated by dialectal research: according to SAND (2005: 13, map 15b) the historically widespread use of *als* in comparatives is preserved in Dutch dialects until today, in fact, *als/as* constitutes the main pattern in dialectal comparatives in the entire Dutch-speaking area except for West and East Flanders.

In Belgian Standard Dutch (BSD), which is used in the Belgian Media and formal situations, the comparison particle used in similatives is *zoals*, a particle formed by univerbation of the correlate *zo*, originally part of the superordinate clause, and the adjacent comparison particle *als* (= (i) in fn. 13). Due to this rise of a new comparison particle in similatives, Belgian Standard Dutch has developed a system with three comparison particles for the three types of comparisons: *zoals* in similatives, cf. (31)(b), *als* in equatives, cf. (30)(a), and *dan* in comparatives, cf. (29)(a) (= type III in Table 1; compare also MHG *also – so – dann*). In most Flemish dialects, as in most Dutch dialects in general, *als/as* has, however, replaced *dan* in comparatives so that *als/as* is uniformly used in comparatives and equatives while *zoals* (*zoas*, *zuas* etc.) or another more recent comparison particle, viz. *(ge)lijk/gelak*, appears in similatives, for instance in the dialect of the city of Antwerp, cf. (29)(b), (30)(b) and (31)(c). The comparison particle *(ge)lijk/gelak* was presumably grammaticalized in similatives in the same way as English *like* (= (ii) in fn. 13, see also sect. 3.1.2). In certain varieties in the region of Antwerp, it even occurs in comparatives, cf. (32).

(29) Dutch

- a. *Mijn kat is een beetje kleiner dan je hondje.* (Stand. Dutch, BSD)  
 b. *Mijn kaet is een bekke klender as awen ond.* (Antwerp Dutch)  
 my cat be:PRS;3SG a bit small:CPD than your dog  
 ‘My cat is a little smaller than your dog.’

(30) Dutch

- a. *Koen is even (/ zo) oud als Antje.* (Standard Dutch, BSD)  
 b. *De Koen is even / zoe oud as Antje.* (Antwerp Dutch)  
 the Koen be:PRS;3SG equally as old as Antje  
 ‘Koen is as old as Antje.’

## (31) Dutch

- a. *Zij zingt als Kylie Minogue.* (Standard Dutch)  
 b. *Zij zingt zoals Kylie Minogue.* (BSD)  
 c. *Zij zingt zuas/gelak Kylie Minogue.* (Antwerp Dutch)  
 she sing:PRS;3SG like Kylie Minogue  
 ‘She sings like Kylie Minogue.’

(32) Sint Lenaarts Dutch (Sint Lenaarts (K209p), DynaSAND)<sup>22</sup>

- ze gelove da wij rijker zijn lijk zullie*  
 they believe:PRS;3PL that we rich:CPD be:PRS;1PL than they  
 ‘They believe that we are richer than them.’

In this particular dialect, we can accordingly observe a recurrent change according to the comparative cycle with *als* replacing *dan* and subsequently *(ge)lijk* replacing *als* in comparatives, both particles shifting from similatives to equatives to comparatives.

### 3.2. Sanskrit

Diachronic distributional shifts of standard markers according to the comparative cycle are not limited to Germanic languages, but are also found in many other Indo-European languages. A syntactic-semantic shift of the comparison particle of this kind presumably already occurred from Vedic to Classical Sanskrit. In Vedic, the particle *ná*, which is of the same origin as the negation particle and is assumed to have developed from it (cf. Pinault 1985; Dunkel 2014: 546f.),<sup>23</sup> is used as a similative/equative particle meaning ‘as/like’, as illustrated in (33). In Classical Sanskrit, however, it appears as a comparative particle meaning ‘than’, cf. (34). According to Zeilfelder (2001: 99) it is not continuously attested, which is why she rather supposes an independent source, taking a semantic change from an equative particle to a comparative particle to be unlikely. However, she also explicitly indicates a similarity to the ‘confusion’ of *als* and *wie* in colloquial German. Viti (2002:77) also

<sup>22</sup> <http://www.meertens.knaw.nl/sand/>

<sup>23</sup> According to Pinault (1985), the Vedic equative particle *ná* is derived from the negation particle *ná* via bridging contexts such as ‘Not (as one might think) A, (but) B is speaking’ > ‘B speaks like A’, see also Dunkel (2014: 546f.: “Der Ausgangspunkt liegt in implizierten Negativvergleichen” ‘The origin lies in implicit negative comparisons’, same origin as negation particle, PIE \**ne* ‘not’).

considers the possibility of classical Sanskrit *ná* in comparatives continuing on from Vedic equative *ná* and again mentions the Non-Standard German use of *wie* in comparatives as a parallel.<sup>24</sup>

(33) Vedic (RV 1.39.10)

<i>īṣum</i>	<i>ná</i>	<i>sṛjata</i>	<i>dvīṣam</i>
arrow:ACC;SG	like	send:PRS;IMP;2PL	hatred:ACC;SG

‘Shoot the hatred like an arrow.’

(34) Classical Sanskrit (Pañcatantra 1.417 (after Viti 2002: 79))

<i>pañḍito</i>	<i>’pi</i>	<i>varam</i>	<i>śatrur</i>
wise:NOM;SG;M	even	better	enemy:NOM,SG

<i>na</i>	<i>mūrkho</i>	<i>hitakāraḥ</i>
than	stupid:NOM;SG;M	ally:NOM;SG

‘Even a wise enemy is better than a stupid ally.’

### 3.3. Baltic languages

If on the basis of Vedic *ná*, one posits PIE *\*né* with the two functions of negative and (secondarily) similative/equative particle, a distributional shift of a standard marker from equative to comparative according to the comparative cycle can also be assumed for Baltic languages: Old Lithuanian as well as Latvian *ne*, which are used as negative particles, also rarely appear in comparatives, as illustrated in (35) and (36), and can be considered to have undergone a development ‘not’ > ‘like’ > ‘than’ (cf. Petit 2021: 123-127).<sup>25</sup> This scenario is supported by the fact that the “Baltic languages show sporadic, but unequivocal traces of purely similative meaning of *\*ne* ‘like’” (Petit 2021: 126), as in (37), which can be traced back to at least the 17th century.

<sup>24</sup> Delbrück (1888: 196) already points out constructions corresponding to Latin *quam* + comparative in the ‘popular language’ (‘Volkssprache’) giving a Late Vedic example including *yác ca* ‘than’ and stating “dass die Erklärung dieser Construction beim Comp. aus der gleichen Construction beim Positiv herzuholen ist, ist wohl klar” (‘It is clear that the explanation of this construction with comparatives is to be found in the same construction with the positive’).

<sup>25</sup> Petit (2021) also discusses an alternative scenario of ‘not’ > ‘than’.

- (35) Latvian (K. Barons & H. Wissendorff, *Latwju dainas* (19th c., BW I 3320d) (after Petit 2021: 123))

*Sche meitas wežakas, ne pate mahte.*  
 here girls:NOM;PL old:CPD;NOM than self:NOM mother:NOM  
 ‘Here, the girls are older than the mother herself.’

- (36) Old Lithuanian (Jacob Brodowski, *Lexicon Germanico-Lithvanicum et Lithvanico-Germanicum* (18th c., I 153, 9) (after Petit 2021: 121))

*Kétures Akkis daugiaus máta ne wiena.*  
 four:NOM eye:NOM;PL much:CPD;ADV see:PRS;3 than one:NOM  
 ‘Four eyes see more than one.’

- (37) Old Lithuanian (Jacob Brodowski, *Lexicon Germanico-Lithvanicum et Lithvanico-Germanicum* (18th c., I 151, 9) (after Petit 2021: 126))

*Auga waikai ne Girroj’ Médžeji.*  
 grow.up:PRS;3 child:NOM;PL like forest:LOC tree:NOM;PL  
 ‘Children grow up like trees in a forest.’

There are further comparison particles in the Baltic languages which are formed on the basis of *ne*, consider for instance the Lithuanian similitive/equative particle *neĩ* and comparative particle *neĩ*, which are almost identical, both formed from *\*né* + particle *\*ei* / *\*i*, only differentiated by intonation (cf. Vine 1978: 183f., see also Petit 2021: 120f., 127), cf. (38) and (39). This similitive/equative particle presumably arose at a time when Lithuanian *né* was not yet a comparative particle but like PIE *\*né* a negative and a similitive/equative particle, in which case we are dealing with another instance of the comparative cycle.

- (38) Lithuanian (after LKŽ 8: 622)

*Grikojai menkesni, nei avižojai.*  
 buckweat-straw:NOM;PL smaller:NOM than oat-straw:NOM;PL  
 ‘Buckweat straw is shorter than oat straw’

- (39) Lithuanian (after LKŽ 8: 624)

*laukia nei gervė giedros.*  
 wait:PRS;3 like crane:NOM;SG good.weather:GEN;SG  
 ‘He/she/they is/are waiting for it like a crane for good weather’

Further instances of the comparative cycle are found in the Baltic languages in the case of Lithuanian *kaĩp*, Latvian *kā* and Old Prussian *kai*. All of these were originally “conjunctions of manner derived from the relative stem *\*ka-* < PIE *\*k<sup>w</sup>o-*“ (Petit 2021: 103) and occur in similatives cf. (40)/(43)/(46) as well as equatives cf. (41)/(44)/(47),<sup>26</sup> which generally do not differ in terms of the standard markers used in the Baltic languages (cf. Petit 2021: 105), but also in comparatives especially after negation, cf. (42)/(45)/(48), since at least the 16th/17th century (for examples of Modern Lithuanian *kaĩp* as a similative, equative and comparative particle see LKŽ 5: 60f.). In fact, Jensen (1934: 124) already explicitly mentions Lithuanian *kaĩp* as a parallel case to German *wie*, Russian *kak* and Romance *que/che* of a comparative particle that was originally only used in similatives/equatives, in other words, that underwent the comparative cycle.

(40) Old Lithuanian (Jacob Brodowski, *Lexicon Germanico-Lithvanicum et Lithvanico-Germanicum* (18th c., II 723, 39641), (after Petit 2021: 103))

*Saufa Málka **kaĩp** Kanklys*  
 dry:NOM wood:NOM like harp:NOM  
 ‘wood dry like a harp’

(41) Old Lithuanian (Jacob Brodowski, *Lexicon Germanico-Lithvanicum et Lithvanico-Germanicum* (18th c., I 115, 9–12) (after Petit 2021: 106))

*taĩp mažas **kaĩp** Pelle*  
 as small:NOM as mouse:NOM  
 ‘as small as a mouse’

(42) Old Lithuanian (Mikalojus Daukša, *Postilla Catholica* (1599: 84, 6) (after Petit 2021: 119))

*Teip’ wel bažnīczia ne túri níeko piktefnio*  
 thus again church:NOM NEG have:PRS;3 nothing:GEN bad:CPD;GEN  
***kaĩp’** mókšlą Hėretikų.*  
 than science:ACC heretic:GEN;PL  
 ‘Thus also the Church has nothing worse than the science of heretics.’

<sup>26</sup> In (47) *gi* constitutes an emphatic particle that is added to *kāi*.

- (43) Old Latvian (Georg Mancel, *Phraseologia Lettica* (1631: Cap. X 16) (after Petit 2021: 103))

*Dfālltāns*                      *ka*    *Wafzka.*  
 yellow:NOM                      like    wax:NOM  
 ‘It is yellow like wax.’

- (44) Old Latvian (Enchiridion (1586: G3A 1, 14) (after Petit 2021: 107))

*Effet*                              *packloufzige* [...] *tha*    *Kunge*    *peetcz / Tick lab*  
 be:IMP;2PL                      obedient:NOM;PL    the:GEN Lord:GEN    behind as    well  
*tam*    *Kōningam/*                      *kha*    *tham*    *Wuerfzenekam*  
 the:DAT    king:DAT                      as                      the:DAT    superior:DAT  
 ‘Be obedient behind the Lord as well to the king as to the superior.’

- (45) Old Latvian (Georg Elger, *Geistliche catholische Gefänge* (1621: 46, 8–9) (after Petit 2021: 122))

*Nāwar*                              *būt*    *faldak*                      *dōmafzan /*    *Ka*    *no*    *Jefu*  
 NEG.can:PRS;3                      be:INF    sweet:CPD;NOM    thought:NOM    than of    Jesus:GEN  
*mufe*    *dwāfels*    *gan.*  
 1PL;GEN    soul:GEN    well  
 ‘There cannot be any sweeter thought than of Jesus, our soul.’

- (46) Old Prussian (Enchiridion (1561: III 49, 6–7) (after Petit 2021: 103))

*Swintai*                              *kai*    *ftai*                      *malnikai*                      *Deiwas*  
 holy:NOM;PL                      like    the:NOM;PL    child:NOM;PL    God:GEN  
 ‘holy like God’s children’

- (47) Old Prussian (Enchiridion (1561: 103, 14-15) (after Petit 2021: 107))

*Titturri*                      *dijgi*    *ftai*                      *wijrai*                      *fwaians*                      *gannans*                      *milijt*  
 so have:PRS;3    also    the:NOM;PL    man:NOM;PL    REFL:ACC;PL    wife:ACC;PL    love:INF  
*kāi-gi*    *swian*                      *fubban*                      *kērmenen.*  
 as-PTCL    REFL:ACC    self:ACC                      body:ACC  
 ‘Men should also love their wives as much as their own body.’

- (48) Old Prussian (Enchiridion (1561 : III 115, 8–9) (after Petit 2021: 118))

*fteimans*    *malnikikamans*    *ni*                      *maffais*    *kai*    *ftēimans*                      *vremmans*  
 ART:DAT;PL    young:DAT;PL    NEG                      less    than    the:DAT;PL    old:DAT;PL  
 ‘to the young ones not less than to the old ones’

### 3.4. Armenian

In Classical Armenian, we also find evidence for a shift of standard markers from equatives to comparatives. According to Zeilfelder (1996: 195f.; 2001: 3) the Classical Armenian comparative case accusative stems from an accusative originally governed by the preposition *i*, which together with an imperative ‘compare!’ formed the equative particle *ibrew*. Kölligan (2021: 72) suggests instead that *ibrew* is originally formed on the basis of *ibr*, an instrumental of ‘thing/which’ and a preposition *ew* which the accusative was originally governed by. At any rate, due to the grammaticalization of *ibrew* into an equative particle, the accusative was reanalysed as a case of comparison. This comparison case as a standard marker was then extended from equatives to comparatives.

Besides, there appears to be another case in point in Classical Armenian, viz. the comparison particle *k’an*. This particle is mostly attested in comparatives, cf. (49), but also appears in comparisons of equality/similarity, as in (51), a type of comparison in which otherwise *ibew* and other particles (see Kölligan 2021: 53) are used, cf. (50). Etymologically *k’an* is generally assumed to be related to Latin *quam / quantus* < PIE *\*kwānt* ‘how much’, which suggests that it originally expressed equality/similarity so that we are dealing with another distributional shift of a standard marker according to the comparative cycle (see also Kölligan 2021: 71f.).

(49) Classical Armenian (Ps. 18(19).11 (after Kölligan 2021: 60))

<i>c'ankali ê</i>	<i>na</i>	<i>k'an</i>	<i>zoski</i>
desirable be:PRS;3SG	3SG;NOM	than	ACC.gold

‘They (God’s commendments) are more precious than gold.’

(50) Classical Armenian (Matt. 28.4, Ms. E (after Kölligan 2021: 55))

<i>eten</i>	<i>ibrew</i>	<i>zme'reals</i>
become:AOR;3PL	like	ACC.dead:ACC;PL

‘They became like dead men.’

(51) Classical Armenian (Matt. 17.19[20] (after Kölligan 2021: 71))

<i>et'ê</i>	<i>ownic'ik'</i>	<i>hawats</i>	<i>k'an</i>	<i>zhat</i>	<i>mananxoy</i>
if	have:PRS;SBJV;2PL	faith:ACC;PL	as (much as)	ACC.seed	mustard:GEN;SG

‘if you have faith like/as small as a mustard seed’



### 3.5. Greek

Ancient Greek also shows a repeated shift of standard markers according to the comparative cycle. The Ancient Greek comparative particle  $\acute{\epsilon}$ , which is used as a standard marker in comparatives (besides comparison case genitive), etymologically derives from  $*\acute{\epsilon}\acute{\epsilon}$  ‘or’ or  $*\acute{\epsilon}\acute{\epsilon}$  ‘how’ (cf. Schwyzer 1950: 565; Zeilfelder 2001: 65). In the latter case, it constitutes a first instance of a comparison particle that shifted from similatives/equatives to comparatives in Greek. Its strengthened form  $\acute{\epsilon}\acute{\upsilon}\tau\epsilon$  underwent the same change, being originally used in comparisons of equality/similarity, as in the similative in (52), but occurring in Homer already also as a comparative particle, cf. (53) (see also Schwyzer 1950: 565).<sup>27</sup>

Finally, the same development can be observed for  $h\acute{o}s$ , a standard marker that was grammaticalized on the basis of interrogative ‘how’ like German *wie* and is mainly used in comparisons of equality. According to Zeilfelder (2001: 297-318)  $h\acute{o}s$  occurs mostly in similatives as in (54), and only secondarily also in equatives. Already during later stages of Ancient Greek, however, it also starts to be used in comparatives, cf. (55). We thus see the same three steps of the development as in German starting from similatives to equatives to comparatives. The similarity of this change to the development of German *wie* was already explicitly pointed out by Hildebrand (1871: 362, fn. 2) and Lerch (1942: 354). As in German, we witness a repeated shift of standard markers according to the comparative cycle in Ancient Greek.

(52) Ancient Greek (Homer Iliad 2, 872)

<i>h\acute{o}s</i>	<i>ka\acute{i}</i>	<i>chryso\acute{n}</i>	<i>\acute{\epsilon}ch\acute{o}n</i>
who:NOM;SG;M	also	gold:ACC;SG;M	have:PTCP;PRS;NOM;SG;M
<i>p\acute{o}lem\acute{o}nd'</i>	<i>\acute{i}en</i>	<i>\acute{\epsilon}\acute{\upsilon}\tau\epsilon</i>	<i>ko\acute{u}r\acute{\epsilon}</i>
battle:ACC;SG;M;ALL	go:PST;3SG	like	girl:NOM;SG;F

‘This one came to the battle dressed in gold like a girl.’

(53) Ancient Greek (Homer Iliad 4, 277f.)

<i>t\acute{o}i</i>	<i>d\acute{\epsilon}</i>	<i>t'</i>	<i>\acute{\alpha}neuthen</i>	<i>e\acute{o}nti</i>
this:DAT;SG;M	PTCL	PTCL	far.away	be:PTCP;PRS;DAT;SG;M
<i>mel\acute{\alpha}nteron</i>	<i>\acute{\epsilon}\acute{\upsilon}\tau\epsilon</i>	<i>p\acute{i}ssa</i>	<i>pha\acute{i}net'</i>	
black:CPD;NOM;SG;N	than	pitch:NOM;SG	appear:IPF;3SG;MID	

‘To him [= a goatherd] standing in the distance, it appeared blacker than pitch.’

<sup>27</sup> De Kreij (2021: 357) treats the example given in (53) as an instance of an equative. However, morphologically (comparative form of the adjective) as well as semantically, it is clearly a comparative.



comparison particle in similatives, on the other hand, is *ut* (cf. Lerch 1942: 330f.; Tarrío 2009: 399f.), as illustrated in (58).

The distribution of *quam* in Classical Latin thus corresponds to that of *als* in 17th century German: just as *als*, which had originated in similatives, had already been largely replaced in this type of comparison by *wie* in the 17th century but still represented the main pattern in equatives (with correlate + parameter) as well as in comparatives, *quam* is hardly used in similatives in Classical Latin any more,<sup>30</sup> but has largely been replaced in this type of comparison by *ut*, whereas it is still used in equatives and at the same time already in comparatives. This suggests that *quam* also changed in three steps from similatives to equatives to comparatives.<sup>31</sup>

(56) Latin (Plaut. Stich. 109 (after Tarrío 2009: 380))

*Facile inuenis peiorem [...] quam illa fuit*  
 easily find:PRS;2SG bad:CPD;ACC;SG than that:NOM;SG;F be:PRF;3SG  
 ‘You’ll easily find [a wife] worse than she was.’

(57) Latin (Cic. Verr. II 4, 126 (after Tarrío 2009: 381))

*tam beati quam iste est*  
 as happy:NOM;PL;M as that:NOM;M be:PRS;3SG  
 ‘as happy as that one is’

(58) Latin (Plaut. Cas. 419 (after Tarrío 2009: 402))

*faciam ut iubes.*  
 do:FUT;1SG as command:PRS;2SG  
 ‘I will do as you bid me.’

(59) Latin (Petron. 38, 15 (after Tarrío 2009: 401))

*Solebat sic cenare quomodo rex.*  
 use.to:IPF;3SG so dine:INF like king:NOM  
 ‘He used to dine like a king.’

<sup>30</sup> Remnants of the use of *quam* in similatives are attested in all stages from Plautus to Late Latin (cf. Tarrío 2009: 387).

<sup>31</sup> According to Small (1924: 55) Latin *et*, which is also used as a comparison particle, represents another instance of the distributional shift from equatives to comparatives.

Besides *ut*, a new standard marker appears in Latin, viz. *quomodo*, cf. (59), which was grammaticalized on the basis of a nominal syntagm (< *quo modo* ‘in which manner’; = (iv) in fn. 13). This comparison particle is restricted to similatives in Latin, cf. Gamillscheg (1957: 743), so that, once more, the innovation starts in similatives. *Quomodo* and its successor forms *como/com/come/comme/cum* etc., which over the course of the development completely replaced *ut* in similatives, play a role as a standard marker in various Romance languages until today and also show a shift according to the comparative cycle in the further diachronic development.

### 3.7. Romance languages

For the standard markers *como/com/come/comme/cum* etc. in the individual Romance languages, continuing on from Latin *quomodo* which had been restricted to similatives (see above), in the next step a distributional shift can be observed again in the same direction as in German and other languages, viz. an extension to equatives. According to Tarrío (2009: 389), this development follows “a general trend of expressions of manner to change into expressions of degree” – a trend which is obviously part of the more general phenomenon of the comparative cycle. The resulting situation can be observed in Modern Spanish, Catalan, Portuguese and Italian, cf. Price (1990), where *como*, *com* and *come*, respectively, are used in similatives as well as in equatives, as illustrated with Italian in (60) and (61)(a). In comparatives, on the other hand, *que/che* appears as a standard marker, cf. (62)(a).<sup>32</sup> This comparison particle and its cognates in various Romance languages are generally assumed to derive from Latin *quam*, cf. Mattoso Camara (1972: 73, 214) on Portuguese, Gamillscheg (1957: 743, 748) and the etymological online dictionary of the Centre National de Ressources

---

<sup>32</sup> In these Romance languages with *como/come/com* in equatives and similatives vs. *que/che* in comparatives we are accordingly dealing with system IIa in Table 1. In equatives a comparison particle that was grammaticalized on the basis of the interrogative/relative ‘how much’, viz. *quanto*, is alternatively used in Portuguese and Italian, cf. (61)(b). In comparatives, besides the comparative particle *que/che* we partly find a functional preposition *de/di*, which is typologically equivalent to a comparison case construction, cf. (62)(b), and in most Romance languages relative constructions with the comparison-case equivalent preposition *de/di* and a pronominal element, which are grammaticalized to varying degrees: Spanish *de lo que*, Portuguese *do que*, French *de ce que*, Italian *di quanto/di quello que/di come*, Romanian *decît*, Catalan *del que* etc. (cf. Price 1990).

Textuelles et Lexicales<sup>33</sup> on French.<sup>34</sup> In these languages, *quam* > *che/que* is accordingly only kept in its historically most recent context of use, viz. comparatives, just as *als* in present-day German. Note, however, that quite a number of Latin elements introducing subordinate clauses phonologically merged in *que/che* in the Romance languages, viz. besides *quam* also *quia*, *quid*, *quod* and *quem* so that *que/che* constitutes a kind of universal complementizer in Romance languages today.<sup>35</sup>

## (60) Italian

*Maria corre come (corre) Anna.*  
 Maria run:PRS;3SG as/like run:PRS;3SG Anna  
 ‘Maria runs like Anna (does).’

## (61) Italian

- a. *Maria è (così) alta come Anna.*  
 Maria be:PRS;3SG as tall:F as Anna
- b. *Maria è alta tanto quanto Anna.*  
 Maria be:PRS;3SG tall:F as.much how.much Anna  
 ‘Maria is as tall as Anna.’

## (62) Italian

- a. *Tu sei più bella che una rosa.*  
 you be:PRS;2SG more beautiful:F than a:F rose
- b. *Tu sei più bella di una rosa.*  
 you be:PRS;2SG more beautiful:F from a:F rose  
 ‘You are more beautiful than a rose.’

The same distribution as observed for *como/com(e)* in the languages discussed above can also be found for *com(e)* in Old French: In contrast to Modern French *comme*, it was not only used in similatives but also in equatives, cf. (63). According to Lerch (1925, I: 232) and Gamillscheg (1957: 748), there is even evidence for a further

<sup>33</sup> www.cnrtl.fr/etymologie/que

<sup>34</sup> Vs. Seuren (1984: 123): *que* < Lat. *quo*, Small (1924: 53f.): French *que*/Ital. *che* etc. < Lat. *quem/quia*, only Romanian *ca* < *quam*.

<sup>35</sup> In Romanian *ca*, however, the distinct vowel of the underlying *quam* is preserved (see below).

distributional shift of this standard marker into comparatives in Old French, cf. (64) – again in line with the comparative cycle.

(63) Old French (Joinville in Paris-Langlois, Chrestom. 220 (after Lerch 1925, I: 228))

*Aussi gros come li bondons d' un tonel*  
as big as the:NOM;PL;M plug:NOM;PL of a barrel:OBL  
'as big as the plugs of a barrel'

(64) Old French (Gaydon 31 (after Gamillscheg 1957: 748))

*Vëis tu onques home plus mal mené*  
see:PST;2SG you ever man more badly behave:PTCP  
*com fu mes sires?*  
than be:PST;3SG my:NOM;SG;M lord:NOM;SG  
'Did you ever see a man more badly behaved than my Lord?'

(65) Old French (Perceval 16 (after Gamillscheg 1975: 750))

*Mais je proverai que li cuens*  
but I prove:FUT;1SG that the:NOM count:NOM  
*vaut miauz que cil ne fist*  
be.worth:PRS;3SG much:CPD than this:NOM;SG;M NEG do:PRS;3SG  
'But I will prove that the count is worth more than this one.'

(66) Old French (Hugues Capet 125 (after Lerch 1925, I: 230))

*Qui est blanche qu' aubespın*  
which:SG;M be:PRS;3SG white:SG;F as hawthorn  
'Which is white as hawthorn'

(67) Modern Standard French

*Elle est plus grande que moi.*  
she be:PRS;3SG more tall:F than I  
'She is taller than me.'

(68) Modern Standard French

*Elle est aussi grande que moi.*  
she be:PRS;3SG as tall:F as I  
'She is as tall as me.'

## (69) Modern Standard French

*Elle fait cela comme moi*  
 she do:PRS;3SG this like I  
 ‘She does this like me.’

Over the course of the further development of French, however, a noteworthy development into the opposite direction can be observed: The standard marker *que*, which in Old French constitutes the regular comparison particle in comparatives, cf. (65), starts to appear very occasionally in equatives, cf. (66) (see Lerch 1925, I: 230). Since the 13th century its use in equatives increases, cf. Gamillscheg (1957: 743) and Buridant (2000: 550), eventually replacing *comme*, which, however, is used in equatives ‘up to the Classical period’ according to Lerch (1942: 331), according to Buridant (2000: 555) until the 17th century and occurs in equatives in Colloquial French and certain dialects occasionally even today, see examples (74) and (75) below. Since the 17th century prescriptive grammarians also demand the use of *que* instead of *comme* after correlates such as *autant*, *tel*, *si* and *aussi*, i.e. in equatives.<sup>36</sup> Thus, the distributional pattern arose that we still find in Modern Standard French with *que* as a uniform particle in [+ degree] comparisons, i.e. comparatives and equatives, cf. (67) and (68),<sup>37</sup> and *comme* in similatives, cf. (69).<sup>38</sup> In contrast to Old French, where it was only the feature [± dissimilarity] that was expressed by the choice of the standard marker, it is now only the feature [± degree] so that we witness a diachronic change from type IIa to IIb in Table 1.

The reason for this development in French since the late Middle Ages in the opposite direction to the crosslinguistically predominant pattern of change captured in the comparative cycle, which we also observed in the earlier development of

---

<sup>36</sup> According to Lerch (1925, I: 228), among the prescriptive French grammarians, during the first half of the 17th century Vaugelas demands *que* instead of *comme* after *autant* and *quel*, but still accepts *comme* after *aussi* und *si*. However, during the second half of the 17th century Ménage, Corneille and Richelet demand *que* instead of *comme* in all of these contexts.

<sup>37</sup> Similarly, Occitan *que* and Walloon *k’* are used as uniform standard markers in comparatives and equatives, cf. Price (1990: 232).

<sup>38</sup> Besides *comme*, a number of other standard markers occur in similatives over the course of the French language history, viz. *ainsi comme*, later also *ainsi que* and *autant que* (16th-18th century) grammaticalized according to pattern (i) in fn. 13, but also standard markers grammaticalized according to pattern (iv), for instance *de même que* (cf. Gamillscheg 1957: 746f.).

French, can be seen in the fact that, as discussed above, numerous Latin expressions introducing subordinate clauses merged in *que*. Therefore, this ubiquitous complementizer is also increasingly generalized in comparison constructions. According to Gamillscheg (1957: 744), it was in particular the homonymy with the relative complementizer *que* in semantically and syntactically similar constructions that contributed to the distributional extension of *que* into equatives, notably the relative construction *le même que* ‘the same as’ as a bridging construction for a reanalysis of *que* as an equative particle.

Similar distributional shifts in the opposite direction, i.e. from comparatives to equatives, can arguably also be found for the cognates of *que* in a few other Romance languages: in Romanian (phrasal) equatives, besides *cum* (< *quomodo*), cognate of French *comme*, the standard marker *ca* (< *quam*), cognate of French *que* is also possible, which otherwise is used as a comparative particle, as illustrated in (70), (71) and (72) (cf. Price 1990: 200, 205). Meyer-Lübke (1899: 304), on the other hand, views Romanian *ca* as the continuation of Latin *quam* in equatives and as an innovation in comparatives, which would correspond to the usual development in the comparative cycle.

(70) Romanian (after Price 1990: 205)

<i>Scriu</i>	<i>tot</i>	<i>aşa</i>	<i>de corect</i>	<i>cum</i>	<i>citeşti</i>
write:PRS;1SG	all:M/N	as	of correctly	as	read:PRS;2SG

*tu.*  
you:NOM  
‘I am writing as correctly as you are reading.’

(71) Romanian (after Price 1990: 202)

<i>Ion</i>	<i>e</i>	<i>mai</i>	<i>mic</i>	<i>ca</i>	<i>mine.</i>
Ion	be:PRS;3SG	more	small	than	I:ACC

‘Ion is smaller than me.’

(72) Romanian (after Price 1990: 205)

<i>E</i>	<i>la</i>	<i>fel</i>	<i>de</i>	<i>înalt</i>	<i>ca</i>	<i>mine.</i>
be:PRS;3SG	after	kind	of	big	as	I:ACC

‘He is as tall as me.’



In Colloquial Italian we partly also find an extension of a standard marker from comparatives to equatives in so far as, in contrast to the standard language, *che* is not only used in comparatives but also in equatives as in (73) (cf. Price 1990: 176). The same is true of regional varieties of Italian such as Friulian and even more so Sardinian where *che* not only appears in comparatives and equatives but optionally also in similatives (cf. Haspelmath & Buchholz 1998: 315).

(73) Colloquial Italian (after Price 1990: 176)

<i>Riesce</i>	<i>tanto</i>	<i>nella pittura</i>	( <i>quanto</i>	/	<i>che</i> )
succeed:PRS;3SG	as.much	in.the painting	how.much	as	
<i>nella</i>	<i>scultura.</i>				
in.the	sculpting				
'He is as good at painting as he is at sculpting.'					

The diachronic development starting from Latin and continuing in individual Romance languages is summarized in Table 3, in which the distribution of standard markers in Latin, Old French and Modern French are given. Further Romance varieties, which of course do not constitute historical periods of French, but potentially correspond to further steps in the development are added in the shaded cells.<sup>39</sup> Thus the situation in Friulian corresponds to a stage between Old and Modern French, while the distribution in Sardinian corresponds to a potential further stage in the development.

Note that while in these particular Romance languages, the original development according to the comparative cycle is turned back, as it were, due to the ubiquitous use of the complementizer *que/che* resulting in diachronic stages with a uniform

---

<sup>39</sup> This scenario holds provided that there was indeed an intermediate diachronic stage at which *como/com(m)e* etc. was used in similatives and equatives as in Old French and Modern Spanish, Portuguese and Standard Italian, and *que/che* in Friulian, Sardinian etc. equatives and partly similatives does not simply constitute a direct continuation of Latin *quam* in these contexts, in which case these languages would not show a development in the opposite direction to the typical comparative cycle, after all. Considering the distribution in Latin, where *quam* was hardly used in similatives any more, this alternative scenario seems unlikely, but as discussed above, it is in fact proposed for Romanian by Meyer-Lübke (1899: 304).

standard marker in [+ dissimilarity] or in [+ degree] comparisons (type IIa and IIb in Table 1, respectively) or even for all three types of comparison as partly in Sardinian (type I in Table 1), this change, even if it takes place in the opposite direction to the usually observed comparative cycle, still evolves in the stepwise manner predicted by the markedness hierarchy of comparisons with equatives taking up an intermediate position between comparatives and similatives.

	similative	equative	comparative
Latin	<i>ut</i> / <i>quomodo</i>	<i>quam</i>	<i>quam</i>
Old French (≈ Spanish, Portuguese, Catalan, Italian)	( <i>quomodo</i> >) <i>come</i>	<i>come</i>	( <i>quam</i> >) <i>que</i> (rarely <i>come</i> )
Friulian (≈ Romanian)	<i>come</i>	<i>come/que</i>	<i>que</i>
Modern French	<i>comme</i>	<i>que</i>	<i>que</i>
Sardinian	<i>comente/che</i>	<i>che</i>	<i>che</i>

Table 3: Development of standard markers in Romance languages.

Interestingly, in Modern non-standard varieties of French, especially in Western parts of France, and in North-American French, for instance in Nova Scotia and Louisiana, *comme* and its equivalents are not only used in similatives but also in equatives, as (74) and (75) illustrate. It remains to be ascertained whether this represents a relict form, as Lerch (1925, I: 228) and Neumann-Holzschuh & Mitko (2018: 740) suggest, or whether this use represents a recent, secondary development. The fact that it is socially very marked as decidedly lower-class or very informal speech seems to support the latter scenario, in which case there would be evidence for a renewed development into the typical direction according to the comparative cycle in French.

(74) North-American French: Nova Scotia (after Neumann-Holzschuh & Mitko 2018: 743)

*I est                    aussi haut        comme    Pierre.*  
 he be:PRS;3SG      as           high        as           Pierre  
 ‘He is as tall as Pierre.’

(75) Modern Non-Standard French (after Gadet 1992: 93)

<i>Il</i>	<i>est</i>	<i>aussi</i>	<i>grand</i>	<i>comme</i>	<i>moi</i>
he	be:PRS;3SG	as	big	as	I

‘He is as tall as me.’

### 3.8. Slavic languages

Slavic languages show evidence of developments of standard markers in comparisons according to the comparative cycle, too. In Polish – as in German and many other European languages – an interrogative-based comparison particle is used in equatives and similatives, viz. *jak*, cf. (76) and (77). In comparatives, on the other hand, the comparison particle *niz*, cf. (78)(a) and (79)(a), or more rarely (and only in phrasal comparatives) the particle *od* in combination with genitive case mark the standard of comparison, cf. (78)(b) and (79)(b).<sup>40</sup> However, in Colloquial Polish the comparison particle *jak* also appears in comparatives if they are negated, cf. (79)(c).<sup>41</sup> Adverbial function of the parameter and the occurrence of negative concord, cf. (80), increase the acceptability of *jak* in comparatives in Polish.

(76) Polish

<i>Anna</i>	<i>idzie</i>	<i>tak</i>	<i>(samo)</i>	<i>szybko</i>	<i>jak</i>	<i>Maria.</i>
Anna:NOM	walk:PRS;3SG	so	same	fast	as	Maria:NOM

‘Anna walks as fast as Maria.’

<sup>40</sup> This goes back to a comparison case construction with the case-equivalent functional preposition *ot* > *od* ‘from’ (corresponding to the Italian comparison-case equivalent preposition *di*, which occurs in comparatives as an alternative to *che*, cf. (62)(b)). For an analysis of phrasal as opposed to clausal comparatives in Slavic see Pancheva (2006; 2010).

<sup>41</sup> “Przyimek *jak* jest używany w konstrukcjach porównawczych, zawierających w pierwszym, zaprzeczonym członie przymiotnik w stopniu wyższym, np. *Nie ma nic lepszego jak mocna herbata*. Natomiast niepoprawne jest używanie w takich porównaniach przyimka *jak* z członem pierwszym niezaprzeczonym.” (‘The pronoun *jak* is used in comparative constructions that contain a comparative form of an adjective in the first, negated clause, e.g. *There is nothing better than* [= *jak* ‘as/like’, A.J.] *strong tea*. However, it is ungrammatical to use the pronoun *jak* if the first clause is not negated.’ Krystyna Długosz-Kurczabowa, University of Warsaw, <https://sjp.pwn.pl/poradnia/haslo/lepszy-niz-i-lepszy-od;10270.html>).

(77) Polish

<i>Anna</i>	<i>idzie</i>	<i>tak</i>	<i>(samo) jak</i>	<i>Maria.</i>
Anna:NOM	walk:PRS;3SG	so	same as	Maria:NOM

‘Anna walks like Maria.’

(78) Polish

a. <i>Anna</i>	<i>idzie</i>	<i>szybciej</i>	<i>niż</i>	<i>Maria.</i>
Anna:NOM	walk:PRS;3SG	fast:CPD	than	Maria:NOM
b. <i>Anna</i>	<i>idzie</i>	<i>szybciej</i>	<i>od</i>	<i>Marii.</i>
Anna:NOM	walk:PRS;3SG	fast:CPD	from	Maria:GEN

‘Anna walks faster than Maria.’

(79) Polish

a. <i>Anna</i>	<i>nie</i>	<i>idzie</i>	<i>szybciej</i>	<i>niż</i>	<i>Maria.</i>
Anna:NOM	NEG	walk:PRS;3SG	fast:CPD	than	Maria:NOM
b. <i>Anna</i>	<i>nie</i>	<i>idzie</i>	<i>szybciej</i>	<i>od</i>	<i>Marii.</i>
Anna:NOM	NEG	walk:PRS;3SG	fast:CPD	from	Maria:GEN
c. <i>Anna</i>	<i>nie</i>	<i>idzie</i>	<i>szybciej</i>	<i>jak</i>	<i>Maria.</i>
Anna:NOM	NEG	walk:PRS;3SG	fast:CPD	than	Maria:NOM

‘Anna doesn’t walk faster than Maria.’

(80) Polish

<i>Nigdy</i>	<i>nie</i>	<i>kupiłam</i>	<i>więcej</i>	<i>jak</i>	<i>5kg ziemniaków.</i>
never	NEG	buy:PST;1SG	many: CPD	than	5kg potato:GEN;PL

‘I never bought more than 5kg potatoes.’

The usual similitive and equative standard marker *jak/kak* ('how', 'like', 'as') is also possible in comparatives in other Slavic languages. Thus in Russian, the particle *kak* also appears in negated comparatives, cf. (81) and (82), where otherwise the standard of comparison is marked by the particle *čem* (in archaic speech also *neželi*) or by comparative case (genitive), see also Jensen (1934: 124).

In Ukrainian, *jak* is already generally acceptable instead of *niž* as a standard marker in comparatives. (83) and (84) illustrate *jak* in a similitive and in a comparative. In this language, negated comparatives presumably also acted as bridging contexts for

the distributional extension of *jak* into comparatives according to the comparative cycle, as early examples of comparative *jak* from around 1800 suggest, cf. (85).

## (81) Russian

*nikto inoj kak ja*  
 nobody:NOM other than I:NOM  
 'none other than me'

## (82) Russian (NKRJa: Izvestija, 2002/09/27)

*on ne pridumal ničego lučše, kak posporit'*  
 he:NOM NEG concoct:PST;SG;M nothing:GEN good:CPD than quarrel:INF  
*s desantnikami*  
 with paratrooper:INSTR;PL  
 '[...] he did not concoct anything better than to quarrel with the paratroopers'

## (83) Ukrainian (Je. Hucalo (after Horodens'ka 2017: 657))

*Ja vže vil'nyj, jak ptaxa v nebi*  
 I:NOM already free:NOM;SG;M like bird:NOM in sky:LOC  
 'I am already free as a bird in the sky.'

## (84) Ukrainian (Ukr. prysliv'ja (after Horodens'ka 2017: 658))

*Lipše ves' vik divuvaty, jak z*  
 good:CPD whole life:ACC be.a.virgin:INF than with  
*neljubom prožyvaty*  
 unloved:INSTR;M live:INF  
 'Better to be a spinster the whole life than to live with an unloved one'

## (85) Ukrainian (Hryhoryj Kvitka-Osnov'janenko (1778-1843) (after Medvedjev 1962: 80))

*Ne bulo j na usij slobodi*  
 NEG be:PST;SG; M even in whole:LOC settlement:LOC  
*durnišoho, jak Parxim Šereverten'*  
 stupid:CPD;GEN;SG;M than Parxim Šereverten'  
 'In the whole settlement there was no-one more stupid than Parxim Šereverten'.'

### 3.9. Hungarian

Even in Non-Indo-European languages, distributional shifts of standard markers according to the comparative cycle can be found, for instance in Hungarian. In all three types of comparison constructions considered here, the comparison particle *mint* is used in Hungarian, cf. (86)-(88) (= type I in Table 1).<sup>42</sup> This particle, too, originates in the interrogative adverb ‘how’ (cf. Heine & Kuteva 2002: 177), and accordingly must have been used in similatives and equatives first and has subsequently been extended to comparatives, similar to *wie* in colloquial and dialectal German.<sup>43</sup>

(86) Hungarian

<i>Néz</i>	<i>mint</i>	<i>egy</i>	<i>bárány</i>
look:PRS;SBJV;3SG	like	a	lamb

‘He looks like a lamb.’

(87) Hungarian

<i>János</i>	<i>olyan</i>	<i>kicsi</i>	<i>mint</i>	<i>Péter</i>
Janos	as	small	as	Peter

‘Janos is as small as Peter.’

(88) Hungarian

<i>János</i>	<i>kisebb</i>	<i>mint</i>	<i>Péter</i>
Janos	small:CPD	than	Peter

‘Janos is smaller than Peter.’

In the history of Hungarian, there is even evidence for a repeated change according to the comparative cycle. The original comparison particle in Old Hungarian is *hogy*, which was beginning to be replaced in equatives by the more recent standard marker

---

<sup>42</sup> The phenomenon that the same comparison particle appears in similatives, equatives and comparatives can also be observed in further Uralic languages, viz. Estonian *kui*, Finnish *kuin*. Whether these particles historically show the same shifts according to the comparative cycle as observed for Hungarian needs to be established in future research.

<sup>43</sup> Besides using a comparison particle, the standard may also be marked by comparison case in Hungarian comparatives, viz. by adessive case, a locative case expressing proximity to an object (e.g. *János kisebb Péter-nél*. - Janos small:CPD Peter:ADESS ‘Janos is smaller than Peter.’)

*mint* already in Old Hungarian, cf. Haader (2003) and Bacskai-Atkari (2014).<sup>44</sup> In historical Hungarian, *hogy* constitutes a kind of uniform comparison particle in all types of comparison, too.<sup>45</sup> Etymologically, it also meant ‘how’ so that *hogy*, too, must have undergone a distributional shift from similatives and equatives to comparatives.

### 3.10. Chinese

Evidence for the comparative cycle can also be found in unrelated Non-European languages such as Chinese. Heine & Kuteva (2002: 256f.) discuss Chinese besides German (comparison particle *wie*) as an example for the development from equatives to comparatives and characterise the respective developments in both languages as directly comparable. The change in Chinese does not only evolve in the same direction as in German but has repeatedly taken place, too, in fact at least three times, which once more corroborates the analysis that we are dealing with an instance of cyclical language change. According to Heine & Kuteva, who base their discussion on Sun (1996), the Late Archaic Chinese and Han Chinese verbal equative marker *bi* ‘to compare with, be like, imitate’, cf. (89), developed into a marker in comparatives in Late Medieval Chinese of the 8th and 9th century with the meaning of ‘more than’, which it still has in Modern Mandarin Chinese today, cf. (90).<sup>46</sup>

The second lexeme that has undergone a development from comparisons of equality/similarity to comparisons of inequality/dissimilarity in Chinese is *ru*. Starting with a meaning of ‘to resemble/be like’, cf. (91), it changed into a standard marker in comparatives in Early Mandarin Chinese, cf. (92). The third instance of this type of syntactic-semantic distributional shift in Chinese concerns the lexeme *xiang*, which again used to mean ‘to resemble/be like’ and later also became used as a standard marker in comparatives.

(89) Old Chinese (Mengzi Gongsun Chou shang (Heine & Kuteva 2002: 256, after Sun 1996: 39))

<i>er</i>	<i>he</i>	<i>ceng</i>	<i>bi</i>	<i>yu</i>	<i>yu</i>	<i>shi?</i>
2SG	how	STRESS	compare	1SG	YU	3SG

‘How (dare) you compare me to him?’

<sup>44</sup> Up to Middle Hungarian, the combination *hogy mint* is attested.

<sup>45</sup> In comparatives, *hogy* is, however, almost only found together with the negative/polarity particle *nem/sem*, later it is also occasionally found without the negation particle in the combination *hogy mint*.

<sup>46</sup> According to Sun (1996: 38f.) the steps of the development are as follows: Old Chinese: verb ‘to compare’ > Middle Chinese: equative preposition ‘as/like’ > comparative particle.

(90) Modern Mandarin Chinese (Heine & Kuteva 2002: 256, after Sun 1996: 38)

*ta bi meimei piaoliang.*  
3SG than sister beautiful  
'She is prettier than (her) sister.'

(91) Early Mandarin Chinese (Yuan kann zaju sanshi zhong Yu Shang Wang (Heine & Kuteva 2002: 256, after Sun 1996: 40))

*xiong-jiujiu de gongren ru hu lang.*  
gallantly PTCL policemen resemble/like tiger wolf  
'Arrogant policemen are like tigers and wolves.'

(92) Early Mandarin-Chinese (Yuan kann zaju sanshi zhong Yu Shang Wang (Heine & Kuteva 2002: 257, after Sun 1996: 40))

*Chi le xie popei chunno sheng ru*  
eat ASP some fermented spirit better than  
*yu xie qiongjiang.*  
jade liquid wine  
'(I) took some fermented wine, better than the best wine.'

Heine & Kuteva (2002: 257) state that further data besides those from the two languages they discuss, viz. German and Chinese, would be required to corroborate the grammaticalization path from equative markers (or 'to resemble/be like') to comparative markers. They speculate that the development in Chinese constitutes an instance of a verb with a salient feature becoming a grammatical marker for that feature. Note, however, that this salient feature would be similarity/equality, whereas the meaning of the resulting grammatical marker is one of dissimilarity/inequality. This can be explained if the development is seen as an instance of the comparative cycle turning a marker in comparisons of equality/similarity (standard markers or other linguistic expressions of similarity such as the discussed Chinese verbs) into a marker in comparisons of inequality/dissimilarity, i.e. comparatives, a phenomenon which is not limited to German and Chinese, but occurs in numerous languages as the discussion in the previous sections has shown.

#### 4. The comparative cycle: Characteristics and causes

As demonstrated above, one can crosslinguistically observe a diachronic shift of comparison markers (notably comparison particles) according to the comparative



cycle. While the development may occasionally proceed in a different way in individual languages due to peculiarities of the language system (e.g. in certain Romance languages) the general direction of change is that from equality to inequality, more specifically from similatives to equatives to comparatives. Similatives not only stand out as a typical context for innovation of standard markers (cf. fn. 13), but are also the least marked type of comparison constructions, which will help to explain the steps and typical direction of change in the comparative cycle.

#### ***4.1. Cyclical language change***

The observed changes constitute a stepwise and repeated development into the same direction, i.e. an instance of cyclical language change. Indeed, as we saw above, in several languages we witness repeated shifts of comparison markers from similatives to equatives to comparatives. The fact that language change may proceed in a cyclical fashion (or spiral) has already been noted by von der Gabelentz (1891: 241ff.) and Meillet (1912) in relation to the future cycle in Romance languages, i.e. the repeated development of future markers from synthetic to analytic to synthetic again etc. (cf. also Hopper & Traugott 2003: 9). Another classical instance of cyclical change is the repeated change of negation particles from clitic negator to clitic + free negator to free negator to clitic negator again, described by Jespersen (1917) and investigated in much subsequent work (van Kemenade 1999, 2000; Rowlett 1998; Jäger 2008; van der Auwera 2010; Willis et al. eds. 2013 among others).

In recent years there has been an increased interest in cyclical change (cf. for instance van Gelderen ed. 2009, 2011, ed. 2016; Breitbarth & Jäger in prep.) because this kind of change brings out the systematic, language-independent side of change and thus allows us to investigate general crosslinguistic principles of language change. There have been in-depth studies of various further cycles besides the future cycle and Jespersen's cycle such as the subject-agreement cycle between subject pronouns and verbal agreement morphology (van Gelderen ed. 2009, 2011, ed. 2016), the copula cycle between a demonstrative or a main verb and a copula (cf. Lohndal 2009), the definiteness cycle from demonstratives to definite articles to nominal case or derivation morphemes (van Gelderen 2007, 2011), or the relative cycle from relative pronoun to relative particle to particle + new pronoun, finally simple pronoun again etc. (van Gelderen 2004).

For several instances of cyclical change it has been suggested that phonological reduction of the original marker and subsequent strengthening and finally replacement by another marker drive the development. While phonological reduction can often also be observed in the diachronic development of comparison particles and other standard markers, phonological reduction and subsequent strengthening does not seem to be the main driving force behind the comparative cycle, for instance for the replacement of German *denn* by *als* in comparatives (*denn* is still used in other functions without being perceived as phonologically too weak) or *als* by *wie* in equatives (again *als* survives in other functions, notably as a comparative particle, i.e. cannot be considered phonologically too weak to be used as an equative particle).

However, phonological reduction is not the only possible cause for cyclical change. Another cause may be a recurrent reanalysis of the same kind. This may for instance be due to markedness reversal as in the case of the repeated reanalysis of plural forms as singular forms and formation of a new plural form e.g. with the noun for ‘pear’ in German (Lat. SG *pirum* – PL *pira* > OHG/MHG SG *bira/bire* – PL *bir(e)n* > Modern German SG *Birne* – PL *Birnen*). Furthermore, desemanticization of an originally pleonastic, emphatic construction turning the respective expression into the regular, non-emphatic marker, i.e. largely semantic-pragmatic developments have been suggested to be the cause behind certain instances of cyclical change.

With respect to the comparative cycle, semantic and syntactic reanalysis as well as loss of emphasis also play a role especially in relation to the grammaticalization of new standard markers, which typically starts in similatives, contributing to the directionality of the change. The primary explanation, however, will be argued to lie in the markedness relations of comparison constructions in combination with linguistic economy at different levels of the language system.

#### **4.2. Chain shifts or functional overload as causes?**

Let us first, however, revise two other causes that have been suggested in the literature specifically with respect to the distributional shift of standard markers in comparisons. The first one is a mechanism of chain shift as it is also assumed in explanations of phonological developments or of semantic changes within so-called word fields. Grimm (DWB 1: 248; 250f.) and Lerch (1942: 349) assume for the development in German that the use of *wie* as a comparison particle caused a push chain of repeated replacement among the comparison particles. However, the

timeline of the developments disproves this hypothesis (for a detailed discussion see Jäger 2018: 401-404). As discussed in sect. 3.1.1, German *als* replaced *denn* in comparatives during the 17th century – well-before it was ‘pushed out’ of equatives by *wie*, which only took place during the 19th century. Moreover, the extension of *wie* into comparatives since the 18th and especially 19th century cannot be linked to any other element pushing it out of similatives and equatives, as it still constitutes the main pattern in these types of comparisons until today, so that the comparative cycle cannot be explained by chain shifts.

The second hypothesis that is occasionally raised in the literature sees the cause of the change within the respective comparison particle itself. It is assumed that a certain item stops to be used as a comparison particle once it develops too many other meanings or functions, i.e. due to ‘functional overload’. Thus the replacement of the German comparative particle *denn* by *als* has been linked to the rise of causal *denn* since the 15th century (DWB 29: 1484; Lerch 1942: 355, 359; Dücker 1961: 215) and the replacement of the equative particle *als* by *wie* to the rise of predicative *als* or an increased use of *als* as a temporal complementizer since late MHG, respectively (DWB 29: 1471; Lerch 1942: 349; Dücker 1961: 207).<sup>47</sup>

A close investigation of the developments, however, shows that this hypothesis does not hold, either (for a detailed discussion see Jäger 2018: 404-418). The increase of German *dann/denn*, for instance, which had already been used in various functions (conditional, temporal etc.) during OHG and MHG, as a causal complementizer took place several centuries before it stopped being used as the main comparison particle in comparatives. *Als* on the other hand considerably increased its functions during MHG and ENHG (modal demonstrative adverb, temporal, conditional, causal and concessive complementizer, relative particle, predicative conjunction, coordinating conjunction) and at the same time as well as during the directly ensuing time period it was even increasingly used as a comparison particle, being extended from one type of comparison to the next. Similarly, *wie* took on more and more functions during ENHG and early NHG (interrogative/relative adverb, complementizer ‘that’, concessive, temporal and causal complementizer, relative particle, coordinating conjunction), yet at the same time it was extended into further types of comparison at the expense of *als*.

---

<sup>47</sup> Similarly, the replacement of the Dutch comparative particle *dan* by *als* during the 16th and 17th century has been considered as a consequence of the rise of *dan* in the sense of ‘but’ (van der Horst 2008: 993).

In fact, the comparative cycle regularly leads to an increase of polysemy and functional load of lexemes as they often take over the function of standard marker in further types of comparison at the same time or directly after an increase of other functions. Lexemes with the single function of equative particle, on the other hand, such as the historical German *alsam*, occasionally attested in MHG, or *inmassen*, sometimes found in ENHG, should be expected to gain importance in this exclusive function and be stably used – instead they quickly disappeared in contrast to the highly polysemous *wie*, which was nevertheless very successful diachronically in all types of comparisons. The fact that polysemy and ‘high functional load’ may even be conducive to an increased use as a comparison particle is also obvious from the increasing use of *que/che* as a comparison particle in various Romance languages (see sect. 3.7). Besides failing with regard to the exact development in concrete cases, these hypotheses fall short of explaining the systematic and crosslinguistic nature of the change by seeking the cause in language-specific and lexeme-specific developments.

#### ***4.3. Economy and markedness as the causes of the comparative cycle***

An explanation of the comparative cycle that captures its lexeme-independent, crosslinguistic and systematic nature must build on universal principles and mechanisms of language change. I would like to argue that markedness and linguistic economy play a crucial role as causes of the observed developments. The directionality and individual steps of the comparative cycle can be argued to result from the characteristics of the different types of comparison constructions, in particular the markedness hierarchy formulated in (2), in combination with linguistic economy at different levels of the language system.

At the level of the lexicon, economy repeatedly leads to the rise of a uniform standard marker for several types of comparison, as can be observed in the diachronic development of many of the languages discussed above. Since the difference between, for instance, [+ dissimilarity] and [- dissimilarity] comparisons is marked by means of the parameter marker (an inflectional morpheme on the adjective/adverb or a free morpheme accompanying it) in the languages under discussion, it need not be marked

by a different standard marker in addition.<sup>48</sup> This is also reflected in the formal semantic analysis of comparison constructions (von Stechow 1984; Heim 1985, 2000), according to which the equative or comparative semantics resides in the respective parameter marker, whereas the comparison particles themselves are assumed to be semantically empty and are therefore deleted at the level of Logical Form. Against this background a recurrent reduction of the functional lexicon w.r.t. lexical entries for comparison particles for reasons of economy is to be expected.

Relevant bridging contexts that facilitate this process of using one particle in several types of comparison are negated equatives or equatives containing expressions of multiples, thus referring to a relation of dissimilarity as a whole, as well as negated comparatives expressing similarity rather than dissimilarity. This is also evident in several of the languages discussed above. Thus a construction such as the equative [*not [as big as]*] may be read as a comparative [*not as big = smaller*] *than*], or [*twice [as big as]*] as [*twice as big = bigger*] *than*] giving rise to the use of the comparative particle in equatives. Conversely, a comparative such as [*not [bigger than]*] may be taken as [*not bigger = as big*] *as*], giving rise to the use of the equative standard marker in comparatives. This can be analysed as an instance of re-bracketing resulting in narrow scope of the respective operator (negation, multiple) over the parameter rather than wide scope over the entire comparison. Indeed, there is evidence for this phenomenon both synchronically as well as historically in various languages such as the Baltic and Slavic languages discussed in sections 3.3 and 3.8, and partly also English cf. sect. 3.1.2. In the history of German, cf. sect. 3.1.1, too, the first uses of the equative particle *als* in comparatives appear in these kinds of constructions, cf. example (11) with a negated comparative, but also the exceptional use of the comparative particle *thanne/dann* in equatives.<sup>49</sup> While this explains how a shift of

---

<sup>48</sup> Note that accordingly the comparative cycle is less likely to occur in languages which only mark the difference between the different types of comparison by means of the standard marker without employing a parameter marker.

<sup>49</sup> An early example from OHG of *thanne* in an equative with an expression of multiples ('twice as ... as') is the following:

*uuánda óuh sélbez taz sáng . nôte stîgen sól [...] únz tára sîn hóhi gât . ih méino uuílôn ióh ze demo áhtodên búohstábe . dér **zuúualt** líutet . **tánne** dér búohstáb . ze démo iz ánafieng*

'because the melody itself shall necessarily rise up to its high point, I mean sometimes even to the eighth tone which sounds twice as high as the tone at which it startet.' (Notker Mus. IV, 16)

markers from one type of comparison to another is possible in principle and will be favoured due to lexicon economy, the typical directionality and steps of the comparative cycle are not yet explained as, on this basis, a shift in both directions may occur.

A central counterpart of lexicon economy that also partly contributes to the directionality of the change is economy at the syntactic and semantic/pragmatic level, which repeatedly leads to the grammaticalization of new standard markers. In this respect, similatives stand out as a primary context of innovation due to their syntactic and semantic characteristics: the typical lack of a parameter expression leads to a frequent adjacency of elements such as the correlate or an element meaning ‘same’ etc. in the superordinate clause and the comparison particle, which may give rise to the grammaticalization of new comparison particles by univerbation with different types of preceding elements (cf. fn. 13). Treating the respective expressions as one syntactic item is more economical syntactically than a syntactic structure in which they are two separate items (see also Weiß 2019: 533-537).

This is further enhanced by semantic/pragmatic bleaching. Part of the semantics of similatives (and equatives) is a certain granularity or tolerance range (cf. Umbach & Gust 2014): For x to be like y it is not necessary that it is identical with respect to all contextually salient features but that the respective feature values are close enough. Therefore, a very close or even exact identity is typically expressed by additional emphatic markers such as ‘fully’ or ‘exactly’ whose emphatic character may be bleached over time so that they are grammaticalized as part of the usual standard marker (cf. (iii) in fn. 13, among others leading to the grammaticalization of German *als* and English *as*). In this respect loss of emphasis also plays a role in the comparative cycle as it does in other types of cyclical change (see sect. 4.1).

Another syntactic factor that contributes to similatives being a preferred context for the grammaticalization of new comparison particles is the fact that similatives are very close syntactically to wh-constructions, notably to free modal (co-)relatives. In many European languages including several languages discussed in sect. 3 above, one can thus observe the grammaticalization of modal interrogative/relative adverbs

---

Similar examples are found in various languages, consider the following example from present-day English:

*The symbol from the last position does not appear anymore in the calculation because the last number is **twice as big than** prior, it is the same as having two symbols in the prior position.*

(Chester Litvin: Advance brain stimulation by psychoconduction (2012: 41); after Google Books)

('how') into comparison particles, which is again driven by syntactic economy: A syntactic head such as a comparison particle is more economical than a full syntactic phrase such as a relative/interrogative adverb (for a detailed discussion of the structural syntactic changes involved in the grammaticalization of a modal *wh*-adverb into a comparison particle as well as those involved in the diachronic development of individual comparison particles during the comparative cycle see Jäger 2018: ch.8.1, 2019: 7-19). The fact that, for these syntactic and semantic reasons, similatives form a primary context for the grammaticalization of new standard markers makes them a typical starting point of the development and thus contributes to the directionality of the comparative cycle. This crucially coincides with the predictions of the markedness hierarchy of comparison constructions given in (2), which not only explains the directionality of the development, starting from similatives, but also explains the individual steps of the change.

Similatives represent the least marked type of comparison constructions, being characterized by [- dissimilarity] as well as [- degree]. This is corroborated by the observation that they show the least specific, least grammaticalized markers crosslinguistically.<sup>50</sup> Data from language acquisition indicates that similatives are also the type of comparisons that is acquired earliest (cf. Hohaus 2015). Moreover, similatives constitute the most frequent type of comparisons in corpora (cf. for instance Zeilfelder 2001: 474 for the early IE languages Hittite, Vedic and Ancient Greek, Hahnemann 1999: 29 for a Modern German newspaper corpus, Jäger 2018: 433 for a historical German corpus covering OHG, MHG, ENHG and early NHG). The pattern used in similatives is thus easily transferred onto other types of comparison. This extension or shift takes place gradually from less marked to more marked contexts in line with standard assumptions of markedness theory on language change (cf. Wurzel 1994: 43f.). Being characterized by [- dissimilarity] and [+ degree], equatives take up an intermediate position between similatives and comparatives regarding markedness. The prediction that the shift of standard markers to equatives constitutes an intermediate step in the comparative cycle is borne out as the discussion in section 3 has shown. The final step in the typical development is the distributional shift of standard markers to comparatives, which constitute the most

---

<sup>50</sup> Similarity also seems cognitively unmarked in contrast to dissimilarity, i.e. similarity comes with less cognitive effort. This fact is employed for instance in perception studies with infants, whose sucking rate is stable if they perceive two items they do not discriminate, i.e. items they perceive as similar/equal, whereas it increases if they perceive an item as different from, i.e. dissimilar to, the previous one.

marked type of comparisons being characterized by [+ dissimilarity] and [+ degree].

## 5. Conclusion

The syntactic-semantic shift of standard markers from comparisons of similarity/equality to those of dissimilarity/inequality, more specifically from similatives to equatives to comparatives, is found in the history of many languages. In several languages, there is even evidence that this kind of change occurred several times and thus constitutes a cycle: the comparative cycle.

The crosslinguistic and systematic nature of the comparative cycle can be explained as an effect of economy at the level of the lexicon leading to the use of one standard marker for several types of comparison, and economy at the syntactic and semantic/pragmatic level leading to the grammaticalization of new standard markers especially in similatives. Their shift into further types of comparisons, viz. equatives and finally comparatives, can be understood against the background of the markedness hierarchy of comparison constructions as an instance of natural language change. In order to deepen our understanding of the comparative cycle, its characteristics and causes, further detailed diachronic investigations would be useful, covering further languages and including data on all three types comparisons discussed, viz. comparatives, equatives and notably similatives.

## Acknowledgements

I would like to thank a number of colleagues for data and/or helpful comments on a number of languages: Anne Breitbarth, Koen van Hooste and Chris de Wulf on BSD and Flemish varieties, Pia-Mareen van de Kerkhof on Standard Dutch, Eugen Hill on PIE and Lithuanian, Daniel Kölligan on Old Armenian and Ancient Greek, Marco Coniglio and Chiara Gianollo on Italian, Martin Becker on Old French, Pierre Larrivé on colloquial and Northern-American French, Barbara Tomaszewicz-Özakin on Polish, Daniel Bunčić on Russian and Ukrainian, Julia Bacskai-Atkari and Melani Wratil on Hungarian, and Barbara Meisterernst on Chinese.

## Abbreviations

1 = 1<sup>st</sup> person

2 = 2<sup>nd</sup> person

GEN = genitive

IMP = imperative

PL = plural

PRF = perfect



3 = 3 <sup>rd</sup> person	INF = infinitive	PRS = present
ACC = accusative	INSTR = instrumental	PST = past
ADESS = adessive	IPF = imperfect	PTCL = particle
ALL = allative	LOC = locative	PTCP = participle
AOR = aorist	M = masculine	Q = question particle
ASP = aspect	MID = middle	REFL = reflexive
AUX = auxiliary	N = neuter	SBJV = subjunctive
CPD = comparative degree	NEG = negation	SG = singular
DAT = dative	NOM = nominative	SPD = superlative degree
F = feminine	OBL = oblique	VOC = vocative
FUT = future	OPT = optative	

### Languages

BSD = Belgian Standard Dutch	MHG = Middle High German	OHG = Old High German
ENHG = Early Middle High German	NHG = New High German	PIE = Proto Indo-European

### References

- Alrenga, Peter. 2007. *Dimensions in the semantics of comparatives*. PhD Dissertation, University of California.
- Bacsikai-Atkari, Julia. 2014. Cyclical change in Hungarian comparatives. *Diachronica* 31. 465-505.
- Breitbarth, Anne & Jäger, Agnes. In prep. Cyclical change. In Bousquette, Joshua & Pickl, Simon (eds.), *The Oxford handbook of the German language*. Oxford : Oxford University Press.
- Buridant, Claude. 2000. *Grammaire nouvelle de l'ancien français*. Paris: Sedes.
- Camara Mattoso, Joaquim Jr. 1972. *The Portuguese language*. Chicago: Univ. of Chicago Press.
- CHEL II = Blake, Norman (ed.). 1992. *The Cambridge history of the English language*. Vol. 2. 1066-1476. Cambridge: Cambridge University Press.
- CHEL III = Lass, Roger (ed.). 2006. *The Cambridge history of the English language*. Vol. 3. 1476-1776. Third edition. Cambridge: Cambridge University Press.
- de Kreij, Nina. 2021. Ancient Greek. In Keydana, Götz & Hock, Wolfgang & Widmer, Paul (eds.), *Comparison and gradation in Indo-European* (= Mouton Handbooks of Indo-European Typology, 1), 349-384. Berlin: de Gruyter.

- Delbrück, Berthold. 1888. *Altindische Syntax*. Halle: Verlag der Buchhandlung des Waisenhauses.
- Dückert, Joachim. 1961. Das geschichtliche Verhältnis des vergleichenden *als* und *wie*. *PBB* 83. 205-230.
- Dunkel, George E. 2014. *Lexikon der indogermanischen Partikeln und Pronominalstämme*. Vol. 1: Einleitung, Terminologie, Lautgesetze, Adverbialendungen, Nominalsuffixe, Anhänge und Indices; Vol. 2: Lexikon. Heidelberg: Winter.
- DWB = Grimm, Jakob & Grimm, Wilhelm. 1854–1961. *Deutsches Wörterbuch*. 33 vols. Leipzig: Hirzel (Reprint: Deutscher Taschenbuchverlag, 1984).
- Gadet, Françoise. 1992. *Le français populaire*. Paris: PUF.
- Gamillscheg, Ernst. 1957. *Historische französische Syntax*. Tübingen: Niemeyer.
- Haader, Lea. 2003. Az ómagyar kor. Mondattörténet: Az összetett mondat. In Kiss, Jenő & Pusztai, Ferenc (eds.), *Magyar nyelvtörténet*, 500-560. Budapest: Osiris Kiadó.
- Hahnemann, Suzan. 1999. *Vergleiche im Vergleich. Zur Syntax und Semantik ausgewählter Vergleichsstrukturen mit "als" und "wie" im Deutschen*. (= Linguistische Arbeiten, 397). Tübingen: Niemeyer.
- Haspelmath, Martin & Buchholz, Oda. 1998. Equative and similitive constructions in the languages of Europe. In Van der Auwera, Johan & O Baoill, Dónall (eds.), *EUROTYP. Adverbial constructions in the languages of Europe, vol. 3* (= Empirical Approaches to Language Typology, 20-3), 277-334. Berlin: de Gruyter.
- Heim, Irene. 1985. *Notes on comparatives and related matters*. Ms. University of Texas, Austin.  
(Online:  
<http://semanticsarchive.net/Archive/zc0ZjY0M/Comparatives%2085.pdf>).
- Heim, Irene. 2000. Degree operators and scope. In Jackson, Brendan & Tanya Matthews (eds.), *Proceedings of Semantics and Linguistic Theory (SALT) 10*, 40-64. Cornell University Ithaca  
(Online:  
<http://journals.linguisticsociety.org/proceedings/index.php/SALT/article/view/3102/>).
- Heine, Bernd & Kuteva, Tania. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.

- Hildebrand, Rudolf. 1871. Rezension zu Ph. Dietz: Wörterbuch zu Dr. Martin Luthers deutschen Schriften. Leipzig: Vogel, 1870. *Zeitschrift für deutsche Philologie* 3. 358-365.
- Hohaus, Vera. 2015. A semantics for degree and property equatives. Ms. Paper presented at the conference “Comparison constructions”, 21/12/2015. University of Cologne.
- Horodens'ka, Kateryna Hryhorivna (ed.) 2017. *Hramatyka sučasnoï ukrains'koï literaturnoï movy*. Kyiv: Vydavničyj dim Dmytra Buraho.
- Hubers, Ferdy & de Hoop, Helen. 2013. The effect of prescriptivism on comparative markers in spoken Dutch. In Aalberse, Suzanne & Auer, Anita (eds.), *Linguistics in the Netherlands 2013* (= *Linguistics in the Netherlands*, 30), 89-101. Amsterdam: Benjamins.
- Jäger, Agnes. 2008. *History of German negation*. (= *Linguistik aktuell*, 118). Amsterdam: Benjamins.
- Jäger, Agnes. 2010. Der Komparativzyklus und die Position der Vergleichspartikel. *Linguistische Berichte* 224. 467-493.
- Jäger, Agnes. 2018. *Vergleichskonstruktionen im Deutschen: Diachroner Wandel und synchrone Variation* (= *Linguistische Arbeiten*, 569). Berlin: de Gruyter.
- Jäger, Agnes. 2019. The syntax of comparison constructions in diachronic and dialectal perspective. *Glossa* 4(1): 70. 1–51.
- Jäger, Agnes & Walkden, George. 2021. West Germanic. In Keydana, Götz & Hock, Wolfgang & Widmer, Paul (eds.), *Comparison and gradation in Indo-European* (= *Mouton Handbooks of Indo-European Typology*, 1), 281-348. Berlin: de Gruyter.
- Jensen, Hans. 1934. Der steigende Vergleich und sein sprachlicher Ausdruck. *Indogermanische Forschungen* 52. 108-130.
- Jespersen, Otto. 1917. *Negation in English and other languages*. (= *Historisk-filologiske Meddelelser*, 1/5). Copenhagen: A. F. Høst.
- Klooster, Willem. 2001. *Grammatica van het hedendaags Nederlands: Een volledig overzicht*. Den Haag: SDU.
- Kölligan, Daniel. 2021. Classical Armenian. In Keydana, Götz & Hock, Wolfgang & Widmer, Paul (eds.), *Comparison and gradation in Indo-European* (= *Mouton Handbooks of Indo-European Typology*, 1), 49-79. Berlin: de Gruyter.
- Lerch, Eugen. 1925. *Historische französische Syntax*. Vol. 1: *Definition der Syntax - syntaktische Methoden - Allgemeines über Satzverknüpfung und Konjunktionen -*

- Beigeordnete Sätze - »Que«-Sätze, Sätze mit »si«, »quand« und »comme«.* Leipzig: Reiland.
- Lerch, Eugen. 1942. Ist schöner wie ein Sprachfehler? Zum Streit um die Berechtigung der Sprachregelung. *PBB* 65. 329-372.
- LKŽ 5 = Ulvydas, Kazys & Jonikaitė, Zuzana & Vosylytė, Bronė & Kruopas, Jonas & Kabelka, Jonas. 1959. *Lietuvių kalbos žodynas. 5: K - klausitys*. Vilnius: Valstybinė politinės ir mokslinės literatūros leidykla.
- LKŽ 8 = Kruopas, Jonas & Jonikaitė, Zuzana & Kėzytė, Sofija & Kučinskaitė, Antanė & Lyberis, Antanas. 1970. *Lietuvių kalbos žodynas. 8: Melūda - ožvilnis*. Vilnius: 'Minties' leidykla.
- Lohndal, Terje. 2009. The copula cycle. In Gelderen, Elly van (ed.), *Cyclical change*, 209-242. Amsterdam: Benjamins.
- Medvedjev, Fedir Pylypovyč. 1962. *Systema spoličnykiv v ukrains'kij movi*. Charkiv: Univ.
- Meillet, Antoine. 1912. L'évolution des forms grammaticales. *Rivista di scienzia* 12. 384-400. (Reprint in: *Linguistique historique et linguistique générale*. Second edition. Paris: Société de Linguistique Paris VIII, 1926. 130-148).
- Meyer-Lübke, Wilhelm. 1899. *Grammatik der romanischen Sprachen*. Vol. 3: *Syntax*. Leipzig: Reiland.
- Middleton, George. 1892. *An essay on analogy in syntax*. London: Longmans.
- Neumann-Holzschuh, Ingrid & Mitko, Julia. 2018. *Grammaire comparée des français d'Acadie et de Louisiane* (GraCoFAL). Berlin: de Gruyter.
- Pancheva, Roumyana. 2006. Phrasal and clausal comparatives in Slavic. In Lavine, James & Franks, Steven & Tasseva-Kurktchieva, Mila & Filip, Hana (eds.), *Formal Approaches to Slavic Linguistics 14: The Princeton meeting 2005*, 236–257. Ann Arbor: University of Michigan Slavic Publications.
- Pancheva, Roumyana. 2010. More Students Attended FASL than CONSOLE. In Browne, Wayles & Cooper, Adam & Kesici, Esra & Predolac, Nikola (eds.), *Formal approaches to Slavic linguistics 18: The second Cornell meeting 2009*, 383-400. Ann Arbor: University of Michigan Slavic Publications.
- Paul, Hermann. 2007. *Mittelhochdeutsche Grammatik*. 25<sup>th</sup> edition; revised by Klein, Thomas & Solms, Hans-Joachim & Wegera, Klaus-Peter & Prell, Heinz-Peter. Tübingen: Niemeyer.

- Petit, Daniel. 2021. Baltic. In Keydana, Götz & Hock, Wolfgang & Widmer, Paul (eds.), *Comparison and gradation in Indo-European*. (= Mouton Handbooks of Indo-European Typology, 1), 99-147. Berlin: de Gruyter.
- Pinault, Georges. 1985. Négation et comparaison en Védique. In *Bulletin de la société de linguistique de Paris*, 103-144. Paris/Louvain: Peeters.
- Postma, Gertjan. 2006. Van “groter dan” naar “groter als”. Structurele oorzaken voor het verval van het comparatieve voegwoord “dan”. *Nederlandse Taalkunde*, 11(1). 2-22.
- Price, Susan. 1990. *Comparative constructions in Spanish and French syntax*. London/New York: Routledge.
- Reinarz, Lukas & de Vos, Hugo & de Hoop, Helen. 2016. Conflicting constraints in the comparative cycle. *Journal of Germanic Linguistics* 28. 403-425.
- SAND = Babiers, Sjef & van der Auwera, Johan & Bennis, Hans & Boef, Eefje & de Vogelaer, Gunther & van der Ham, Margreet. 2005ff. *Syntactische Atlas van de Nederlandse Dialecten*. Amsterdam: Amsterdam University Press.
- Schwyzler, Eduard. 1950. *Griechische Grammatik*. Vol. 2: *Syntax und syntaktische Stilistik*. München: C.H. Beck.
- Seuren, Pieter A. M. 1984. The comparative revisited. *Journal of Semantics* 3. 109-141.
- Small, George William. 1924. *The comparison of inequality. The semantics and syntax of the comparative particle in English*. Greifswald: Abel.
- Small, George William. 1929. *The Germanic case of comparison. With a special study of English* (= Language Monographs, 5). Philadelphia: Linguistic society of America.
- Stassen, Leon. 1985. *Comparison and universal grammar*. Oxford/New York: Blackwell.
- Stassen, Leon. 2005. Comparative constructions. In Haspelmath, Martin & Dryer, Matthew S. & Gil, David & Comrie, Bernard (eds.), *The world atlas of language structures*, 490-493. Oxford: Oxford University Press.
- Sun, Chaofen. 1996. *Word-order change and grammaticalization in the history of Chinese*. Stanford, California: Stanford University Press.
- Tarriño, Eusebia. 2009. Comparative clauses. In Cuzzolin, Pierluigi & Baldi, Philip (eds.), *New perspectives on historical Latin syntax*. Vol. 4: *Complex sentences, grammaticalization, typology*, 373-426. Berlin: de Gruyter.
- Thurmair, Maria. 2001. *Vergleiche und Vergleichen. Eine Studie zu Form und Funktion der Vergleichsstrukturen im Deutschen* (= Linguistische Arbeiten, 433). Tübingen: Niemeyer.

- Umbach, Carla & Helmar, Gust. 2014. Similarity demonstratives. *Lingua* 149, 73-93.
- van der Auwera, Johan. 2010. On the diachrony of negation. In Horn, Larry (ed.), *The expression of negation*, 73-101. Berlin: de Gruyter.
- van der Horst, Joop. 2008. *Geschiedenis van de Nederlandse syntaxis, vol. 1*. Leuven: Universitaire Pers Leuven.
- van Gelderen, Elly. 2004. *Grammaticalization as economy*. Amsterdam: Benjamins.
- van Gelderen, Elly. 2007. The definiteness cycle in Germanic. *Journal of Germanic Linguistics* 19. 275-305.
- van Gelderen, Elly (ed.). 2009. *Cyclical change*. Amsterdam: Benjamins.
- van Gelderen, Elly. 2011. *The linguistic cycle: Language change and the language faculty*. Oxford: Oxford University Press.
- van Gelderen, Elly (ed.). 2016. *Cyclical change continued*. Amsterdam: Benjamins.
- van Kemenade, Ans. 1999. Sentential negation and clause structure in OE. In Tieken-Boon van Ostade, Ingrid & Tottie, Gunnel & van der Wurff, Wim (eds.), *Negation in the history of English*, 147-165. Berlin: Mouton de Gruyter.
- van Kemenade, Ans. 2000. Jespersen's cycle revisited. In Pintzuk, Susan & Tsoulas, George & Warner, Anthony (eds.), *Diachronic syntax*, 51-74. Oxford: Oxford University Press.
- von der Gabelentz, Georg. 1891. *Die Sprachwissenschaft: Ihre Aufgaben, Methoden und bisherigen Ergebnisse*. Leipzig: Weigel. (Reprint: Narr, 1972).
- von Stechow, Arnim. 1984. Comparing Semantic Theories of Comparison. *Journal of Semantics* 3. 1-77.
- Vine, Brent. 1978. On the metrics and origin of Rig-Vedic *ná* 'like, as'. *Indo-Iranian Journal* 20. 171-193.
- Viti, Carlotta. 2002. Comparazione e Individuazione: uno studio sugli equativi rigvedici *iva* e *ná*. *Archivio Glottologico Italiano* 87. 46-87.
- Weiß, Helmut. 2019. Rebracketing (Gliederungsverschiebung) and the Early Merge Principle. *Diachronica* 36. 509-545.
- Weiss, Michael. 2009. *Outline of the historical and comparative grammar of Latin*. Ann Arbor: Beech Stave Press.
- Willis, David & Lucas, Christopher & Breitbarth, Anne. 2013. Comparing diachronies of negation. In Willis, David & Lucas, Christopher & Breitbarth, Anne (eds.), *The development of negation in the languages of Europe, Vol. 1: Case studies*, 1-50. Oxford: Oxford University Press.

- Wurzel, Wolfgang Ullrich. 1994. *Grammatisch initiiertes Wandel*. Unter Mitarbeit von Andreas und Dagmar Bittner. Bochum: Brockmeyer.
- Zeilfelder, Susanne. 1996. Zur Syntax des Komparativs im Altarmenischen. *Historische Sprachforschung* 109. 175–198.
- Zeilfelder, Susanne. 2001. *Steigern und Vergleichen in indogermanischen Sprachen*. Ms. Habilitation Dissertation, University of Jena.
- Zifonun, Gisela & Hoffmann, Ludger & Strecker, Bruno. 1997. *Grammatik der deutschen Sprache*. (= Schriften des Instituts für Deutsche Sprache 7). Berlin: de Gruyter.

### Sources

- Brian Holton: *The Mossflow*. [= Scottish corpus of texts & speech: document 1628].
- JBang = *Thüringische Chronick oder Geschichtbuch (...) anfanglich auss einem alten geschriebenen (...) Exemplar colligirt (...) vnd ferner durch weiland Ern Friderichen Schmidt Pfarherrn zu Grossen Beruingen reuidiert vnd vermehret / dann (...) von andern (...) continuirt. Jetzo aber (...) in Druck geben / durch Johan Bangen*. Mühlhausen, 1599. Druck: Andreas Hantzsch. [= Bonner Frühneuhochdeutschkorpus: text 255].
- JMath = *Passionale Mathesij, das ist / christliche vnd andechtige Erklerung vnd Außlegung des zwey und zwanzigsten Psalms / vnd drey vnd funffzigsten Capitels des Propheten Esaia geprediget durch M. Johannem Mathesium / weyland Pfarrer in S. Joachimsthal*. Leipzig 1587. Druck: Johann Beyer. [= Bonner Frühneuhochdeutschkorpus: text 145].
- Lil = *Die Lilie*. Hrsg. von Paul Wüst. Berlin: Weidmann 1909. [= Referenzkorpus Mittelhochdeutsch: text M327].
- Notker Mus. = *Notker Labeo: De musica*. Edition, Übersetzung und Kommentar von Martin van Schaik. Utrecht: Galapagos, 1995.
- Notker Ps. = *Die Werke Notkers des Deutschen*. Neue Ausgabe. Begonnen von Edward H. Sehr und Tayler Starck. Fortgesetzt von James C. King und Petrus W. Tax. Vol. 8: Der Psalter. Psalm 1-50. Hrsg. von Petrus W. Tax. Tübingen: Niemeyer, 1979.
- Vol. 9: Der Psalter. Psalm 51-100. Hrsg. von Petrus W. Tax. Tübingen: Niemeyer, 1981.
- Pañcatantra = *Panchatantra I*. Edited, with notes, by F. Kielhorn. Sixth Edition. Bombay: Government Central Book Depôt, 1896.

Phys = *Der altdeutsche Physiologus. Die Millstätter Reimfassung und die Wiener Prosa (nebst dem lateinischen Text und dem althochdeutschen Physiologus)*. Hrsg. von Friedrich Maurer. Tübingen: Niemeyer, 1967. [= Referenzkorpus Mittelhochdeutsch: text M157].

RV = *Rig Veda. A metrically restored text with an introduction and notes*. Ed. by Barend A. Van Nooten and Gary B. Holland. Cambridge, Massachusetts and London: Harvard University Press, 1994.

SalH = *Salomonis hūs*. Hrsg. von Johann Valentin Adrian. In: *Mittheilungen aus Handschriften und seltenen Druckwerken*. Frankfurt a.M.: Sauerländer, 1846, 417-455. [= Referenzkorpus Mittelhochdeutsch: text M337].

Tatian = *Die lateinisch-althochdeutsche Tatianbilingue Stiftsbibliothek St. Gallen Cod. 56*. Achim Masser (ed.) Göttingen: Vandenhoeck & Ruprecht, 1994.

TrHL = *Das St. Trudperter Hohe Lied*. Hrsg. von Hermann Menhardt. (Rheinische Beiträge und Hilfsbücher zur germanischen Philologie und Volkskunde 21). Halle a.d. Saale, 1934. [= Referenzkorpus Mittelhochdeutsch: text M113].

WRal = *Walter Ralegh: Amerika*. Frankfurt am Main 1599. Druck: Matthäus Becker. [Bonner Frühneuhochdeutschkorpus: text 245].

## Corpora

*Bonner Frühneuhochdeutschkorpus*

<https://korpora.zim.uni-duisburg-essen.de/Fnhd/>

*NKRJa = Russian National Corpus/Nacional'nyj korpus russkogo jazyka*

<http://ruscorpora.ru/>

*Referenzkorpus Mittelhochdeutsch*

<https://www.linguistics.rub.de/rem/>

*Scottish corpus of texts & speech*

<https://www.scottishcorpus.ac.uk>

## CONTACT

[agnes.jaeger@uni-jena.de](mailto:agnes.jaeger@uni-jena.de)



# Comparative of inferiority: marking and aspects of use

VALERIA MODINA<sup>1, 2</sup>

<sup>1</sup> SAINT PETERSBURG STATE UNIVERSITY

<sup>2</sup> INSTITUTE FOR LINGUISTIC STUDIES, RUSSIAN ACADEMY OF SCIENCES

Submitted: 16/11/2020 Revised version: 6/6/2021

Accepted: 17/03/2021 Published: 31/08/2021

## Abstract

In this paper the comparative of inferiority ('A is less tall than B') is discussed in regards to its coding and functioning. The classification of the marking is based on the connection of the marking of inferiority to the marking of other constructions of comparison. Thus, two main types of the marking are distinguished: specific and derived. The discussion of some problematic issues connected to the comparative of inferiority accompanies the description of the marking. The findings in the marking are interpreted as signs of the markedness of the comparative of inferiority. The remaining part of the paper is devoted to the description of the functions of the comparative of inferiority and its aspects of use as suggested by the data from Russian.

**Keywords:** Comparison; comparative constructions; comparative of inferiority; language typology; rivalry; Russian

## 1. Introduction

This paper examines comparative constructions of inferiority. Comparative constructions of inferiority are used to describe the referent that has some property to a lower degree through the comparison with the other referent.

(1) Russian (Indo-European)

<i>Petja</i>	<i>meneje</i>	<i>vysokij</i>	<i>čem</i>	<i>Vanja</i>
Petja	less	tall	than	Vanja

'Petja is less tall than Vanja'

Different aspects of comparison have attracted the attention of many linguists and such topics as cross-linguistically attested types of constructions of comparison and

the rivalry of different constructions of comparison in a specific language have been explored in a number of papers. However, comparison of inferiority is only rarely mentioned in those works and almost never gets discussed in depth.

The notion of the comparative of inferiority is not particularly new as it appears in grammars dating back at least to the end of the 19th century (see, for example, this grammar of Somali, Larajasse & Sampont 1897: 64-65). This fact is probably connected with the presence of markers of inferiority in a number of European languages (e.g. English *less* or Spanish *menos*). Nevertheless, comparative constructions of inferiority are rarely even mentioned in descriptive grammars and, if mentioned, the information given is quite scarce and non-uniform most of the times. Therefore, the task of constructing a typology of comparative of inferiority is fairly difficult due to the lack of information on these constructions. In this paper, I present a preliminary outline of the types of constructions that express the meaning of inferiority. Despite the fact that this sketch is by no means comprehensive, I still hope that this description can be of some value for the study of comparison.

The other problem addressed here is the rivalry of the comparative of superiority and the comparative of inferiority within one language. Though the issue of the rivalry of several comparative constructions has been addressed previously (see, for example, Hilpert (2008) and Kosheleva (2016) on the rivalry of synthetic and analytical comparative constructions) it has never included the comparative of inferiority. The conditions of the use of these constructions and their functions are of particular interest since the comparative of inferiority may be seen as unnecessary due to the possibility of using an antonym (i.e. '*Petja is shorter than Vanja*' for (1)) or switching the referents (i.e. '*Vanja is taller than Petja*'). This rivalry is analyzed in this paper using the Russian data. Such analysis of distribution of the comparative constructions in Russian potentially gives new perspectives on the issue that can be addressed in the descriptions of other languages as well.

The paper is organized as follows. I begin by introducing and clarifying some of the terms used in the research (Section 2). Section 3 is devoted to the classification of the attested marking of the inferiority. In Section 4 the opposition of superiority and inferiority is discussed in terms of markedness. The analysis of the rivalry of comparative constructions in Russian is presented in Section 5. Section 6 contains conclusions.

## 2. Terms and notions

Construction of comparison is defined here as a construction which is used when one referent is described through the comparison of its degree on a gradable scale with the degree of another referent.

I adhere to the practice of identifying constitutive elements of the comparison construction which is used consistently through the publications on the subject (Ultan 1972; Stolz 2013; Haspelmath et al. 2017; Treis 2018). Still, the terms used and definitions given in the works on the comparison can be quite different, therefore it is important to present the terminology as it is used throughout this paper. Consider the following example.

(1') Russian (Indo-European)

Comparee	Parameter Marker	Parameter	Standard Marker	Standard
<i>Petja</i>	<i>meneje</i>	<i>vysokij</i>	<i>čem</i>	<i>Vanja</i>
Petja	less	tall	than	Vanja

'Petja is less tall than Vanja'

It is possible to identify five key components of the construction of comparison. Three of those components are primary and constitute any comparison construction, explicitly or implicitly:

*Comparee* – the referent which is described through the comparison.

*Standard of comparison* – the referent to which the comparee is being compared.

*Parameter of comparison* – the property of comparison. It is worth noting that the term parameter is applied here to only one member of an antonymic pair rather than to the common basis of the antonyms. For example, *high* and *low* are considered to represent two parameters, not one parameter of height.

The other two components of the constructions of comparison are used to mark the comparison.

*Standard marker* – marker of comparison closely associated with the standard of comparison.

*Parameter marker* – marker of comparison closely associated with the parameter.

Note that for the component called here *parameter marker* the terms *degree marker* or *degree* are employed sometimes (Ultan 1972; Stolz 2013; Haspelmath 2017; Treis 2018). Definitions themselves also may contain the notion of degree. In fact, it seems

like the actual relation of degrees between comparee and standard (i.e. ‘more’ or ‘less’) does not have to be marked exclusively on the parameter. Conjoined comparatives (formed by two juxtaposed clauses, Stassen 1985: 37-38), exceed comparatives (comparative constructions where the comparee is the subject of an ‘exceed’ verb and the standard is the direct object, Stassen 1985: 42)<sup>1</sup> and comparative constructions with sole standard marker carry the semantics of the relation of degrees as well, suggesting that the semantic notion of degree is useless for defining any of the constituents of comparative construction. Consequently, I argue here that the term *parameter marker* is a more appropriate one than *degree marker* and that the definition of this element should not refer to the notion of degree.<sup>2</sup> Therefore, the definitions used here are strictly structural (following Haspelmath et al. 2017: 11).

The act of comparison can result in two possible outcomes, namely in the assertion of equality or in the assertion of inequality of the items regarding some parameter. Inequality, in turn, can be encoded both in the comparative construction of superiority and in the comparative construction of inferiority. *Comparative of superiority* is defined here as a construction in which the referent that has the property to a higher degree appears in the position of the comparee (e.g. *Horses are bigger than dogs*). *Comparative of inferiority* is a construction in which the referent that has the property to a lower degree appears in the position of the comparee (e.g. *Dogs are less big than horses*).

It needs to be mentioned here that the notion of inferiority is largely absent not only from the descriptive grammars of particular languages but also from the very basic terms used in the literature on the subject. In other words, the study of

---

<sup>1</sup> The case of exceed comparatives is a telling one as the ‘exceed’ verb has been analysed both as the standard marker (in Ultan 1972: “In Sotho, a verb meaning ‘surpass, excel’ assumes the marking function in the comparative (...) Here again, applying the criterion of immediate constituency, *fēta* ‘surpass’ must be regarded as a standard marker since it is in constituency with its object which is equivalent to the standard of comparison in spite of the fact that it obviously contains a semantic feature of degree”) and as the parameter marker (or index in Dixon 2012: 343-375). This kind of comparative construction was also entirely excluded from the opposition of parameter marker and standard marker in Haspelmath (2017): “Like equative constructions, comparative constructions usually have a standard-marker, or otherwise they may have a verb (‘exceed’) expressing the relationship between the comparee and the standard. (...) If the construction has a standard-marker rather than using a verb, it may also have a degree-marker”.

<sup>2</sup> Though it seems to be true that if a construction contains both a parameter marker and a standard marker the difference in degree (‘more’ or ‘less’) is reflected in the parameter marker.

comparison is skewed to the superior side of the issue. The terms *comparative* and *superlative*<sup>3</sup> themselves are often defined as constructions that express superior meaning (e.g. Shvedova 1980: 545; Mel'čuk 1998: 117; Cuzzolin & Lehmann 2004: 1213, Treis 2018: iii).<sup>4</sup>

As long as such definitions are employed expressions parallel to English 'less' and 'least' cannot be included in the discussion of comparison. Therefore, I opt here for less biased definitions following the ones given in Ultan (1972):

Comparative – the construction used to express the situation when the comparee differs in the degree of the parameter from the standard, where the latter does not contain all the members of the class to which the former belongs (i.e. *John is taller than the brothers* – the referent set of NP *brothers* does not contain John).

Superlative – the construction used to express the situation when the comparee differs in the degree of the parameter from the standard, where the latter contains all the members of the class to which the former belongs (*John is the tallest of the brothers* – the referent set on NP *brothers* does contain John).

This bias is not random as it reflects the asymmetry of superiority and inferiority that is discussed further in Section 4. It is important to mention that this bias in terminology may have partially led to the scarcity of data on the comparative of inferiority in descriptive grammars. Therefore, a more precise terminology that incorporates both superiority and inferiority is a desideratum.

### 3. Marking of the comparative of inferiority

In this section, I sketch out the types of markers of inferiority and give examples of them. Other constructions of comparison have been studied in depth in regards to their marking (Ultan 1972; Stassen 1985; Heine 1997; Dixon 2012; Gorshenin 2012; Bobaljik 2012; Stolz 2013; Haspelmath et al. 2017). These typologies were based on a variety of grounds such as the source of the marker and the morphosyntactical traits of constructions. The basic distinction made here is whether the marker of inferiority

---

<sup>3</sup> An anonymous reviewer points out that while the term *comparative* is neutral with respect to the direction of comparison, the term *superlative* is etymologically connected to the notion of superiority. Therefore, it may be useful for the theory of comparison to come up with a neutral term that will bracket *superlative* and “inferlative” (*the least tall*).

<sup>4</sup> It is still worth noting that at the same time at least in the last two papers mentioned above the comparative of inferiority is described as a kind of comparative construction despite the given definition.

is derived from other constructions of comparison or not. The following is not a classification per se but rather an outline of attested variation in marking accompanied by the discussion of the issues connected to it.

### 3.1. Problematic points

Firstly, I would like to mention some problematic points regarding the description of the comparative of inferiority. One of them is related to the fact that, apparently, in some languages there are several markers of the comparative of superiority whose distribution is based on the semantics of the parameter. The “negative” member of an antonymic pair (e.g. *small, light*) in those cases is marked with a marker different from the one applied to the “positive” member (e.g. *big, heavy*) in comparative constructions. For example, this kind of opposition is present in Murui (2), where the meaning of the standard marker refers to distance, interiority, and vertical position.

(2) Murui (Witotoan; Wojtylak 2018: 175)

- |    |  |                 |                                     |                           |
|----|--|-----------------|-------------------------------------|---------------------------|
| a. | Comparee   | Standard        | Standard Marker                     | Parameter                 |
|    | <i>[kai jo-fo]<sub>vcs</sub></i>   | <i>[oo-ie</i>   | <i>aa-fe-mo]<sub>np:perf</sub></i>  | <i>aare<sub>vcc</sub></i> |
|    | 1PL house-CLF:CAVITY   | 2SG-GEN         | above-CLF:SIDE-LOC                  | long                      |
|    | ‘Our house is taller than yours (lit. our house, yours on the top side, long).’                    |                 |                                     |                           |
| b. | Comparee   | Standard        | Standard Marker                     |                           |
|    | <i>pila-ji = disubj</i>  | <i>[mechera</i> | <i>foo-fe-mo]<sub>np:perf</sub></i> |                           |
|    | battery.SP-CLF:SMALL.ROUND = SUBJ/A.TOP  | lighter.SP      | inside-CLF:SIDE-LOC                 |                           |
|    | Parameter  |                 |                                     |                           |
|    | <i>jano-re-d-e<sub>pred</sub></i>  |                 |                                     |                           |
|    | small-ATT-LK-3   |                 |                                     |                           |
|    | ‘The battery is smaller than the lighter (lit. the battery, the lighter on the inside, is small).’ |                 |                                     |                           |

The author explicitly refers to *foofemo* as to the marker of inferiority (instead of the marker of superiority for the “negative” antonym) and, at the same time, underlines that the notion “inferiority” is not understood as expressing ‘less’ but rather ‘higher degree of a “negative” adjective’. The same meaning of “inferiority” seems to be

adopted by A. Aikhenvald for the analysis of one of the comparative strategies in Yalaku (3).

(3) Yalaku (Ndu; Aikhenvald 2018: 4)

- a. [*semi*= *de-te*]                      [*wore-I de-te*]  
tall/long = 3MASC.SG-stay go.up-go 3MASC.SG-be  
'He is tall, he goes up (in height)' (lit. Go up go he is)
- b. [*foi*= *de-t*],                      [*tada-d*]  
short = 3MASC.SG-be go.down-3MASC.SG  
'He is short, he is shorter (than the other child)' (lit. He goes down)

Though this kind of opposition of the markers is definitely worth attention of the researchers it must be distinguished from the opposition discussed here. Therefore, it is suggested to investigate this phenomenon further and to develop alternative terminology for the semantically induced opposition of the markers of superiority. This kind of difference is not discussed here further.<sup>5</sup>

The other problem is connected to the biclausal constructions of comparison which come in a variety of flavours (see Dixon 2012: 358-360). One of them is conjoined comparative (in terms of Stassen's typology) formed by juxtaposition of two independent clauses with antonyms (*A is big, B is small*) or two independent clauses, one of which includes negation (*A is big, B is not big*) as in Samoan (4).

(4) Samoan (Oceanic; Cuzzolin & Lehmann 2004: 1214)

*Ua loa lenei va'a, ua puupuu lena*  
is long this boat, is short that

'This boat is longer than that.'

As both clauses have the same grammatical structure, it is difficult to identify the comparee and the standard of such comparative construction. As long as identification of the components stays problematic, it is not possible to describe such constructions as instances of comparative of superiority or comparative of inferiority.

Nevertheless, such comparatives are sometimes analysed in regards to superiority or inferiority. As demonstrated in the description of the comparative constructions in

---

<sup>5</sup> This opposition is also relevant for Urarina (Olawsky 2006: 208-209), Afar (Bliese 1977: 90).

Kanoê, the construction expresses inferiority in the case when the first clause represents the assertion of the lower degree (5).

(5) Kanoê (Kapixana; Bacelar 2004: 249, 269)

- a. *ñā tyj ej-turo-e-re pja tyj ej-turo*  
 POSS1SG house big-place-DECL-AUX POSS2SG house big-place  
*k-e-re*  
 NEG-DECL-AUX  
 ‘My house is bigger than yours’ (Lit.: ‘My house is big, your house is not big’)
- b. *ñā vae-nake ā-nake k-e-re [aj ā-kȳj*  
 POSS1SG cousin-FEM tall-FEM NEG-DECL-AUX 1SG tall-MASC  
*ō-e-re]*  
 1-DECL-AUX  
 ‘My cousin is less tall than me’ (Lit.: ‘My cousin is not tall, I am tall’)

Therefore, the order of clauses is relevant and it is suggested that in Kanoê it is the first clause where the comparee is located, but the reason for this analysis is not specified. One may argue for some kind of iconicity in the ordering of the clauses in a way that the subject of the first clause is considered to have a higher communicative rank and should, therefore, be considered the comparee. However, it is not clear if this interpretation is not arbitrary (i.e. if there is actual communicative inequality of the referents and the ordering is truly iconic) and if this kind of ordering is systematic within a single language and cross-linguistically.

Another possibility is to assume the parallelism of the basic word order and order of constituents in comparative constructions (where comparee corresponds to S, and standard to O), as suggested by Romero-Figueroa (1986). Romero-Figueroa argues that in Warao, where the basic word order is OSV, the order of clauses in the comparative construction follows the basic word order of the language, which means that the first clause contains the standard of comparison, not the comparee (Romero-Figueroa 1986: 105-106). The examples (5a, 5b) from Kanoê corroborate this analysis as well, because the basic word order in Kanoê is SOV (Bacelar 2004: 228). At the same time, an opposite point of view is reflected in the grammar of Ayutla Mixe: since the comparative construction is biclausal, it is not relevant for the word order correlations (Romero-Méndez 2008: 443).



Considering that there is no universal rule to identify the comparee in biclausal constructions and that the grammars largely remain silent on the basis of the particular interpretation, it is unclear if the opposition of superiority and inferiority is relevant for this type of constructions. Thus, biclausal constructions are not discussed henceforth. The issue of the ordering of clauses in a biclausal comparative construction requires further study.

### 3.2. *Specific markers of inferiority*

“Specific markers of inferiority” are defined here as ones that are not derived from other constructions of comparison. This kind of markers can be found in comparative constructions of structurally different types. The subdivision in this category is based on the degree of parallelism of constructions of superiority and inferiority.

#### 3.2.1 *Parallel specific markers of inferiority*

First of all, structurally specific markers of inferiority sometimes reflect the features of the ones that express superiority, i.e. the constructions are identical except for difference in the relation marker (i.e. the unit that expresses ‘more’ or ‘less’). For example, in Amis the opposition of superiority and inferiority is possible only in one type of comparative construction, in which the relation is expressed by a predicate that is followed by the parameter of comparison. The only difference present is the predicate itself (6).

(6) Amis (East Formosan; Kuo & Sung 2010: 32-33)

a. *Ø-ikaka ku su'su' ni mama aku tisuwanan*  
AF-more NOM fat NCM.SG.GEN father 1SG.GEN 2SG.OBL  
'My father is fatter than you.'

b. *Ø-isafa ku takaraw nira takuwanan*  
AF-less NOM tall 3SG.GEN 1SG.OBL  
'He is less tall than me.'

Specific parallel markers are also found in comparative constructions of other morphosyntactical nature. In Lizu the comparative of superiority is formed with prefix *jæ-* and the comparative of the inferiority is marked by the negator *mw-*, exemplifying the case of parallel affixes (7).

(7) Lizu (Tibeto-Burman; Chirkova 2019: 29-30)

- a.  $\alpha = \hat{t}$  *jênv* *le*  $ne = \hat{t}$  *jênv* *pv* *jæ-mbɔ̃*  
 $\alpha = \hat{t}$  *jênv* *le*  $ne = \hat{t}$  *jênv*  
 1SG = GEN younger.brother TOP 2SG = GEN younger.brother  
*pv* *jæ-mbɔ̃*  
 like **more**-be.tall  
 ‘My brother is taller than your brother.’
- b.  $\alpha = \hat{t}$  *jênv* *le*  $ne = \hat{t}$  *jênv* *pv* *mw-mbɔ̃*  
 $\alpha = \hat{t}$  *jênv* *le*  $ne = \hat{t}$  *jênv*  
 1SG = GEN younger.brother TOP 2SG = GEN younger.brother  
*pv* *mw-mbɔ̃*  
 like **NEG**-be.tall  
 ‘My brother is not as tall as your brother.’

This marker of inferiority is analysed here as a non-derived one because, firstly, the comparative of inferiority does not contain the marker of superiority *jæ-*, and, secondly, the marker *jæ-* is not described as optional (i.e. it is not the case that the comparative of inferiority is formed by the negation of the comparative of superiority with the omitted marker *jæ-*). The same analysis applies to the comparative constructions of Xuwen (Li & Thompson 1983: 20-21).

The relation between the referents may be realized in the standard marker (giving us another example of the uselessness of the notion “degree” when defining elements of the comparative construction). A special marker of inferiority appears as a marker of standard in Teribe (8) and Navajo (9).

(8) Teribe (Talamanca; Quesada 2000: 139, as cited in Dixon 2012: 362)

- a. [*Bor* *u*] *kégué* *bopoya* *kinmo*  
 1SG.POSS house old 2SG.POSS **above**  
 ‘My house is older than yours’

- b. *Kwe*                      *kégué*                      *bop dorko*  
 DEM                      old                      2SG **under**  
 ‘This one is less old than you’

(9) Navajo (Athapaskan; Bogal-Allbritten & Coppock 2020: 127)

- a. *Alice*    (*Ben*)    *yi-lááh*                      ’*áníłnééz.*  
 Alice    Ben    3OBJ-**beyond**    3SUBJ.tall  
 ‘Alice is taller than Ben/him/her/it.’
- b. *Alice*    *shi-’oh*                      ’*áníłnééz.*  
 Alice    1OBJ-**short.of**    3SUBJ.tall  
 ‘Alice is less tall than me.’

Some languages demonstrate usage of non-derived markers of inferiority parallel to the “exceed” comparatives of superiority. This type of coding can be found in Amharic (Leslau 1995: 788) and Hausa (10).

(10) Hausa (Chadic; Newman 2000: 93-96)

- a. *Kanò*                      *tā*                      *fi*                      *Kādūna*    *yawàn*                      *mutānē*  
 Kano                      it                      **exceeds** Kaduna    quantity.of                      people  
 ‘Kano is bigger (i.e., more populous) than Kaduna’
- b. *Gidānā*                      *yā*    *gazā̀*                      *nākà*                      *girmā*  
 House.of.my                      it    **falls short**    of yours                      size  
 ‘My house is not as large as yours’

### 3.2.2 Non-parallel specific markers of inferiority

Nevertheless, sometimes comparative of superiority and comparative of inferiority are not quite parallel in structure. Consider an example from Arapaho (11). The comparative of superiority is expressed by the use of /cebe’ei/ (‘more, beyond’) or /wo’ow/ (‘farther, more’) with a verb while the comparative of inferiority is expressed quite differently with the use of a negative verb and the particle *wootíí* ‘like’.

(11) Arapaho (Algonquian; Cowell & Moss 2008: 228-230)

- a. *ceebe'eitéi'éłht*  
*cebe'ei-tei'eihi-t*  
 IC.beyond-strong(AI)-3S  
 'S/he is stronger.'
- b. *hoow(u)téi'éłh*      *wootíí nenééninoo*  
*ihoowu-tei'eihi*      *wootii*      *neeni-noo*  
 NEG-strong(AI)      **like**      IC.to be(AI)-1SG  
 'S/he is not as strong as me.' (lit. 'S/he is not strong like me.')

Given that the particle *wootíí* does not occur in the comparative of superiority or the comparative of equality, it, accompanied by a negator on the verb, is considered to constitute a specific marker of inferiority that is structurally different from the marker of superiority.

In Central Alaskan Yupik some roots can be expanded by verbalizing suffixes to express inferiority. The marker of superiority is of different origin: note that the structures of comparative are different and comparative of superiority is actually attached to one of the verbalizing suffixes that are in opposition with the *-kelli-* (12).

(12) Central Alaskan Yupik (Eskimo; Miyaoka 2012: 278)

- qas-tu-uq* 'it is loud'  
*qas-kit-uq* 'it is quiet, less loud'  
*qas-kelli-uq* 'it is getting quieter, less loud'

(13) Central Alaskan Yupik (Eskimo; Miyaoka 2012: 1305)

- a. *tuner-tu-nrurt-uq*  
 'it is getting more powerful'.  
 b. *tuner-kelli-uq*  
 'it is getting less powerful'

### ***3.3. Derived markers of inferiority***

At the same time, quite often languages employ markers of inferiority that are derived from other constructions of comparison such as equative and comparative of superiority.

### 3.3.1 Negated equative

One of the most frequently attested ways of expressing inferiority is a negated equative construction. This kind of derived marking seems to be the only way to mark inferiority in a number of languages of different origin (14)-(16).

(14) Fongbe (Kwa; Lefebvre & Brousseau 2002: 435)

- a. *Kòkú sù òì / sò Àsíbá.*  
 Koku be.tall resemble / equal Asiba  
 ‘Koku is as tall as Asiba.’
- b. *Kòkú sù sò Àsíbá ǎ.*  
 Koku be.tall equal Asiba NEG  
 ‘Koku is not as tall as Asiba.’ (while the translation does not contain less, the construction is referred to as a way to express inferiority)

(15) Kadiwéu (Guaicuruan; Sandalo 1997: 74-75)

- a. *Maria dawé alikyagi nGijo lyone:Ga.*  
*Maria y-d:-awe alikyagi nGijo lyone.Ga*  
 Mary 3SG.SUBJ-theme-be.fast like DEM young.man  
 ‘Mary is as fast as this boy.’ (Lit.: ‘Mary is fast like this boy.’)
- b. *Maria adawé alikyagi nGijo lyonerGa*  
*Maria aG-y-d:-awe alikyagi nGijo lyone.Ga*  
 Mary NEG-3SG.SUBJ-theme-be.fast like DEM young.man  
 ‘Mary is less fast than this boy.’ (Lit.: ‘Mary is not fast like this boy.’)

(16) Nuosu (Burmese-Lolo; Gerner 2013: 447-449)

- a. *vit gga a hni su si nip vit gga a shy su ngex ngep nrat.*  
 clothes red NOM with clothes yellow NOM similarly nice  
 ‘The red clothes are as beautiful as the yellow clothes.’
- b. *zze ti cyx ma li a zzyx ma ngex ngep a-ap-du.*  
 table DEM.PROX CL TOP DEM.DIST CLF similarly thick <NEG >  
 ‘This table is less thick than that table.’

The negation of the equative is introduced by a special predicate *jokodu* ‘to be true’ in Bororo (17).

(17) Bororo (Bororoan; Nonato 2008: 101-102)

a. *ime erijore areme kori*

<i>ime</i>	<i>e</i>	<i>rico</i>	<i>re</i>	<i>areme</i>	$\emptyset$	<i>kori</i>
men	3PL	(to be) tall	ASSERT	women	3SG	<b>comparison</b>

‘Men are as tall as women’

b. *areme erijojokodukare ime ekori*

<i>areme</i>	<i>e.rijo.jokodu.ka.re</i>	<i>ime</i>	<i>e.kori</i>
women	3PL.(to be)tall.(to be>true.not.ASSERT	men	3PL.comparison

‘Women are less tall than men’ (lit. ‘it is not true that women are as tall as men’)

It is important to point out that this kind of construction is not actually compositional, because the negation of equation is in fact ambiguous: inequality can be both of superior and inferior nature. This non-compositionality gives us ground to consider the negated equative construction a distinct type of derived marker of inferiority. The crosslinguistic regularity of this interpretation of such constructions is of particular interest (see also for Zaar, Caron 2017: 170, Kambaata, Treis 2018: 15-16, Somali, Evangeliste & Cyprien 1897: 64-65).

The issue that is important to note here is that the scope of negation may vary. In the case of the negated equative the fact that it is not only the parameter that lies in the scope of negation is sometimes underlined. For the example (18), the author highlights: “Note that the infinitival VP is included in the scope of the negative, otherwise the translation would be ‘I equal him in (extent of) not eating’”.

(18) Koyra Chiini (Songhay; Heath 1999: 318-319)

<i>ay</i>	<i>si</i>	<i>gaa [ka too ga]</i>
ISG.S	IMPF.NEG	eat [INF <b>attain</b> 3SG.O]

‘do not equal him in eating’.(= ‘I eat less than he [does].’)

### 3.3.2 Negation and the comparative of superiority

This difference in the interpretation induced by the scope of negation does not seem to be problematic in the case of the negated equative. However, in some languages comparative of inferiority is described as expressed by the presence of the negation

in the environment typical not for equative, but for the comparative of superiority. In these cases the role of the scope of negation can be a bit trickier, as the difference in interpretation is not as straightforward as in the case of the negated equative. One of the possible scope-induced differences is discussed on the data from Mbyá Guaraní in Thomas (2017). In Mbyá Guaraní “the order of suffixes on a predicate determines the respective semantic scope of the operators that they denote” (Thomas 2017: 251). Thus, the comparative of inferiority and the negated comparative of superiority are comprised of the same components (parameter, marker of superiority, negation), but these components are put together in different orders. In the case of negated comparative of superiority it can be schematically modelled as [[adjective] + ‘more’] + negation]. The ordering specifies truth conditions: while (19b) is true if Juan and Pedro have the same height, (19a) is not.

(19) Mbyá Guaraní (Tupian; Thomas 2017: 251)

- a. *Juan Pedro gui nda-i-jyvate-i-ve.*  
Juan Pedro from NEG-B3-tall-NEG-ve  
‘Juan is less tall than Pedro.’
- b. *Juan Pedro gu nda-i-jyvate-ve-i.*  
Juan Pedro from NEG-B3-tall-ve-NEG  
‘Juan is not taller than Pedro.’

Therefore, the comparative of inferiority should be distinguished from the negation of the comparative of superiority.

The other problem worth consideration in this regard is the issue of the comparative constructions with morphological antonyms such as *unhappier*. Morphological antonym is analyzed here as a separate lexeme, as the negation forms an opposite parameter rather than functions as a part of the mark of inferiority. An argument in favour of this analysis is that the negation is already present in the positive degree of an adjective (*unhappy*), making it a lexeme feature, not a feature of the comparative construction. Thus, constructions like *more unhappy* are schematically organized as [[adjective]<sub>LEXEME</sub> + negation]<sub>NEW LEXEME</sub> + ‘more’] and not analyzed as instances of comparative of inferiority but rather as instances of comparative of superiority with antonymic parameter. Strictly speaking, the Mbyá Guaraní example (19a) could be an example of this scheme, but the circumfix *n...-i*

is used exclusively on predicates (while in attributive position the negation is marked by the suffix *-(e)ʔ*, Thomas 2017: 249), making the antonymic analysis unlikely.

This kind of distinction is not always made in the descriptions of comparative constructions. An example that was described as an instance of comparative of inferiority comes from Sonora Yaqui (20).

(20) Sonora Yaqui (Cahita; Dedrick & Casad 1999: 111)

- a. *če'a huni'i tu'ii*  
 more even good  
 'it is even better'
- b. *če'a huni'i kaa-tu'ii*  
 more even **not**-good  
 'it is even worse'

In this case the authors explicitly state that “the negative is used to derive an antonym” (note the translation as well), therefore this example is not qualified here as a comparative of inferiority. Other examples of the appearance of negation on the marker can be found in Guajiro (Álvarez 2005: 25), Eastern Geshiza (Honkasalo 2019: 525-526).

An actual “inferior” interpretation of constructions with negation and marker of superiority is possible in three cases.

The first one is when the negation is not utilized for the formation of morphological antonym and the scope of negation in the comparative construction does not include ‘more’ (the case of (19a)).

The second possibility can be represented as [adjective+ [‘more’ + negation]] (“not-more happy”). This kind of construction can be identified as a comparative of inferiority only if it gets non-compositional interpretation of inferiority, otherwise it is ambiguous. Apparently, an example of this construction may be found in Komi (21) (for discussion see also Bobaljik 2012: 217-218).

(21) Komi (Permic; Lytkin 1955: 168-169)

- a. *mičja-dʒik*  
 beautiful-**cmpr**  
 ‘more beautiful’
- b. *abu-dʒik mičja*  
**neg-cmpr** beautiful



‘less beautiful’

The third possibility is when the negated comparative of superiority (“not happier”) discussed earlier regularly gets the same unambiguous (i.e. non-compositional) interpretation of inferiority. It seems like this kind of marking might be present in Mongsen Ao: “An alternative strategy (to express inferiority – V.M.) is for a comparative proposition ‘X is bigger than Y’ to be globally negated by a clause final negative particle *nun*” (Coupe 2007: 261). Also see the literal translation of an example from Purépecha (22).

(22) Purépecha (Tarascan; Chamoreau 2007: 478)

*Maria sani = taru no wiŋapi-f-ti eski thu.*

Maria few = **more** NEG be strong-AOR-ASSERT3 SUB 2

‘Maria is weaker (less strong) than you are.’ (Maria is not stronger than you are)

The negation is found on the standard of comparison in the case of Kashibo-Kakataibo, but it seems to negate the whole predication. The non-literal translation suggests unambiguous interpretation of this construction:

(23) Kashibo-Kakataibo (Panoan; Biondi 2011: 343)

a. *Roberto ka mas xuá ki Emilio ‘iken*

*Roberto ka mas xuá ki Emilio ‘iken*

Roberto.ABS NAR.3PL more fat than Emilio be.3PL.NON.PAST

‘Roberto is fatter than Emilio.’

b. *Emilio ka mas xua ki Robertoma ‘iken*

*Emilio ka mas xua ki Roberto = ma*

Emilio.ABS NAR.3PL more fat than Roberto = NEG

*‘iken*

be.3PL.NON.PAST

‘Emilio is not fatter than Roberto (i.e. is less fat).’

Thus, the data suggests that the comparative of inferiority in theory can be based on the negation of comparative of superiority. Nevertheless, the description of these constructions needs more attention in regards to the scope of negation and regularity

of non-compositional interpretation. Given the lack of the uniform terminology in the subject field and the scarcity of information on the scope of negation, it might well be that some of examples actually do not fall into the comparative of inferiority as it is defined here. These issues require further investigation.

### 3.3.3 Other derived means

The type of marking of inferiority that is present in the vast amount of the Indo-European languages – that is a marker like English *less*, Russian *meneje*, Spanish *menos* – is also seen as a kind of derived marking here, though this point may seem debatable. The reasoning behind this analysis lies in the fact that these markers seem to contain the meaning of superiority in them (but no negation involved). This may not be obvious due to the suppletion as in the case of *little-less* in English, but in some languages the parts ‘more’ and ‘few, little’ are overt. Some examples like German *wenig-er* ‘less’ = *wenig* ‘little, few’ + *-er* ‘more’ were discussed in Bobaljik (2012: 215-217). Other examples may be found in Romanian (24), Albanian and several other Indo-European languages.

(24) Romanian (Indo-European; Dobrovie-Sorin & Giurgea 2013: 444-457)

- a. *Ion e mai înalt decât George.*  
 Ion is more tall than George  
 ‘Ion is taller than George.’
- b. *Maria e mai puțin înaltă decât Andreea.*  
 Maria is more little tall than Andreea  
 ‘Maria is less tall than Andreea.’

Another example of the marking of inferiority that involves marker of superiority and no negation is found in Paraguayan Guarani, where the marking of inferiority involves diminutive suffix *-i-* and the marker of superiority *-ve* (25).

(25) Paraguayan Guarani (Tupian; Estigarribia 2020: 249)

- a. *Che amba'apove ndehegui.*  
*che a-mba'apo-ve ndehegui*  
 I 1SG.ACT-work-more from.you  
 ‘I work more than you.’

- b. *omba'apo'ive ñande hígado ha pitikiri'i kuéra*  
*o-mba'apo-'i-ve*                      *ñande-hígado*                      *ha pitikiri'i = kuéra*  
3.ACT-work-DIM-more                      1PL.INCL.ACT-liver and kidney = PL  
'our liver and kidneys work less'

It is possible for a language to make use of several types of the comparative of inferiority. In that case it seems very likely that the language that has a non-derived construction will also have a derived one. It is also possible that some of the derived types are employed in those languages for which there is no information on the comparative of inferiority in the grammars (perhaps, derived markers were seen as simply compositional). The classification presented here is rather coarse but still may be of value to the theory of comparison.

#### 4. The asymmetry of superiority and inferiority

The connection of marking of inferiority to other constructions of comparison was chosen as a basis for classification of the types of comparatives of inferiority for one rather simple reason: this distinction clearly demonstrates that in some cases the marking of the comparative of inferiority is derived from other constructions of comparison. This may seem trivial, but at the same time no languages were found to form other constructions of comparison on the base of comparative of inferiority. These facts suggest that the comparative of inferiority is the marked member in the opposition of superiority and inferiority. Though the notion of markedness itself may be problematic, I will still use this term in the sense of multidimensional correlation of different properties pointing to the complexity, difficulty or abnormality of one of the members of opposition (Haspelmath 2006: 37-38).

The direction of the derivation is not the only sign of the markedness of the comparative of inferiority. Some languages appear to have no way to express 'less' other than switching the referents or using an antonym (e.g. in Mualang, Tjia 2007: 120, Mazatec and Chiquihuitlán, Jamieson 1988: 167-168). The fact that some languages have no markers of inferiority while having the comparative of superiority (and never the other way around) suggests that the comparative of inferiority is simply unnecessary for expressing the meaning of inequality.

Another indicator of the asymmetry is connected to the optionality of the parameter marker in some languages. It seems that, in those languages, this

optionality is only applied to the comparative of superiority, and for expression of inferiority the parameter marker is obligatory. An example is found in Turkish (26).

(26) Turkish (Turkic; Lewis 1967: 54)

- |    |                        |               |             |             |
|----|------------------------|---------------|-------------|-------------|
| a. | <i>kurşun-dan</i>      | <b>(daha)</b> | <i>ağır</i> |             |
|    | lead-ABL               | <b>(more)</b> | heavy       |             |
|    | 'heavier than lead'    |               |             |             |
| b. | <i>kurşun-dan</i>      | <b>(daha)</b> | <i>az</i>   | <i>ağır</i> |
|    | lead-ABL               | <b>(more)</b> | little      | heavy       |
|    | 'less heavy than lead' |               |             |             |

The opposite situation, when the construction with no parameter marker is interpreted as comparative of inferiority and the marker of inferiority is optional was not attested.

The phenomenon of “lesslessness” – the cross-linguistic absence of synthetic markers of inferiority – indicates the markedness of inferiority as well. This fact was pointed out without thorough discussion numerous times (Mel’čuk 1998: 119; Cuzzolin & Lehmann 2004: 1213; Treis 2018: ix), but in Bobaljik (2012) it got attention as an argument for the proposed Complexity Condition which states that “certain types of meanings are complex in ways that the resources of UG cannot pack into a single morpheme” (Bobaljik 2012: 212). It is suggested that inferiority in fact consists of superiority plus reversing operator and, therefore, the meaning of inferiority is considered to be too complex to be expressed by a single morpheme. This theory is reinforced by the aforementioned fact that in some languages the item in which the relation is expressed (e.g. English *less*) in fact contains the comparative of superiority thus indicating the derivation of the comparative of inferiority. Still, it has to be mentioned that, as discussed above, some languages do have non-derived markers of inferiority, some of which are even described as morphemes (see examples for Lizu (7) and Navajo (9)). Nevertheless, the absence of the synthetic marker of inferiority in the presence of the synthetic marker of superiority (even if relevant only for some languages) still may be considered as a sign of the markedness of the comparative of inferiority.

It has been argued that the marked member of the opposition is cognitively more complex: it requires more mental effort, attention and time to be processed (Givón

1991: 337). Psycholinguistic studies have shown that the asymmetry of superiority and inferiority in terms of usage, interpretation and judgement is validated experimentally: “People therefore use ‘more than’ statements more frequently, agree with them more, more readily believe them, and like them better” (Hoorens & Bruckmüller 2015: 765). The effects of this cognitive asymmetry are described in terms of saliency: perceptually it is easier to perceive the presence of an attribute rather than its absence (Hoorens & Bruckmüller 2015: 754).

One of the most prominent properties of comparatives of inferiority is their low frequency. It sometimes gets underlined in grammars (for example, for Arapaho in Cowell & Moss, 2008: 230) and definitely shows up in Russian.<sup>6</sup> The data from Russian National Corpus demonstrates that the comparative of inferiority constitutes only about 4.7% of all the cases of comparatives with adjectives (both synthetic and analytic). Considering that the structural markedness is, arguably, a result of frequency asymmetries, this distinction may be considered the most important one (Haspelmath 2006: 48-49).

Therefore, a number of features make comparative of inferiority the marked member in the opposition of superiority and inferiority. The conditions of use of the marked member in a pair of interchangeable items are of special interest. These conditions are of even higher interest in the case of comparative constructions since there is a possibility of using non-marked comparative of superiority through the switching of referents or using an antonym (*i.e.* *A horse is taller than a dog* and *A dog is shorter than a horse* instead of *A dog is less tall than a horse*), which seems to be enough to express comparison for some languages. The investigation of the distribution of the constructions of inequality will make it possible to describe the functions of the comparative of inferiority in those languages that have a way to express it.

## **5. Aspects of the use of the comparative of inferiority in Russian**

In this section, I provide a discursive, semantic, lexical and pragmatic account for the usage of the comparative of inferiority in Modern Russian language. Here I present only some of the findings based on the data from the Russian National Corpus (RNC)

---

<sup>6</sup> One example of an alleged preference of the comparative of inferiority is found in Bura (Chadic): “Bura adjectives and adverbs are compared ‘down’ rather than ‘up’. Instead of having ‘black, blacker’, they speak of ‘black, less black’” (Blench 2009: 12).

and data collected in the acceptability judgement experiment, for the in-depth discussion see Modina (2020).

Russian has both comparatives of superiority and inferiority. Superiority may be marked both synthetically and analytically and inferiority is marked only analytically, therefore illustrating the common phenomenon of “lesslessness”. The negation of equative construction is regularly interpreted as expressing ‘less’. Consequently, in this paper equatives under negation in Russian are considered to express inferiority. Therefore, there are four possible constructions to express inequality in Russian in terms of marking the relations of degrees: synthetic comparative of superiority (27.a), analytical comparative of superiority (27.b), comparative of inferiority with *meneje* ‘less’ (27.c) and negation of equality (27.d).

(27) Russian (Indo-European)

- a. *Vanja vyš-e čem Petja*  
 Vanja tall-‘more’ than Petja  
 ‘Vanja is taller than Petja’
- b. *Vanja boleje vysokij čem Petja*  
 Vanja more tall than Petja  
 ‘Vanja is taller than Petja’
- c. *Petja meneje vysokij čem Vanja*  
 Petja less tall than Vanja  
 ‘Petja is less tall than Vanja’
- d. *Petja ne takoj vysokij kak Vanja*  
 Petja NEG so tall as Vanja  
 ‘Petja is less tall than Vanja’ (lit. ‘Petja is not as tall as Vanja’)

Taking into account that either of compared items can, theoretically, appear in the position of comparee (e.g. *Horses are bigger than dogs* vs. *Dogs are less big than horses*) and the possibility of usage of any of the antonyms as the parameter (“bigness” or “smallness”) there is a rivalry of eight expressions with the same assertion of inequality.

In the light of the markedness of comparative of inferiority described in the previous section the comparative constructions of superiority are considered to be the default comparative constructions. Comparative of inferiority is seen as a “last resort”

that is used when the usage of comparative of superiority is impeded (or “blocked”) by factors of diverse nature. There are two types of blocking: “blocking of the conversive” (the need for the appearance of the particular referent in the position of comparee) and “blocking of the antonym” (the need for the usage of a particular member of an antonymic pair as the parameter of comparison). Essentially, it means that the comparative of inferiority is used when a particular referent needs to be described through a particular parameter. Blocking of a conversive is mostly related to the factors of a discursive nature and blocking of an antonym is mostly induced by the semantic features of an antonymic pair.

### 5.1. Discursive functions

Comparative may be seen as a type of valency-increasing operation because a new participant, namely, the standard of comparison, is added to the situation expressed by a positive (Plungyan 2011: 208). Comparative, then, is a source of the regular conversive formation. Jurij D. Apresyan distinguishes two kinds of conversives: lexical and grammatical ones. Both of them represent a shared category that reflects differences in the “logical accentuation” (Apresyan 1995: 257), i.e. reflects the distribution of the communicative ranks of the participants without changing the event structure (similar to the notion of function-changing operations in terms of Haspelmath, 2010: 236-237, but also includes non-grammatical oppositions). After the addition of a new participant the redistribution of communicative ranks becomes available. It can be employed either lexically with an antonym (28.b) or grammatically, with the same lexeme, through the comparative of inferiority (28.c).

(28) Russian (Indo-European)

- a. *Vanja* ***vyš-e*** *čem* *Petja*  
Vanja **tall-‘more’** than Petja  
‘Vanja is taller than Petja’
- b. *Petja* ***niž-e*** *čem* *Vanja*  
Petja **short-‘more’** than Vanja  
‘Petja is shorter than Vanja’
- c. *Vanja* ***meneje*** ***vysokij*** *čem* *Petja*  
Vanja **less tall** than Petja  
‘Vanja is less tall than Petja’

In this sense, the opposition of superiority and inferiority resembles the voice oppositions. Semantic, syntactic and communicative linking for the voice opposition in Russian can be conveniently demonstrated with the following tables (Table 1 and Table 2).

(29) Russian (Indo-European)

- a. *Raboč-ije*            *strojat*            *škol-u*  
 worker-NOM.PL        build                school-ACC.SG  
 ‘The workers build the school’

Table 1		
Semantic role	Agent	Patient
Syntactic function	Subject	(Nondirect) object
Communicative rank	Higher	Lower

Table 1: ‘The workers build the school’.

- b. *Škol-a*                *stroit-sja*            *raboč-imi*  
 school-NOM.SG        build-REFL            worker-INS.PL  
 ‘The school is built by workers’

Table 2		
Semantic role	Patient	Agent
Syntactic function	Subject	(Nondirect) object
Communicative rank	Higher	Lower

Table 2: ‘The school is built by workers’.

However, it is rather difficult to describe the arguments of a comparative construction in terms of either semantic roles or syntactic functions. For example, in Russian an adjective in comparative construction may appear both in an attributive and in a predicative function, and in the latter case both the comparee and the standard appear in the nominative case (30). In the attributive function an adjective can modify a



noun that is not the subject of the sentence. In this case the comparee is not the subject and is not marked by nominative (the comparee is marked by dative in (31)), while the standard still appears in nominative.

(30) Russian (Indo-European)

*Vmeste s tem taktičeskije raznoglasija na etot sčet*  
Together with this tactic **disagreements** on this issue  
*gorazdo meneje važny dlja nas čem obščnost' strategičeskix*  
far less important for us than **community** strategic  
*interesov Rossii i SŠA.*  
Interests Russia and USA

'At the same time the tactic **disagreement** on this issue is far less important for us than **common strategic interests** of Russia and the USA' [RNC]

(31) Russian (Indo-European)

*Noxo rešyl čto prežnije ugodja volki ostavili komu-to boleje*  
*Noxo decided that old land wolves left someone more*  
*sil'nomu, čem golod.*  
*strong than hunger*

'Noxo decided that wolves left the old land for **someone** who is stronger than **hunger**' [RNC]

Therefore, the notions of semantic roles and syntactic functions are not particularly helpful in the case of the comparative constructions. What is important here is the redistribution of communicative ranks behind the redistribution of syntactic functions (Plungyan 2011: 185–191) – a procedure that is applicable to the comparative constructions, given that the comparee and the standard are defined in communicative terms. The comparee (defined here as the referent which is described through the comparison) has the highest communicative rank in a comparative construction, while the standard has the lowest. Parallel to the agent-backgrounding operations, comparative of inferiority may be defined as a “comparee-backgrounding” operation. It can be illustrated by the tables parallel to the ones for the voice alternations (Table 3 and Table 4).

(32) Russian (Indo-European)

- a. *Vanja boleje vysokij čem Petja*  
 Vanja **more tall** than Petja  
 ‘Vanja is more tall than Petja’

Table 3		
Referent	Has the property to a <b>higher</b> degree	Has the property to a <b>lower</b> degree
Component of the comparative construction	Comparee	Standard
Communicative rank	Higher	Lower

Table 3: ‘Vanja is more tall than Petja’.

- b. *Petja meneje vysokij čem Vanja*  
 Petja **less tall** than Vanja  
 ‘Petja is less tall than Vanja’

Table 4		
Referent	Has the property to a <b>lower</b> degree	Has the property to a <b>higher</b> degree
Component of the comparative construction	Comparee	Standard
Communicative rank	Higher	Lower

Table 4: ‘Petja is less tall than Vanja’.

Consequently, the comparative of inferiority may serve functions similar to that of a derived voice. For example, the comparative of inferiority may be used for the change of topic (topic understood here as the referent that is “intended to persist in the

subsequent discourse” Givón 2020: 113), as in examples (33) (previous topic – other kinds of Lagomorphs) and (34) (previous topic – another type of transport).

(33) Russian (Indo-European)

*Meneje izvestnyje predstaviteli zajtseobraznyx piščuxi Oni mel'č-e*  
Less known specimen Lagomorphs pikas they small-‘more’

*zajtsev ushki i lapki u nix koroč-e.*  
hares ears and legs at they short-‘more’

‘Pikas are a less known kind of Lagomorphs. They are smaller than hares and their ears and legs are shorter’ [RNC]

(34) Russian (Indo-European)

*Vozdušnyj že transport v etom smysle meneje privlekatelen.*  
Air PTCL transportation in this sense less appealing

‘Air transportation is less appealing in that regard’ [RNC]

At the same time, the comparative of inferiority may serve an opposite function, i.e. it may be used for keeping the topic. This is the “blocking of the conversive” mentioned before: the comparative of inferiority is used to keep the particular participant in the position of the comparee because of the participant’s high topicality. The following examples are retrieved from encyclopedias: in an encyclopedia the topic of a particular article is identified clearly and should not be changed (example (35) is from an article about anaerobiosis, (36) from an article on rheas, (37) from an article about Germany).

(35) Russian (Indo-European)

*poetomu anaerobioz kak tip obmena veščestv*  
therefore anaerobios as type exchange substances

*meneje effectiveň čem aerobioz*  
less effective than aerobiosis

‘[...] this is why anaerobiosis as a type of metabolism is less effective than aerobiosis’ [RNC]

## (36) Russian (Indo-European)

*Nandu ptitsy vneshne poxožie na strausov no sostavljajuščie*  
 Rheas birds externally similar on ostriches but constituting  
*samostojatel'nyj otrad. Oni meneje jarkije i pomen'sh-e*  
 separate order **they less** colorfull and small-‘more’  
*strausov*  
 ostriches

‘rheas look like ostriches, but are considered to be a separate order. They are less colorful and smaller than ostriches’ [RNC]

## (37) Russian (Indo-European)

*Torgovlya aktsijami kompanij gorazdo meneje razvita*  
 Trading stock companies far less developed  
*čem v anslosaksonskix stranax*  
 than in Anglo-Saxon countries

‘The stock trading (in Germany) is far less developed than in the Anglo-Saxon countries’ [RNC]

For example, it would be quite questionable to make Anglo-Saxon countries the comparee and to use the comparative of superiority in (37) (‘The stock trading in the Anglo-Saxon countries is more developed than in Germany’) when the topic of the article is Germany.<sup>7</sup>

The usage of the comparative of inferiority in coreferential comparison (defined as comparative construction where “the comparee and the standard are the same object described at different stages” (Knjazev 2007: 208) can be motivated by topicality as well, as there may be a need for accentuating the earlier or the later state of affairs. The topic in (38), the condition after drinking alcohol, not before, while in (39) it is the state of affairs earlier in history.

---

<sup>7</sup> The comparative of superiority would be appropriate if there were no need to maintain the topicality of Germany (for example, if the article were about stock trading and the author would like to move on to description of the stock trading in other countries). In this case the comparative of superiority would serve to change the topic – the function, discussed earlier for the comparative of inferiority. An anonymous reviewer points out that these functions of keeping and changing the topic, therefore, are not exclusive for the comparative of inferiority, which is absolutely correct. Nevertheless, the goal of this Section is to explore the distribution of the comparative of inferiority rather than to pinpoint its unique features.

(38) Russian (Indo-European)

<i>Nemedlennoe</i>	<i>dejstvije</i>	<i>alkogola</i>	<i>na</i>	<i>organizm</i>	<i>čeloveka</i>
Immediate	effect	alcohol	on	organism	human
<i>xorošo izvestno.</i>	<i>Posle prinatii 2-3</i>	<i>portsyi</i>	<i>alkogola</i>	<i>zrenije</i>	
well known	after taking	portions	alcohol	<b>vision</b>	
<i>stanovitsja</i>	<i>menee</i>	<i>četkim i</i>	<i>jasnym</i>	<i>a reč</i>	
becomes	<b>less</b>	<b>precise and clear</b>		and speech	
<i>nevnatnoj</i>	<i>uxudšaetsja</i>	<i>koordinatsija</i>	<i>dviženij</i>		
unintelligible	worsens	coordination	actions		

‘The immediate effect of the alcohol on the human body is well-known. After 2-3 drinks the vision becomes less clear, the speech becomes unintelligible and the coordination worsens’ [essilor.ru]

(39) Russian (Indo-European)

<i>Narod</i>	<i>prosveščalsja</i>	<i>imenno</i>	<i>sozertsaja</i>	<i>steny</i>	<i>knižnaja</i>	<i>miniatura</i>
People	educated	exactly	looking	walls	<b>book</b>	<b>miniature</b>
<i>takže</i>	<i>kak</i>	<i>i</i>	<i>sama</i>	<i>kniga</i>	<i>byla</i>	<i>gorazdo meneje dostupnoj</i>
as.well as	as	and	itself	book	was far	<b>less accessible</b>

‘People educated themselves by looking at the walls because book miniatures, just like books, were far less accessible (than nowadays)’ [RNC]

Thus, on one hand, comparatives of superiority and inferiority reflect different communicative ranks of the participants, and, therefore, the comparative of inferiority may be used for changing the topic. On the other hand, the comparative of inferiority may function as a way to keep the topic. Either way, the discursive factors play a crucial role in the use of comparative constructions.

## 5.2. Semantic reasons

Here and further, I present some results of the statistical analysis performed on a sample of texts from Russian National Corpus (all texts were written after 1945).

The usage of comparative of superiority may be impeded by the features of the antonymic pair. Firstly, the most obvious reason for the blocking of the use of an antonym is the absence of a readily available antonym (40).

## (40) Russian (Indo-European)

*Mne kažetsja v otličije ot kotov koški meneje*  
 To.me seems in difference from male.cats female.cats less  
*svobodolubivye bol'se strematsja k domašnemu ujutu*  
**freedom-loving** more seeks to indoor coziness

'I feel like female cats, unlike male cats, are less freedom-loving, they seek the home coziness' [RNC]

## (40') Russian (Indo-European)

*Koški ?-ee / boleje ?*  
 Female.cats ?-'more' / more ?  
 'female cats are ?-er/ more ?'

Therefore, the absence of an antonym is analyzed as a possible condition in which the usage of the comparative of inferiority rises. As shown in Table 5, corpus data showed that this prediction is supported in the case of the negation of equative ( $\chi^2$ ,  $p=0.028$ ), but not in the case of the comparative of inferiority with *meneje*.<sup>8</sup>

Table 5		
	Comparative of superiority	Negated equative
Adjectives with no antonym	318 (95.50%)	15 (4.50%)
Adjectives with an antonym	1766 (97.62%)	43 (2.38%)
Sum	2084 (97.29%)	58 (2.71%)

Table 5: Adjectives with no antonym.

Secondly, pairs of morphological antonyms were analysed using the same data. These kind of pairs are quite heterogeneous in respect to their semantic features (Joshi 2012). While some antonyms seem to be quite interchangeable in comparative constructions (41), others do not (42).

<sup>8</sup> A readily available antonym was considered absent if there were no antonym in the AntonymsDictionary (database comprised of four Russian antonyms dictionary): <http://web-corpora.net/wsgi/antonyms.wsgi/antonyms>.

(41) Russian (Indo-European)

*Osnovnaja           zadača           sdelat'   vstavki   vozmožno   meneje*  
Main               goal           make   insertions possible   less  
*zametnymi*  
visible

'the main goal is to make the insertions as less visible as possible' [RNC]

(41') Russian (Indo-European)

*sdelat'   vstavki           vozmožno   boleje   nezametnymi*  
make   insertions       possible   more   invisible

'to make the insertions as more invisible as possible'

(42) Russian (Indo-European)

*vozmožno iz-za                   meneje   kačestvennogo   snaraženija*  
possible because.of       less   high-quality   equipment  
*nam   prixoditsja ispytyvat'   bol'sije   fizičeskije   nagruzki   čem*  
we   have.to   undergo   big.more   physical   excertion   than  
*našim   kollegam   na   Zapade*  
our   colleagues on   East

'it is possible that, because the quality of our equipment is lower (lit. because of less high-quality equipment) we undergo more physical exertion than our western colleagues' [RNC]

(42') Russian (Indo-European)

? *iz-za                   boleje   nekačestvennogo   snaraženija*  
because.of       more not-high-quality   equipment

Lit. 'because of more not-high-quality equipment'<sup>9</sup>

---

<sup>9</sup> There is no equivalent morphological antonymic pair in English for Russian *kačestvennyj/nekačestvennyj* 'high-quality'/ 'not-high-quality', but consider the difference between *less pleasant music* and *more unpleasant music*.

Meanwhile, corpus data shows that in pairs of morphological antonyms the proportion of comparative of inferiority is higher than in other cases ( $\chi^2$ ,  $p = 0.002$ ; cf. Table 6).<sup>10</sup>

Table 6		
	Comparative of superiority	Comparative with <i>meneje</i> and negated equative
Adjectives with morphological antonym	830 (90.41%)	88 (9.59%)
Other adjectives	1254 (93.93%)	81 (6.07%)
Sum	2084 (92.50%)	169 (7.50%)

Table 6: Adjectives with morphological antonym (number of lexemes)

This fact is interesting in the light of the derivation of the comparative of inferiority proposed in Bobaljik (2012): the author suggests that the comparative of inferiority is formed when the reversing operator (or negator) fails to merge with the adjective. If it were the case, the pair of antonyms like *pleasant* and *unpleasant* would not be expected to appear in the comparative of inferiority: the operator is overtly merged with the adjective. The analysis shows that Russian data does not support this derivation as these kinds of pairs, on the opposite, appear to be one of the conditions for the higher rate on the comparative of inferiority.

Thirdly, the influence of the class of adjective pair in terms of the classification proposed in Bierwisch (1989) on the choice of the construction was analysed. The author makes a distinction between dimensional (*tall*, *narrow*) and evaluative (*smart*, *lazy*) adjectives. Pairs of evaluative adjectives cannot be used for describing the same situation because they localize their referents on different scales. The pair in (43) seems to be more interchangeable than the pair in (44).

(43) Russian (Indo-European)

- a. *Vanja* *vyš-e* *čem* *Petja*  
*Vanja* tall-‘more’ than *Petja*

<sup>10</sup> The adjective was defined as only having a morphological antonym if, firstly, no lexical antonyms were found in the AntonymDictionary database and, secondly, a morphological antonym was found in more than five texts in the main corpus of RNC (so as to filter the cases of occasional word formation).



‘Vanja is taller than Petja’

- b. *Petja niž-e čem Vanja*  
 Petja short-‘more’ than Vanja  
 ‘Petja is shorter than Vanja’

(44) Russian (Indo-European)

- a. *Vanja umn-eje čem Petja*  
 Vanja smart-‘more’ than Petja  
 ‘Vanja is smarter than Petja’
- b. *Petja glup-eje čem Vanja*  
 Petja stupid-‘more’ than Vanja  
 ‘Petja is stupider than Vanja’<sup>11</sup>

These limitations are considered to be the cases of potential blocking of an antonym. Therefore, it is predicted that pairs of evaluative adjectives appear in the comparative constructions of inferiority more frequently than pairs of parametric adjectives (Table 7).<sup>12</sup> Corpus data support this prediction (Fisher’s exact test,  $p = 0.01252$ ).<sup>13</sup>

Table 7		
	Comparative of superiority	Comparative with <i>meneje</i> and negated equative
Dimensional adjectives	1176 (99.32%)	8 (0.68%)
Evaluative adjectives	315 (95.74%)	14 (4.26%)
Sum	1491 (98.54%)	22 (1.46%)

<sup>11</sup> Note a similar comment in a grammar of Basque when discussing the comparative of inferiority and the possibility of the use of the comparative of superiority with an antonym: “I will close this subsection with the important observation that Basque possesses no comparatives denoting inferiority in quality. In particular, there is no straightforward way of rendering the English sentence *Mary is less pretty than Eve*. Since *Miren Eba baino itsusiagoa da* ‘Mary is uglier than Eve’ won’t do, the only possibility is to resort to a negated equative: *Miren ez da Eba bezain polita* ‘Mary is not as pretty as Eve’” (de Rijk 2008: 713-714).

<sup>12</sup> This prediction in other terms is also formulated in Apresyan (1995: 266).

<sup>13</sup> The evaluation of the pair of adjectives regarding the opposition of evaluative and dimensional adjectives is rather convoluted. The details of this process can be found in Modina (2020).

**Table 7:** Parametric and evaluative adjectives (number of tokens)

To sum up, it can be said that the semantic features of a pair of antonyms (availability of an antonym, its morphological structure, its semantic class) influence the choice of a comparative construction.

### ***5.3. Lexical factors***

At the same time, it is possible that the factors that cannot be described in terms of blocking have an impact on the choice of the comparative construction as well. In particular, the high frequency of constructions of superiority may partially stem from the connection of those constructions to the high-frequency lexemes.

Firstly, the analysis of the corpus data showed that the distribution of constructions by lexemes is not homogeneous, i.e. different constructions are used with different lexemes. Though this finding may seem trivial, in fact it is not: if the low frequency was just a feature of the comparative of inferiority, it would mean that every adjective appears in the comparative on inferiority in about 5% of the cases (as the whole proportion of the comparative of inferiority is about 4.7%). Non-homogeneity shows that the lexeme itself is a factor that affects the distribution of the constructions.

Secondly, the comparative of superiority establishes closer links with lexemes than comparative of inferiority because even the lexemes with the highest proportions of appearance as the part of the comparative of inferiority (such as *zametnyj* 'notable' that has the highest rates for the construction of inferiority and the negated construction of equality) are used in those constructions in less than 50% of the cases. At the same time, the lexemes with the highest rates for the comparative of superiority are used in them almost exclusively, having the percentage of appearance close to 100%. The results, of course, are limited by the data analysed, but even these preliminary results demonstrate the differences between the constructions in question.

The distribution of constructions is connected to the frequency of the lexeme in such a way that the most frequent adjectives<sup>14</sup> tend to appear in the constructions of superiority ( $\chi^2$ ,  $p < 0.001$ , Table 8).

---

<sup>14</sup> Top 100 adjective lexemes by frequency based on the frequency dictionary Lyashevskaya & Sharoff (2009).

Table 8		
Table	Comparative of superiority	Comparative of inferiority
Top 100 adjective lexemes by frequency	2363 (97.93%)	50 (2.07%)
Other lexemes	2086 (92.51%)	169 (7.49%)
Sum	4449 (95.31%)	219 (4.69%)

**Table 8:** Top 100 lexemes by frequency (number of tokens).

This connection of the comparative of superiority to the high frequency of a lexeme partially explains why comparison of superiority appears in texts more often than the comparison of inferiority. Nevertheless, even without high-frequency lexemes the proportion of the comparative of inferiority is still much smaller than the proportion of the comparative of superiority (7.49% vs 92.51%).

#### 5.4. Pragmatic features

In a number of researches on the comparative in Russian it has been argued that the analytical comparative of superiority has a positive degree in the presupposition in contrast to the synthetic comparative of superiority (e.g. Mel'čuk 1998: 123; Grashchenkov & Lyutikova 2017: 124-125). That is 'Vanja is more tall than Petja' have the presupposition 'Petja is tall'.

Those investigations do not specify any pragmatic features of the comparative of inferiority (only a brief comment on the comparative with *meneje* is given in Grashchenkov & Lyutikova 2017: 125). It is worth noting that in the literature on the presuppositions equative is mentioned as one of the triggers of the presupposition (Levinson 1983: 183), i.e. 'Vanja is as tall as Petja' has the presupposition 'Petja is tall'. If it is true, the presupposition should hold for the negated equative which is considered here as an instance of the comparative of inferiority. The presence of a presupposition can influence the relative frequency of the usage of different comparative constructions, because presupposition naturally narrows down the amount of the contexts in which the utterance is possible.

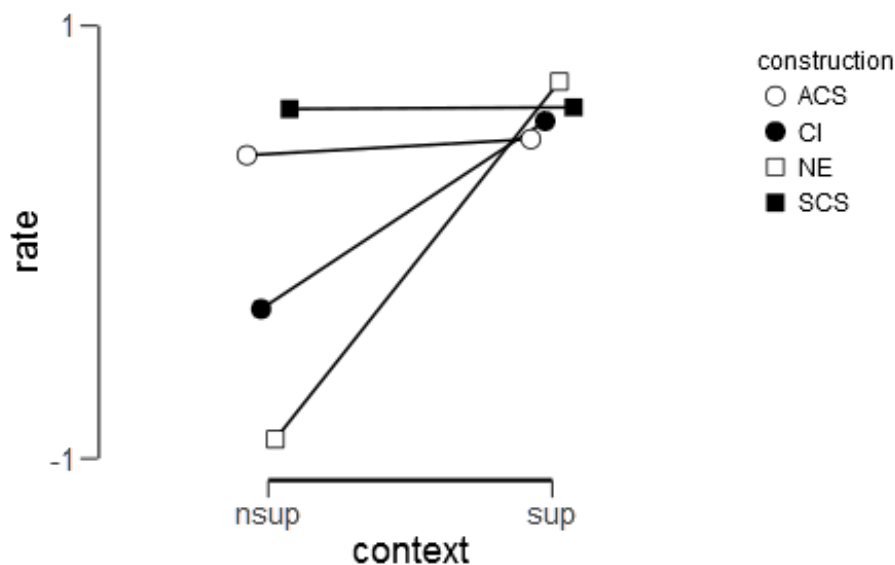
Aforementioned suggestions have not been tested on any experimental data. To test the hypothesis of pragmatic opposition of comparative constructions in Russian an acceptability judgment experiment was carried out. The hypothesis behind the test was that if a construction has a presupposition then an utterance with such a construction will have significantly lower acceptability rates in the case that the context does not support this presupposition. E.g. if the negated equative has presupposition of the positive degree for the standard, than the utterance (45a) should get lower rates than the utterance (45b).

(45) Russian (Indo-European)

- a. *Pervoje zdanije ne vysokoje. Vtoroe zdanije ne takoje vysokoje*  
 first building NEG tall second building NEG so tall  
*kak pervoje*  
 as first  
 ‘The first building is **not** tall. The second building is not as tall as the first one’
- b. *Pervoje zdanije vysokoje. Vtoroje zdanije ne takoje vysokoje*  
 first building tall second building NEG so tall  
*kak pervoje*  
 as first  
 ‘The first building is tall. The second building is not as tall as the first one’

The details of the experiment and thorough discussion can be found in Modina (2020). Here I report only the results relevant to the topic of the current paper (see Fig. 1). 146 respondents completed the survey, where they were asked to rate the acceptability of the utterances (32 target sentences and 32 fillers) on a Likert scale (from 1 to 5).

The analysis showed that synthetic and analytical comparative of superiority do not have a positive degree for the standard of comparison in presupposition: the rates of the utterances in the verifying and falsifying context do not differ significantly (Student’s t-test,  $p = 0.781$  and Mann-Whitney U-test,  $p = 0.052$ ). Nevertheless, the p-value in the case of analytical comparative of superiority is pretty close to the threshold of 0.05, reflecting the mixed intuitions found in the literature.



**Figure 1:** Mean rates of the target sentences. The rates are z-transformed. ACS – analytical comparative of superiority (*boleje vysokij* ‘more tall’), CI – comparative of inferiority with *meneje* (*meneje vysokij* ‘less tall’), NE – negated equative (*ne takoj vysokij kak* ‘not as tall as’), SCS – synthetic comparative of superiority (*vyše* ‘taller’); nsup – falsifying context, sup – verifying context.

Since the rates of utterances in verifying and falsifying context differ significantly in the case of comparative of inferiority (Mann-Whitney U-test,  $p < 0.001$ ) it can be said that the comparative of inferiority has the positive degree for the standard of comparison in presupposition. Consequently, the comparative of inferiority can be used in a smaller amount of contexts. This fact can partially account for the rarity of the comparative of inferiority in Russian.

In the situation of rivalry, members of the opposition can develop additional aspects of differentiation (Andersen 1983: 119). Marked members can start to serve additional functions that will keep them used in certain situations. From this perspective, the existence of the presupposition in the case of comparative of inferiority may motivate speakers to use these constructions. By using them, a speaker can give additional information about the standard of comparison.<sup>15</sup>

<sup>15</sup> Note a similar observation on the pragmatic features of one type of the comparative constructions in Amis: “Another intriguing observation on *ikaka/isafa* comparatives is their pragmatic implication. The usage of *ikaka/isafa* comparatives not only asserts the truth condition of the comparison, but also provides the speaker’s judgment regarding the degree of the property of comparison for the compared entities” (Kuo & Sung 2010: 35).

In the following example the information “two previous lines were wide” is conveyed by the usage of the comparative of inferiority *ne takije širokije* ‘less wide’ instead of the comparison of superiority such as *boleje uzkiye* ‘more narrow’.

(46) Russian (Indo-European)

<i>Neskol'ko raz</i>	<i>doroga peresekala ešče</i>	<i>polosy černoj</i>	<i>Gobi soveršenno</i>
Several times	road crossed	again lines black	Gobi completely
<i>besplodnoj</i>	<i>no ne</i>	<i>takije širokije</i>	<i>kak dve uže</i>
barren	but NEG	so wide	as two already
<i>projdennyje</i>	<i>tak čto</i>	<i>my proxodili ix</i>	<i>v odin nočnoj</i>
<b>crossed</b>	so that	we crossed	them in one night
<i>perexod</i>			
crossing			

‘The road crossed the lines of black and totally barren Gobi Desert several more times, but these lines were not as wide as the two previous ones, so we passed through them in one night’ [RNC]

Apparently, pragmatic characteristics of comparative construction come from an interaction of different factors. The pragmatic opposition can be a feature of constructions, types of adjectives or classes of antonym pairs and the overlap of these aspects needs to be investigated further. Nevertheless, it may be argued that the pragmatic features of the comparative of inferiority are different from those of the comparative of superiority and can both limit and provoke the usage of the former.

## 6. Conclusion

To conclude, it is important to underline that the description of the comparative of inferiority is a challenge even within one language due to the low frequency of the phenomenon. Most of the issues connected with the comparative of inferiority are not exhaustively discussed even on the data of well-known languages.

It is no wonder that this type of comparison hardly gets described in grammars, as the low frequency makes the elicitation difficult even in the case of the specific markers of inferiority. The markers that are derived are naturally quite often analysed as compositional, and the subtlety of semantic differences between the comparative of inferiority and other units (other constructions of comparison with negation,

morphological antonyms) do not serve the description of comparison well. Nevertheless, the relations between the comparative of inferiority and other phenomena allow us to ask a great amount of questions about the comparative of inferiority and comparison in general, so the researchers are encouraged to devote more attention to it.

The aspects of use of comparative constructions, even when examined thoroughly, do not seem to be easily identifiable and clear-cut. Still, it is precisely the rarity of the comparative of inferiority that makes it interesting and important: if so rare and seemingly redundant, why at all present? This is the question that was partially examined in this paper and that seeks further investigation.

## Acknowledgements

I am grateful to two anonymous reviewers for their insightful feedback. I am also greatly indebted to Dmitry Gerasimov and Artemii Kuznetsov for their valuable comments.

## Abbreviations

1 = first person	DECL = declarative	NAR = narrative register
2 = second person	DEM = demonstrative	NCM = non-common name
3 = third person	DEM.DIST = distal	marker
A = subject of transitive	demonstrative	NEG = negation
verb	DEM.PROX = proximal	NOM = nominative
ABL = ablative	demonstrative	NP = noun phrase
ACT = active	DIM = diminutive	OBJ = object
AF = actor focus	FEM = feminine	OBL = oblique
AI = animate subject,	GEN = genitive	PAST = past
intransitive verb stem	IC = initial change	PERF = peripheral
AOR = aorist	IMP = imperfective aspect	(argument)
AUX = auxiliary	INCL = inclusive of the	PL = plural
ASSERT = assertive	addressee(s)	POSS = possessive
ATT = attributive	INF = infinitive	PRED = predicate
B = cross-referenced	INS = instrumental	PTCL = particle
argument, class b	LK = linker	REFL = reflexive
CLF = classifier	LOC = locative	SUB = subordinating
CMPR = comparative	MASC = masculine	conjunction

SUBJ = subject of  
intransitive verb  
SG = singular

SP = Spanish loanword  
TOP = topic

VCC = verbless clause  
complement  
VSC = verbless copula

## References

- Aikhenvald, Alexandra Y. 2018. Comparison, Contrast and Similarity in Yalaku. *Linguistic Discovery* 16(1). 162-182.
- Álvarez, José. 2005. Comparative constructions in Guajiro/Wayuunaiki. *Opción* 21(47). 9–36.
- Andersen, Paul K. 1983. *Word Order Typology and Comparative Constructions*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Апресян, Юрий Д. (Апресян, Юрий Д.) 1995. *Izbrannye trudy, tom I. Leksicheskaja semantika* (Избранные труды. Том 1. Лексическая семантика. Синонимические средства языка) [Selected papers, Vol. 1, Lexical Semantics]. Moscow: Jazyki Russkoj Kul'tury.
- Bacelar, Laércio Nora. 2004. *Gramática da língua Kanoê*. PhD dissertation, Katholieke Universiteit Nijmegen.
- Bierwisch, Manfred. 1989. The semantics of gradation. In Manfred Bierwisch & Ewald Lang (eds.), *Dimensional Adjectives*, 71-261. Berlin: Springer.
- Biondi, Roberto Zariquiey. 2011. *A grammar of Kashibo-Kakataibo*. PhD Dissertation, LaTrobe University.
- Blench, Roger. 2009. *Bura grammar for a beginner*. Bura Language and Bible Translation Project.
- Bliese, Loren F. 1977. *A generative grammar study of Afar*. PhD Dissertation, University of Texas at Arlington.
- Bobaljik, Jonathan D. 2012. *Universals in Comparative Morphology Suppletion, Superlatives, and the Structure of Words*. Cambridge MA: The MIT Press.
- Bogal-Allbritten, Elizabeth & Coppock, Elizabeth. 2020. *Quantification, Degrees, and Beyond in Navajo*. In Peter Hallman (ed.), *Interactions of Degree and Quantification*, 121–162. Leiden: Brill.
- Caron, Bernard. 2017. Comparison, similarity and simulation in Zaar. In Yvonne Treis & Martine Vanhove (eds.), *Similitive and Equative Constructions: A Cross-linguistic Perspective*, 169–188. Amsterdam/Philadelphia: John Benjamins Publishing Company.



- Chamoreau, Claudine. 2007. Grammatical borrowing in Purepecha. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 465–480. Berlin: Mouton De Gruyter.
- Chirkova, Katia. 2019. Comparison constructions in Lizu (Tibeto-Burman). *Faits de Langues - Les Cahiers*. Paris: Ophrys.
- Coupe, Alexander R. 2007. *A Grammar of Mongsen Ao*. Berlin: Mouton de Gruyter.
- Cowell, Andrew & Moss, Sr. Alonzo. 2008. *The Arapaho language*. Boulder: University Press of Colorado.
- Cuzzolin, Pierluigi & Lehmann, Christian. 2004. Comparison and gradation. In Geert Booij, Christian Lehmann, Joachim Mugdan & Stavros Skopeteas (eds.), *Morphologie. Ein internationales Handbuch zur Flexion und Wortbildung, 1857-1882*. Berlin, New York: W. de Gruyter.
- Dedric, John M. & Casad, Eugene H. 1999. *Sonora Yaqui language structures*. Tucson: University of Arizona Press.
- Dixon, R. M. W. 2012. *Basic Linguistic Theory. Volume 3. Further Grammatical Topics*. Oxford: Oxford University Press.
- Dobrovie-Sorin, Carmen & Giurgea, Ion. 2013. *A Reference Grammar of Romanian. Volume 1: The noun phrase*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Estigarribia, Bruno. 2020. *A Grammar of Paraguayan Guarani*. London: ULC Press.
- Gawron, Jean Mark. 1995. Comparatives, Superlatives and Resolution. *Linguistics and Philosophy* 18. 333-380.
- Gerner, Matthias. 2013. *A Grammar of Nuosu*. Berlin: Mouton De Gruyter.
- Givón, Talmy. 1991. Markedness in grammar: Distributional, communicative and cognitive correlates of syntactic structure. *Studies in Language* 15(2). 335-370.
- Givón, Talmy. 2020. *Coherence*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Gorshenin, Maksym. 2012. The crosslinguistics of the superlative. In Cornelia Stroh (ed.), *Neues aus der Bremer Linguistikwerkstatt: Aktuelle Themen und Projekte* 31, 55-160. Bochum: Brockmeyer.
- Grashchenkov, Pavel V. & Lyutikova Ekaterina A. (Гращенко, Павел В., Лютикова, Екатерина А.) 2017. О синтаксисе сравнительного в русском языке (О синтаксисе компаратива в русском языке) [Syntax of the comparative in Russian]. *Русский язык в научном освещении (Russian Language and Linguistic Theory)* 1(33). 116–141.

- Haspelmath, Martin. 2006. Against markedness (and what to replace it with). *Journal of Linguistics* 42. 25–70.
- Haspelmath, Martin & Sims, Andrea D. 2010. *Understanding Morphology*. 2nd edn. London: Hodder Education.
- Haspelmath, Martin and the Leipzig Equative Constructions Team. 2017. Equative constructions in world-wide perspective. In Yvonne Treis & Martine Vanhove (eds.), *Similitive and Equative Constructions: A Cross-linguistic Perspective*, 9-32. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Heath, Jeffrey. 1999. *A Grammar of Koyra Chiini, the Songhay of Timbuktu*. Berlin: Mouton de Gruyter.
- Heine, Bernd. 1997. *Cognitive Foundations of Grammar*. Oxford: Oxford University Press.
- Hilpert, Martin. The English comparative – language structure and language use. *English Language and Linguistics* 12(3). 395-417.
- Honkasalo, Sami. 2019. *A Grammar of Eastern Geshiza: A Culturally Anchored Description*. PhD Dissertation, University of Helsinki.
- Hoorens, Vera & Bruckmüller, Susanne. 2015. Less is more? Think again! A cognitive fluency-based more–less asymmetry in comparative communication. *Journal of Personality and Social Psychology* 109(5). 753-766.
- Jamieson, Carole. 1988. *Gramática Mazateca: Mazateco de Chiquihuitlán de Juárez*. México: Instituto Lingüístico de Verano.
- Joshi, Shrikant. 2012. Affixal Negation – Direct, Indirect and their Subtypes. *Syntaxe & Sémantique* 13. 49-63.
- Князев, Юрий П. (Князев, Юрий П.) 2007. *Grammaticheskaja semantika* (Грамматическая семантика) [*Grammar Semantics*]. Moscow: Jazyki Russkoj Kul'tury.
- Kosheleva, Daria. 2016. Rival Forms of Comparatives in Russian. *Poljarnyj vestnik, Norwegian Journal of Slavic Studies* 19. 42–50.
- Kritka, Manfred. 2007. Negated Antonyms: Creating and Filling the Gap. In Uli Sauerland & Penka Stateva (eds.) *Presupposition and Implicature in Compositional Semantics*, 163–177. Palgrave Macmillan, London.
- Kuo, Cheng-Chuen and Sung, Li-May. 2010. On the syntax of Amis comparative constructions. *Concentric: Studies in Linguistics* 36. 25–57.

- Larajasse, Evangeliste de & Sampont, Cyprien de. 1897. *Practical grammar of the Somali language, with a manual of sentences*. London: Kegan Paul, Trench, Trübner & Co.
- Lefebvre, Claire & Brousseau, Anne-Marie. 2002. *A Grammar of Fongbe*. Berlin: Mouton de Gruyter.
- Leslau, Wolf. 1995. *Reference Grammar of Amharic*. Wiesbaden: Otto Harrassowitz.
- Levinson, Stephen C. 1983. *Pragmatics*. Cambridge: Cambridge University Press.
- Li, Charles & Thompson, Sandra. 1983. A Grammatical description of Xuwen: A colloquial dialect of Lei-zhou Peninsula (Part I). *Cahiers de Linguistique Asie Orientale* 13(1). 3–21.
- Lewis, Geoffrey L. 1967. *Turkish Grammar*. Oxford: Clarendon Press.
- Lyashevskaya, Olga N. & Sharoff, Sergej A. (Ляшевская, Ольга Н. & Шаров, Сергей А.) 2009. *Novyj častotnyj slovar' russkoj leksiki* (Новый частотный словарь русской лексики) [New frequency dictionary of Russian lexicon]. Available online at: <http://dict.ruslang.ru/freq.php>.
- Lytkin, Vasily I. (Лыткин, Василий И.) (ed.) 1955. *Sovremennyj komi jazyk* (Современный коми язык) [*Modern Komi*]. Syktyvkar: Komi knizhное izdatel'stvo.
- Mel'čuk, Igor A. (Мельчук, Игорь А.) 1998. *Kurs obščej morfologii. T. II* (Курс общей морфологии, Т. II) [General morphology course of. Vol. II]. Moscow–Vienna: Языки русской культуры–Wiener Slawistischer Almanach.
- Miyaoka, Osahito. 2012. *A grammar of Central Alaskan Yupik: An Eskimo Language*. Berlin: Mouton de Gruyter.
- Modina, Valeria V. (Модина, Валерия В.) 2020. *Inferiornye konstruktsii v kontekste konkurentsii komparativnyh konstrukcij s prilagatel'nymi v sovremennom rusском языке* (Инфериорные конструкции в контексте конкуренции компаративных конструкций с прилагательными в современном русском языке) [Comparative of inferiority in the context of the rivalry of comparative constructions with adjectives in Modern Russian]. Master's Thesis, St. Petersburg State University.
- Newman, Paul. 2000. *The Hausa Language: An Encyclopedic Reference Grammar*. New Haven: Yale University Press.
- Nonato, Rafael B. 2008. *Ainore Boe egore: um estudo descritivo da língua bororo*. Master's thesis, Universidade Estadual de Campinas
- Olawsky, Knut. 2006. *A Grammar of Urarina* [Mouton Grammar Library 37]. Berlin: Mouton de Gruyter.

- Plungian, Vladimir A. (Плунгян, Владимир А.) 2011. *Vvedenie v grammatičeskiju semantiku* (Общая морфология: Введение в проблематику) [Introduction to the grammar semantics]. Moscow: Russian State University for the Humanities.
- Quesada, J. Diego. 2000. *A grammar of Teribe*. Munich: Lincom Europa.
- Rijk, Rudolf P. G., de. 2008. *Standard Basque: A Progressive Grammar*. Cambridge, Massachusetts: MIT Press.
- Romero-Figueroa, Andrés. 1986. Warao\* Comparatives. *Studies in Language* 10. 97-108.
- Romero-Mendez, Rodrigo. 2008. *A reference grammar of Ayutla Mixe*. PhD Dissertation, University at Buffalo.
- Sandalo, Maria Filomena. 1995. *A Grammar of Kadiweu*. PhD Dissertation, University of Pittsburgh.
- Shvedova, Natalia Yu. (Шведова, Наталия Ю.) (ed.) 1980. *Russkaja grammatika. T. I* (Русская грамматика. Т. I) [Russian Grammar. Vol. I]. Moscow: Nauka.
- Stassen, Leon. 1985. *Comparison and Universal Grammar*. Oxford: Blackwell.
- Stolz, Thomas. 2013. *Competing Comparative Constructions in Europe*. Berlin: Akademie Verlag.
- Thomas, Guillaume. 2017. Comparison and the expression of degree in Mbyá Guaraní. In: Bruno Estigarribia & Justin Pinta (eds.). *Guarani Linguistics in the 21st Century*, 231–258. Leiden: Brill.
- Tjia, Johnny. 2007. *A Grammar of Mualang*. Utrecht: LOT.
- Treis, Yvonne. 2018. Comparative Constructions: An Introduction. *Linguistic Discovery* 16(1). i-xxvi.
- Treis, Yvonne. 2018. Comparison in Kambaata: Superiority, equality and similarity. *Linguistic Discovery* 16(1). 64–99.
- Ultan, Russell. 1972. Some features of basic comparative constructions. *Working Papers on Language Universals (Stanford)* 9. 117–162.
- Wojtylak, Katarzyna I. 2018. Comparative Constructions in Murui (Witotoan, Northwest Amazonia). *Linguistic Discovery* 16(1). 162-182.

**CONTACT**

[st072157@student.spbu.ru](mailto:st072157@student.spbu.ru)

## Small events.

# Verbal diminutives in the languages of the world

JENNY AUDRING<sup>1</sup>, STERRE LEUFKENS<sup>2</sup>, EVA VAN LIER<sup>3</sup>

<sup>1</sup>LEIDEN UNIVERSITY, <sup>2</sup>UTRECHT UNIVERSITY, <sup>3</sup>UNIVERSITY OF AMSTERDAM

Submitted: 16/11/20 Revised version: 16/05/21

Accepted: 21/05/21 Published: 31/08/2021

### Abstract

Diminutives are typically nouns. However, verbs can also be diminutivised, i.e. marked for reduced intensity, duration, seriousness or success of the action or event. This paper is a first attempt at a typology of verbal diminutives, based on a balanced sample of 248 languages. We discuss the analytical and terminological challenges that arise from the study of a category that is not widely recognised and does not have an established place in grammatical descriptions. Our sample shows that verbal diminutives occur across the world, with a slightly higher predominance in the Americas and somewhat fewer cases in Africa. Among the language families, Austronesian has the highest percentage of verbal diminutives in our sample. We present our results for the various formal exponents of verbal diminution on the one hand and the array of semantic effects on the other. Meanings are separated into three categories: attenuation in quantity, attenuation in quality and affective meanings. In many cases, markers of verbal diminution encode additional meanings, some of which contradict the core meaning of attenuation by expressing intensity, durativity or iteration. Such apparent paradoxes have parallels in nominal diminutives. The paper closes with recommendations for further research.

**Keywords:** verbal diminutive; diminutive; evaluative morphology; attenuation; reduplication; morphology.

## 1. Introduction

Verbal diminutives are grammatical constructions indicating that an action or event is ‘smaller’ than usual. A truly classic example is (1), a line from Horatius’ *Ars Poetica*:

(1) Latin (Horatius, *Ars Poetica*, line 359)

<i>indignor</i>	<i>quandoque</i>	<i>bonus</i>	<i>dormi-ta-t</i>	<i>Homerus</i>
indignant	whenever	good	sleep-DIM-3SG.PRS	Homer
'I am indignant when worthy Homer nods.'				

The verb *dormitare* ‘to nod, to snooze’ is the diminutive of the verb *dormire* ‘to sleep’: Homer sleeps lightly and for a brief time.

Like all diminutives, verbal diminutives belong to the domain of evaluative morphology. As such, they can be considered an implicit form of comparison, expressing reduction or attenuation with respect to some standard (Jurafsky 1996: 551, Dressler & Merlini Barbaresi 1994: 153, Körtvélyessy 2015: 4, 32, 41-42). This view is explicitly emphasised by Körtvélyessy, who defines evaluative morphology as “a continuum in which prototypical cases express the meaning of quantity under or above the default value” (2015: 4). In her model, the standard or default for verbal diminutives is anchored in the cognitive category “quantity of action”.

Conceptually speaking, the quantity of actions can be reduced in a variety of dimensions.<sup>1</sup> The most obvious are intensity, time and space (but see Dressler & Merlini Barbaresi 1994: 132, for whom intensity is the only gradable dimension for verbs). As we will see (§2.3 and §3.3.2), not all of these are used for diminution in equal measure: attenuation in the spatial dimension is found more rarely than attenuation in time or intensity. Furthermore, actions could potentially be graded according to their frequency, but the lowest end for this parameter, a single event, is the norm rather than the exception for most verbs.<sup>2</sup> As we will see later, frequency does make an appearance in our data, but not in the sense of low frequency. Instead,

---

<sup>1</sup> A similar variety of dimensions is also found with nominal diminutives. While concrete nouns are usually diminutivised in the spatial domain, abstract nouns can be reduced in non-spatial domains such as time, strength or scale, as in Italian *sinfonia* > *sinfonietta* ‘small-scale symphony’, *cena* > *cenetta* ‘small supper’, *pioggia* > *pioggerella* ‘light rain, drizzle’ (Taylor 2003: 173). In a similar spirit, Jurafsky (1996: 559) points out that different semantic shades of the diminutive involve different “scales”, such as amount, time, space or illocutionary force.

<sup>2</sup> In cases where languages employ singulative, punctual or semelfactive markers to mark the single occurrence of a normally repeated action, it is questionable whether these should be analysed as diminutives. This type of marking resembles nominal singulatives, which are not normally considered diminutives, just the way plurals are not seen as augmentatives. Therefore, such cases were not included in this study.

diminution of the verbal action may go hand in hand with iterativity; this is shown and discussed in §3.3.2.

We extend Körtvélyessy's definition of verbal diminution by also including attenuation in qualitative dimensions, such as seriousness and/or effort (diminution implies nonchalant, pretended or playful execution of the action) and completeness and/or success (diminutive marks incomplete or unsuccessful actions). In addition, verbal diminutives, like nominal diminutives, can have a variety of affective meanings such as endearment or contempt. The three functional domains – reduction in quantity, reduction in quality and affective functions – will be discussed further in §2.3 and §3.3.

Verbal diminutives such as example (1) are common and frequent in some language families, e.g. in Slavic and Romance, but in general do not constitute a well-known category and are cross-linguistically understudied (though see the explorations in Olsson 2012 and Makarchuk 2020). While accounts exist for individual languages, e.g. Finnish (Armoskaite & Koskinen 2008), Italian (Grandi 2009, Tovená 2011), Hebrew (Greenberg 2010), French (Amiot & Stosic 2014), German (Weidhaas & Schmid 2015), Russian (Makarova 2014) and Czech (Káňa 2017), the typologically oriented literature on evaluative morphology such as Jurafsky (1996), Bauer (1997), Grandi & Körtvélyessy (2015) and Körtvélyessy (2015) clearly shows that the primary domain of the diminutive, both in cross-linguistic attestation and in linguists' awareness, is the noun.<sup>3</sup>

The present study is the first attempt at a substantial typology of verbal diminution, covering 248 languages. In our sample, we found 112 verbal diminutives, with some languages displaying more than one marker or strategy. We begin by providing our methodology of sampling in §2.1. In §2.2 we situate the category among other grammatical properties and address terminological concerns. Our reasons to include or exclude individual cases are motivated in §2.3. Section §3 constitutes the heart of the paper: §3.1 summarizes the results of our cross-linguistic study, §3.2 focusses on the formal exponents of verbal diminution, and §3.3 discusses the semantics and

---

<sup>3</sup> Various sources show that nominal diminutives are more common than verbal or adjectival diminutives (see e.g. Dressler & Merlini Barbaresi 1994: 131, Bauer 1997 and the sources cited there, as well as Körtvélyessy 2015: 102). Occasionally, verbal or adjectival diminutives are thought to be metaphorical or metonymical extensions of nominal diminutives (e.g. Makarova 2015: 16). We do not pursue this hypothesis, since, as just shown, verbs possess diminishable semantics of their own.

pragmatics of the constructions in question. §4 outlines some issues of further interest, and §5 concludes the paper.

## 2. Methodology

### 2.1 Sampling

This paper is based on a combination of two variety samples of 102 and 223 languages, respectively. The samples were composed by applying the Diversity Value technique to the language classification of Ethnologue 2015 (Lewis et al. 2015). This technique, developed by Rijkhoff & Bakker (see Rijkhoff et al. 1993, Rijkhoff & Bakker 1998, Bakker 2011) and validated by Miestamo and colleagues (Miestamo et al. 2016), involves computing diversity values for the nodes in a language family tree. These values reflect the degree of internal complexity of the subgroups under the nodes and thereby determine the number of languages that should be sampled from each subgroup, given the desired sample size.

For a pilot study on verbal diminutives (henceforth: VDims), we used the minimal sample (one language per family) of 102 languages (see Audring et al. 2019). In order to include more variation and to consolidate our preliminary results, we extended the coverage by adding Mattiola's (2020) sample, constructed according to the same technique applied to the same classification.<sup>4</sup> Elimination of overlap yielded a final sample of 248 languages.

Merging Mattiola's sample with our pilot sample led to a number of cases where different representatives had been chosen within a particular genetic grouping. When this happened, we used the following criteria to decide between the languages: (i) general quality of the available description; (ii) if relevant, status of the (description of the) VDim construction. VDim constructions were occasionally harder to classify in one language than in the other, in which case we favoured the language with the

---

<sup>4</sup> As Mattiola (2020) explains, his 223-language sample is in fact a reduced version of a larger 424-language sample constructed by adding all isolates, as well as a number of pidgin/creole languages and sign languages, to an initial 350-language sample. While Miestamo et al. (2016) suggest selecting just one isolate and one unclassified language, Mattiola views such languages as essentially representing a family of their own. Moreover, pidgins, creoles, and sign languages tend to be underrepresented in linguistic typology. Hence, including them also adds to overall diversity. Eventually, however, a number of languages in these categories were dropped for practical reasons, resulting in a 223-language sample.



clearest case. If none of this was decisive, we chose, if available, a language with a VDim rather than one without a VDim, in order to maximize the diversity of our coverage.

In a number of cases, we were not able to find usable descriptions for the representative of a particular group. If we could not replace the language with another from the same group (either because it was an isolate or because no - suitable - descriptions were available for any other languages in the group), this meant that the group had to be omitted. This was the case for 13 groups.

The sample languages, together with their genealogical classifications and the source(s) we used can be found in the online Appendix (Audring et al. 2021).

## ***2.2 Verbal diminutives and their place in the grammar***

Since verbal diminutives are not a standard category in descriptive grammars, they are described – if at all – as part of various grammatical subsystems. The most obvious, as mentioned in §1, is the domain of evaluative morphology. However, diminution is usually discussed in the context of nominal morphology; the extension of the category to verbs and other parts of speech may be mentioned in passing. More often, grammars report attenuative markers in a chapter or section on verbal aspect (or TAM morphology in general), alongside habituais, iteratives and intensives (which, with Körtvélyessy 2015, might be considered cases of verbal augmentatives). In other cases, instances of verbal diminution were found in the context of pluractionality, as meanings of attenuation and repetition appear in tandem in various languages. The apparent paradox between reduction (attenuation) and increase (iteration), which we might describe as ‘less is more’, will be addressed briefly in §3.3.2. Finally, we sometimes found relevant descriptions in sections about the semantic effects of verbal reduplication. Since this is a formal rather than a functional criterion, it sidesteps the question of where in the grammar verbal diminutives belong.

The practical challenge of identifying relevant cases in descriptive sources reflects theoretical difficulties in situating the phenomenon in a particular grammatical subsystem. These issues are exacerbated by a great variation in terminology. The term “verbal diminutive” does not commonly appear in grammars, especially for languages outside Europe. A telling example is the grammar of Chukchi (Dunn 1999), where one of the suffixes in (2), *-qeet*, is glossed as DIM (diminutive), while the prefix *mec-* in the same example is glossed as APPR (approximative), although its description

suggests a highly similar function: “The prefix *mec-* indicates that the action/event of the verb occurs slightly or incompletely” (Dunn 1999: 266).<sup>5</sup>

(2) Chukchi (Chukotko-Kamchatkan; Dunn 1999: 268)

*e-mec-pintəqet-qeet-lin*

PRF-APPR-show.self-DIM-3SG

‘It showed itself slightly.’

Other candidate cases are referred to as “delimitatives” (Mandarin Chinese; Li and Thompson 1981: 232), “incompletives” (Alto Perené; Mihas 2015: 250), “de-intensification” (Epena Pedee; Harms 1994: 40), or “attenuatives” (Toqabaqita; Lichtenberk 2008: 186).

This variation in terminology has repercussions for a cross-linguistic investigation of VDims. In addition, we will see that the category is heterogeneous, especially in its semantics. Our considerations and criteria for including or excluding potential cases will be explained next.

### 2.3 Selection and coding

As outlined in §1, we recognize three sets of functions of VDims: a) reduction in quantity of the action, b) reduction in quality of the action and c) affective meanings. Functions a) and b) can be considered semantic, while functions of type c) are pragmatic in nature.

Quantitative reduction can be seen in example (3) from Jarawara, where the reduplication of the first syllable of the verb (here: *joko* ‘push’) changes the meaning to ‘V a bit’.

(3) Jarawara (Arauan/Jamamadi; Dixon 2004: 275)

*Okomobi awa jo.joko na-ka*

name stick RDP.push AUX-DECL

‘Okomobi is giving the stick a little push.’

Reduction in the qualitative sense can be seen, for example, in the Austronesian language Ibatan (Maree 2007), which has a “pretense mode” indicating that “the

<sup>5</sup> The similarity of the meanings is pointed out explicitly by the author (Dunn 1999: 266).

action is pretended or is performed playfully or hypocritically” (Maree 2007: 209). An example is (4); the markers are a prefix glossed as PRT ‘pretense’, plus the reduplication of the first stem syllable and vowel lengthening.

- (4) Ibatan (Austronesian, Malayo-Polynesian; Maree 2007: 209, example shortened and glosses modified)

*Naysin-ta:tanyis*      *saw*      *adedekey*

PRT-RDP-cry              3PL.NOM      children

‘The children pretended to cry.’

A second example was found in the Australian (Bunaban) language Gooniyandi (McGregor 1990), which has so-called affixal iteratives. For some of the examples, the description says “It may be that the process was attempted, unsuccessfully, a number of times” (McGregor 1990: 243). More examples are given in §3.3.2 below.

Affective meanings have a slightly different status, as not every expression of affection or deprecation on the verb qualifies as a verbal diminutive. Rather, such functions were included only when they occurred alongside semantic functions of reduced quantity and/or quality. A relevant example is shown in (5), from the isolate language Karok. Verbal diminution is marked by the suffix *-ač* as well as by changes in the stem consonants, here a shift from *r* to *n*. The semantic effect is ‘to V a little’, but this type of marking can also “[indicate] a speaker's familiar or affectionate attitude towards a situation”.

- (5) Karok (isolate, North America; Bright 1957: 114)<sup>6</sup>

*ikrémyahtih*

*iknémyahtiháč*

‘(wind) to be blowing’

‘(wind) to be blowing a little’

Several grammars described constructions that resembled verbal diminutives, but were not included in our study. The clearest case was diminutive markers appearing on the verb, but not actually pertaining to the verb but rather to one of its arguments. This situation can be seen in the Algonquian language Passamaquoddy, where diminutive verb forms typically indicate that one of the arguments of the verb is “small, cute, or an object of affection or pity” (LeSourd 1995: 133). For example, the form *mehci-né-hs-o*, from *mehci-né* ‘he/she dies, is dead’, does not mean ‘he/she dies a

---

<sup>6</sup> Where examples lack glossing, the source did not provide them.

little' (if such a meaning makes sense at all) but 'the little one is dead' (LeSourd 1995: 108). A similar situation is discussed for the four Northern American languages Lakhota, Creek, Chicasaw, and Maricopa in Munro (1988), and for the Arawak language Mojeño in Rose (2018). However, participant diminution can sometimes be observed in true verbal diminutives, for example in the language Iquito, which has a "cumulative diminutive" (CUM) with a variety of meanings; see example (6).

(6) Iquito (Zaparoan; Lai 2009: 533)

*Nu = capi-juu-yaa-Ø.*

3S = cook-CUM-IPFV-EC

- a. 'She is cooking slowly.'
- b. 'She is cooking in small portions (i.e. in a small pot) many times in order to make enough food for a meal.'
- c. 'A little child is cooking a small portion.'

Among these meanings, (6a) is a case of attenuated action, i.e. verbal diminution, while (6b) and (6c) indicate small participants (the object in b., subject and object in c.). If participant diminution was the only reported meaning (unlike in Iquito), we did not include the construction as a case of a VDim.

A second type of excluded cases can be seen in the Torricelli language Walman (Dryer 2016). In this language, verbs show diminutive agreement with diminutive nouns. This has no consequences for the semantics of the verb, which is why it is not a verbal diminutive according to our typology.

Another case excluded on semantic grounds was the category sometimes glossed as "frustrative" or "avertive", which encodes events intended to happen or on the verge of happening, but not actually occurring. For an action to be attenuated, it is necessary that it has at least been initialised; hence, we distinguish actions that are incomplete or unsuccessful from events that have failed to take place altogether.

In addition, we chose not to consider instances of imperfective or so-called incompletive aspect. Incompletive marking usually does not mean that the action or event is literally incomplete, but that the perspective is on the progress of the event, with the end point out of view. Moreover, including such cases would bloat the category of VDims in an unhelpful and un insightful way. This is not to say that there is no relation between (im)perfectivity and verbal diminution. For example, Grandi (2009) points out that evaluative suffixes in Italian are more frequently combined

with imperfective verbs. We will briefly discuss the relation between verbal diminution and aspect in §4.

Last but not least, we were careful about cases where we could not be sure that the diminution is truly verbal rather than adjectival. This is the case whenever languages express property meanings by verbal lexemes. For example, in the Nilo-Saharan language Lango (Noonan 1992) initial CV reduplication (including tone) has a diminutive function, but is restricted to property-denoting lexemes, which can be used either attributively or predicatively, e.g. *ràc* '(be) bad' → *ràràc* '(be) sort of bad' (Noonan 1992: 174). Such cases, where a diminutive strategy is limited to stative verbs, were excluded from our data.

All potential instances in the grammars were recorded and coded for the type of formal exponent(s) (§3.2) and their functional effect(s) (§3.3). Among the functional effects, we did not only include those that were directly indicative of verbal diminution, but also any other effect described for the markers in question. This resulted in a penumbra of associated meanings, which will be discussed in §3.3. We made special note of cases where attenuation coincided with iterativity (the 'less is more' diminutives, see §3.3.2).

We also included cases that were mentioned in the text but for which no examples were provided. If there was a discrepancy between the gloss and the idiomatic translation, we followed the gloss. Glosses and idiomatic translation in grammars were generally left unchanged. When a source failed to provide glosses, as in example (5) above, we did not construe them ourselves.

### **3. Results**

#### ***3.1 General distribution of VDims***

Our sample yielded 112 VDims in 85 languages. The number of VDims exceeds the number of languages, because some languages have more than one VDim construction. In order to find out whether VDims are especially common in particular language families and/or linguistic areas, we investigated the geographical and genealogical spread by computing the percentage of languages with VDims in the six macro-areas and in the ten language families in our sample from which we included more than one language. The results for the areal distribution are shown in Table 1, ordered by proportion of VDims. Comparing the proportion of VDims per macro-area

to the proportion of VDim(s) in the total sample, it is clear that VDim(s) are relatively frequent in South-America and relatively uncommon in the languages of Africa (the line in the table separates the areas above average from the areas on/below average).

Macro-area	Number of languages in sample	Number of languages with VDim(s)
South-America	54	22 (41%)
North-America	42	16 (38%)
Eurasia	36	13 (36%)
Papunesia	68	23 (34%)
Australia	6	2 (33%)
Africa	42	9 (21%)
Total	248	85 (34%)

**Table 1:** Areal distribution of VDim languages.

Table 2 provides an overview of the genealogical distribution of VDim(s) in our sample. The table only includes language families represented by two or more languages in the sample (note that some of the figures are more informative than others due to the higher or lower number of languages in the sample). The percentages show that the number of VDim(s) is relatively high in Austronesian languages, while they are relatively uncommon in the Trans-New Guinean family, in Sino-Tibetan, Nilo-Saharan as well as in Indo-European.

Family	Number of languages in sample	Number of languages with VDim(s)
Austronesian	28	18 (64%)
Creole	2	1 (50%)
Australian	6	2 (33%)
Afro-Asiatic	7	2 (29%)
Niger Congo	21	5 (24%)
Nilo-Saharan	6	1 (17%)
Sino-Tibetan	6	1 (17%)
Indo-European	7	1 (14%)
Trans-New Guinea	12	1 (8%)
Total	99	32 (32%)

**Table 2:** Genealogical distribution of VDim languages.

In the next two sections, we discuss the formal strategies used in VDim marking and elaborate on their semantic and pragmatic functions.

### 3.2 Form

Languages make use of a wide variety of strategies to express verbal diminution. These strategies are:

- Affixation (mostly suffixation, occasionally pre- or infixation)
- Reduplication
- Cliticisation
- Freestanding element/particle
- Verbal element (auxiliary, post-verb, serialised verb)<sup>7</sup>
- Base modification (segmental or suprasegmental change)

While Körtvélyessy mentions compounding as a diminution strategy for nouns (the diminutiviser usually being the word for ‘child’ or ‘young’, Körtvélyessy 2015: 20), we did not encounter any cases of verbal diminution by compounding.

Before discussing the frequency of the various strategies, we illustrate each with examples.

Affixation is by far the most frequent strategy. Example (7) from Central Alaskan Yup’ik shows suffixation, the examples in (8) from Georgian illustrate prefixation.

(7) Central Alaskan Yup’ik (Eskimo-Aleut; Miyaoka 2012: 653)

*tang-cuar-tuq*

see-DIM-3SG

‘He sees a little bit.’

Georgian has three prefixes (called “preverbs”) with diminutive-like semantics, plus a fourth translated roughly as ‘to V (aimlessly) around’, a meaning commonly found with VDims, but analysed as intensification in Hewitt (1995: 164). Example (8a)

---

<sup>7</sup> Verbal diminutives expressed by serial verb constructions may be more common than our data indicates, as grammars are often unable to exhaustively cover the varied semantic effects of verb serialisation.

shows the prefixes *mo-* and *šer-*, (8b) illustrates the use of *c'a*, and (8c) demonstrates that *c'a-* and *mo-* can occur in combination.

(8) Georgian (Kartvelian; Hewitt 1995: 162-167)

a.	<i>darbilda</i>	<i>mo-rbilda/ šer-bilda</i>
	‘X softened.’	‘X softened a bit.’
b.	<i>isauzmeb</i>	<i>c'a-isauzmeb</i>
	‘You will have breakfast.’	‘You will snatch a bit of breakfast.’
c.	<i>ic'vimebs</i>	<i>c'a-mo-c'vims</i>
	‘It will rain.’	‘There will be a short rainfall.’

Reduplication is also highly common. We already saw a case in §2.3 (Jarawara, example (3)), illustrating partial reduplication. Full reduplication is shown in example (9) from Waray.

(9) Waray (Austronesian; Oyzon & Payne in prep., Thomas Payne p.c.)

a.	<i>káon</i>	<i>káon-kaon</i>
	‘eat’	‘eat a little/ playfully’
b.	<i>lakat</i>	<i>lákát-lakat</i>
	‘walk’	‘walk a little/ playfully/randomly’

Other languages mark VDimS by means of cliticisation; one of them is Tariana (example (10)).

(10) Tariana (Arawakan; Aikhenvald 2003: 366, 193)

*nha kida = tuki = sina*  
 they ready = DIM = REM.PST.INF  
 ‘They must be a little bit ready.’



The Malayo-Polynesian language Toqabaqita marks verbal diminution by means of a particle, a freestanding non-verbal element.<sup>8</sup> There are three such particles; example (11) shows two of them, glossed as ATTN, i.e. attenuative.

(11) Toqabaqita (Austronesian, Malayo-Polynesian; Lichtenberk 2008: 169)

*Nia kai {thafa/thafeqe} qono naqa*

3SG 3SG.IPFV ATTN/ATTN sit PRF

‘He/she is feeling better now.’ (lit.: ‘He/she is sitting a little now.’)

(Said about a person recovering from an illness.)

The same language also has a verbal marker, *sukani*, with a similar meaning, shown in (12).

(12) Toqabaqita (Austronesian, Malayo-Polynesian; Lichtenberk 2008: 168)

*Nau ku sukani mataqi*

1SG 1SG.NFUT be.of.little.degree be.sick

‘I am a little sick.’

Free-standing verbal markers of verbal diminution are also found in other languages, where they are described as auxiliaries, postverbs or serial verb constructions. Example (13) shows the attenuative auxiliary *gi* in Bauzi, which has a connotation of ‘just, only’ and therefore fits the qualitative dimension we will call NON-SERIOUS in §3.3. This language encodes other aspectual distinctions with free words, so *gi* is part of a set, which prompts us to treat it as a grammaticalised, not just a lexical element.

(13) Bauzi (Geelvink Bay; Briley 1976: 8)

*em gi la lo*

I ATTN go doing

‘I’m just going.’ (nothing special in mind)

The Turkic language Uyghur has post-verb constructions, i.e. verbs following another verb and bearing a generalised meaning (some sources, e.g. Hahn 1998: 390, describe

---

<sup>8</sup> A reviewer pointed out that “evaluative morphology” should imply morphological, i.e. bound markers. In this paper, we use the term only to situate our research amongst other relevant work, without committing ourselves to such limitations.

them as auxiliaries). VDim meanings can be found with posture verbs such as *oltu(r)*- (lit. ‘sit’) (Lopnor dialect; Abdurehim 2014: 160) or *tur*- (lit. ‘stand’) (Hahn 1998: 392).

Last but not least, we find verbal diminution expressed by base modification. Two striking examples come from Beja (North Cushitic) and Huave (Huavean), shown in (14) and (15). In Beja, the alveolar trill /r/ shifts to an alveolar lateral approximant /l/ to form diminutives of nouns, adjectives and verbs. (14) is a verbal example in which the ‘smallness’ of the action is beautifully evident.<sup>9</sup>

(14) Beja (North Cushitic; Vanhove & Ahmed 2018: 67)

<i>birʔik</i>	<i>biɫik</i>
‘fly’ (of birds)	‘flutter about’ (of butterflies)

In Huave, “[d]iminutivization essentially involves raising of all non-high root vowels to high, plus palatalization of any eligible (i.e. coronal) root-final consonants” (Kim 2008: 320). An example is (15).

(15) Huave (Huavean; Kim 2008: 322)

<i>jajybij</i>	<i>jujyuj</i>
‘shake’	‘shake gently’

The sound symbolism in diminutives expressed by high vowels has long been recognised as typical (see Jurafsky 1996: 534 for early references).

In rare cases we found combined strategies of suffixation plus reduplication and/or base modification. For example, the Hmong Mien language Western Mien, also known as Xong, shows a combination of full reduplication of the verb, interspersed with two morphemes, *lib* and *daod*, which are not glossed, but characterised as some kind of sound-symbolic units. The process is productive with stative verbs, but also applies to some dynamic verbs and is described as “always having an attenuating effect” (Sposato 2015: 504). (16) shows an example with a stative (a) and a dynamic verb (b).

---

<sup>9</sup> The process shown here is reminiscent of, or even analysable as, a phonaesthetic alternation. Such processes are also associated with evaluative notions like diminutivity and augmentativity; see Willemsen & Miltersen (2020).

(16) Western Mien (Hmong Mien; Sposato 2015: 504-505)

- a. *Zheit-zhauf giand-lib-giand-daod*  
 outside-door icy-LIB-RDP-DAOD  
 ‘It’s a little bit icy outside.’
- b. *wel sheit-lib-sheit-daod*  
 1SG write-LIB-RDP-DAOD  
 ‘I’ve written a little bit.’

The Niger-Congo language Wawa has four suffixes described as pluractionals, iteratives or distributives. They attach to a tonally modified stem, sometimes with reduplication. The example in (17) shows the suffix *-kəkā*, which “stresses that the actions are done a little bit each time, it can also refer to actions being done jokingly, like pretending to hit someone or doing so only lightly” (Martin 2012: 305) and was therefore identified as a verbal diminutive in the qualitative as well as quantitative domain. The stem vowel *u* is the locus of the tone modification.

(17) Wawa (Niger-Congo; Martin 2012: 305)

- gǔ-* *gù-kəkā*  
 ‘fall’ ‘fall a little bit many times’

Table 3 indicates the frequencies of the strategies used in the languages of our sample. Note that the sum of the numbers is, again, higher than the number of VDim-languages in the sample because some languages employ more than one strategy, either combined or as individual constructions.

Strategy	Number of constructions in the sample	Number of languages using strategy for one or more VDim constructions
Affixation	54	42
Reduplication	28	27
Cliticisation	14	10
Freestanding element	7	6
Verbal element	8	7
Base modification	4	4

Table 3: Exponence of verbal diminution.

The overview shows that affixation and reduplication are by far the most common strategies of verbal diminution across the languages of the world. Diminution by reduplication is particularly interesting, since, as noted in the literature (e.g. Moravcsik 1978: 317), it may appear counterintuitive: principles of iconicity decree that increase in form indicates increase in meaning. We will return to this issue in §4. First, however, we take a look at the various functions of verbal diminutive markers, which show more interesting complexities.

### 3.3 Functions

#### 3.3.1 Earlier observations

The examples above show that the functions of verbal diminutives vary. This does not come as a surprise: since Jurafsky (1996), it has been noted that diminutives form a “radial category” of related meanings (though see Dressler & Merlini Barbaresi 2001, Fortin 2011, and Mutz 2015 for criticism and alternative proposals). Therefore, it is interesting to explore the semantic effects of verbal diminutives in greater detail.

According to Jurafsky (1996), diminution encompasses a variety of notions, with ‘small’ and ‘child’ at the center and meanings such as ‘imitation’ ‘related-to’, ‘partitive’ and ‘approximation’ at the periphery. In addition, diminutives tend to come with affective meanings, both positive and negative. Figure 1 from Jurafsky shows the proposed universal structure of the diminutive in a semantic map, which also illustrates how the more peripheral meanings relate to the more central ones.

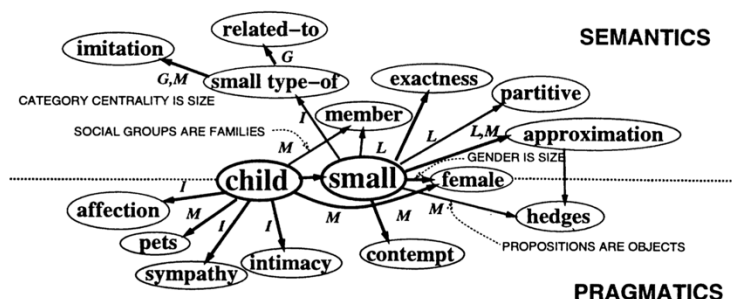


Figure 1: Proposed universal structure for the semantics of the diminutive (Jurafsky 1996: 542).

Whereas Jurafsky's examples are almost exclusively nominal or adjectival, Weidhaas & Schmid (2015) find that German verbal diminutives in *-el* can be arranged in a similar configuration, with the following list of meanings (Weidhaas & Schmid 2015: 203):

- Semantic attenuation:
  - low intensity
  - iterative
  - small pieces
  - playful-tentative and playful-pretentive
- Pragmatic attenuation:
  - language of proximity
  - contempt
  - affection and sympathy
  - trivialisation
  - euphemism

As we will show in the following, these observations for German are largely corroborated by our sample of languages and other languages discussed in the literature.

### 3.3.2 *Cross-linguistic functions of verbal diminutives*<sup>10</sup>

In our sample, we observe eleven semantic and pragmatic categories, listed below, that occur with a certain regularity, as well as a few more idiosyncratic meanings. Assuming the gradable dimensions of verbal meaning introduced in §1, quantity and quality, plus affective meanings,<sup>11</sup> we can identify six categories directly related to attenuation. These are LOW INTENSITY, SHORT TIME, NON-SERIOUS, INCOMPLETE, EMOTION (AFFECTION/CONTEMPT) and POLITENESS.

---

<sup>10</sup> In this section we will ignore more grammatical functionalities of VDims, such as transitivity or perfectivisation (see Katunar 2013: 3), effects on telicity, etc.

<sup>11</sup> Affective meanings can be called pragmatic, as we did in §2.3, but it should be noted that semantics and pragmatics cannot always be distinguished with certainty. For example, in the category NON-SERIOUS we find both objective meanings such as 'pretend' and affective meanings such as 'careless', as well as in-between cases such as 'playful'.

The categories referred to as “associated meanings”, i.e. SMALL ARGUMENT, ITERATIVE, DISTRIBUTIVE, DURATIVE and INTENSIVE, are not themselves indicative of verbal diminution, but occur as additional effects of the VDim markers. Some of these meanings are surprising, as they are downright contradictory to attenuation.

● LOW INTENSITY	}	quantitative domain
● SHORT TIME		
● NON-SERIOUS	}	qualitative domain
● INCOMPLETE		
● EMOTION (AFFECTION/CONTEMPT)	}	affective meanings
● POLITENESS		
● SMALL ARGUMENT	}	associated meanings
● ITERATIVE		
● DISTRIBUTIVE		
● DURATIVE		
● INTENSIVE		

We briefly discuss each group in turn.

The meanings LOW INTENSITY and SHORT TIME represent reduction in the quantitative domain. Example (18) illustrates LOW INTENSITY in the Niger-Congo language Noni.

(18) Noni (Niger-Congo; Hyman 1981: 35)

<i>kám</i>	<i>kám-cé</i>
‘squeeze’	‘squeeze a little’

SHORT TIME is also frequent, often with descriptions such as ‘limited duration’ or ‘for a while’. Interestingly, we find temporal attenuation both manifested as ‘rapidly’ (Urarina; Olawsky 2006: 472) and ‘slowly’ (Toqabaquita; Lichtenberk 2008: 169, see example (30) in §4). Another special case is the Muskogean language Creek, which uses the diminutive suffix to express closeness in time, i.e. ‘just now’ (Martin 2011: 234). The basic temporal meaning is nicely illustrated in example (19) from Kolyma Yukaghir.

(19) Kolyma Yukaghir (Yukaghir; Maslova 2003: 276)

<i>morie-</i>	<i>morie-s’í:-</i>
‘wear’	‘wear for a short time’

As anticipated in §1, our catalogue of meanings does not list SMALL SPACE, as reduction in a spatial sense is rare; our sole example is from Yurakaré, as shown in (20). Note that this example could also be interpreted as reduction in intensity.

(20) Yurakaré (Yurakaré; van Gijn 2006: 120, cited in Körtvélyessy 2015: 78)

*ana-ja-lë*            *baja-nñu-ø*

DEM-MEA-AMP    subside-DIM-3

‘The water has subsided a little.’

Within the qualitative domain, we see reduction in the dimension of seriousness and/or effort, which we labeled NON-SERIOUS, and in the dimension of completeness and/or success, for which we use the term INCOMPLETE.

The former category encompasses a variety of meanings, the most common being ‘playful’ and ‘pretend to’ (this includes the meanings described by Weidhaas & Schmid 2015: 203 as playful-tentative and playful-pretentive). Other variants are described as ‘relaxed attitude’, ‘less purposeful’, ‘careless’ or ‘aimless’. We also included meanings translated as ‘just’, ‘simply’ or ‘merely’, which fit the general impression of low dedication and effort. An example of the ‘aimless’ meaning is given in (21). The diminutive is formed by partial reduplication (plus, in this particular verb, the addition of the vowel *o*).

(21) Palauan (Austronesian, Malayo-Polynesian; Josephs 1975: 236)

*mɛraɛl*

‘walk’

*mɛreroaɛl*

‘walk aimlessly’

The category INCOMPLETE subsumes meanings of the type ‘partially done’ or ‘accomplished to a lesser extent than expected’. In addition, we included cases where the action is attempted but not completed. For example, the Papuan language Imonda shows verb complexes with an extra root following the stem. There are several such roots; the forms and their semantics are illustrated in (22).

(22) Imonda (Papunesia; Seiler 1985: 103-104)

a. *nagtõ* ‘incomplete’ (as in ‘I have chopped the tree halfway through’)

b. *sabeha* ‘pretend’, “indicates that the event depicted by the lexical verb is in some

sense not ‘the real thing’, that it was only begun but not finished.” (Seiler 1985: 104)<sup>12</sup>

- c. *səlōh* ‘in vain’ (without success)
- d. *nòg* ‘incomplete’ (like *sabeha*) or ‘in vain’ (like *səlōh*)

All of these fit the wider category we are describing here.

Affective meanings manifested themselves in two basic ways, one we refer to as EMOTION (AFFECTION/CONTEMPT) the other as POLITENESS. As to the former group, grammars sometimes mention ‘speakers’ feelings’ or qualify the action or event as ‘close to the speaker’, without specifying a positive or negative thrust of the emotions in question. In most cases, however, affective meanings were explicitly identified as positive (fondness, approval or compassion) or negative (disparagement, trivialisation). It is worth noting that emotions could be linked either to the action or to a participant. The former situation can be seen in German, which has a number of verbs for touching events, such as *kuscheln* ‘to cuddle’, *streicheln* ‘to stroke’, *hätscheln* ‘to pet’ or *tätscheln* ‘to pat’. All of these contain the diminutive suffix *-el*, which here transports positive connotations of intimacy.<sup>13</sup> The latter situation, which appears to be more common, can be seen in Chukchi, where the diminutive can express “both fondness and disparagement” towards a participant (Dunn 1999: 268; examples in (23) are shortened).

(23) Chukchi (Chukotko-Kamchatkan; Dunn 1999: 268)

- a. *ilu-ke*            *q-ə-twa-qaat-ə-rkən*  
       move-NEG     INT-E-be-DIM-E-PROG  
       ‘Stop it you little [idiot]!’
- b. *jəlqet-qeet-yʔi*  
       sleep-DIM-TH  
       ‘He fell asleep, the poor little thing.’

Such cases were only included if they occurred alongside meanings of reduced quantity and/or quality, as outlined in §2.3. This is true for Chukchi; witness example

<sup>12</sup> The verb root *sabeha* also shows the attenuative meaning NON-SERIOUS.

<sup>13</sup> Jurafsky (1996) lists “intimate” as a separate category. As this meaning only occurred once in our dataset, we included it under EMOTION.



(2), repeated here for convenience, which contains the marker *-qeet-* used as a true verbal diminutive.

(2) Chukchi (Chukotko-Kamchatkan; Dunn 1999: 268)

*e-mec-pintəqet-qeet-lin*

PRF-APPR-show.self-DIM-3SG

‘It showed itself slightly.’

In addition, verbal diminutives can be used as POLITENESS markers, e.g. to soften imperatives or requests or to express self-deprecation. The first can be seen in Mandarin Chinese (24),<sup>14</sup> the second in Zulu (25), both expressed by reduplication.

(24) Mandarin Chinese (Sino-Tibetan; Li & Thompson 1981: 235)

*qǐng nǐ bǎ mén kāi-kāi*

please 2SG BA door open-open

‘Please open the door.’

(25) Zulu (Niger-Congo, Atlantic Congo; Van der Spuy & Mjiyako 2015: 520)

*ngi-ya-cul-a-cul-a*

‘I’m just singing a bit.’

Among the associated meanings, which are not themselves indicative of verbal diminution but were described as meanings for the same markers, we see semantic effects related to the participant, in other words, to a verbal argument. This situation is mentioned in §2.3 above and illustrated with example (6) from Iquito; example (23) from Chukchi is another case in point. This category, which we call SMALL ARGUMENT here, comes in various shades of meaning, from ‘small participant’, especially ‘child’, to ‘toy’ and ‘small pieces’ (a category also identified by Weidhaas & Schmid 2015). An example for the latter is found in the Austronesian language Nias Selatan, where “initial-syllable-reduplication indicates that the action is done many times, often with the sense that the actions are small ones, or result in many small pieces” (Brown 2001: 529). The ‘small pieces’, though not the ‘small action’ meaning can be seen in (26). This case was included by virtue of the verbal ‘small action’ meaning mentioned in the source.

---

<sup>14</sup> We owe this example to Olsson (2012).

(26) Nias Selatan (Austronesian; Brown 2001: 529)

*I-ta-taba*            *geu*

3S.RLS-RDP1-cut wood:MUT

‘He cut up the wood (into many small pieces).’

In addition, we see argument-related affective meanings, such as in Chukchi. Taken broadly, this category also encompasses verbs used mostly or exclusively towards or in reference to children. An example can be found in German (27).

(27) German (Indo-European; personal knowledge)

*fremd*

*fremd-el-n*

‘strange’

‘be shy with strangers’

While shyness with strangers can also be felt by adults, the verb *fremdeln* is specifically used for children, often referring to a typical developmental phase in early childhood.

Of the other associated meanings, three are aspectual in nature, i.e. ITERATIVE, DISTRIBUTIVE and DURATIVE. Iterativity is a common meaning accompanying verbal diminutives. Being a more conspicuous or simply better-known category, iterativity is sometimes taken to be the central meaning of the relevant forms, which are then glossed as “iteratives”, “frequentatives” or even “pluractionals”, while their semantic profile also contains evidence of attenuation. We saw an example in the language Wawa (example (17), repeated for convenience).

(17) Wawa (Niger-Congo; Martin 2012: 305)

*gǔ-*

*gù-kə̀kə̀*

‘fall’

‘fall a little bit many times’

The marker *-kə̀kə̀* is referred to as a “pluractional” in the grammar, while – repeating the quote from §3.2 – “it can also refer to actions being done jokingly, like pretending to hit someone or doing so only lightly” (Martin 2012: 305), which means that it doubles as a verbal diminutive.

The polysemy between verbal diminution and pluractionality suggests that there is a connection between the two (see Mattiola 2017, 2019 for a semantic map of pluractionality, which can be linked to the conceptual space we are discussing here). Similar relations, i.e. between verbal diminution and iterativity, are discussed in

various places in the literature, e.g. by Amiot & Stosic (2014) for French, Katunar (2013) for Croatian, Grandi (2009) and Tovená (2011) for Italian, and Weidhaas & Schmid (2015) for German. Such polysemy patterns are somewhat paradoxical, as they involve a contradiction which Tovená (2011) refers to as “small is many” and we like to call “less is more”: diminution equals reduction, while iterativity equals increase. Explanations offered in the literature are mostly along the lines of Cusic (1981), who argues that event plurality comes in two kinds, a) repetition of the event and b) repetition *within* the event. The latter constitutes the link to verbal diminution, as introducing repetition within the boundaries of an event reduces the size of the subevents (see Tovená 2011, Amiot & Stosic 2014, François 2004 and Kouwenberg & LaCharité 2005 for similar arguments).

A category that is easier to reconcile with attenuation is DISTRIBUTIVE, i.e. ‘little by little’, ‘here and there’ or ‘one by one’. For example, the language Nias Selatan we saw in (26) can use initial-syllable-reduplication not only to indicate small actions or small pieces, but also for distributive actions, as in (28).

(28) Nias Selatan (Austronesian; Brown 2001: 529, glosses slightly adjusted)

*La-ta-taru*                      *zinaanö*  
3PL.RLS-RDP-plant      seedling:MUT

‘One plants the seedlings one by one (in a wet rice field).’

The intuitive relation between attenuation and distributivity is nicely explained by Hyman (1981), who writes that in the Niger-Congo language Noni “30% [of the attenuative verb forms attested] have the meaning ‘here and there’ or ‘little by little’ [...]. This interpretation derives from the attenuative meaning: instead of performing an action all at once as a single event, one attenuates the action into a sequence of smaller events” (Hyman 1981: 36; Hyman analyses this as verbal plurality and reserves the term “distributive” for a different suffix, but translates both the same way, as ‘several times’).

Finally, the less-is-more paradox manifests itself in two further variants, i.e. the categories DURATIVE (or ‘continuous’) and INTENSIVE, which quite clearly contradict the core meanings of VDimS, SHORT TIME and LOW INTENSITY, but nevertheless occur with a certain frequency in our dataset. A striking example can be found in the language Westcoast Bajau, where full reduplication can have contradictory meanings, even

with the same verb (example (29) shows reduplication of the verb *keet* ‘to glow’; meaning a. corresponds to our INTENSIVE).

(29) Westcoast Bajau (Austronesian; Miller 2007: 81)

***keet-keet***

- a. ‘to burn brightly’ (emphatic meaning)
- b. ‘to burn dimly’ (diminutive meaning)
- c. ‘to burn over a period of time’ (continued action meaning).

The fact that these patterns are found specifically with reduplication is probably motivated by iconicity. We refer again to Kouwenberg & LaCharité (2005), who address the less-is-more paradox in relation to reduplication in three Caribbean Creole languages.

In addition to the meanings discussed in this section, we found individual cases of meanings like ‘incipient action’, ‘habitual’, ‘counterexpectation’, and a particularly expressive function called “diminuendo” (‘less and less’, the language is Palauan, shown in example (21)).

Quantitative meanings	Qualitative meanings	Affective meanings	Associated meanings	N of VDim constructions per combination
+	-	-	-	31
-	+	-	-	12
-	-	+	-	(0)
-	-	-	+	(0)
+	+	-	-	14
+	-	+	-	5
+	-	-	+	13
-	+	+	-	0
-	+	-	+	10
-	-	+	+	(0)
+	+	+	-	4
+	+	-	+	12
+	-	+	+	7
-	+	+	+	0
+	+	+	+	4
<b>Total</b>				<b>112</b>

**Table 4:** Frequency of combinations of meanings.

The zeroes in brackets follow from our methodology: affective and associated meanings were not considered unless they accompanied quantitative or qualitative meanings. To round off this section, Table 4 illustrates the frequency in which meanings occurred in combination in our sample.

The overview shows that purely quantitative meanings, i.e. LOW INTENSITY or SHORT TIME, were clearly the most common scenario. All combinations with ten or more instances are highlighted in grey. It should be noted that the meanings summarised under “associated meanings” are highly heterogeneous, which explains the fairly high number of signs + in this column.

#### 4. Further issues of interest

Before concluding, we would like to offer a brief look at a number of issues we could not address, but which seem worthwhile avenues for further research.

First of all, we mentioned a number of contradictory meanings involving attenuation on the one hand and pluractionality, durativity or intensity on the other, subsumed under “less is more”. This is not the only paradox in the realm of verbal diminution (and in diminutives in general). Another such paradox lies in the fact that diminutives can be associated both with positive and negative affective meanings. Jurafsky (1996) proposes a way out by assuming two core meanings, “child” and “small”, and analysing “affection” as derived from “child” by means of inference (see Figure 1 in §3.3.1), while “contempt” is seen as metaphorically related to “small”. It is not so evident how this solution might be transferred to the VDims. Instead, the contradictory effects may simply be a consequence of whether a reduced event or activity inspires fondness or annoyance, given the lexical semantics of the verb. A second paradox, also mentioned in Jurafsky (1996: 535), holds between ‘approximation’ and ‘exactness’, which both occur in the penumbra of nominal diminutives. In our data, we mostly see approximation (which we distributed over two categories, NON-SERIOUS and INCOMPLETE). In one case, however, the opposite semantics is found: the two VDim markers in Toqabaquita, *thafa* and *thafeqe* (shown in example (11) in §3.2), can also signal that “an event is (to be) performed or is taking place in a calm, slow, quiet, careful, measured manner” (Lichtenberk 2008: 169). Example (30) illustrates this use.

(30) Toqabaqita (Austronesian, Malayo-Polynesian; Lichtenberk 2008: 169)

*Qoko thafa ngata.*

2SG.SEQ ATTN speak

‘Speak calmly/slowly/quietly.’

However, this semantic effect appears to be rare.

A second general issue we have not covered is the interplay between VDimS and Aktionsart or lexical aspect. Various sources point out that VDimS show preferences for particular verbal subclasses and/or produce different semantic effects from one class to another. For example, the Mandarin “delimitatives” described in Li & Thompson (1981) occur only with volitional activity verbs. The authors explain this as follows: “Since the delimitative aspect means that the subject does something a little bit, it follows that only volitional verbs, that is, those expressing events over which one has some control, can be reduplicated to show delimitative aspect” (Li & Thompson 1981: 235). Related observations are discussed for other languages, e.g. in Kouwenberg & LaCharité (2005), Armoskaite & Koskinen (2008), Grandi (2009), Németh & Sörös (2018), and Makarchuk (2020). We have also not touched on issues such as transitivity or telicity, which appear to be relevant as well.

A third issue worth discussing is where VDimS are situated between inflection and derivation. Various grammars list them as aspectual distinctions, i.e. as inflection, while other cases might be considered derivational or are expressed by syntactic strategies. Alternatively, as mentioned in the §1, VDimS can be seen as instances of evaluative morphology, which is sometimes analysed as a third category between inflection and derivation (see Scalise 1984 and, for a different view, Stump 1993; Bauer 1997: 12 provides a summary). This has sparked a lively and controversial debate, to which the verbal cases could add new evidence.

Moreover, it would be interesting to take a closer look at the relation between verbal and nominal diminutives. In some languages, nouns and verbs take the same diminutive marker, for example in Tariana, which has a “floating enclitic” *tuki*, which can attach to “any focussed constituent”, e.g. nouns, verbs and adjectives (Aikhenvald 2003: 366). Another relevant case is German verbal *-el*, mentioned in (27) above, which might have arisen by reanalysis of a homophonous nominal suffix (Schmuck 2018). In many cases, however, the verbal diminutive is distinct from the nominal diminutive. For those languages where the marker is the same, we face the interesting issue that a grammatical feature is grafted onto a different part of speech. This touches

on a more general issue, i.e. the extent to which grammatical functions can cross category boundaries.<sup>15</sup> A well-known case is, of course verbal number, i.e. pluractionality, another is nominal tense (Lecarme 2012, Bertinetto 2020). Velupillai (2012: 125) mentions Mwotlap (François 2005), “where nouns may take tense, mood and aspect markers”. The Niger-Congo language Klao/Kru, described by Rickard (1970), marks completive/incompletive aspect in the pronominal paradigm. Various other instances can certainly be found. If diminution is considered a primarily nominal feature, the verbal diminutives provide interesting material in this field.

Last but not least, it can be observed that verbal diminutives are found in at least a number of creole languages (Kouwenberg & LaCharité 2005, Ponsonnet 2018), which shows that VDims are not necessarily a “mature” phenomenon in the sense of Dahl (2004), i.e. typical of later stages of linguistic life cycles.

## 5. Conclusions

This study offers the first substantial typology of verbal diminutives, based on a balanced sample of 248 languages. We have framed diminution as an instance of comparison, as it always involves attenuation with respect to some implicit standard. The dimensions along which verbal meanings can be reduced were identified as:

- Quantitative: LOW INTENSITY OR SHORT TIME
- Qualitative: NON-SERIOUS OR INCOMPLETE

In addition, verbal diminutives – like nominal diminutives – can have affective meanings like fondness or deprecation or can be used as politeness markers, e.g. to soften requests.

As additional meanings encoded by the same markers, we found ‘iterative’, ‘distributive’ and ‘durative’, plus meanings associated with a verbal argument (mostly ‘child’ or ‘small pieces’). Some of these are particularly interesting, as they contradict the more central meanings by expressing higher rather than lower intensity or a longer instead of a shorter duration of the action. Such paradoxical semantic effects have also been noted for the semantics of nominal diminutives, as well as in the

---

<sup>15</sup> Note that these categories can be semantic in nature (i.e. concern the difference between events and objects or individuals) or they can be related to issues of word class flexibility (i.e. the ability to use different types of lexemes in multiple syntactic functions).

semantic effects of reduplication (see e.g. Moravcsik 1978, Regier 1994, Kiyomi 1995, Rubino 2005, Abraham 2005, Körtvélyessy 2016).

Regarding formal exponence, our study shows that verbal diminutives can be expressed by a wide variety of morpho-syntactic strategies, the most frequent being affixation. We have not systematically studied the relation between form and function(s) of verbal diminutives. However, this would be an interesting next step to take, especially since it may shed light on the diachronic development of this category, which as yet is largely uncharted territory.

Our study points out various other avenues of further research. Of special interest is the place of verbal diminutives amongst other categories applying across parts of speech or ontological categories prototypically associated with them, such as actions, properties, and individuals. We hope that the typology outlined in this paper contributes as a starting point for research in this area.

## Acknowledgements

All three authors have contributed equally to the data collection and the content of the paper. The writing was done by Jenny Audring. Many thanks to Cerys Clarke for helping with the ‘grammar crunch’. We are grateful to Bruno Olsson, Kees Hengeveld, Jan Rijkhoff, Alexandre François, Jeroen Willemsen, two anonymous reviewers, the members of the Typology Research Group at the University of Amsterdam, and the audiences at SLE 2019 and the Dutch Annual Linguistics Day 2020 for advice, comments and suggestions. We dedicate this paper to the late Pieter Muysken, whose eyes lit up when he heard about the verbal diminutives.

## Abbreviations

1 = 1 <sup>st</sup> person	E = epenthetic schwa	PRF = perfect
2 = 2 <sup>nd</sup> person	EC = extended current tense	PROG = progressive
3 = 3 <sup>rd</sup> person	INF = infinitive	PRS = present
AMP = amplification	INT = intentional	PRT = pretense mode
APPR = approximative	IPFV = imperfective	PST = past
ATTN = attenuative	MEA = measure	RDP = reduplication
AUX = auxiliary	MUT = mutated nominal	REM = remote
CUM = cumulative	NEG = negation	RLS = realis
DECL = declarative	NFUT = nonfuture	SEQ = sequential
DEM = demonstrative	NOM = nominative	SG = singular
DIM = diminutive	PL = plural	TH = thematic suffix



## References

- Abdurehim, Esmael. 2014. *The Lopnor dialect of Uyghur: A descriptive analysis*. PhD Dissertation, University of Helsinki.
- Abraham, Werner. 2005. Intensity and diminution triggered by reduplicating morphology: Janus-faced iconicity. In Bernhard Hurch (ed.), *Studies on reduplication*, 547-568. Berlin: Mouton de Gruyter.
- Aikhenvald, Alexandra Y. 2003. *A Grammar of Tariana, from Northwestern Amazonia*. Cambridge: Cambridge University Press.
- Amiot, Dany & Stosic, Dejan. 2014. When evaluative morphology, pluractionality and aspect get tangled up: a case study of French suffixed verbs. In Zoe Gavriilidou & Anthi Revithiadou (eds.), *Mélanges offerts à Anna Anastassiades-Syméonides à l'occasion de sa retraite*, 16-33. Kavala: Saita Publishers.
- Armoskaite, Solveiga & Koskinen, Päivi. 2008. Diminutive verbal suffixes in Finnish. In *Proceedings of the 2008 Annual Conference of the Canadian Linguistic Association*, 1–13.
- Audring, Jenny, Leufkens, Sterre & van Lier, Eva. 2019. A typology of verbal diminutives. Paper presented at SLE52, Leipzig, Germany, August 23, 2019.
- Audring, Jenny, Leufkens, Sterre & van Lier, Eva. 2021. Verbal diminutives - Appendix [online file]. <https://doi.org/10.24416/UU01-C5PE8E>
- Bakker, Dik. 2011. Language sampling. In Jae Jung Song (ed.), *The Oxford handbook of linguistic typology*, 100-127. Oxford: Oxford University Press.
- Bauer, Laurie. 1997. Evaluative morphology: In search of universals. *Studies in Language* 21(3). 533-575.
- Bertinetto, Pier Marco. 2020. On nominal tense. *Linguistic Typology* 24(2). 311-352.
- Bright, William. 1957. *The Karok Language*. Berkeley/Los Angeles: University of California Press.
- Briley, Joyce E. 1976. An overview of the Bauzi verb phrase. *Irian* 5. 3-17.
- Brown, Lea. 2001. *A grammar of Nias Selatan*. PhD Dissertation, University of Sydney.
- Cusic, David. 1981. *Verbal plurality and aspect*. PhD Dissertation, University of Stanford.
- Dahl, Östen. 2004. *The growth and maintenance of linguistic complexity*. Amsterdam: John Benjamins.

- Dixon, R. M. W. 2004. *The Jarawara language of Southern Amazonia*. Oxford: Oxford University Press.
- Dressler, Wolfgang & Merlini Barbaresi, Lavinia. 1994. *Morphopragmatics: Diminutives and intensifiers in Italian, German and other languages*. Berlin: Mouton de Gruyter.
- Dressler, Wolfgang & Merlini Barbaresi, Lavinia. 2001. *Morphopragmatics of diminutives and augmentatives*. In István Kenesei & Robert M. Harnish (eds.), *Perspectives on Semantics, Pragmatics, and Discourse: A Festschrift for Ferenc Keifer*, 43-58. Amsterdam: John Benjamins.
- Dryer, Matthew S. 2016. Crosslinguistic categories, comparative concepts, and the Walman diminutive. *Linguistic Typology* 20(2). 305-331.
- Dunn, Michael J. 1999. *A grammar of Chukchi*. PhD Dissertation, Australian National University.
- Fortin, Antonio. 2011. *The morphology and semantics of expressive affixes*. PhD Dissertation, University of Oxford.
- François, Alexandre. 2004. *La réduplication en mwotlap: les paradoxes du fractionnement*. In Elizabeth Zeitoun (ed.), *Les langues austronésiennes (Faits de langues, 24)*, 177-194. Paris: Ophrys.
- François, Alexandre. 2005. A typological overview of Mwothlap, an Oceanic language of Vanuatu. *Linguistic Typology* 9(1). 115-146.
- Grandi, Nicola. 2009. Restrictions on Italian verbal evaluative suffixes: The role of aspect and actionality. *York Papers in Linguistics* 2(10). 46-66
- Grandi, Nicola & Körtvélyessy, Lívia (eds.). 2015. *Edinburgh handbook of evaluative morphology*. Edinburgh: Edinburgh University Press.
- Greenberg, Yael. 2010. Event internal pluractionality in Modern Hebrew: A semantic analysis of one verbal reduplication pattern. *Brill's Annual of Afroasiatic Languages and Linguistics* 2(1). 119-164.
- Hahn, Reinhard F. 1998. Uyghur. In Lars Johanson & Éva Á. Csató (eds.), *The Turkic languages* (Routledge Language Family Descriptions), 379-396. London: Routledge.
- Harms, Philip L. 1994. *Epena Pedee Syntax*. Dallas: The Summer Institute of Linguistics and the University of Texas at Arlington.
- Hewitt, George B. 1995. *Georgian: A Structural Reference Grammar*. Amsterdam: John Benjamins.
- Horatius Flaccus, Quintus. 18 BC. *Ars Poetica*. Accessed online at <http://thelatinlibrary.com/horace/arspoet.shtml>

- Hyman, Larry M. 1981. Noni grammatical structure, with special reference to verb morphology. *Occasional Papers in Linguistics* 9. Los Angeles: University of Southern California.
- Josephs, Lewis S. 1975. *Palauan reference grammar*. Honolulu: University of Hawaii Press.
- Jurafsky, Daniel. 1996. Universal tendencies in the semantics of the diminutive. *Language* 72(3). 533–578.
- Káňa, Tomáš. 2017. Verbale Diminutive – ihre Abgrenzung und Bedeutung. *Linguistica Pragensia* 28(2). 58-82.
- Katunar, Daniela. 2013. Diminutives in Action: A cognitive account of diminutive verbs and their suffixes in Croatian. *Suvremena Lingvistika* 39(75). 1-23.
- Kim, Yuni. 2008. *Topics in the Phonology and Morphology of San Francisco del Mar Huave*. PhD Dissertation, University of California.
- Kiyomi, Setsuko. 1995. A new approach to reduplication: a semantic study of noun and verb reduplication in the Malayo-Polynesian languages. *Linguistics* 33(6). 1145-1167.
- Körtvélyessy, Lívia. 2015. *Evaluative morphology from a cross-linguistic perspective*. Cambridge: Cambridge Scholars Publishing.
- Körtvélyessy, Lívia. 2016. Reduplication from a cross-linguistic perspective. *BAS British and American Studies* 22. 153-164.
- Kouwenberg, Silvia & LaCharité, Darlene. 2005. Less is more: Evidence from diminutive reduplication in Caribbean Creole languages. In Bernhard Hurch (ed.), *Studies on Reduplication*, 533-546. Berlin: Mouton de Gruyter.
- Lai, I-Wen. 2009. *Time in the Iquito Language*. PhD Dissertation, University of Texas at Austin.
- LeSourd, Philip S. 1995. Diminutive verb forms in Passamaquoddy. *International Journal of American Linguistics* 61(1). 103-134.
- Lecarme, Jacqueline. 2012. Nominal tense. In Robert Binnick (ed.), *The Oxford handbook of tense and aspect*, 696-720. Oxford: Oxford University Press.
- Lewis, M. Paul, Simons, Gary F., & Fennig, Charles D. 2015. *Ethnologue: Languages of the world*. Dallas: SIL international.
- Li, Charles N. & Thompson, Sandra A. 1981. *Mandarin Chinese: A functional reference grammar*. Berkeley: University of California Press.
- Lichtenberk, Frantisek. 2008. *A grammar of Toqabaqita*. Berlin: Mouton de Gruyter.

- Makarchuk, Ilya. 2020. *Typology of verbal attenuative*. MA Thesis, National Research University Higher Economics, Moscow.
- Makarova, Anastasia. 2014. *Rethinking diminutives: A case study of Russian verbs*. PhD Dissertation, University of Tromsø.
- Makarova, Anastasia. 2015. One type of verbal diminutives in Russian: Verbs ending in *-n'kat'*. *Russ Linguist* 39(1). 15-31.
- Maree, Rundell D. 2007. *Ibatan: A grammatical sketch of the language of Babuyan Claro Island*. Manila: Linguistic Society of the Philippines.
- Martin, Jack B. 2011. *A Grammar of Creek (Muskogee)*. Lincoln: University of Nebraska Press.
- Martin, Marieke. 2012. *A grammar of Wawa: An endangered language of Cameroon*. PhD Dissertation, University of Kent.
- Maslova, Elena. 2003. *A Grammar of Kolyma Yukaghir*. Berlin: Mouton de Gruyter.
- Mattiola, Simone. 2017. The conceptual space of pluractional constructions. *Lingue e Linguaggio* 16(1). 119-146.
- Mattiola, Simone. 2019. *Typology of pluractional constructions in the languages of the world*. Amsterdam: John Benjamins.
- Mattiola, Simone. 2020. Two language samples for maximizing linguistic variety. Bologna: AMS Acta – Alma Mater Studiorum – Università di Bologna (<http://amsacta.unibo.it/6504/>).
- McGregor, William. 1990. *A Functional Grammar of Gooniyandi*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Miestamo, Matti, Bakker, Dik & Arppe, Antti. 2016. Sampling for variety. *Linguistic Typology* 20(2). 233–296.
- Mihas, Elena. 2015. *A Grammar of Alto Perené (Arawak)*. Berlin: Mouton de Gruyter.
- Miller, Mark T. 2007. *A Grammar of West Coast Bajau*. PhD Dissertation, University of Texas.
- Miyaoka, Osahito. 2012. *A grammar of Central Alaskan Yupik: An Eskimo language*. Berlin: Mouton de Gruyter.
- Moravcsik, Edith 1978. Reduplicative constructions. In Joseph H. Greenberg (ed.), *Universals of Human Language, Vol. 3: Word Structure*, 297-334. Stanford: Stanford University Press.
- Munro, Pamela. 1988. Diminutive syntax. In William Shipley (ed.), *In honor of Mary Haas*, 539-555. The Hague: Mouton.

- Mutz, Katrin. 2015. Evaluative morphology in a diachronic perspective. In Nicola Grandi & Lívia Körtvélyessy (eds.), *Edinburgh Handbook of Evaluative Morphology*, 142-154. Edinburgh: Edinburgh University Press.
- Németh, Boglárka & Sőrés, Anna. 2018. Evaluative morphology in the verbal domain. Complex verbs suffixed with *-kVdik* in Hungarian. *Studies in Language* 42(1). 202-225.
- Noonan, Michael. 1992. *A Grammar of Lango*. Berlin: Mouton de Gruyter.
- Olawsky, Knut. 2006. *A Grammar of Urarina*. Berlin: Mouton de Gruyter
- Olsson, Bruno. 2012. *Verbal diminutives. A cross-linguistic pilot study*. Unpublished manuscript.
- Oyzon, Voltaire Q. & Thomas E. Payne. In preparation. *Pagsantop han Winaray [Understanding Waray Grammar]*.
- Ponsonnet, Maïa. 2018. Do linguistic properties influence expressive potential? The case of two Australian diminutives (Gunwinyguan family). *Anthropological Linguistics* 60(2). 157-190.
- Regier, Terry. 1994. *A preliminary study of the semantics of reduplication*. Berkeley: International Computer Science Institute. (Technical Report TR-94-019.)
- Rickard, David T. 1970. *Kru grammar: A preliminary study*. Monrovia: Department of Literacy Work and Literature Production of the United Methodist Church.
- Rijkhoff, Jan & Bakker, Dik. 1998. Language Sampling. *Linguistic Typology* 2(3). 263–314.
- Rijkhoff, Jan, Bakker, Dik, Hengeveld, Kees & Kahrel, Peter. 1993. A method of language sampling. *Studies in Language* 17(1). 169–203.
- Rose, Françoise. 2018. The rise and fall of Mojeño diminutives through the centuries. *Studies in Language* 42(1). 146-181.
- Rubino, Carl. 2005. Reduplication: Form, function and distribution. In Bernhard Hurch (ed.), *Studies on Reduplication*, 11–29. Berlin: Mouton de Gruyter
- Schmuck, Mirjam. 2018. Verbal diminutive. Diachron & diatopisch. Unpublished presentation, October 11, 2018.
- Scalise, Sergio. 1984. *Generative Morphology*. Dordrecht: Foris.
- Seiler, Walter. 1985. *Imonda, a Papuan Language*. Canberra: Australian National University.
- Sposato, Adam. 2015. *A grammar of Xong*. PhD Dissertation, State University of New York at Buffalo.

- Van der Spuy, Andrew, & Mjiyako, Lwazi. 2015. Evaluative morphology in Zulu. In Nicola Grandi & Lívía Körtvélyessy (eds.), *Edinburgh Handbook of Evaluative Morphology*, 515-522. Edinburgh: Edinburgh University Press.
- Stump, Gregory T. 1993. How peculiar is evaluative morphology? *Journal of Linguistics* 29(1). 1-36.
- Taylor, John R. 2003. *Linguistic categorization*. Oxford: Oxford University Press.
- Tovena, Lucía M. 2011. When small is many in the event domain. *Lexis online* 6.
- van Gijn, Rik. 2006. *A grammar of Yurakaré*. PhD Dissertation, Radboud University Nijmegen.
- Vanhove, Martine & Ahmed, Mohamed-Tahir Hamid. 2018. Diminutives and augmentatives in Beja (North-Cushitic). *Studies in Language* 42(1). 51-80.
- Velupillai, Viveka. 2012. *An introduction to linguistic typology*. Amsterdam: John Benjamins.
- Weidhaas, T., & Schmid, H. J. 2015. Diminutive verbs in German: Semantic analysis and theoretical implications. *Morphology* 25(2). 183-227.
- Willemsen, Jeroen & Miltersen, Ehm Hjort. 2020. The expression of vulgarity, force, severity and size. Phonostemic alternations in Reta and in other languages. *Studies in Language* 44(3). 659-699.

#### CONTACT

Jenny Audring [j.audring@hum.leidenuniv.nl](mailto:j.audring@hum.leidenuniv.nl) (corresponding author)

Sterre Leufkens [s.c.leufkens@uu.nl](mailto:s.c.leufkens@uu.nl)

Eva van Lier [E.H.vanLier@uva.nl](mailto:E.H.vanLier@uva.nl)

# Diachronic evolution of Russian standard markers *kako* and *aky*

IRINA KOBOZEVA<sup>1</sup>, NATALIA SERDOBOLSKAYA<sup>2</sup>

<sup>1</sup>LOMONOSOV MOSCOW STATE UNIVERSITY, <sup>2</sup>INSTITUTE OF LINGUISTICS, RUSSIAN ACADEMY OF SCIENCES/PUSHKIN STATE RUSSIAN LANGUAGE INSTITUTE

Submitted: 16/11/20 Revised version: 27/06/21

Accepted: 01/07/21 Published: 31/08/2021

## Abstract

Modern Russian uses the same marker *kak* ‘how’ to introduce the standard in equative and similitive constructions. Historical grammars claim that the same polysemy is found in Old Russian, where three markers, *kako*, *aky* and *jako*, are used interchangeably. Based on the analysis of chronicles and documents of the 11<sup>th</sup>–15<sup>th</sup> centuries and queries in the Russian National Corpus, we show that it is not the case: the markers *kako* and *aky* are distributed functionally. In the 11<sup>th</sup>–15<sup>th</sup> centuries *kako* is predominantly used in specific contexts which we propose to call “implicit parameter equative” (IPE), while *aky* is the main standard marker in similitives. In the 16<sup>th</sup> and 17<sup>th</sup> centuries *kako* expands onto similitive and equative constructions. The 18<sup>th</sup> century sees the complete loss of *aky* and the fossilization of the equative construction including the correlative pairs such as *tak(oj)... kak, stol’... kak* and others. As for the marker *jako*, it is a general subordinator that can be used in all the relevant contexts.

**Keywords:** Old Russian; manner; comparison; equative construction; similitive construction

## 1. Introduction

In Modern Russian (MR), the comparative constructions of equality and similitude are introduced by the same manner question word and subordinator *kak* ‘how/as/like’:

(1) MR

*Pëtr takoj že vysokij, kak Ivan.*  
 Peter such.M.NOM.SG PTCL smart.M.NOM.SG how Ivan  
 ‘Peter is as tall as Ivan.’

(2) MR

*On poj-ot, kak solovej.*  
 he sings-PRS.3SG how nightingale  
 ‘He sings like a nightingale.’

*Kak* is a standard marker (STM) in both equative (1) and similitive (2) constructions. Thus, MR presents a non-differentiated system typical of SAE languages (see Haspelmath & Buchholz 1998).

However, equative constructions also include an indication of identity, i.e. an adjectival demonstrative *takoj* ‘such’ (1) or a demonstrative adverb *tak* ‘in such a way’ (3), or much more rare *stol’* and *nastol’ko* ‘to such an extent’ (4) with the emphatic particle *že*.

(3) MR

*Pëtr begaj-et tak že bystro, kak Ivan*  
 Peter runs-PRS.3SG so PTCL quickly how Ivan  
 ‘Peter runs as fast as Ivan.’

(4) MR

*Ona stol’ že krasiva, kak jejë mat’*  
 she to.such.extent PTCL beautiful how her mother  
 ‘She is as beautiful as her mother.’

In equatives like (1), (3) and (4) the PAM and the STM form a correlative pair or construction which is attested as a prominent type of equative constructions in languages of Europe (Haspelmath & Buchholz 1998: 285).

The construction ‘demonstrative + emphatic particle’ is the main explicit marker of identity used in equatives (Kobozeva & Inkova 2018: 193–199); in case it is absent the construction is ambiguous and mostly interpreted as a similitive. For example,



*Petr vysokij, kak Ivan* is understood not as ‘as tall as Ivan’, rather as ‘Peter is tall, and he is similar to Ivan in exceeding the neutral height of people of his age’.

Thus, the possibility to insert *takoj/tak/stol’/nastol’ko zhe* can be used as a test distinguishing between equatives and similatives in Modern Russian.

In Old Russian, both meanings could be introduced by the same marker *aki/aky*:

(5) OR (PVL)

...людіє съкуще, аки траву.

*ljudi-je sĕk-uščę, aki trav-u*

people-ACC.PL cut-PTCP like grass-ACC

‘...cutting people like grass’

(6) OR (RNC. XVI)

есть же мѣсто то на земли ниско съз(д)ано аки вдрѣ.

*jest’ že mĕsto to na zeml-i nisko*

be.PRS.3SG EMPH place that.N.NOM.SG on soil-LOC.SG low

*sъzda-n-o aki odrĕ*

create-PTCP-N.SG like couch

‘And that place is settled on soil as low as a couch.’

As for the predecessor of *kak*, which is *kako/kakъ*, examples can also be found of both the equative (8) and the similative (7) meanings:

(7) OR (Sreznevskij 1893: 1180)

Тамо естъ Давыдовъ домъ какъ городокъ.

*tamo jest’ Davyd-ovъ dom-ъ какъ gorod-ok-ъ*

there be.PRS.3SG David-ADJ(M.SG) house-NOM.SG how town-DIM-NOM.SG

‘There is David’s house there, [which is] **like** a small fortress.’ 1392.

(8) OR (Vass. XVI)

А вы бы, бояре..., служили бы есте моему сыну, как есте мне служили прямо.

*a vy by, bojar-e... služ-i-l-i by este mojemu*

and you IRR gentlemen-NOM.PL serve-PST-PL IRR be.PRS.2PL my.DAT

*syn-u, kak este mne služil-i prjamo.*  
 son-DAT how be.PRS.2PL I.DAT serve-PST-PL faithfully

‘And you, gentlemen, should serve my son as faithfully as you served me.’

The question arises, first, whether actually both *aki/aky* and *kako/kakъ* could be used interchangeably in both contexts, or they were distributed according to some semantic parameters. Second, the question is how exactly the process of the loss of *aki/aky* has taken place, namely, what meanings and contexts were the first to disappear.

The paper is aimed at answering these two questions. Specifically, we claim that in the earlier period, i.e. 11<sup>th</sup> to 15<sup>th</sup> centuries, *kako/kakъ* was chiefly used in adverbial manner clauses and was extremely rare in comparative constructions, while *aki/aky* covered both relevant contexts, that of equality and that of similarity. We suppose that the expansion of *kako/kakъ* onto comparative constructions started from equative constructions, and the switch context from manner clauses to equatives was a specific meaning, which we propose to term as implicit equality.

The material of the study comes from Old Russian manuscripts from 11<sup>th</sup> to 16<sup>th</sup> century (see the List of sources), the Old Russian and the Middle Russian subcorpora of the Russian National Corpus, and from historical dictionaries of Old and Middle Russian.

For the sake of simplicity, we ignore the difference between phonetic variants and use the labels *aky* and *kako*. The labels OR and MR are used for Old Russian and Modern Russian, respectively.

In section 2, we provide background information about comparative constructions in Modern and Old Russian, and introduce some relevant terms. Section 3 is devoted to the quantitative distribution of *kako* and *aky* in OR. In section 4, we analyze their functions in OR and discuss possible paths of their evolution.

## 2. Russian comparative constructions in typological perspective and in reference grammars

### 2.1. Comparison in Modern Russian: typological perspective

Taking the typological perspective outlined in Haspelmath & Buchholz (1998) and Treis & Vanhove (2017) the realm of comparison in MR is divided into two parts with

respect to the marking of the STAN. In comparative constructions in narrow sense (those of inequality) the STAN is marked either by genitive case (9a) or by the linker *chem* (< instrumental case of interrogative-relative pronoun *čto* ‘what’) (9b).

(9) MR

- a. *Slivk-i*            *vkusn-eje*        *molok-a*.  
 cream-NOM.PL    tasty-PAM        milk-GEN.SG  
 ‘Cream is tastier than milk’.
- b. *Slivk-I*            *vkusn-eje,*        *čem*        *molok-o*.  
 cream-NOM.PL    tasty-PAM        what.INS    milk-NOM.SG  
 ‘Cream is tastier than milk’.

The second part comprises equative and similitive constructions. As shown in (1) and (2) from the Introduction, they both make use of the same STM *kak* to introduce the standard (STAN).

Modern Russian has a specific construction to differentiate equality from similarity (Kobozeva & In’kova 2018: 193), described as a ‘relative-based equative construction with both parameter and standard marker’ in Haspelmath & Buchholz (1998: 297), cf. (10a,b) and (11a,b). The equative involves the demonstrative adjective *takoj* or the adverb *tak* and the emphatic particle *že*.

(10) MR

- a. *Ivan tak-oj*            *že*        *vysok-ij,*            *kak (i) jeho otec*.  
 Ivan such-M.SG.NOM    EMPH    tall-M.SG.NOM    how PTCL his        father
- b. *Ivan vysok-ij,*        *tak-oj*                            *že*        *kak (i) jeho otec*.  
 Ivan tall-M.SG.NOM    such-M.SG.NOM        EMPH    how PTCL his        father
- a. = b. ‘Ivan is as tall as his father.’
- c. *Ivan vysok-ij,*        *kak Ejfelev-a*            *bašnj-a*.  
 Ivan tall-M.SG.NOM    how Eiffel-F.SG.NOM    tower-NOM.SG  
 ‘Ivan is tall, like the Eiffel tower.’

(11) MR

- a. *Ivan begaj-et*        *tak*    *že*        *bystr-o,*        *kak (i) jeho otec*.  
 Ivan run-PRS.3SG    so    EMPH    fast-ADV    how PTCL his        father

- b. *Ivan begaj-et bystr-o, tak že kak (i) jeho otec.*  
 Ivan run-PRS.3SG fast-ADV so EMPH how PTCL his father  
 a. = b. 'Ivan runs as fast as his father.'
- c. *Ivan begaj-et bystr-o, kak veter.*  
 Ivan run-PRS.3SG fast-ADV how wind  
 'Ivan runs fast like the wind.'

Thus, both examples (10a,b) and (11a,b) are interpreted as the indication of an equal extent of the parameter (PARA) 'height' and 'speed', while (10c) and (11c) are interpreted as characterization of Ivan's height and speed as exceeding the certain norm and being thus similar to the Eiffel tower and the wind, which also exceed these norms.

It is noteworthy that in MR the word order in equatives is not fixed as it is in English, cf. (10a), (11a) vs. (10b), (11b) in both examples. The difference between (a) and (b) in both (10) and (11) is minimal, and primarily consists in presentation of information: in (10a) and (11a) the focus is on the comparison, while in (10b) and (11b) the PARA (fast) is focused, and the comparison is added as additional information.

This statement can be verified if we put the figurative phrase under comparison:

(10) MR

- d. # *Ivan tak-oj že vysok-ij, kak Ejfelev-a*  
 Ivan such-M.SG.NOM EMPH tall-M.SG.NOM how Ejfel-NOM.SG.F  
*bašnj-a.*  
 tower-NOM.SG

- e. # *Ivan vysok-ij, tak-oj že kak Ejfelev-a*  
 Ivan tall-M.SG.NOM such-M.SG.NOM EMPH how Ejfel-NOM.SG.F  
*bašnj-a.*  
 tower-NOM.SG

(11) MR

- d. # *Ivan begaj-et tak že bystro, kak veter.*  
 Ivan run-PRS.3SG so EMPH fast-ADV how wind
- e. # *Ivan begaj-et bystr-o, tak že kak veter.*  
 Ivan run-PRS.3SG fast-ADV so EMPH how wind.

Indeed, figurative comparison does not really presuppose equality of the PARAs, and would not be welcome in the equative construction. This is what happens in (10d) and (10e): these examples sound awkward, because they have a pragmatically strange interpretation, namely, that Ivan is exactly as tall as the Eifel tower. The same is true of (11d) and (11e) but they are even worse because they presuppose that the wind runs. Thus, the presence and acceptability of the correlative pair in MR can be used as a test for distinguishing between equatives and similatives while translating the OR examples.

Taking the presence of *tak(oj) že* as a test, we introduce another context that seems to be close to equality constructions:

(12) MR

*Vo vremja zatmenija solnce stalo **takim že, kak** trëxdnevnyj mesjac.*

‘During the eclipse, the sun became of same size as the moon on its third night.’ (lit. became such as...)

(13) MR

*On prodal drova **tak že, kak** včera.*

‘He sold firewood on the same conditions / paying the same tax etc. as yesterday.’ (lit. in such a way as yesterday)

These constructions seem to be close to equatives because they have the same marking (the correlative pairs discussed above). Furthermore, they involve the same semantic pattern, that is, the exact equality of the PARA(s) of the comparee (size in (12) and conditions/tax in (13)) to the same PARA(s) of the STAN. It must be emphasized that these constructions are not about similarity, they express exact equality. The only difference between them and true equative is the absence of the explicit expression indicating the PARA). One step further we find the same correlative constructions with *tak(oj) že*, where, however, the exact nature of the omitted PARA is neither quantitative nor gradable unlike in (12) and (13). Still, they seem to presuppose equality, or at least identity of objective properties. Cf.:

(14) MR

*V Saksonii našli **takuju že** rudu, **kak** na Jaximovskix mestoroždenijax v Čexii.*

‘In Saxony the same ore was found, as in Jachimov deposits in the Czech Republic.’ (Adopted example from <https://cyberleninka.ru/article/n/razrabotka-uranovyh-mestorozhdeniy-i-radiatsionno-ekologicheskaya-reabilitatsiya-rayonov-saksonii-i-tyuringii>)

(15) MR

*Znak na dveri snizu **takoj že, kak** tatuirovka u mojej sestry.*

‘The sign of the downstairs door is the same as that tattooed on my sister.’

(<https://opus.nipl.eu/OpenSubtitles2016.php>).

The example (14) asserts the sameness of chemical mixture and (15) the sameness of form, which can be measured on objective grounds, and thus, it does not have the meaning of similarity, rather that of equality. In the following example the PAM of comparison is binary: either the subjects participate in the parliamentary life or not, the manner of their participation is not discussed. Thus, the two situations show equality in their truth value.

(16) MR

*Jejě predstaviteli v parlamente učastvujuť v parlamentskoj žizni **tak že, kak** oni delali eto raneje.*

‘Its representatives in parliament participate in parliamentary life **just as** they did before.’ (<https://conferences.unite.un.org>).

We deliberately took the examples (15) and (16) from the texts translated by professional translators to show that the English equivalents of these examples have STMs *same as* and *just as* containing *as*, which is the prototypical STM of equatives. If (15) and (16) were similatives, their English equivalent would have been (*just*) *like*.

All the examples (12) – (16) are unified by the following: they involve an objective PARA that is not expressed overtly (size, price and conditions, chemical mixture, form and truth value) and that can be measured / counted / evaluated on objective grounds. Thus, the semantics of these constructions is the indication of the equality of an implicit PARA.

It may be concluded that various constructions with *tak(oj) že* in Russian encode identity. Equality of degree of an explicit property (as tall as...) is merely a particular case of identity. For the sake of brevity we shall later use the term “equative” in a

broader sense than in Haspelmath & Buchholz (1998) and Treis (2018) and apply it for all examples that either contain “demonstrative + *že*” construction or can be transformed into such a construction in a given context without the change of meaning. To avoid any misunderstanding, we propose to call the constructions in question *implicit parameter equatives* (IPE).

It is important for our discussion to emphasize that the discussed constructions are different from manner adverbial clauses as the following:

(17) MR

*Oformi-l zakaz. Vsë dostavi-l-i kak dogovori-l-i-s'.*  
register-PST(M.SG) order all deliver-PST-PL how arrange-PST-PL-REFL  
'I placed the order. They delivered everything as it was arranged.'  
(<https://medtyla.ru> › page\_6).

Even if they seem similar to the constructions of equality of manner like in (13), they may be distinguished by their semantic pattern. In IPE the arguments with the same semantic role are compared, the manner (including conditions, price and tax) of selling the firewood today and yesterday. By contrast, in (17) the speaker is not comparing the manner of arranging to the manner of delivering the wares; rather, the manner of delivering is the goal/theme argument of the verb “arrange”. Note that manner adverbial clauses are most often encoded by *kak* without the correlative pair as above.

Thus, in what follows we are going to focus on STM in the following types of constructions: equative, similitive, and IPE.

## ***2.2. Background on description of comparative constructions in Modern and Old Russian***

Descriptive grammars of MR label all the types (1-4), as well as inequality comparatives, as comparative constructions, and classify them according to various formal and semantic criteria. There is not a unanimous classification, the number and the labels of the classes introduced in different grammars do not coincide. All MR grammars distinguish between inequality vs. other types of comparative constructions (1-2). However, they do not capture the semantic difference between quantitative and qualitative comparison, i.e., equality (1) and similarity (2).

The classifications used in reference grammars of MR are mostly focused on the distinction of real vs. unreal comparison (in terms of *dostovernoje* ‘credible’ vs. *nedostovernoje* ‘incredible’ in Švedova 1980). This parameter is also considered among others by Letuchiy (2015) in the most recent Russian Corpus grammar. He also examines word order, syntactic type and referential properties of the STAN, realis vs. irrealis clauses and other parameters. However, these parameters are not relevant for the semantic types we are considering in this paper, and thus, we do not refer to them in what follows.

The main means of marking STAN in equative and simulative constructions is *kak*, as shown in (1); equative constructions also include demonstratives with the emphatic particle, see section 2.1. Being the main STM in equative and simulative constructions, *kak* has several other functions in Modern Russian. It can introduce subordinate clauses and phrases of various other kinds: 1) event complement clauses; 2) accord clauses 3) role phrases (in terms of Haspelmath & Buchholz 1998); 4) temporal adverbial clauses; 5) small clauses of reason; 6) conditional clauses, see Kobozeva & In’kova (2018) for the details.

Historical grammars of Russian hardly ever address comparative constructions, for example, they are not described at all in Borkovskij & Kuznecov (1965). Lomtev (1956) only considers constructions of inequality. Bulaxovskij (1958) and Stecenko (1972) analyze inequality constructions and some sub-types of equative and simulative constructions without making a distinction between them.

All historical grammars testify that in constructions of inequality the STAN is marked with the genitive case:

(18) OR (Voskr. XVI)

Юрьи Кончаковичь бѣ бол-ій вс-ѣхъ Половецъ.

*Jurii Končakovičь bĕ bol-ij vs-ĕxъ Polovecъ*

Jurii Konchakovich was.AOR.IPFV.3SG big-PAM all-GEN.PL Polovets.GEN.PL

‘Jurii Konchakovich had a bigger army than the Cumans.’

The subordinators *neželi* (< *ne* ‘not’ + *že* EMPH + *li* Q) and starting from the 16<sup>th</sup> century also *čem* (< *čto*-INS) are used almost exclusively in comparative-oppositive and substitution constructions (Krys’ko 2020: 399) containing the comparative form with the meaning ‘better’, exemplified in (Bulaxovskij 1958: 377) by *И чем было сосуды ковати, ино лучше бы шуба переменити* ‘And **instead of** forging tableware,



it would be better to change the fur coat' (XVI). Neither of the two, *kako* or *aky*, are used in contexts like (18) until the 17<sup>th</sup> century when *kako* expands its use onto these contexts, as well; this use is, however, not possible in the modern language.

As for similitive and equative constructions, they are analyzed in the same terms in all the sources. It is argued that they could be introduced by any of the three subordinators *aky* (with its phonetic variant *aki*), *kako* (with the variant *kakъ*) and *jako*, see also Sreznevskij (1893).

The marker *jako* is used as a general subordinator in OR, covering nearly all the functional types of complement and adverbial clauses, including eventive and propositional complements, adverbial clauses of time, reason, concession, and others. Therefore, we limit our research to *aky* and *kako*.

The first work mentioning the semantic difference between *aky* and *kako* is Stecenko (1972: 288). He notes that *aky* mainly introduces «figurative» similarity, as in the following:

(19) OR (The Moscow chronicle, cit. after Stecenko 1972, 124)

...сниде огнь съ небесе, акы облакъ велик над ручаи Лыбеди.

*snid-e*                      *ognъ съ*              *nebes-e,*              *aky* *oblakъ velik*  
come.down-AOR.3SG fire from sky-GEN.SG like cloud big(M.NOM.SG)  
*nad*    *rucha-i*              *Lybed-i*  
above river-ACC.SG Lybed'-ACC.SG

'A fire from heaven went down, like a big cloud above the river Lybed'.

In the recent encyclopedic dictionary "Historical Grammar of Russian Language" (Krys'ko 2020) the observation of A.N. Stecenko turns into a categorical statement: «The conjunction *aky* (> *aki*) was used to form only similitive sentences, containing figurative simile»<sup>1</sup>.

However, examples like (20) show that *aky* also introduces similitives that do not represent a «figure of speech»:

---

<sup>1</sup> The term *simile* is used for a subtype of similitives that have generic STAN, denoting a prototypical possessor of the high degree of the PAM (Kölbel 1993; Fortescue 2010).

## (20) OR (HG)

любити ближняго своего аки себе.

*ljubi-ti bližnj-ago svoj-ego aki seb-e*

love-INF neighbor-ACC.SG REFL.ADJ-ACC.M.SG like REFL-ACC.SG

‘to love one’s neighbor as oneself’

At the same time, neither in Stecenko (1972), nor in Kryś’ko (2020) it is not specified, whether *kako* can be used in such contexts.

Note that *jako* and *aky* are considered to be stylistically marked as belonging exclusively to written language (Stecenko 1972: 288; Bulaxovskij 1958: 368–372).

Thus, historical grammars do not answer the question whether OR has a differentiated system of comparative constructions. In the present paper we argue that *aky* and *kako* were semantically distributed in the 11<sup>th</sup>–15<sup>th</sup> centuries, and later *kako* has expanded onto *aky*-contexts.

### 3. Basic distribution of *aky* and *kako* in OR

This section is focused on the quantitative distribution of *aky* and *kako*. Given that *aky* is not used already in 19<sup>th</sup> century, we expected to find the gradual loss of *aky* and the expansion of *kako* onto the relevant contexts during the 11<sup>th</sup>–18<sup>th</sup> centuries.

At first, we estimated the frequency of both markers based on simple queries for *kako/kakъ/kak* and *aky/aki* in the Russian National Corpus (RNC). It turned out that *kako* has a larger use than *aky* even in OR, resulting in a gradual decrease and loss of the use of *aky* in 18<sup>th</sup> and 19<sup>th</sup> centuries. The exact figures are given in Table 1.

Time period	11 <sup>th</sup> –16 <sup>th</sup> cent.	17 <sup>th</sup> cent.	18 <sup>th</sup> cent.
<i>kako</i>	777	767	33820
<i>aky</i>	546	313	760
Ratio	1,4	2,4	44,5

**Table 1:** Number of examples with *kako* and *aky* from 11<sup>th</sup> to 18<sup>th</sup> century in RNC.

Table 1 shows that the proportion of *kako* to *aky* slightly changes in the 17<sup>th</sup> century (with *kako* attested 2,4 times more frequently than *aky*) showing sharp increase in the 18<sup>th</sup> century (with *kako* attested 44,5 times more often).

It is obvious that the use of simulative and equative markers is highly influenced by the genre of the text, so we also provide the distribution of *kako* and *aky* in texts

belonging to the same genre and devoted to the same topic, i. e. non-fiction historical texts. Table 2 shows that the proportion is slightly different for these texts; however, it still reveals the sharp decrease of *aky* in the 18<sup>th</sup> century and its complete loss by the beginning of the 19<sup>th</sup> century. For example, all the 4 occurrences of *aky* in Karamzin’s “Istoriya gosudarstva rossijskogo” (History of the Russian State) are found in citations of Old Russian proverbs. In archbishop Theophan’s texts, written in the beginning of the century, the ratio is already different than in older chronicles, while in the beginning of the 19<sup>th</sup> century *aky* is nearly absent.

Time period	Chronicles written in 11 <sup>th</sup> –16 <sup>th</sup> centuries (our sample)	Archbishop Theophan, beginning of 18 <sup>th</sup> century (RNC)	Historical texts of 18 <sup>th</sup> century (RNC) <sup>2</sup>	“Istoriya...” of Karamzin, beginning of 19 <sup>th</sup> century (RNC)
<i>kako</i>	140	207	1734	2185
<i>aky</i>	85	108	351	0 (4 citations)
Ratio	1,65	1,92	4,94	n/a

**Table 2:** Distribution of *kako* and *aky* in non-fiction historical texts of RNC and in our sample.

Thus, it seems reasonable to compare the distribution of *kako* vs. *aky* up to the 18<sup>th</sup> century, when the latter was completely lost. It can be expected that the major changes in the functions of both markers would be found in the texts of the 16<sup>th</sup>-17<sup>th</sup> centuries, as it was the time period of increase of the usage of *kako*.

#### 4. The functional distribution of *kako* and *aky* in OR

As discussed in section 2.2, historical grammars and dictionaries of OR do not reveal any semantic and syntactic difference between *aky* and *kako*. Examples can be found of both of them in both equative and similitive constructions. However, it is noteworthy that the relevant sections of historical grammars do not provide a detailed periodization, giving together examples dating from 11<sup>th</sup> century and 17<sup>th</sup> century (the relevant data are mostly discussed in sections dedicated to syntax, and they are not as detailed as the ones on historical phonetics and morphology). Another important question is the one of frequency. Even if a meaning is attested once or twice, it might

<sup>2</sup> A query of documents created in the 18<sup>th</sup> century and tagged as “text type: non-fiction; topic: historical” in RNC.

be an occasional example, and the expression in question might allow such use only marginally. To reveal the main uses of words and constructions in historical corpora, recent works on historical syntax count the frequency of various uses in a limited sample of texts (Hilpert & Gries 2009; Diessel & Hilpert 2016). This allows to determine, which meaning or use is the basic meaning of the examined word/construction at a given historical period, and which use is diminishing or increasing.

Thus, to reveal the functional distribution of *aky* and *kako* in OR, we have compiled several samples of their instances attested in documents belonging to various time periods. We made a special focus on the 18<sup>th</sup> century, making a randomized search query of 400 examples with *aky* and 400 examples with *kako* texts in RNC. For the 16<sup>th</sup> and the 17<sup>th</sup> century we made randomized search queries of 150 examples with *aky* and 150 examples with *kako*. For older texts the RNC does not give the possibility of specifying a query for a given period. Therefore, we compiled our own sample of texts starting from the first available manuscripts of the 11<sup>th</sup> century up to the manuscripts of 15<sup>th</sup> century. To make our sample consistent from the point of view of language varieties, we limited our research, first, to documents created in the north or center of Russia, second, to original documents (rejecting translations). The list of sources is given at the end of the present article.

The 11<sup>th</sup>–15<sup>th</sup> centuries sample has 81 occurrences of *aky* and 164 occurrences of *kako* in it. All the examples (our sample + RNC queries) with the markers in question have been manually annotated according to the meaning represented and the presence of other markers (since both *aky* and *kako* are largely used with other subordinators and particles to form complex subordinators and markers, for example *kako ti*, *kako to*, *aki by*). Note that we have left *kako i*, since this complex marker seems to have the same distribution in comparative constructions as *kako* without *i*. Sticking to the terminology discussed in section 2.1 we use the following semantic tags: equative, similitive, inequality comparative, manner adverbial clause, IPE. The examined markers have a number of functions outside these semantic domains, namely, additive, role, approximation, ‘hearsay’ and irrealis. The lexeme *kako* is a manner question word largely used in independent questions (21) and exclamations and subordinate adverbial clauses of manner (22), temporal adverbial clauses, parentheticals, manner and eventive complement clauses.

(21) OR (NovgB 49, XV)

оу мене бориса в животъ нѣтъ. какъ се господо мною попецалуєте и  
моими дѣтми.

*u mene boris-a v život-ě něť kakъ se*  
around I.GEN Boris-GEN.SG in life-LOC.SG NEG.COP how REFL

*gospod-o mnoju popecaluj-ete i moi-mi dět-mi*  
gentleman-VOC.PL I.INS care-PRS.2PL and my-INS.PL children-INS.PL

‘My [husband] Boris passed away (lit. At/by me, there is no Boris in life).

Dear Sirs, **how** are you going to care for me and my children?’

(22) OR (NovgB 359, XIV)

осподине како љсьмъ порадилесе тако и живу а василке село пустоши.

*ospodine kako ěsťmъ porědi-l-e-se tako i živ-u*  
gentleman how be.PRS.SG arrange-PST-PL-REFL so and live-PRS.1SG

*a vasilke selo pustoši*

while Vasilke village ruin.PRS.3SG

‘Sir, I live **as** we have arranged (**according to the conditions** we have discussed), while Vasilke brings the village to ruin.’

The results of our tagging are as follows. Up to the 16<sup>th</sup> century, *aky* and *kako* divide the functions in the following way: *aky* marks the STAN in similitive (23) and equative constructions (24), approximation (25), role (26), and irrealis complement clauses (27), while *kako* is only attested in IPEs (28) and adverbial clauses of manner (22).

(23) OR (PVL)

постомъ явишася отци наши акы свѣтила в миръ

*post-omъ javi-ša-sja otc-i naš-i akы svět-il-a v mir-ě...*

fast-INS appear-AOR.3PL-REFL father-PL our-PL like star-PL in world-LOC

‘our fathers appeared in the world **like** stars, through fasting...’

(24) OR (PVL)

Нѣциї видѣша рано възходящю солнцю бысть на 3 углы яко и коврига,

потомъ мнѣи бысть аки звезда.

*potomъ mněi bystъ aki zvězd-a*  
 then small.CMPR be.AOR.3SG as star-NOM.SG

‘Early in the morning some people saw, as the sun was rising that it was triangle-formed, like a loaf and then became smaller, **as** a star’.

(25) OR (RNC: Chudesa Nikoly. XII)

чюдахоуса гл҃юще кѣ себѣ. акы гласъ ѣсть нашего соуѣда дѣмитриѣа.

*čudj-axu-sja glagolju-šče kъ seb-ě aki glasъ jestъ*  
 surprise-IMPV.3PL-REFL say-PTCP to REFL-DAT.SG like voice be.PRS.3SG  
*naš-ego susěd-a dьmitrij-a*  
 our-GEN.SG neighbor-GEN.SG Dmitriy-GEN.SG

‘[The neighbours] were surprised and told to themselves: it is, **apparently**, a voice of our neighbor, Dmitriy. (However, he has left yesterday.)’

(26) OR (PVL)

сію бо хвалѣть Рустіе сынове, акы началницю.

*sij-u bo xvalj-atъ Rusti-je synov-e, aki*  
 this.F-GEN.SG because praise-PRS.3PL Russian-NOM.PL son-NOM.PL like  
*načalnic-u*  
 initiator-ACC.SG

(She was the first Russian who entered the kingdom of heaven.) ‘This one is praised by sons of Russia **as** their initiator.’

(27) OR (PVL)

мнѣти же всѣмъ человекомъ зряче, акы кровь прольѣна на снѣгу.

*mně-ti že vsěmъ čelovek-omъ zrja-če, aki krovъ*  
 think-INF EMPH all.DAT.PL person-DAT.PL see-PTCP like blood  
*prol’ja-n-a na sněg-u*  
 spill-PTCP-F.SG on snow-LOC.SG

‘(There was a sign from God.) All the people saw **as if** blood was spilled over the snow.’

The STAN of IPEs may be marked by both *aky* and *kako*:

(28) OR (Novg1. XIII-XIV)

солнче... бы акы въ 5 ноціи мѣсяць.

*solnč-e by akы въ 5 nocii měsjačь*

sun-NOM.SG be.AOR.3SG like in 5 night moon.GEN.PL

‘(Describing the sun eclipse.) The sun became of the same size as a moon on its fifth night’.

(29) OR (NovgB 364. XIV)

даи намъ ржи на полтину какъ людомъ поцнешь давать.

*dai namъ rž-i na poltin-u kakъ ljud-otъ*

give.IMP.2SG we.DAT rye-GEN.SG on 50.kopejka-ACC.SG how people-DAT.PL

*počn-ešъ dava-tъ*

begin-PRS.2SG give-INF

‘Give us rye that costs 50 kopejka’ price, [on the same conditions] as you are going to give to other people.’

(28) is arguably an example of the IPE, since the STAN is given a detailed description allowing to identify its exact size (the new moon as it is on its fifth day). It seems hard to interpret this example as a similitive: the speaker’s aim is to describe the size of the sun, rather than make a comparison (small like a moon). As for (29), it makes use of *kako* to refer to the exact conditions of selling rye, which is explicitly indicated in the context. Thus, it is not about similarity, rather the speaker aims at identifying the extent of an implicit PARA (conditions).

Thus, the IPE context is the only one where both *kako* and *akы* can be used interchangeably. As for other comparison contexts, they are chiefly introduced by *akы*. To verify this claim, we made randomized searches of *kako* in the Old Russian subcorpus of the Russian National Corpus. The results follow the pattern observed in our sample. This does not mean that *kako* could never be used as a STM of similarity or equality (historical dictionaries Sreznevskij (1893) and Slovar russkogo jazyka... (from 1975) give two similarity examples dating from the 14<sup>th</sup> and 15<sup>th</sup> centuries and one equality example dating from the end on the 15<sup>th</sup> century); however, its frequency in these constructions was low, as reflected in our sample. It may be concluded that *kako* started to expand onto similitives and equatives approximately in the 14<sup>th</sup>–15<sup>th</sup> centuries. By contrast, the IPE reading was widespread much earlier.

Further history of *aky* and *kako*, as it appears in our samples, is summarized in Table 3. For the sake of brevity, we omit non-relevant uses of both markers, such as the subordinator use and use of *kako* as a question word. The percentage is, therefore, calculated not from all uses, rather from the relevant sample. The exact figures are given in the last column. Empty cells in the table represent zero values.

Marker	Time period (cent.)	Adv manner clauses	Implicit PARA equative (IPE)	Equative	Similarative	Role	Approximator	Additive	Parentetical	Inequality compar.	Total relevant examples /whole sample
<i>aky</i>	11–15 <sup>th</sup>		10,13		83,54	6,33					79/81
<i>kako</i>	11–15 <sup>th</sup>	66,67	33,33								15/164
<i>aky</i>	16 <sup>th</sup>		0,93		78,7	3,7	15,74	0,93			99/150
<i>kako</i>	16 <sup>th</sup>	10,26	56,41	5,13	28,21						35/150
<i>aky</i>	17 <sup>th</sup>			2,34	82,03	7,81	6,25	1,56			124/150
<i>kako</i>	17 <sup>th</sup>	23,81	33,33	4,76	28,57				9,52		21/150
<i>aky</i>	18 <sup>th</sup>				76	12,57	10,28	1,14			237/278 <sup>3</sup>
<i>kako</i>	18 <sup>th</sup>	7,14	14,29	11,61	29,46	19,64	3,57	3,57	2,68	8,04	604/621

**Table 3:** Distribution of comparative and related meanings of *aky* and *kako* in the relevant part of the sample.

The results of the Table 3 are represented in Figure 1.

<sup>3</sup> The whole number of examples of *aky* in this time period.



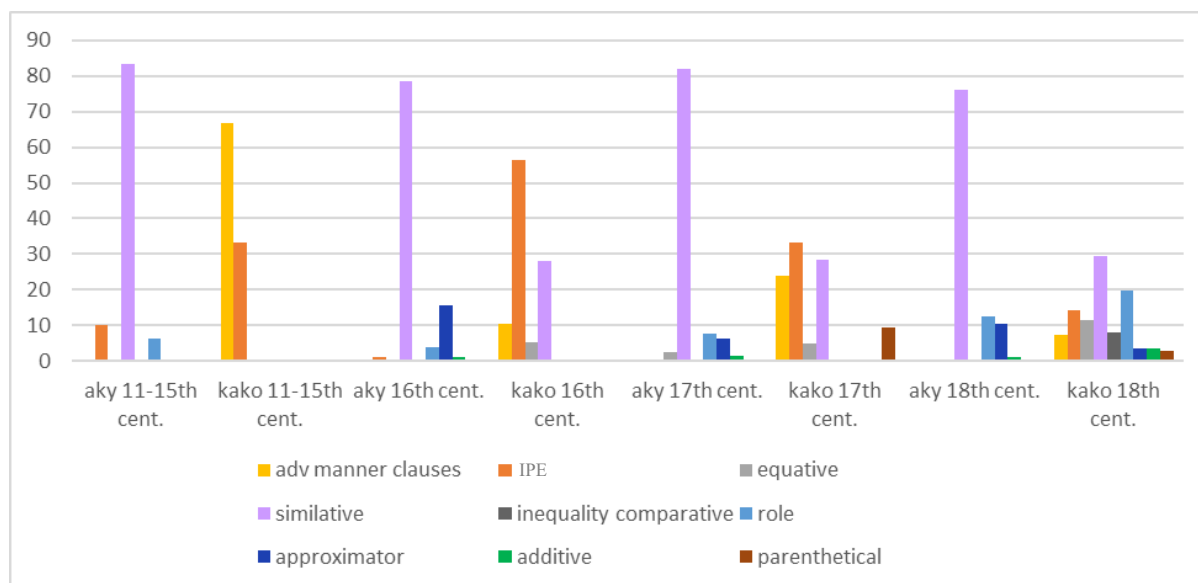


Figure 1: Functional distribution of *aky* and *kako* in OR and Middle Russian.

The distribution of meanings seems to be similar in the 16<sup>th</sup> and the 17<sup>th</sup> centuries, so further we treat them on a par. The results show that the main function of *aky* was the simulative (30), since the percentage is nearly the same (about 80%) in all the time periods.

### (30) OR (RNC. XVII)

молот же истребляет железо аки плевели.

*molot že istrebljaj-et želez-o aki plevel-i*

hammer EMPH destroy-PRS.3SG iron-NOM.SG like weed-NOM.PL

‘As for the hammer, it destroys iron like weed.’

The equative and IPE meanings are not infrequent; however, they seem to show something that looks like complementary distribution: equative is absent in the earliest period, while IPE constructions decrease in frequency in the 17<sup>th</sup>–18<sup>th</sup> centuries. We will return to this below.

At large, *aky* did not significantly change its distribution among simulative and semantically related contexts during the whole time period. The only significant change was a decrease in use in IPEs.

By contrast, *kako* was hardly ever used in any contexts of comparison in the earliest period. The only relevant context it could cover is the IPE, where it competed with *aky*. Sixteenth and seventeenth centuries, however, see the expansion of *kako* onto

contexts of comparison, both similitive (31) and equative (32). The IPE is still well-attested (33). The parenthetical use arises in the 17<sup>th</sup> century.

(31) Middle Russian (RNC. XVI)

79-ть ложекъ серебряны золочены... на концѣхъ какъ коруночки литые.  
*na konc-ěxъ kakъ korunočk-i lityj-e*  
 on end-LOC.PL how crown-NOM.PL casted-NOM.PL  
 ‘79 gilded silver spoons... which have at their end as if casted crowns.’

(32) Middle Russian (RNC. XVII)

Да за стрелетцкие хлебные запасы только взять против государева указу,  
как збирано Устюжские чети з городов во 149-м и во 150-м, и во 151-м,  
 и во 152-м году за стрелетцкие хлебные запасы по 168 рублев с сохи...  
*tolьko vzja-tь protiv gosudarev-a ukaz-u, kak*  
**as.many** take-INF according Sire-GEN.SG law-GEN.SG how  
*zbird-n-o [...] po 168 rublev s sox-i*  
 collect-PCTP-N.SG.NOM DISTR 168 rouble.GEN.PL from plough-GEN.SG  
 ‘According to the Sire’s law, the streletses’ (a Russian army unit) bread stock  
 should be paid 168 roubles each plough, **as much as** it was taken from  
 the towns in years 149, 150, 151 and 152 by in Ustjug’s area.’

(33) Middle Russian (RNC. XVII)

вы бѣ съ тѣхъ лавокъ тѣмъ лавочнымъ сидѣлцомъ, въ которыхъ лавкахъ  
 сидятъ по земь, велѣли платити къ Ивану Предтечи на годъ  
 по полуполтинѣ, или какъ иные сидѣлцы платятъ.  
*ili kakъ inyj-e sidělc-y platj-atъ*  
 or how other-NOM.PL leaser-NOM.PL pay-PRS.3PL  
 ‘You should tell the sellers, who are earth leasers, to pay at John the Baptist’s  
 holiday 25 kopejka a year, or **on the conditions** kept by other leasers.’

In the 18<sup>th</sup> century *kako* (already in the form *kak*) is largely used in all contexts of comparison (34-36) and closely related meanings such as additive, approximation, role etc. Note that it can also mark STAN of inequality equatives (37); this use, however, was lost in MR.

(34) Middle Russian (RNC. XVIII)

*Živ-ut*        *v les-ax,*        *kak dikij-e*        *zver-i.*  
 live-PRS.3PL in forest-LOC.PL how wild-NOM.PL animal-NOM.PL  
 ‘They live in forest, like wild animals do.’

(35) Middle Russian (RNC. XVIII)

*prežn-ije*        *gllov-n-yje*        *ubor-y*        *bud-ut*        *tam*  
 former-NOM.PL head-ATTR-NOM.PL dress-NOM.PL be.FUT-3PL there  
*stol’*        *že*        *redk-i,*        *kak drevn-ije*        *stroenij-a*  
 same.extent EMPH rare-NOM.PL how ancient-NOM.PL building-NOM.PL  
*egiptjan.*  
 Egyptians.GEN.PL  
 ‘The old hats will be there as rare as ancient Egyptian buildings.’

(36) Middle Russian (RNC. XVIII)

Доношу же вам, что мы между Верхотурья и Чусовой нашли железной  
 руды такой же, как у Демидова на Тагиле.  
*naš-l-i*        *železn-oj*        *rud-y*        *tak-oj*        *že,*        *kak*  
 find-PST-PL iron-ATTR-F.GEN.SG ore-GEN.SG such-GEN.SG EMPH how  
*u Demidov-a*        *na Tagil-e.*  
 at Demidov-GEN.SG on Tagil-LOC.SG  
 ‘I hereby report to you that we have found between Verxoturje and Chusova  
 iron ore, the same as Demidov has in Tagil.’

(37) Middle Russian (RNC. XVIII)

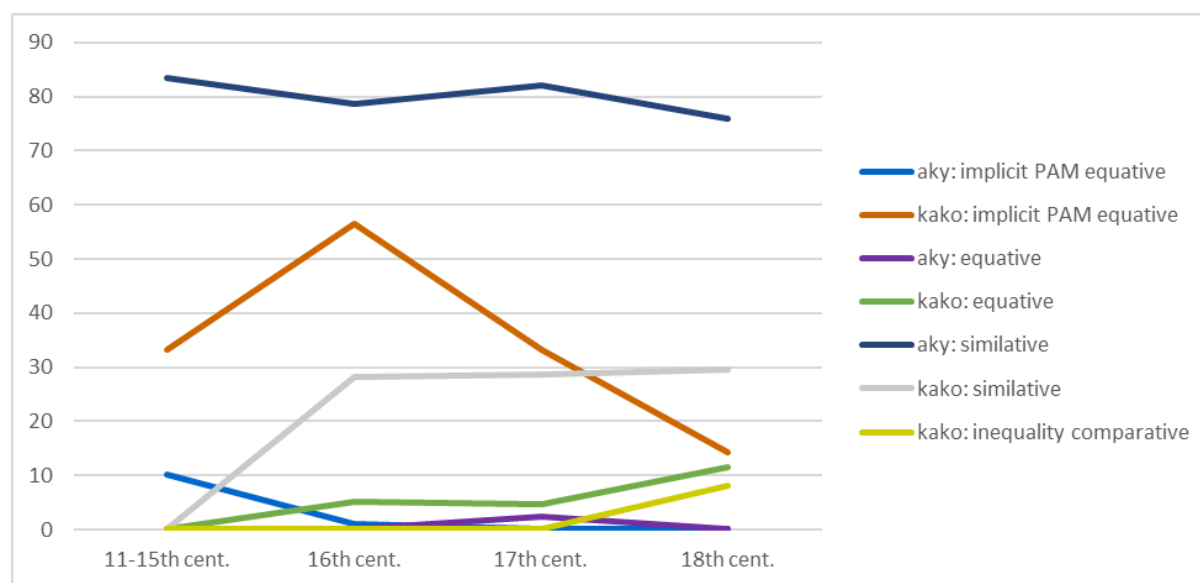
*k dostiž-enij-u*        *blagopoluchij-a*        *net*        *bliže*        *i*  
 to attain-VN-DAT.SG prosperity-GEN.SG NEG.COP close.CMPR and  
*priličn-eje*        *sredstv,* *kak prjam-aja dobrodetel’*        *i prjam-oj*        *razum...*  
 decent-CMPR means how right-F.SG virtue        and right-M.SG mind  
 ‘There is no closer and more decent means of attaining prosperity than the right  
 virtue and the right mind’.

The marker *aky* could also cover a number of meanings closely related to comparison, such as approximation, additive, role etc. These uses are exclusively covered by *aky* until the 18<sup>th</sup> century, while in the 18<sup>th</sup> century they can be introduced by *kako*, as

well. The parenthetical use is only attested with *kako* and only in the 17<sup>th</sup> and 18<sup>th</sup> centuries.

Thus, we suppose that the similarity/equality use of *kako* develops in the 16<sup>th</sup> century and slowly increases afterwards. In the 18<sup>th</sup> century *kako* shows an expansion onto all the contexts closely related to comparison, while *aky* preserves the same functions it had in older texts. The period of expansion coincides with the major increase of *kako* in contexts of comparison.

Now let us examine individual types of comparative constructions. The results on the basic comparison constructions are demonstrated in Figure 2.



**Figure 2:** Functional distribution of *aky* and *kako* in the basic contexts of comparison.

The similative meaning is the majoring type of *aky* in all time periods; *kako* expands onto these contexts already in the 16<sup>th</sup> century. Thus, this type involves competition of both markers, observed in historical grammars and dictionaries.

The inequality comparative use (the yellowish green line) is only attested with *kako* in the 18<sup>th</sup> century. The development of this use happened after the complete disappearance of *aky*.

The equative and the IPE contexts show the following pattern. True equative examples are not found at all in the earliest period, neither with *aky* nor with *kako*. Yet, this does not mean that they were totally impossible. We have made several corpus queries of the type “adjective/adverb + *aky/aki/kako/kakъ*” and “*aky/aki/kako/kakъ* + adjective/adverb”, with a distance from 1 to 3 words. Single

examples of equatives are found, see (6); however, this construction seems to be extremely infrequent. We also tried similar queries with *jako* (the general subordinator); however, only solitary examples are found as the following:

(38) OR (RNC. XIII)

и тако далече живоутъ ѿ римланъ, ѿко нбо ѿ земља.

*i tako daleče živ-utъ oтъ rimljan-ъ jako nebo oтъ*  
and so far live-PRS.3PL from Roman-GEN.PL SUB heaven from  
*zemlj-a.*

earth-GEN.SG

‘And they live as far from Romans as heaven from earth.’

Note that this example comes from the translation of “The Jewish War”, authored by Flavius Josephus, and could be influenced by the original Greek text.

The rise of equative use is observed in the 16<sup>th</sup> century, when it is mostly introduced by *kako* as STM (the grey line); *aky* is also attested in these constructions, although very infrequently (the violet line). Hence, we assume that *kako* is the main means of encoding STAN of equative in the 16 – 18<sup>th</sup> centuries.

It is intriguing that the context of IPE decreases in frequency parallelly with the rise of the true equative. Namely, in the earliest period this context is covered by both *aky* and *kako* (*kako* being more frequent), in the 16<sup>th</sup> century it is quite widespread; however, it is predominantly covered by *kako* (39). Examples with *aky* (40) are not numerous. The 17<sup>th</sup> and 18<sup>th</sup> centuries see the decrease of this use.

(39) Middle Russian (RNC. XVI)

...on u nix korm svoj i kon-sk-oj kup-it  
he at they.GEN forage REFL.M.SG and horse-ATTR-NOM.SG buy-PRS.3SG  
*po cen-e, kak jemu prodad-ut.*  
according price-DAT.SG how he.DAT sell-PRS.3PL

‘He will buy from them forage for himself and for horses **at the price they will sell it.**’

(39) is arguably not a similitive, rather an IPE, as can be inferred by the context: it involves the instructions that the price be equal, not similar. This meaning can also be encoded by *aky*, see (40), where the size of the sun is determined by equality to

the size of the moon on its third night. Hence, the implicit PARAs here are the price and the size.

(40) Middle Russian (RNC. XVI)

*a solnc-e že osta-sja aki molod-ъ*  
 and sun-NOM.SG EMPH remain.AOR.3SG-REFL like young-NOM.SG  
*měsjac-ъ treju dnej.*  
 moon-NOM.SG three.GEN day.GEN.PL

‘The sun became just as the new moon on its third night.’

We explain the observed asymmetry in frequency of the two constructions in the following way. The equative as a construction was nearly absent and only started to develop in the 11–15<sup>th</sup> centuries period; the earliest examples (coming from corpus queries (6) and historical dictionaries) date from the 13<sup>th</sup> century and are solitary. The equative meaning at this time was encoded by the construction we termed as “implicit PAM equative”. Indeed, many relevant examples contain an adjective that occurs after (and not before) the STAN, see (19) and the following:

(41) OR (RNC. XIII)

и выросло у него и(з) чела трояндифило(въ) цвѣ(т) ѿ сыропустную н(д)лю.  
аки сыръ бело.

*trojandofil-ov-ъ cvět-ъ [...] aki syr-ъ bel-o.*  
 rose-ATTR-NOM.SG.M flower-NOM.SG like cheese-NOM.SG white-NOM.SG.M

‘And during the week before Lent a rose flower grew out of his forehead white as cheese / similar to the white cheese.’

(42) OR (HG. XIII)

Си же благочъстивии князи рязаньстии концяшася мѣсяця июля въ 20,  
 на святого пророка Илии, и прияша вѣнца от Господа Бога, и съ своею  
 дружиною, акы агньцы непорочьни прѣдаша душа своя Богови.

*aky agnъc-y neporočъn-i předa-ša duš-a*  
 like lamb-NOM.PL innocent-NOM.PL give-AOR.3PL soul-NOM.DU  
*svoj-a Bogovi.*  
 REFL.GEN-NOM.DU God.DAT

‘They gave their souls to God, as innocent as lambs / innocent like lambs.’

All these examples are ambiguous from the point of view of their syntactic structure: they can be analyzed as “X, like a an [N Adj]” or “X, as/like a N, Adj”. Given that punctuation marks are lacking these structures cannot be differentiated. Even if the syntactic structure is [N Adj], it seems that semantically they should be analyzed as constructions of equality “big as a cloud” / “white as cheese” / “innocent as lambs”. The alternative interpretation seems to be problematic: it would be strange to compare a rose to a white cheese, and fire to a big cloud, while the equative interpretation seems to be pragmatically more appropriate. We termed these examples as “implicit parameter equatives”; however, we suppose that they could have the same meaning as English equatives like *as innocent as lambs*.

Note that in our sample *kako* in IPE is already attested at the turn of the 12<sup>th</sup>–13<sup>th</sup> centuries, and it is well represented afterwards. By contrast, the equative use of *kako* only starts in the 16<sup>th</sup> century and comprises a single example, see (8) in section 1. Our explanation clarifies why the equative construction is so poorly attested in the earlier period: apparently, it developed from the IPE.

Thus, we assume that the IPE constructions gave rise to the true equative constructions with the parametric adjective preceding the STAN like in (6). This explains the fall of frequency of implicit PAM constructions in the 17<sup>th</sup> century, parallel to the rise of true equatives. Perhaps, the word order illustrated in (41-42) and (19) is reflected in the later variation of word order in MR equatives, as shown in (10).

The similarity use of *kako* emerged from the IPE constructions, as well. We suggest the following pattern of semantic change for *kako*:

subordinator in adverbial clauses of manner	→	STM in IPEs	→	STM in equatives and similatives
(17)		(40)		(6)
“payed <i>kako</i> we have decided”		“payed <i>kako</i> other holders” (on the same conditions)		EQU: “payed much <i>kako</i> (as) other holders” SIM: “payed very much, <i>kako</i> (like) other holders”

Figure 3: The pattern of semantic change for *kako*.

Given that similatives encode sameness of manner (Haspelmath & Buchholz 1998: 278), it can be hypothesized that manner has served as a starting point for the expansion of the similative use in OR. We suggest that the context of IPE has served as a switch context, where PARA was a kind of manner, including quantitative PARAs, price, form, structure and other. In the 16<sup>th</sup> century this construction expanded onto subjective PARAs such as beauty, love (beautiful like a flower, loved them like his children etc.). At this point, however, equatives are hard to differentiate from similatives, since it is often unclear whether the author intended to mean “paid as much as X” or “paid a lot, like X did”.

Another issue is the emergence of the correlative pair constructions illustrated above (3-4). We made several queries of the type “*aky/aki/kako/kakъ/jako... tako/takъ*” with different word order and different word-to-word distance in RNC. The queries’ results show that the correlate started to appear regularly in the 15<sup>th</sup> century, while solitary examples may be occasionally found earlier, see (44) dating from the 13–14<sup>th</sup> century.

(43) OR (Novg1. XIII–XIV)

тма бысть в солнци, съ запада акы мѣсяць бысть въ 5 ноци.

*tma bystʹ v solnc-i sʹ zapad-a aky mĕsjac-ъ*  
 darkness be.AOR.3SG in sun-LOC.SG from east-GEN.SG like moon-NOM.SG  
*bystʹ vʹ 5 nočii.*  
 be.AOR.3SG in 5 night.LOC.PL

‘There was darkness in the sun, [looking] from the east it was **just as** the moon on its fifth night.’

(44) OR (Novg1. XIII-XIV)

тма бысть тако же акы мѣсяць 5 ноци.

*tma bystʹ tako že aky mĕsjac-ъ 5 nočii.*  
 darkness be.AOR.3SG so EMPH like moon-NOM.SG 5 night.GEN.PL

‘There was darkness, [the sun became] just as the moon on its fifth night.’

Note that the correlate was not obligatory, since similar examples are found in the same document, describing the same situation of the sun eclipse, with and without *tako že*.



However, the majority of examples with correlative pairs attested before the 15<sup>th</sup> century include *jako* (38) or *jako že*, rather than *aky* or *kako*. We made some counts among the equatives and IPEs in our samples. In the 16<sup>th</sup> and 17<sup>th</sup> centuries, *aky* never cooccurs with a correlate. As for *kako*, the percentage of examples with correlates rises to 39% in the 16<sup>th</sup> century and 50% in the 17<sup>th</sup> century, and in the 18<sup>th</sup> century all our examples with equative and IPE include the correlate *tako*, *takoj*, *stol'* and others.

Thus, we conclude that the correlative pairs with *aky* and *kako* spread in the 15<sup>th</sup> century, and in the 18<sup>th</sup> century the construction “*kak... tak*” / “*tak... kak*” got fossilized in the equative meaning.

In conclusion, the expansion of *kako* onto *aky*-contexts happened parallelly with the decrease of *aky* in frequency, namely, in the 16<sup>th</sup> and the 17<sup>th</sup> centuries. At the same time the correlative constructions got grammaticalized as a main means of encoding the equative meaning.

As for inequality constructions, they do not include neither *kako* nor *aky* in the considered time. This meaning is encoded by the genitive case both in OR and MR. However, in the 18<sup>th</sup> century *kako* is largely used as STM in inequality constructions. In MR this function is taken over by *chem*, the instrumental case of the question word *chto* ‘what’.

## 5. Conclusions

The present paper studies the functional distribution of two OR markers used in the domain of comparison, *aky* and *kako*. Contrary to what is stated in historical grammars and dictionaries, we demonstrate that the two markers are not synonymous. The frequency counts show that they divide the functional domain in the following way: in 11<sup>th</sup>–15<sup>th</sup> centuries the main function of *aky* is similitive (although it could also occur in IPEs), while *kako* is mostly used as STM in IPEs.

In the 16<sup>th</sup>–17<sup>th</sup> centuries both markers could be used interchangeably in similitive contexts, while *kako* (unlike *aky*) expanded onto equatives.

In 18<sup>th</sup> century *aky* was lost, concurrently with the expansion of *kako* onto similitive contexts and onto comparative contexts. The marker *kako* then remained as a solitary device used for all types of comparison. At this stage the correlative pair construction has grammaticalized to encode the equative meaning. Thus, the similitive and the equative became differentiated, and this differentiation is observed

until nowadays. The comparative function of *kak* was completely lost in the 20<sup>th</sup> century (Švedova 1980, v. 2: 489, § 2812).

Equative contexts are hardly ever observed in OR texts of the 11 – 15<sup>th</sup> centuries. We hypothesize that the emergence of equative use of *aky* and *kako* started from the IPE contexts like “we sold rye to Ivan as to you (on the same conditions)”. These contexts started to include adjectives and adverbs (‘as much as’; ‘as tall as’) roughly in the 15<sup>th</sup> century and later this use was reserved to *kako*, unlike *aky*. This latter showed a strong preference towards the similative in all time periods.

### Acknowledgements

This paper is written with support from the project “Parametric description of the languages of Russian Federation” realized at Pushkin State Russian Language Institute.

### Abbreviations

1/2/3 = 1st/2d/3d person	GEN = genitive	PAM = parameter
ACC = accusative	IMP = imperative	PARA = parameter
ADJ = adjective	IMPF = imperfect	PL = plural
ADV = adverb	INF = infinitive	PRS = present
AOR = aorist	INS = instrumental	PST = past
ATTR = attributive	IPE = implicit parameter	PTCL = particle
CMPR = comparative suffix	equative	PTCP = participle
COP = copula	IPFV = imperfective	REFL = reflexive
DAT = dative	IRR = irrealis	SG = singular
DIM = diminutive	LOC = locative	SIM = similative
DISTR = distributive	M = masculine	STAN = standard of comparison
DU = dual	MR = Modern Russian	STM = standard marker
EMPH = emphatic particle	N = neuter	SUB = subordinator
EQU = equative	NEG = negation	VN = verbal noun
F = feminine	NOM = nominative	VOC = vocative
FUT = future	OR = Old Russian	

### References

Bulaxovskij, Leonid A. 1958. *Istoričeskij kommentarij k ruskomu literaturnomu jazyku* [Historical comments to Standard Russian]. Kiev: Radjans’ka škola.

- Borkovskij, Viktor I. & Pëtr S. Kuznecov. 1965. *Istoričeskaja grammatika russkogo jazyka*. [Historical Grammar of Russian.] Moscow: Nauka.
- Diessel, Holger & Martin Hilpert. 2016. Frequency effects in grammar. In Mark Aronoff (ed.), *Oxford Research Encyclopedia of Linguistics*. New York: Oxford University Press.
- Hilpert, Martin & Stefan Th. Gries. 2009. Assessing frequency changes in multi-stage diachronic corpora: Applications for historical corpus linguistics and the study of language acquisition. *Literary and Linguistic Computing* 24(4). 385-401.
- Kölbel, Marianne. 1993. To grin like a Cheshire cat. *Arbeiten aus Anglistik und Amerikanistik* 18. 3–22.
- Fortescue, Michael. 2010. Similitude: A conceptual category. *Acta Linguistica Hafniensia* 42(2). 117–142.
- Haspelmath, Martin & Oda Buchholz. 1998. Equative and similative constructions in the languages of Europe. In Johan van der Auwera (ed.), *Adverbial Constructions in the Languages of Europe*, 277–334. Berlin: Mouton de Gruyter.
- Kobozeva, Irina M. & Olga YU. In'kova. 2018. “Kak” i ego dvuhmestnye varianty. *Semantika konnektorov: kontrastivnoe issledovanie*. Moscow: TORUS PRESS. 169–239.
- Krys'ko, Vadim B. 2020. *Istoričeskaja grammatika russkogo jazyka: encyklopedičeskij slovar'* [Historical grammar of Russian: encyclopedia]. Moscow: Azbukovnik.
- Letuchiy, Alexander B. 2015. *Sravnitel'nye konstrukcii* [Comparative constructions]. Materials for the project of the Russian Corpus Grammar (<http://rusgram.ru>). Ms. Moscow.
- Lomtev, Timofej P. 1956. *Očerki po istoričeskomu sintaksisu russkogo jazyka: učebnoje posobije dlja gosudarstvennyx universitetov* [Sketches on historical syntax of Russian: a manual for state univerisities]. Moscow: MSU.
- Slovar russkogo jazyka XI~XVII vv.* [Dictionary of Russian language of the XI-XVIIth centuries]. Starting from 1975. Moscow: Nauka.
- Sreznevskij, Izmail I. 1893. *Materialy dlya slovarya drevnerusskogo yazyka* [Materials for the dictionary of the Old Russian language]. Saint-Petersburg: edition of the Department of Russian language and literature of the Imperial Academy of Sciences (1893–1902).
- Švedova, Natalia YU. 1980. *Russkaja grammatika* [Russian grammar], vol. I–II. Moscow: Nauka.

- Stecenکو, Aleksej N. 1972. *Istoricheskij sintaksis Russkogo yazyka* [Historical syntax of Russian.] Moscow: Vysshaja Shkola.
- Treis, Yvonne & Martine Vanhove (eds.). 2017. *Similitive and Equative Constructions: A cross-linguistic perspective*. Amsterdam/Philadelphia: John Benjamins.
- Treis, Yvonne. 2018. Comparative Constructions: An Introduction. *Linguistic Discovery* 16(1). i–xxvi.

## Sources

From Polnoje sobranije russkix letopisej (The complete collection of the Russian chronicles). Saint-Petersburg: tipografija Eduarda Praca, 1841:

- *PVL*: Povest' vremennyx let (Primary Chronicle). Lavrent'jevskij spisok (copy). XI–XII<sup>th</sup> cent.
- *NovgI*: The first Novgorod synodal chronicle. XIII–XIV.
- *Suzd*: The Suzdal chronicle. XII–XIV.
- *Voskr*: The Voskresenskaya chronicle (up to page 172). XVI.

From the electronic collection of birchbark letters <http://gramoty.ru/birchbark/>:

- *NovgB*: Novgorod birchbark letters. XII–XV<sup>th</sup> centuries.

From the electronic collection <http://lib.pushkinskijdom.ru/>:

- *Izb*: Izbornik (a collection of spiritual and moral texts). 1076.
- *RusP*: Russkaya Pravda (the Russian code of laws). XIV.
- *HG*: Hagiographies and sermons. XII–XIV.
- *Vass*: A story of sickness and death of Vassilij the III. XVI.
- *ANik*: A. Nikitin. Journey beyond three seas. XV. (Cit. after Xoženije za tri morja Afanasija Nikitina / Edited by Ja. S. Lur'je and L. S. Semenov. 3-e edn. Leningrad: Nauka, 1986.)

Subcorpora of the Russian National Corpus (RNC):

- The Old Russian subcorpus of the Russian National Corpus at [https://ruscorpora.ru/new/search-old\\_rus.html](https://ruscorpora.ru/new/search-old_rus.html)
- The Middle Russian subcorpus of the Russian National Corpus at [https://ruscorpora.ru/new/search-mid\\_rus.html](https://ruscorpora.ru/new/search-mid_rus.html) (We excluded from the results

of the queries examples from chronicles, since most of them were copied from earlier sources.)

- The Main subcorpus of the Russian National Corpus.

#### CONTACT

Irina Kobozeva [kobozeva@list.ru](mailto:kobozeva@list.ru) (corresponding author)

Natalia Serdobolskaya [serdobolskaya@gmail.com](mailto:serdobolskaya@gmail.com) (corresponding author)

# Comparative constructions across the German minorities of Italy: a semasiological approach

LIVIO GAETA<sup>1</sup>

<sup>1</sup>DEPARTMENT OF HUMANITIES, UNIVERSITY OF TURIN

Submitted: 25/11/2020 Revised version: 16/06/2021

Accepted: 22/06/2021 Published: 31/08/2021

## Abstract

Comparative constructions of inequality display a recurrent pattern throughout all Germanic languages, which is partially inherited from the Indo-European mother tongue. This common semasiological format consists in a copulative construction in which the adjective expressing the quality carries a comparative suffix and is accompanied by a particle introducing the standard. For the latter, a morpheme coming from various onomasiological domains is generally recruited. After a general overview of the construction within the Germanic family, the paper will focus on its consistency in the German linguistic islands of Northern Italy, where a remarkable variety is found, which is only partially due to the long-standing contact with Romance languages. Besides an overview of the Bavarian islands of the North-East, particular attention is devoted to the Walser German islands of the North-West, where a number of peculiar patterns are found, which partially reflect structural possibilities attested in earlier stages of the German-speaking territory, but also display unique developments such as for instance the comparative particle *ŝchu* 'so' found in Rimella.

**Keywords:** comparative construction; semasiology; onomasiology; language minority; linguistic island; language contact.

## 1. Introduction

Comparative Constructions of Inequality (= CCI) display a recurrent pattern throughout all Germanic languages, which is partially inherited from the Indo-European mother tongue and corresponds to the other cognates of the family. This

generally consists in a copulative construction in which the adjective expressing the quality also carries a comparative suffix and is accompanied by a particle introducing the standard. In addition, as is also typical of other Indo-European languages, the suffixes show to be diachronically replaced by comparative adverbs preceding the adjectives. A similar trend towards the increase of analytic coding also concerns the expression of the standard, which in Proto-Germanic used to display case-marking in the absence of any comparative particle. From this perspective, it does not come out as a surprise to observe that the new analytic pattern displays a variety of particle types coming from a set of source morphemes distributed from North to South in a consistent way.

In this paper, basically relying on Stolz's (2013) impressive typological investigation on CCIs in the European languages, I will discuss in Section 2 the theoretical premises of my work suggesting a semasiological approach to CCIs as a convenient way for doing typological comparison. On this basis, I will review in Section 3 the comparative constructions as they are found throughout the Germanic family, including varieties found outside Europe. Then in Section 4, I will focus on the German linguistic islands of Northern Italy. We will see that they display interesting and partially unprecedented developments within the Germanic family. In Section 5 the systematic distinction between the semasiological and the onomasiological level will be shown to shed light on some inconsistencies emerging in Stolz's approach, especially with regard to the diachronic perspective opened by grammaticalization. Section 6 draws the conclusion.

## 2. CCIs and the semasiological approach

One of the substantial merits of Stolz's (2013) impressive monograph on comparative constructions is the development of a clear conceptual apparatus, which allows us to investigate CCIs on safe methodological grounds. Accordingly, Stolz (2013: 9) assumes the following possible components of a constructional schema for CCIs:

(1) [The comparison]<sub>COMPAREE</sub> is [more]<sub>DEGREE</sub> [natural]<sub>QUALITY</sub> [than]<sub>TIE</sub> [the contrast]<sub>STANDARD</sub>

- the COMPAREE (= C) is the participant in a situation of comparison whose share of the QUALITY (= Q) is measured against the STANDARD (= S);

- the STANDARD is the participant in a situation of comparison which serves as yardstick for the measurement of the QUALITY with the COMPAREE;
- the QUALITY is the property in which the COMPAREE and the STANDARD partake;
- the DEGREE (= D) expresses the (in)difference of the shares the COMPAREE and the STANDARD have of the QUALITY;
- the TIE (= T) is the relation connecting the STANDARD to the COMPAREE and the QUALITY.

On this basis, we can construct what we can label as the semasiological format of a CCI, i.e. the sign-oriented set of its possible components. This has to be kept distinct from the onomasiological content, which points to the semantic domain to which the involved signs actually refer. This distinction is based on Geeraerts' (2010: 23) classical formulation (see also Glynn 2015 for further discussion):

[S]emasiology takes its starting point in the word as a form, and charts the meanings that the word can occur with; onomasiology takes its starting point in a concept, and investigates by which different expressions the concept can be designated, or named. Between the two, there is a difference of perspective: semasiology starts from the expression and looks at its meanings, onomasiology starts from the meaning and looks at the different expressions.

In Gaeta (2013), this basic distinction, traditionally applied to lexical entries, is extended to semasiological formats which are distinct from the onomasiological contents connected with them.

### ***2.1. The semasiological format of CCIs***

The semasiological format is not based on the postulation of a common semantic value – i.e., an onomasiological format defined a priori – but it rather generalizes over single formal components of a construction, in our case involving a comparative procedure. The latter constitutes a specific situation type – for convenience exemplified by the English example in (1) above – for which “the morphosyntactic construction(s) or strategies used to encode” (Croft 2003: 14) are investigated cross-linguistically:



Although categories (and constructions) are language-specific as morphosyntactic structures, categories and constructions may be compared across languages according to their function ... The formulation of cross-linguistic universals is in fact dependent on identifying categories and constructions across languages in terms of shared function (Croft 2001: 51).

Thus, the semasiological approach takes advantage of Croft's radical constructional procedure, which allows us to concretely identify language-specific constructions which display cross-linguistically a shared function. On the other hand, "the semasiological approach aims to provide a typology of the source constructions which give rise to [CCIs]. This typology allows us to reconstruct those cognitive processes of meaning extension and generalization which are at the heart of the genesis of grammar" (Gaeta 2013: 478-479). In this perspective, the typology opens a diachronic window on the possible sources of the morphemes recruited in the language-specific constructions via common processes of grammaticalization.

To illustrate the semasiological format, I will use two examples from Gothic, which at the same time show the two constructions that are likely to be postulated for Proto-Germanic:<sup>1</sup>

(2) Gothic (East-Germanic; Stolz 2013: 244, Harbert 2007: 174)

a. <i>unte þái</i>	<i>[sun-jos</i>	<i>þis</i>	<i>áiwis]</i> <sub>C</sub>	<i>[frod]</i> <sub>Q</sub> <i>[-oz]</i> <sub>D</sub> <i>-ans</i>
and then	son(M)-PL.NOM	this.GEN	time.GEN	wise-COMP-M.PL.NOM
<i>[sun[-um]</i> <sub>T</sub>	<i>liuhadis]</i> <sub>S</sub>	<i>in kunja</i>	<i>seinamma</i>	<i>sind</i>
son(M)-PL.DAT	light.GEN	in kind.DAT	their.DAT	are.3PL

---

<sup>1</sup> Besides the standard abbreviations, the translations provided in the glosses correspond to the general semantic content of the morphemes in the languages at stake. Accordingly, the Gothic morpheme *þau* is glossed as 'but' because this is its general value, although in this context it rather corresponds to the value of the English particle *than*. Moreover, I will make general reference to the BUT-particle, because this also corresponds to its etymological value. This latter need not coincide with the actual semantic content attested in a language. For instance, I will gloss the German morpheme *weder* as 'neither', although I will refer to it in terms of a WHETHER-particle, because the latter represents its etymological value. Finally, I will also use the caps to refer to the onomasiological domain of a sign occurring in a semasiological format. Accordingly, the Gothic particle pertains to the domain of CONTRAST.

‘For the children of this world are in their generation wiser than the children of light.’

- b. *frijondans*                      [*wiljan*      *seinana*]<sub>C</sub>              [*mais*]<sub>Q+D</sub> [*þau*]<sub>T</sub>  
 love.PRS.PTCP.M.PL.NOM    will(M).ACC    POSS3.M.ACC    more            but  
 [*gub*]<sub>s</sub>  
 God(N).ACC  
 ‘lovers of their will more than lovers of God.’

In both CCIs, the QUALITY and the DEGREE are expressed by means of the comparative form of an adjective. On the other hand, they differ in the strategy adopted for expressing the TIE: in the first case (2a), a synthetic coding is found consisting of a suffix for dative case on the STANDARD, while in the second case the TIE consists of a particle preceding the STANDARD.<sup>2</sup> As shown by the examples, this is an undeniable advantage of Stolz’s approach, which is fairly well accommodated into the semasiological format: the single ingredients may take different forms (affixes or analytic particles) encoding the same constructional role. Thus, the semasiological format results from the (language-)specific constructions concretely instantiating a certain general pattern which can be taken to correspond cross-linguistically to a similar situation type encoded via grammatical means, in our case the CCIs. The semasiological approach focuses on the signs entering the construction as well as on their general role within the particular language.

The difference observed in the two Gothic examples is likely to be due to the different internal structure of either CCI. In particular, in (2b) the difference is due to the particular morphosyntactic environment in which the STANDARD is placed, i.e. “the standard of comparison is always introduced by *þau* when the two things being compared are not (understood) subjects” (Harbert 2007: 174). For this reason, “the dative as TIE might render the construction difficult to parse and thus the disjunctive conjunction is an alternative solution” (Stolz 2013: 244). Thus, the synthetic coding of the TIE is dispreferred when it is likely to lead to syntactic opacity, i.e. when the

---

<sup>2</sup> An anonymous reviewer contends that the two examples are different because in the second one the QUALITY is incorporated into the particle expressing the DEGREE. While this is only partially true, because the form *mais* can be further segmented in a suppletive base *ma-* to which a comparative suffix *-iz-* is added (cf. Braune 2004: 125), this does not affect my point that the STANDARD is expressed in different ways in the two examples. A similar coding is theoretically conceivable and in fact attested in Old Icelandic (cf. Harbert 2007: 175).

case-marking blurs the syntactic coding of the STANDARD if the latter differs from the arguably default case, i.e. the nominative. In this case, the analytic construction is employed as a viable alternative recruiting a particle which is normally used with a disjunctive value. Thus, the rise of the analytic alternative is held to respond to a constraint of a different nature (parsing ease), while the particle is recruited on the basis of independent factors.

The independence of the formal aspects from their semantic content is well captured by the semasiological format and qualifies as a further advantage of this approach because it does not contain any reference to the range of possible semantic values, nor to the formal aspect (particle, clitic, affix, etc.) covered by the single components. In addition, not every piece of the semasiological format needs to be concretely realized in a language-specific construction. This is shown, for instance, by Wolof, in which the CCI lacks the overt expression of the TIE, and by Dhaasanac, where the DEGREE is absent or inferable from the TIE:

- (3) Wolof (Senegambian, Niger-Congo; Malherbe & Sall 1989: 37)

[suma kër]<sub>C</sub> moo [gën]<sub>D</sub> [mag]<sub>Q</sub> [sa kër]<sub>S</sub>  
 my house 3PS.SBJ.FOC more big your house  
 ‘My house is the one which is bigger than your house.’

- (4) Dhaasanac (Cushitic, Afro-Asiatic; Tosco 2001: 293)

[máa = l = a]<sub>C</sub> [ye]<sub>S</sub> [du]<sub>T</sub> [dér]<sub>Q</sub>  
 man = DEM = DET 1SG.OBL upon tall  
 ‘This man is taller than me.’

## 2.2. The onomasiological content of CCIs

The semasiological format is distinct from the conceptual level expressed by means of event schemas as maintained by Stolz (2013) in Heine’s (1997) shade. For instance, the Gothic particle used as TIE in (2b) above is held to pertain to the event schema of CONTRAST. Basically, Heine’s event schemas reflect what I label as onomasiological domains from which their sign-components – the semasiological formats – are recruited, that concretely encode a CCI. As demonstrated by Stolz (2013: 264), “[t]he choice of event schema is largely independent from that of the morpho-syntax of the construction type and vice versa”. On the other hand, this conclusion, which gives

support to a strict separation between the semasiological format and the onomasiological content, forces us to discard the confusing term ‘event schema’ adopted by Stolz.

A first reason for doing this is that in many cases no event *stricto sensu* is involved as shown by the Wolof and Dhaasanac examples, but rather a situation type in Croft’s sense mentioned above. In this sense, a given situation type, which involves one or more entities, provides the general accommodation for a detailed semantic representation referring to one specific onomasiological domain such as for instance CONTAINER or GIVE, in whose connection processes of naming are likely to take place. The onomasiological process of naming consists in adopting a certain onomasiological domain to encode a certain situation type. Accordingly, the onomasiological domain – which originally refers to a basic situation type – adopted for encoding a different situation type (in our case the CCI) carries along its original semasiological format, which is recruited for the new function via metaphorical processes of meaning generalization and semantic bleaching typical of grammaticalization processes.

A second reason for speaking of onomasiological domains rather than of event schemas is that they allow us to express their complementary role with regard to the semasiological formats, and in fact to account for their peculiar status, which can also have an impact on the latter forcing a certain concrete output. For instance, in Goemai the onomasiological domain pertaining to OUTDO does not involve any explicit TIE or QUALITY – the latter is implicit in the COMPAREE – as it involves a transitive verb encoding the DEGREE, while in Hausa the QUALITY is represented as a post-verbal adjunct:

- (5) Goemai (Chadic, Afro-Asiatic; Dixon 2012: 357)

[k<sup>l</sup>oom muk]<sub>C</sub> [ma]<sub>D</sub> [m-mak]<sub>S</sub>  
 strength 3SG.POSS surpass NMLZ-2SG.M.POSS  
 ‘He is stronger than you’, lit. ‘His strength surpasses yours.’

- (6) Hausa (Chadic, Afro-Asiatic; Dixon 2012: 356)

[Bàlaa yaa]<sub>C</sub> [fi]<sub>D</sub> [Muusaa]<sub>S</sub> [karfi]<sub>Q</sub>  
 Bala 3M.SG.COMPL exceed Musa strength  
 ‘Bala is stronger than Musa’, lit. ‘Bala exceeds Musa in strength.’

On the other hand, the same onomasiological domain can be expressed by means of different semasiological formats as in the case of the onomasiological domain pertaining to LOCATION in the following two examples, from Swahili and Malto:

(7) Swahili (Bantu, Niger-Congo; Heine 1997: 123)

- a. [Juma]<sub>C</sub> ni [m-refu]<sub>Q</sub> [ku-li-ko]<sub>T</sub> [Ali]<sub>S</sub>  
 Juma COP CLF1-long INF-be-LOC Ali  
 ‘Juma is taller than Ali’, lit. ‘Juma is long there being Ali.’
- b. [Juma]<sub>C</sub>ni [m-refu]<sub>Q</sub> [ku-shinda]<sub>D</sub> [Ali]<sub>S</sub>  
 Juma COP CLF1-long INF-defeat Ali  
 ‘Juma is taller than Ali’, lit. ‘Juma is long defeating Ali.’

(8) Malto (Dravidian; Stolz 2013: 16)

- [sardareh]<sub>C</sub> [majyen]<sub>S</sub>[-te]<sub>T</sub> [beḍoh]<sub>Q</sub>  
 sardar village.chief-LOC high  
 ‘A sardar is higher than a village-chief.’

In the example (7a) from Swahili I interpreted the verbal infinitive carrying a locative marker as TIE, paralleling in this way the Malto example in (8), where the locative marker is suffixed to the noun serving as STANDARD. In fact, *kuliko* is normally glossed as corresponding to (*more*) *than* (cf. Burt 1910: 191, Brauner & Bantu 1964: 124). Exactly the same structure is employed in Swahili in the example (7b) in connection with the onomasiological domain pertaining to OUTDO and actually shows that they “are constructed essentially in the same way” (Heine 1997: 123). This is because in its etymology *kuliko* “is the infinitive of a verbal stem *-liko* ‘to be at, to be present’” (Stassen 1985: 170), where *-ko* is a locative suffix. From this perspective the verbal infinitive of (7a) might also be interpreted as a DEGREE, paralleling the example in (7b).

### 2.3. Distinguishing the levels

The clear-cut distinction between a semasiological “skeleton” and its complementary onomasiological “flesh” helps us avoid a conceptual flaw which lurks in Stolz’s approach and has severe consequences on the whole picture. In fact, Stolz interprets

Heine's event schemas, i.e. in our terms: onomasiological domains, as prototypically connected with particular constructions, i.e. semasiological formats. For instance, with regard to the LOCATION schema Stolz (2013: 17) says that “[p]rototypically, the TIE is represented by spatial adpositions or spatial cases (e.g. locative, essive, inessive, superessive, etc.)”.

While this alignment of the semasiological format (adpositions or case-marking) and the onomasiological content pertaining to LOCATION might generally be true, it actually runs into troubles with what Stolz and Heine term PARTICLE COMPARATIVE in which “[t]he TIE marker is a particle, i.e., usually a free invariable monomorphemic element which does not govern (morphologically) the elements it combines with”. The crucial point is that in this type the particle “is etymologically non-transparent, or opaque” (Heine 1997: 120). In Heine's and Stolz's typology, this type simply flanks the other ones which are based on a precise onomasiological content. As is well known, the issue of opacity and of the relevance of the source domains for synchronic typological comparison – especially with regard to “the crucial problem of determining how far back in (pre-)history one has to look to determine the etymological origin of a given item” – is in fact “a general problem of grammaticalization research” (Stolz 2013: 23).

In my view, while it is true that the onomasiological domain is often captured only in etymological terms as it is synchronically blurred, this is not a problem of grammaticalization studies, but rather an advantage. In fact, as pointed out by Heine (1997: 111) “like other grammatical expressions, comparative markers tend to be derived from other, more concrete, entities”. The onomasiological content has to be understood as the range of possible meanings to which the source morphemes composing the semasiological format can be traced back in etymological terms. In this light, no zero option is admitted where the source morphemes are opaque. In principle, any ingredient of a semasiological format should be traced back to a source morpheme. Then, opacity can only be due to the limits of our knowledge, but cannot be immanent in a morpheme. For this reason, the usage of a PARTICLE COMPARATIVE on a par with the other onomasiological contents cannot be accepted, unless it is used as it is, namely as referring to a specific semasiological format, i.e. an analytic construction, which is based on an element recruited for serving as TIE. We will see in the next sections that the recruitment has not necessarily to be seen in terms of the direct grammaticalization of a certain morpheme pertaining to a given onomasiological domain.

Thus, in contrast to Stolz's mixed (and to a certain extent confusing) approach, it is more convenient to adopt a strictly semasiological representation which accounts for the whole range of constructions types which can give rise to CCIs (cf. Dixon 2012: 346, and Gaeta 2013: 483 for a view on existential constructions):<sup>3</sup>

<b>Mono-clausal CCIs</b>	
i.	ENT1 <sub>C</sub> (COP) PART <sub>D</sub> PROP <sub>Q</sub> PART <sub>T</sub> ENT2 <sub>S</sub>
ii.	ENT1 <sub>C</sub> PRED <sub>D</sub> ENT2 <sub>S</sub> PROP <sub>Q</sub>
<b>Bi-clausal CCIs</b>	
i.	ENT1 <sub>C</sub> (AND) ENT2 <sub>S</sub> (COP) / ENT1 <sub>C</sub> (COP) PROP <sub>Q</sub>
ii.	ENT1 <sub>C</sub> (COP) PROP <sub>Q</sub> / ENT2 <sub>S</sub> (COP) ¬PROP <sub>Q</sub>

**Table 1:** Semasiological formats for CCIs.

First, there are mono-clausal CCIs of the type seen above. A first type (i) of mono-clausal CCIs in Tab. 1 is basically encoded by means of a copulative construction (COP) in which the copula is present as in Gothic (cf. (2) above) or not as in Wolof (cf. (3) above), the subject refers to the COMPAREE (ENT1<sub>C</sub>), while the QUALITY is explicitly expressed by means of a specific morpheme (PROP<sub>Q</sub>). In this first type, the DEGREE and the TIE are encoded respectively via analytic particles (resp. PART<sub>D</sub> and PART<sub>T</sub>) as in Wolof (cf. (3) above) or via affixal elements attached to the STANDARD (ENT2<sub>S</sub>) as in Gothic (cf. (2a) above) or via a combination of the two as in Gothic (cf. (2b) above).<sup>4</sup> A second type (ii) of mono-clausal CCIs exploits a predicative construction in which the verb encodes the DEGREE as in Goemai (cf. (5) above) while the QUALITY is possibly represented by an adjunct as in Hausa (cf. (6) above).

<sup>3</sup> For brevity, no indications referring to word order in the CCIs are provided in the semasiological formats in the Tab. 1, although this is an important parameter of variation which should also be taken into consideration. This also means that the linearization of the abstract components given in the Tab. 1 for convenience does not exclude that the opposite orders are also possible and remains a matter for further investigation. Given the parasitic nature of CCIs with regard to other syntactic constructions (copula- or verb-centered, adposition- or complementizer-based, etc.), the null hypothesis is that their word order reflects that of their source constructions. At any rate, this deserves a specific investigation which cannot be undertaken here.

<sup>4</sup> In other words, the semasiological formats given in Tab. 1 as analytic constructions can be rephrased according to the specific morphosyntactic properties of a language, for instance by means of a suffixal representation as shown in (2a) above for Gothic: ENT1<sub>C</sub> COP PROP<sub>Q</sub>-SUFF<sub>D</sub> ENT2<sub>S</sub>-SUFF<sub>T</sub>.

In addition, in Tab. 1 the bi-clausal CCIs are also reported in which the comparative value emerges inferentially from the juxtaposition of two clauses. The first type (i) of bi-clausal CCIs in Tab. 1 has been called by Heine (1997: 120) TOPIC SCHEMA because the COMPAREE (ENT1<sub>C</sub>) and the STANDARD (ENT2<sub>S</sub>) are paired in a coordinated conjunction (AND) and serve as the topic for the subsequent clause which implicitly profiles the COMPAREE against the STANDARD on the basis of the QUALITY encoded by a specific morpheme (PROP<sub>Q</sub>) as in the following example from Nyanja:

(9) Nyanja (Bantu, Niger-Congo; Heine 1997: 120)

[*madzi*]<sub>S</sub> *ni* [*čakudia*]<sub>C</sub> [*komo*]<sub>Q</sub> [*čakudia*]<sub>C</sub>

Water and food good food

‘Food is better than water’, lit. ‘As for water and food, food is good.’

The second type (ii) of a bi-clausal CCI in Tab. 1 is termed POLARITY SCHEMA by Heine (1997: 117) because the second clause contains the STANDARD (ENT2<sub>S</sub>), which stands either in an antonymic or in a negative relation with regard to the QUALITY (¬PROP<sub>Q</sub>) predicated for the COMPAREE (ENT1<sub>C</sub>), as in the following examples respectively from Monumbo and Hixkaryana:

(10) Monumbo (Torricelli, Papuan; Stassen 1985: 185)

[*tsek*]<sub>C</sub> [*angam*]<sub>Q</sub> [*ek*]<sub>S</sub> [*put*]<sub>¬Q</sub>

you tall I short

‘You are taller than me.’

(11) Hixkaryana (Cariban; Stassen 1985: 186)

[*kaw-ohra*]<sub>¬Q</sub> *naha* [*Waraka*]<sub>S</sub> [*kaw*]<sub>Q</sub> *naha* [*Kaywerye*]<sub>C</sub>

tall-NEG 3M.SG.COP Waraka tall 3M.SG.COP Kaywerye

‘Kaywerye is taller than Waraka.’

With this general typology of possible semasiological formats for CCIs in mind, in the next section we will turn to the Germanic family, in which a peculiar type of CCI is actually attested.



### 3. CCIs in the Germanic family

The Germanic family consistently adopts for the CCI the type (i) of the mono-clausal semasiological formats in Tab. 1 above. Relevant points of variation concern the different onomasiological domains from which the particles expressing the TIE are recruited as well as the form in which the DEGREE and the TIE are encoded. In the Gothic example in (2b) the TIE is taken from the onomasiological domain of CONTRAST – which also parallels the Greek original – while in (2a) the dative case can be reconstructed as reflecting an old ablative and refers insofar to the domain pertaining to SOURCE. The latter used to be quite widespread across the old Indo-European languages settled in Europe and outside (cf. Stolz 2013: 278).

#### 3.1. CCIs in the modern Germanic languages

In the modern Germanic languages, case-marking for the TIE – in dative – is only found in Icelandic (12a) where it is flanked by a second possibility also mirroring the Gothic construction with the TIE expressed by a particle pertaining to the domain of CONTRAST (12b):

(12) Icelandic (West-Scandinavian, North-Germanic; Stolz 2017: 47, 57)

- a. [hún]<sub>C</sub> var [hver[-ri]<sub>T</sub> kon[-u]<sub>T</sub>]<sub>S</sub> [fríð]<sub>Q</sub>[-ari]<sub>D</sub>  
 she was each-DAT woman-DAT beautiful-COMP  
 'She was more beautiful than each woman.'
- b. [Harry]<sub>C</sub> var [fljót]<sub>Q</sub>[-ari]<sub>D</sub> [en]<sub>T</sub> [Higgs]<sub>S</sub>.  
 Harry was fast-COMP but Higgs  
 'Harry was faster than Higgs.'

Besides Icelandic, the North-Germanic languages, i.e. Danish, Faroese, Norwegian and Swedish, all converge in showing a source morpheme for the TIE belonging to the domain of CONTRAST, while the North-Sea Germanic group excluding Frisian, i.e. English and Dutch, recruits for the TIE a particle pertaining to the onomasiological domain of SEQUENCE as exemplified by Dutch *dan* 'then, than':

(13) Dutch (Low Franconian, West-Germanic)

[Harry]<sub>C</sub> was [snell]<sub>Q-er</sub><sub>D</sub> [dan]<sub>T</sub> [Hilarius]<sub>S</sub>.  
Harry was fast-COMP then Hilarius  
'Harry was faster than Hilarius.'

Moving towards the South, the rest of the West-Germanic family mostly recruits for the TIE particles pertaining to the SIMILARITY domain, coming in particular from two different source morphemes, namely Luxembourgish *wéi* 'how' (cf. German *wie* 'how') and *als* 'as, when' shown by the standard German example:

(14) Luxembourgish (Central Franconian, West-Germanic; Freimann et al. 2010: 18)

[China]<sub>C</sub> ass [méi]<sub>D</sub> [grouss]<sub>Q</sub> [(e)wéi]<sub>T</sub> [Lëtzebuerg]<sub>S</sub>.  
China is more big how Luxembourg  
'China is bigger than Luxembourg.'

(15) Standard German (High German, West-Germanic)

[Harry]<sub>C</sub> war [schnell]<sub>Q-er</sub><sub>D</sub> [als]<sub>T</sub> [Higgs]<sub>S</sub>.  
Harry was fast-COMP as Higgs  
'Harry was faster than Higgs.'

Note that in Limburgian a small transition area is observed, in which the THEN-type alternates with the HOW-type (cf. Stolz 2013: 49).<sup>5</sup> The two particles found on the German territory are distributed unevenly insofar as the AS-particle is mostly found – besides Standard German – in the North: West Frisian (Tiersma 1999: 47), North Saxon (Goltz & Walker 1990: 45), North Central Westphalian (Durrell 1990: 78), East Low German (Stavenhagen, Schönfeld 1990: 111), and in some varieties along the Rhine border, namely Alsatian (Colmar, Philipp & Bothorel-Witz 1990: 321), Central Swabian (Russ 1990: 351) and High Alemannic (in particular the varieties spoken in Bern and Zürich, Russ 1990: 373).

---

<sup>5</sup> Examples of the THEN-type are also found in Standard German in particular environments, namely in combination with the particle *je* 'ever' and to avoid the repetition of *als* (DUDEN: 372):

(i) [Online-Tauschbörsen]<sub>C</sub> sind [beliebt]<sub>Q[-er]</sub><sub>D</sub> [denn]<sub>T</sub> [je]<sub>S</sub>.  
online-swap.meet.PL are liked-COMP then ever  
'Online swap meets are more popular than ever.'

Note that the AS-particle coming from Standard German is reaching a wider distribution as shown by the city dialect of Zürich in (16a), at the expense of the older particle *wöder* also found in Zürich (16b), which corresponds to Standard German *weder* ‘neither’ (see Old High German (*h*)*wedar* ‘which of the two, whether’) and pertains to the domain of CONTRAST:

(16) Swiss German (Alemannic, West-Germanic; Reese 2007: 75, SI: s.v. *wöder*)

- a. *das äigentlich [di alerelteschten Uufname]<sub>C</sub> ... [fascht besser]<sub>Q+D</sub> sind*  
 that in.fact DET oldest recordings almost better are.3PL  
*[als]<sub>T</sub> [die, wo dän spòt-er ygschpilt woorde sind]<sub>S</sub>*  
 as DEM.PL REL then late-COMP record.PST.PTCP become.PST.PTCP are.3PL  
 ‘that in fact the oldest recordings ... were almost better than those which were taken later.’
- b. *[’s]<sub>C</sub> isch<sup>t</sup> nid [vil [größ]<sub>Q-er</sub>]<sub>D</sub> g’si<sup>n</sup> [wöder]<sub>T</sub> [dā, wo-n-i<sup>ch</sup> g’lège<sup>n</sup> bin]<sub>S</sub>*  
 it is not much big-COMP been neither there where-LE-1SG  
 posed am  
 ‘It is not much bigger than where I am placed.’

Besides Luxembourgish (13c), the HOW-particle is mainly found in central and south-eastern German varieties as well as in Austria, namely South Hessian (Durrell & Davies 1990: 231), Palatinate Franconian (Kaulbach, Green 1990: 252), Thuringian (Unterellen, Spangenberg 1990: 279), Upper Saxon (Friedersdorf, Bergmann 1990: 304), Bavarian (Wiesinger 1990: 489) and West Tyrolean (Wiesinger 1990: 508). Note that in Tyrolean the AS- and the HOW-particles are combined (cf. Stolz 2013: 49).<sup>6</sup>

- 
- (ii) *Heutzutage sind Mediziner [wenig]<sub>Q-er</sub>]<sub>D</sub> [als Heiler]<sub>C</sub> [denn]<sub>T</sub> [als Berater]<sub>S</sub>*  
 nowadays are.3PL physician.PL little-COMP as healer.PL then as advisor.PL  
*gefragt.*  
 ask.PST.PTCP

‘Nowadays physicians are in great demand less as healers than as advisors.’

<sup>6</sup> This possibility is deemed to be obsolete in German, but it is still found in authors like Thomas Mann (from *Der Zauberberg*) and in non-standard or colloquial registers (ii) (cf. DUDEN: 372):

- (i) *Es ist [hier]<sub>C</sub> [anders]<sub>Q</sub> [als]<sub>T1</sub> [wie]<sub>T2</sub> [zu Hause]<sub>S</sub>.*  
 It is here different as how to home  
 ‘Here it’s different from home.’

Besides the variation relating to the onomasiological domains which provide the source morphemes for the particle used as TIE, another issue concerns the synthetic / analytic expression of the DEGREE, which in the Gothic examples in (2) above used to be strictly synthetic by means of a suffix attached to the adjective encoding the QUALITY. This state-of-affairs is also likely to hold for Proto-Germanic (cf. Hopper 1975: 75, Lehmann 1975: Section 5.1.2). Recall that the Gothic example in (2b) also contains the adverb *mais* ‘more’ used for cases in which no explicit QUALITY carrying the DEGREE occurs in the morphosyntactic environment.

In the modern languages a general tendency towards the usage of an analytic particle for the DEGREE cognate with Gothic *mais* and preceding the QUALITY is observed, as shown for instance by the particle *méi* in the Luxembourgish example in (13b) above. However, this usage is not fully generalized nor uniformly distributed across the family. The picture emerging from Stolz (2013: 51-53) shows that within the Germanic family the suffixal marking of the DEGREE is still quite robust, although in several languages there is a more or less pronounced tendency for polysyllabic, morphologically complex and/or non-native adjectives to display the analytic particle for the DEGREE. At any rate, this tendency affects the Scandinavian as well as the North-Sea group of the Germanic family, leaving apart German and its varieties where the analytic expression of the TIE is mostly unknown (cf. Šticha 2011).<sup>7</sup> A true exception to this homogeneous picture found in the area covered by the German-speaking territory is provided by a variety for which language contact can be argued to play a major role, as reconstructed in details by Stolz (2013: 50), namely Luxembourgish with regard to French, where only analytic coding is found. On the other hand, “[n]one of the other Germanic varieties which are heavily exposed to

- 
- (ii) [Das]<sub>C</sub> *schmeckt doch* [viel besser]<sub>Q+D</sub> [als]<sub>T1</sub> [wie]<sub>T2</sub> [Sprudel ohne Geschmack]<sub>S</sub>.  
 this taste.3SG yet much better as how fizz without flavor  
 ‘This tastes much better than fizz without flavor.’

<sup>7</sup> Notice, however, that also in German the analytic particle is used with participles which have a “heavily verbal character” (i) or with adjectives whose synthetic comparative is “hard to form” as for instance with compounds (ii) (cf. Helbig & Buscha 1991: 307):

- (i) [Diese Straße]<sub>C</sub> ist [mehr]<sub>D</sub> [befahren]<sub>Q</sub> [als]<sub>T</sub> [die Nebenstraße]<sub>S</sub>.  
 this.F.SG street(F) is more congested as the.F.SG side.street(F)  
 ‘This street is more congested than the back road.’
- (ii) [Hans]<sub>C</sub> ist [mehr]<sub>D</sub> [be-mitleiden-s-wert]<sub>Q</sub> [als]<sub>T</sub> [du]<sub>S</sub>.  
 Hans is more PREF-commiserate-LE-worth as you.2SG  
 ‘Hans is more pitiful than you.’

influence from French attest to the remodeling of their [CCI] according to the Romance model” (Stolz 2013: 54). Thus, it is crucial that Luxembourgish lies at the margin of the German-speaking territory, and is therefore less exposed to normative influences than the varieties spoken in Germany.

### 3.2. *The role of contact: the case of Yiddish*

Language contact can also be made responsible for the use of a variety of particles for encoding the TIE which characterizes Yiddish, historically subject to intense contact with the Slavic languages. In Yiddish up to five different particles are found, with the addition of the possible use of *als*, considered however a Germanism (“daytshmerish”):

(17) Yiddish (High German, West-Germanic; Jacobs 2005: 183)

- a. [er]<sub>C</sub>            iz [rajx]<sub>Q</sub>[-ər]<sub>D</sub> [vi / ejdər / vidər]<sub>T</sub> [der  
3SG.M.NOM is rich-COMP how / before / again DET.M.SG.NOM  
man]<sub>S</sub>.  
man(M).SG  
'He is richer than the man.'
- b. [er]<sub>C</sub>            iz [rajx]<sub>Q</sub>[-ər]<sub>D</sub> [far / fun]<sub>T</sub> [dem            man]<sub>S</sub>.  
3SG.M.NOM is rich-COMP before / from DET.M.SG.DAT man(M).SG  
'He is richer than the man.'

In (17a), besides the particle *vi* ‘how’, which pertains to the SIMILARITY domain, the particles *ejdər* and *vidər* – which mean respectively ‘before’ and ‘again’ in Yiddish – pertain to the CONTRAST domain, because they are etymologically connected respectively with an EITHER- and a WHETHER-particle, although in the latter case a merge with the particle meaning ‘again’ (cf. respectively German *jeder* and *wieder*, Old High German *eogiwedar* and *widar*) has taken place, but see the form *jetvidər* ‘each, every’.

Furthermore, in (17b) we also find particles pertaining to GOAL – *far* ‘for, before’, see German *vor* ‘before’<sup>8</sup> – and to SOURCE – *fun* ‘from’, see German *von* ‘from, of’. In

<sup>8</sup> This is the account suggested by Stolz (2013: 56), although the reference to the domain of LOCATION might appear more appropriate. At any rate, I leave the issue open for further investigation.

particular, Stolz emphasizes the role of contact with Slavic for the usage of the particle *fun*, which corresponds straightforwardly to the use of ablative particles for the TIE in Polish and Macedonian (*od* ‘from’), Bulgarian (*ot* ‘from’), Ukrainian (*vid* ‘from’), etc. (cf. Stolz 2013: 65), and has no direct matching within the Germanic family. On the other hand, in spite of the rich attestation of analytic coding for the DEGREE found throughout the Slavic languages (cf. Stolz 2013: 177–179), the synthetic expression of the DEGREE is usually preserved in Yiddish.<sup>9</sup>

### 3.3. The role of contact: German varieties outside Europe

That contact can have a strong impact enhancing the general tendency towards an analytic expression of the DEGREE is also shown by cases of contact involving Germanic varieties outside Europe. A first example is Afrikaans, which continues and expands the tendency already present in Dutch of using the analytic particle with polysyllabic, complex (especially converted from participles) and non-native adjectives:

(18) Afrikaans (Low Franconian, West-Germanic; Donaldson 1993: 177-178)

[Sy]<sub>C</sub> is [(nog) [lang]<sub>Q</sub>[-er]<sub>D</sub> / [meer]<sub>D</sub> [tevrede]<sub>Q</sub> [as]<sub>T</sub> [ek]<sub>S</sub>.  
3F.SG.S is even long-COMP / more satisfied as 1SG.S  
‘She is (even) taller / more satisfied than I.’

Moreover, Afrikaans clearly departs from Dutch because the particle *as* is normally used, while “[d]an is a very formal synonym of *as* in this sense and if used at all, is used to avoid confusion with other *as*’s” (Donaldson 1993: 177).

Further peculiar examples are provided by two varieties exposed to strong contact, namely Pommersch or Pomeranian spoken in Brazil and Unserdeutsch or Rabaul Creole spoken in Papua New Guinea. Pommersch results from the migration of Lutheran settlers in the state of Espírito Santo in Brazil from Farther Pomerania (*Hinterpommern* or *Ostpommern*) around 1850. They spoke Ostpommersch, a variety of Low Saxon, and in this light we are not surprised to observe that the AS-particle is used for encoding the TIE:

---

<sup>9</sup> However, in the superlative an analytic construction is found involving the particle *samə*: “*di samə grojsə štot* = *di grestə štot* ‘the biggest city’” (Jacobs 2005: 183).

(19) Pommersch (Low Saxon, West-Germanic; Postma 2019: 210)

[*Kaie*]<sub>C</sub> *ka* [*beeter*]<sub>Q+D</sub> [*as*]<sub>T</sub> [*ik un mijn uldsch*]<sub>S</sub>.  
 no.one can better as 1SG.S and my wife  
 ‘No one is more apt than me and my wife.’

On the other hand, Unserdeutsch is the only German-based relexified creole of the world developed towards the end of the 19th century by children who usually spoke Tok Pisin (New Guinea Pidgin English) when they were hosted in the orphanage of the Vunapope Catholic Mission on the Gazelle Peninsula of New Britain (then called *Neu-Pommern*, New Pomerania). Given the presence of English in the speakers’ repertoire, we are not surprised to observe the use of the THEN-particle for encoding the TIE in (20a), besides the AS-particle typical of (Low) German in (20b):

(20) Unserdeutsch (West-Germanic; Maitz, Lindenfesler & Volker in press)

- a. [*mehr*]<sub>D</sub> [*dunkel*]<sub>Q</sub> [*than*]<sub>T</sub> [*me*]<sub>S</sub>  
 more dark then 1SG.ACC  
 ‘darker than me.’
- b. [*ganz mehr*]<sub>D</sub> [*jüing*]<sub>Q</sub>[-*er*]<sub>D</sub> [*als*]<sub>T</sub> [*i*]<sub>S</sub>  
 very more young-COMP as 1SG  
 ‘much younger than me.’

Note that for the DEGREE we observe analytic coding by means of a particle (20a) as well as the simultaneous combination of the synthetic and of the analytic construction (20b). The same examples are also found in Pommersch:

(21) Pommersch (Low Saxon, West-Germanic; Postma 2019: 94)

- a. *åwer wen* [*dai eir*]<sub>C</sub> [*meir*]<sub>D</sub> [*hard*]<sub>Q</sub> *is, ...*  
 but when DET earth more hard is  
 ‘but when the soil is harder, ...’
- b. *wen* [*dai farken*]<sub>C</sub> [*meir*]<sub>D</sub> [*gröt*]<sub>Q</sub>[-*er*]<sub>D</sub> *sin, ...*  
 when DET pig.PL more big-COMP are.3PL  
 ‘when the pigs are bigger, ...’

The cases of double marking are fairly well known in first and second language acquisition and in several non-standard varieties (cf. English examples like *more happier*) as well as in several other languages (typically with suppletive comparatives like non-standard Italian *più migliore* ‘more better’) and “suggest for the stepwise replacement of the synthetic degree-marking strategy by the analytic strategy” (Stolz 2013: 53). At any rate, contact seems to play a crucial role in this connection.

### 3.4. *Intermezzo: The overall typology of CCI in Germanic*

In sum, considering the semasiological elements (SemElem) forming the semasiological format and their onomasiological content (OnomCont) we obtain the following space of variation for CCIs in the actual Germanic family:

SemForm = ENT1 <sub>C</sub> COP PROP <sub>Q</sub> -SUFF <sub>D</sub> PART <sub>T</sub> ENT2 <sub>S</sub>		
SemElem	Germanic branch	OnomCont
TIE	†East Germanic	ENT2 <sub>S</sub> -T: SOURCE (DAT) PART <sub>T</sub> : CONTRAST (OR)
	North Germanic	ENT2 <sub>S</sub> -T: SOURCE (DAT) PART <sub>T</sub> : CONTRAST (BUT)
	North-Sea West Germanic	PART <sub>T</sub> : SEQUENCE (THEN)
	Continental West Germanic	PART <sub>T</sub> : SIMILARITY (AS, HOW), CONTRAST (WHETHER)
	Isolates: Yiddish	PART <sub>T</sub> : CONTRAST ( <i>ejdər, vidər</i> ), GOAL ( <i>far</i> ), SOURCE ( <i>fun</i> ), SIMILARITY ( <i>vi, as</i> )
	Afrikaans	PART <sub>T</sub> : SIMILARITY ( <i>as</i> )
	Pommersch Unserdeutsch	PART <sub>T</sub> : SIMILARITY ( <i>as</i> ) PART <sub>T</sub> : SEQUENCE ( <i>than</i> ), SIMILARITY ( <i>als</i> )
DEGREE	Luxembourgish	PART <sub>D</sub> : ( <i>méi</i> )
	North/North-Sea Germanic	PART <sub>D</sub> : (MORE) when PROP <sub>Q</sub> is [polysyllabic], [converted], [- native]
	Isolates: Afrikaans	PART <sub>D</sub> : ( <i>meer</i> ) when PROP <sub>Q</sub> is [polysyllabic], [converted], [- native]
	Pommersch Unserdeutsch	PART <sub>D</sub> : ( <i>mehr</i> ) (before PROP <sub>Q</sub> -SUFF <sub>D</sub> ) PART <sub>D</sub> : ( <i>meir</i> ) (before PROP <sub>Q</sub> -SUFF <sub>D</sub> )

Table 2: The overall typology of CCIs through the Germanic family.



As for the source particles for the TIE, there is a clear division of labor among the three groups within the Germanic family. Apart from cases like Icelandic and Zürich German where an older expression is being replaced by the newer particle, one single source domain is normally selected, namely CONTRAST in North-Germanic, SEQUENCE in North-Sea Germanic and – mostly – SIMILARITY in Continental Germanic.<sup>10</sup> In contrast to this, Yiddish, isolated from the German-speaking territory and exposed to massive contact, displays a certain richness of possible analytic strategies and – besides particles attested in the rest of the family – shows peculiar developments which are essentially different from what is observed in the rest of the family.

As for the encoding of the DEGREE, Luxembourgish stands alone in Continental Germanic as for the full adoption of the analytic strategy. In the other varieties, a similar trend towards the use of a particle with complex adjectives is observed, which is even more pronounced in varieties found outside Europe as shown by Afrikaans, Pommersch and Unserdeutsch. In Yiddish, on the other hand, this trend is not observed in spite of the analytic constructions displayed by the neighboring languages.

#### **4. German minorities in Italy**

Besides the South-Tyrolean region, in which German, including its varieties, is in fact the majority language, two distinct groups of language minorities are also present in Northern Italy, in the West and in the East, belonging respectively to the Alemannic and to the Bavarian branches of the German dialects. They share a similar origin, as they both result from Low Middle-Age migrations of settlers looking for better pasture and farming conditions.

##### ***4.1. CCI in the linguistic islands of North-Eastern Italy***

Several Bavarian enclaves survive in the north-eastern Italian territory, which are immersed as linguistic islands in a Romance-speaking environment, including

---

<sup>10</sup> However, it cannot be excluded a priori that also in these branches of the Germanic family other types are possibly attested besides those listed above and in Stolz's (2013) detailed investigation. More research is needed here in order to answer the question whether the contact with non-Germanic languages and the isolation from the West-Germanic territory really account for the rise of analytic particles, or whether it is rather the low codification of these varieties that provides the key to really understand the phenomenon.

varieties like Venetian, Ladin and Friulan. Cimbrian villages are found in the North-West of Veneto – especially in the provinces of Verona and Vicenza – and in the South of Trentino (cf. Bidese et al. 2005, Bidese 2008, 2010, Tyroller 2003), while Mòcheno is spoken in three towns of the Bersntol ‘Valley of the river Fersina’ in Trentino (cf. Bidese & Cognola 2013, Rowley 2010). Other sparse Bavarian enclaves are found in other villages of Friuli, namely Sappada, Sauris and Timau:



Figure 1: The Bavarian islands in Veneto, Trentino and Friuli.

These islands result from different migration waves which started around the year 1000, coming either from Germany or from Austria. Since they are placed in different environments and partially display different origins, their actual sociolinguistic condition is not homogeneous across the different villages and places.

As for the CCIs, synthetic coding is the most widespread way of encoding the DEGREE in Luserna Cimbrian (22) and in Mòcheno (23):

(22) Luserna Cimbrian (Bavarian, West-Germanic; WDS: 127)

*Soinante khalt, lege-mar à [a [sber]<sub>Q</sub>-ar]<sub>D</sub>-na franéla]<sub>C</sub>*  
 be.PRS.PTCP cold put.1SG-1SG.DAT on DET heavy-COMP-SG.ACC flanel  
*[alz]<sub>T</sub> [da bombasate]<sub>S</sub>.*  
 as DEM cotton

‘Being cold I put on a heavier sweater than this of cotton.’

(23) Mòcheno (Bavarian, West-Germanic; Rowley 2010: 123)

*[hait]<sub>C</sub> is [khelt]<sub>Q</sub>[-er]<sub>D</sub> [as]<sub>T</sub> [gester]<sub>S</sub>.*  
 today is cold-COMP as yesterday

‘Today is colder than yesterday.’

On the other hand, the usage of an analytic particle is also found, especially – but not exclusively – with polysyllabic and/or non-native adjectives (cf. Kranzmayer 1981: 259, Tyroller 2003: 150 and Rowley 2010: 123), in Mòcheno (24) as well as in Seven Communities Cimbrian (25) and in Thirteen Communities Cimbrian (26):

(24) Mòcheno (Bavarian, West-Germanic; Rowley 2010: 123)

*de hom [mear]<sub>D</sub> [naturalet]<sub>Q</sub> galòt [de plent]<sub>C</sub>*  
 DEM.PL have.3PL more natural let.PST.PTCP DET polenta  
 ‘They left the polenta more natural.’

(25) Seven Communities Cimbrian (Bavarian, West-Germanic; WDS: 127)

*Sodar benne 's machet khalt ich leghe mich au*  
 so.there when 3N.SG make.3SG cold 1SG.NOM put.1SG 1SG.ACC also  
*[an biüllana]<sub>C</sub> [meeront]<sub>D</sub> [sbear]<sub>Q</sub> [bon]<sub>T</sub> [doi bon bombaas]<sub>S</sub>.*  
 DET husk more heavy from DEM from cotton  
 ‘Because it’s cold, I put on a heavier sweater than this of cotton.’

(26) Thirteen Communities Cimbrian (Bavarian, West-Germanic; WDS: 127)

*Tort iz machat kalt, I leige-an [a majùn]<sub>C</sub>*  
 because 3N.SG make.3SG cold 1SG.NOM put.1SG.NOM-on DET sweater  
*[mearur]<sub>D</sub> [sbèr]<sub>Q</sub> [mun]<sub>T</sub> [daz 'un bombolje]<sub>S</sub>.*  
 More heavy when DEM from cotton  
 ‘Because it’s cold, I put on a heavier sweater than this of cotton.’

Besides the question of the analytic particle for the expression of the DEGREE, Mòcheno and Cimbrian, as well as the other Bavarian enclaves of North-Eastern Italy, are interesting also because they display a certain variety for the encoding of the TIE. In the Cimbrian villages of the Seven and of the Thirteen Communities, the particles *bon* (25) and *mun* (26) are respectively found which are etymological cognates of German *von* ‘from, of’ and *wenn* ‘when, if’ and pertain to the domain of SOURCE and SEQUENCE. Moreover, besides the AS-particle found in Luserna Cimbrian (22) and in Mòcheno (24), Cimbrian also displays the THEN-particle (27a) pertaining to SEQUENCE and usually attested in the North-Sea Germanic, as well as the particle *ödar* (27b)

pertaining to CONTRAST (see *weder* above) which corresponds to German *entweder* ‘either’ (see Swiss German *eiter* – SI, s.v. – from *eindeweder* ‘either’):<sup>11</sup>

(27) Cimbrian (Bavarian, West-Germanic; Stolz 2013: 54–55)

a. *s'ist* [*pessor*]<sub>Q+D</sub> [*an stuke marmelada bon bostanajen un proat*]<sub>C</sub> [*dan*]<sub>T</sub>  
 it = isbetter a piece marmalade from carrots and bread then  
 [*brioss un andere gaplettarach*]<sub>S</sub>.

brioche and other things

‘A slice of carrot-marmalade and bread is better than brioche or such like.’

b. [*beelz jaar iste*]<sub>C</sub> *gabest* [*mear*]<sub>D</sub> [*hungar*]<sub>Q</sub> [*ödar*]<sub>T</sub> [*hemmest*]<sub>S</sub>.  
 many years ago exist.PST more hunger either today

‘In the past, there was more hunger than today.’

In Mòcheno we also record the complex particle *abia* (28) resulting from the combination of the AS- and of the HOW-particle as observed for West Tyrolean above:

(28) Mòcheno (Bavarian, West-Germanic; Rowley 2010: 123)

[*der mai<sup>n</sup> hunt*]<sub>C</sub> *is* [*greas*]<sub>Q</sub>[-*er*]<sub>D</sub> [*abia*]<sub>T</sub> [*der dai<sup>n</sup>*]<sub>S</sub>.  
 DET 1SG.POSS dog is big-COMP how DET 2SG.POSS

‘My dog is bigger than yours.’

Similar examples of complex particles are also found in the Friulian enclaves of Sappada (29), Sauris (30) and Timau (31):

(29) Sappada (Bavarian, West-Germanic; WDS: 127)

*Bail's kòlt is, leigimer* [*ana dick*]<sub>Q</sub>[-*ar*]<sub>D</sub>-*a*  
 because = 3N.SG cold be.3SG put.1SG.NOM.1SG.DAT DET thick-COMP-SG.ACC

<sup>11</sup> This is likely to be the oldest particle in Cimbrian, because it is mentioned in Slaviero’s (1760) grammatical sketch: *Du pist reichor öder ich* ‘You are richer than me’. The account suggested by Stolz (2013: 55) in which *öder* as well as the Yiddish cognate *ejdar* seen above are held to correspond to German *ehar* ‘earlier, rather’ and accordingly to pertain to SEQUENCE does not stand the etymological reconstruction and has to be rejected. In this connection, it is noteworthy to observe that in the 19th century the particle *bedar* – corresponding to German *weder* ‘neither’ – is also found for encoding the TIE: *ear ist grözor bedar ich* ‘he is taller than me’ (cf. SWB, s.v. *bedar*).

*fanèlla*]<sub>C</sub> on [*a bi*]<sub>T</sub> [*dei pambullina*]<sub>S</sub>.  
 flanel on as how DEM cotton  
 ‘Because it’s cold, I put on a heavier sweater than this of cotton.’

(30) Sauris (Bavarian, West-Germanic; WDS: 127)

[*d*]<sub>C</sub> ’*ist* [*dikh*]<sub>Q</sub>[-*ar*]<sub>D</sub> [*assbie*]<sub>T</sub> [*de sele va pamböle*]<sub>S</sub>.  
 DEM = be.3SG thick-COMP as.how DET that from cotton  
 ‘This is heavier than that of cotton.’

(31) Timau (Bavarian, West-Germanic; WDS: 168)

[*Dar peton*]<sub>C</sub> *meik hakli sain, ovar nit* [*haklig*]<sub>Q</sub>[-*ar*]<sub>D</sub>  
 DET cement may frangible be.INF but not frangible-COMP  
 [*a bia*]<sub>T</sub> [*dar glos*]<sub>S</sub>!  
 as how DET glass  
 ‘The cement can be frangible, but not more frangible than the glass!’

Finally, in Mòcheno as well as in Luserna Cimbrian the particles *bos* (32) and *baz* (33) are respectively found, which correspond to German *was* ‘what’:

(32) Mòcheno (Bavarian, West-Germanic; WDS: 127)

’*S ist kòlt, alura I leig me u’* [*an*  
 3N.SG be.3SG cold then 1.SG.NOM put.1SG 1SG.ACC onDET  
 [*dick*]<sub>Q</sub>[-*er*]<sub>D</sub>-*en jack*]<sub>C</sub> [*bos*]<sub>T</sub> [*der sell va bombasch*]<sub>S</sub>.  
 thick-COMP-M.SG.ACC jacket(M) what DET that from cotton  
 ‘It’s cold, then I put on a heavier sweater than this of cotton.’

(33) Luserna Cimbrian (Bavarian, West-Germanic; WDS: 168)

[*Dar zemént*]<sub>C</sub> *möse lai prèchan, ma*  
 DET cement must.SUBJ.PRS.3SG PTC break.INF but  
*nètt* [*pell*]<sub>Q</sub>[-*ar*]<sub>D</sub> [*baz*]<sub>T</sub> [*’z glass*]<sub>S</sub>!  
 not soon-COMP what DET glass  
 ‘The cement might also break, but not easier than the glass!’

## 4.2 The Walser German area

The label ‘Walser German’ identifies a group of dialects belonging to the Highest Alemannic branch of Upper German, originally spoken in the most south-western province of Switzerland, the Wallis ‘the (Rhone) Valley’, whence Wal(li)ser. At the outset of the last millennium groups of settlers left the Wallis and migrated south- and eastwards in search of better conditions for life and founded villages on the higher segments of the alpine valleys characterized by a common architectural landscape, hallmarked by the Städl, the typical Walser house made of wood and stone (cf. Rizzi 1993). A number of villages were also founded on the south side of the Monte Rosa massif, in which they were in contact with the local Romance-speaking population for centuries without losing, however, their relations and contacts with the native homeland as well as with the southern regions of Germany. Nowadays, the Walser islands on the Italian territory, which are placed in Aosta Valley and Piedmont (see fig. 2), are losing their linguistic identity with the last speakers of the Walser German variety mostly using the other varieties of their repertoire, namely Piedmontese and Standard Italian as well as French and Franco-Provençal for the varieties spoken in Aosta Valley, although the process of language shift is not yet completed (cf. Dal Negro 2004).

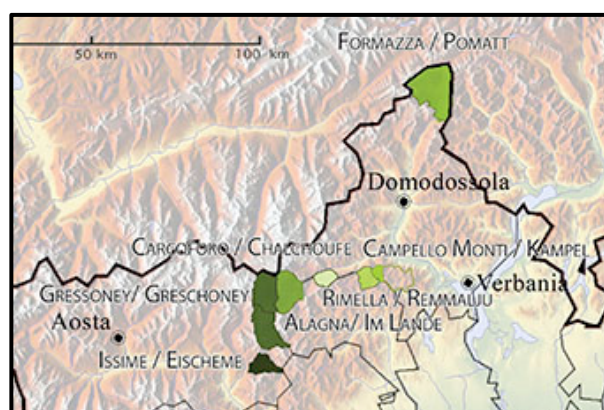


Figure 2: The Walser German islands in Piedmont and Aosta Valley.

In recent years, a number of projects were started to preserve this identity, to collect data on the Walser German varieties in order to make it available for future generations as well as for research. Thanks to these projects, the data presented in this section could be collected into a digital archive and carefully analyzed (cf. Angster et al. 2017, 2020, Gaeta in press for details).

Given their position at the southern edge of the Upper German area, Walser German dialects are traditionally known for their conservative character typical of

such marginal areas (cf. Bohnenberger 1913, Russ 1990: 367, and Eufe & Mader 2018 for a recent survey). This is for instance reflected in the retention of adjective agreement in predicative position as well as of distinct classes for weak verbs in neat contrast to all other dialectal varieties found in Germany and Switzerland (see Gaeta et al. 2019, Gaeta 2020 and Russ 1990: 383 for a survey). In this connection, however, one should not forget the role of language contact enhancing for instance the retention of final unstressed vowels which is a crucial factor for preserving those morphological traits (cf. Moulton 1941: 39, Zürcher 2011: 105).

#### 4.2.1. CCIs in Walser German

While CCIs are expectedly based on the general semasiological format typical of the Germanic family summarized in Tab. 2 above, we observe an astonishing variety of the particles used for encoding the TIE. As for the two villages of Aosta Valley, Gressoney, which lies across the Swiss border, displays the AS-particle belonging to the SIMILARITY domain (34) also found in German, while in the near village of Issime the THEN-particle belonging to the SEQUENCE domain occurs (35):<sup>12</sup>

(34) Gressoney (Alemannic, West-Germanic; DOK\_0441)

[*De Gnid*]<sub>C</sub> *éscht* [*schterch*]<sub>Q</sub>[-*or*]<sub>D</sub> [*als*]<sub>T</sub> [*ds Bedure*]<sub>S</sub>  
 DET envy is strong-COMP as DET compassion  
 ‘The envy is stronger than compassion.’

(35) Issime (Alemannic, West-Germanic; SW: 96)

[*d sunnu*]<sub>C</sub> *ischt gsing* [*schoarh*]<sub>Q</sub>[-*ur*]<sub>D</sub> [*den*]<sub>T</sub> [*is*]<sub>S</sub>  
 DET sun is been strong-COMP then 3M.SG  
 ‘The sun was stronger than him.’

Note that in Formazza, which is found in Piedmont on the western side of Monte Rosa, in spite of its distance from Issime the THEN-particle is found, too:

<sup>12</sup> The Walser examples indicated by DOK come from CLiMAIp, a digital archive which is freely accessible online at the website [www.climalp.org](http://www.climalp.org) (see Angster et al. 2017, 2020, Gaeta et al. 2019).

(36) Formazza (Alemannic, West-Germanic; SW: 94)

*Un äso der wén het erchent das [t sunna]<sub>C</sub>*  
 and thus DET wind(M) has admit.PST.PTCP that DET sun(F)  
*éscht [schterch]<sub>Q</sub>-er]<sub>D</sub> [de]<sub>T</sub> [är]<sub>S</sub>.*  
 is strong-COMP then 3M.SG.S  
 ‘And thus the wind admitted that the sun was stronger than itself.’

It must be added that – similarly to what we have seen above for the Zürich variety of High Alemannic – the AS-particle found in Gressoney should be treated as a Germanism modeled after the standard variety, because the older particle used for the TIE is *wan*, which corresponds to German *wann* ‘when’ and can be held to pertain to the SEQUENCE domain:

(37) Gressoney (Alemannic, West-Germanic; Zürrer 1982: 83)

*[d χue]<sub>C</sub> iš [gross]<sub>Q</sub>[-ur]<sub>D</sub> [wan]<sub>T</sub> [ds χalb]<sub>S</sub>.*  
 DET.F.SG cow(F) is big-COMP when DET.N.SG calf(N)  
 ‘The cow is bigger than the calf.’

In this connection, note that the other Walser island surrounded by Romance varieties but placed in the Swiss Ticino, namely Bosco Gurin, displays the same WHETHER-particle shown in (16b) above in Zürich as the older alternative with regard to the actual AS-particle:

(38) Bosco Gurin (Alemannic, West-Germanic; PALWaM: 41)

*wen-sch grian-s fress-en, éscht [dar chaašch]<sub>C</sub> [fell [galw]<sub>Q</sub>-ar]<sub>D</sub>*  
 when-3PL green-N.SG eat-3PL be.3SG DET.M.SG cheese(M) very yellow-COMP  
*[widar]<sub>T</sub> wenn [wenn-sch héww fress-en]<sub>S</sub>.*  
 neither when when-3PL hay eat-3PL  
 ‘When (the cows) eat green (grass), the cheese is much yellower than when they eat hay.’

In the other Walser villages of Piedmont other ways for encoding the TIE are found, which come from disparate source morphemes and are partially unprecedented in the Germanic family. In Alagna (39) and Macugnaga (40) the FROM-particle representative of the SOURCE domain is found:



## (39) Alagna (Alemannic, West-Germanic; SW: 93)

*Der schturmwind het miessä erchennä [di šchonna]<sub>C</sub> šchi*  
 DET storm.wind(M) has must.INF admit.INF DET sun(F) 3F.SG  
*šchige [schterch]<sub>Q</sub>[-ur]<sub>D</sub> [fam]<sub>T</sub> [šchi]<sub>S</sub>.*  
 be.SUBJ.3.SG strong-COMP from 3F.SG.S

‘The heavy gale had to admit that the sun was stronger than itself.’

## (40) Macugnaga (Alemannic, West-Germanic; SW: 96)

*Der gruos wind hetschi šchi móssó zeichu, [t sunna]<sub>C</sub>*  
 DET big wind(M) has.3F.SG 3F.SG must.INF show.INF DET sun(F)  
*hettschi si ksid [schterch]<sub>Q</sub>[-er]<sub>D</sub> [fan]<sub>T</sub> [ém]<sub>S</sub>.*  
 had.3F.SG be.INF been strong-COMP from 3M.SG.DAT

‘The big wind had to indicate that the sun had been stronger than itself.’

Paralleling the case of Yiddish seen in (17b) above with respect to Slavic, the particle *fan* (or *van*) ‘from, of’ – as well as the particle *bon* found in Seven Communities Cimbrian seen in (25) above – is likely to result from contact with the surrounding Romance varieties, in particular Italian *di*, Valdotain *de*, etc. (cf. Stolz 2013: 60).

## 5. Grammaticalization and the semasiological approach

In the other two Romance languages normally belonging to the Walser speakers’ repertoire, the particle used for expressing the TIE is “an element *que/ca* that, synchronically, has a wide range of functions beyond the [CCI] such that it can be understood as a desemanticized general subordinator” (Stolz 2013: 58). This is true both for French *que* and Piedmontese *che*, but in fact in other Romance varieties this particle is widely used as an – in several cases obligatory – alternative for encoding the TIE. For instance, in Italian the situation is quite complex – as also recognized by Stolz (2013: 142) – and reflects both the employment of *di* and of *che*, depending on the type of STANDARD:

## (41) Standard Italian (Romance)

a. [*Andare in bici*]<sub>C</sub> è [*più*]<sub>D</sub> [*faticos-o*]<sub>Q</sub> [*che*]<sub>T</sub> / \**[di]*<sub>T</sub> [*andare*  
 go.INF in bike is more strenuous-M.SG that of go.INF

*in moto*]<sub>S</sub>.

in motorcycle

‘Biking is more strenuous than riding a motorcycle.’

b. [La bici]<sub>C</sub> è [più]<sub>D</sub> [faticos-a]<sub>Q</sub> [della]<sub>T</sub> / (coll.) [che]<sub>T</sub> [la  
 DET bike(F) is more strenuous-F.SG of.DET that DET  
*moto*]<sub>S</sub>.

motorcycle

‘A bike is more strenuous than a motorcycle.’

Note that also in the case of a STANDARD consisting of a NP the alternative use of *che* is largely possible, although it might sound slightly colloquial to some speakers’ ears. Moreover, it should be added that also *di* serves a large variety of functions, while the SOURCE value is in fact only possible in sentences like *Teo è di Roma* ‘Teo is from Rome’ containing a place name which identifies the birth place of the subject, but does not have a true ablative value: *Teo viene da / \*di Roma* ‘Teo comes from Rome’. On this basis, one might argue that also *di* serves as a general desemanticized subordinator although it introduces a different class of subordinated elements with regard to *che*.

This brief discussion shows that, while the question of the desemanticized general subordinator is marginal for the Germanic family, it raises an important theoretical question which lies behind Stolz’s (2013: 58) general conclusion that “[m]ost probably, constructions with *que/ca* are the closest one can get to the ideal form of a PARTICLE COMPARATIVE”. As already hinted at in Section 2.3 above, here is where the semasiological format is mixed with the onomasiological content. Recall that on the basis of such a desemanticized subordinator, Stolz identifies an autonomous type of PARTICLE COMPARATIVE for encoding the TIE on a par with the other types which are strictly connected with a specific onomasiological content referring to a basic cognitive domain. But in fact, the alleged onomasiological content which can be associated with the autonomous type of PARTICLE COMPARATIVE in the European languages basically coincides with the *que/ca* particle found in the Romance languages.<sup>13</sup> Besides a certain circularity in the reasoning, the philosophy adopted in this paper cannot share Stolz’s conclusion because the semasiological approach crucially relies on a principled distinction between the form and the content of CCIs, and no ‘ideal comparative particle’ can be envisaged.

<sup>13</sup> In Stolz’s (2013: 84) sample 22 of the 32 languages belonging to the alleged type of PARTICLE COMPARATIVE belong to the Romance family which on the whole features 44 languages.

Instead, we adopt the vantage point provided by grammaticalization, which allows us to build a diachronic bridge between the conventionalized forms observed in a language and their onomasiological source domain. This does not mean, however, that for any single semasiological format a certain onomasiological source domain can be directly identified. In fact, particles recruited for encoding parts of a CCI can result from processes which are not immediately connected to grammaticalization channels of the type suggested by Heine (1997), i.e. via gradual semantic bleaching from a well-defined cognitive domain. In fact, they may also result from the generalization of certain morphemes already grammaticalized in a given language for certain functions. The functional motivation is similar to what we have already observed for the Gothic example in (2b) above in which the particle allows to encode as STANDARD any possible syntactic configuration (NPs, subordinate clauses, etc.). Especially when the latter is complex, languages can resort to employ general subordinators in order to overcome the possible structural difficulty and the resulting syntactic opacity. This is especially the case in languages where such general subordinators are widespread in a whole range of syntactic contexts like in the Romance languages. The Italian examples seen in (41) above illustrate pretty well this state-of-affairs. In this light, it is misleading to adopt a specific type called PARTICLE COMPARATIVE on a par with the other onomasiological domains, because the former identifies cases which result from a different diachronic mechanism than the latter ones.

Moreover, among the examples included by Stolz under PARTICLE COMPARATIVE we should distinguish cases where the original meaning of the particle is “irretrievable synchronically”, as maintained by Stolz (2013: 80), from cases where we have really to do with the probable generalization of a desemanticized subordinator. The Romance languages provide a good example of this second case, to which the Albanian particle *se* can be added as the latter displays a rather wide range of functions and overlaps with that of English *that, than, because, since, unless*, etc. (Stolz 2013: 80). Their usage for encoding the TIE unveils a different diachronic mechanism of generalization. For this reason, they might be called generalized subordinators (= SUB) in order to express the neat contrast to the other particles which are related to a full-fledged onomasiological domain via a process of grammaticalization. Notice that the etymology of these SUB-particles is not obscure at all as they go back respectively to the Latin pronoun *quid* and to the Proto-Albanian pronoun *\*tšīā* (cf. Orel 1998, s.v. *se*), both going back ultimately to Proto-Indo-European *\*k<sup>u</sup>id*.

The case of SUBS where the particle displays a full-fledged variety of functions has to be kept fully distinct from examples where no clear etymological source can be identified (yet). These latter examples are better treated with caution, also because further research might provide the correct account in the future. For instance, in Breton two different particles are used for encoding the TIE, namely  *eget*  and  *evit* , with the former giving ground to the latter in colloquial registers. While this latter displays a large variety of usages, including the introduction of the benefactive or of the topic role as well as of concessive and final clauses,  *eget*  is apparently limited only to CCIs. Accordingly, Stolz (2013: 81) assigns  *evit*  to the onomasiological domain of GOAL, while  *eget*  is assigned to the type of PARTICLE COMPARATIVE in the light of its etymological opacity. But this example is clearly different from the Albanian and Romance cases discussed above and cannot be considered a SUB-particle. Rather, we have to conclude that  *eget*  defies a precise categorization and requires more research in the future.<sup>14</sup> Thus, the type SUB assumed here does not qualify as a sort of  *Restklasse* , but points to an important channel for recruiting morphemes used as TIE, which is of a different nature with regard to grammaticalization.

### 5.1. The SUB-particle as an alternative to grammaticalization

The relevance of this brief discussion becomes tangible when we consider data coming from another Germanic variety exposed to a long-standing contact. In particular, in Pennsylvania German, which results from the migration from the 17th to the 19th centuries of German settlers coming from the Upper Rhine valley and speaking a variety of Central Franconian, both the HOW-particle (42a) and the AS-particle are found (42b):

(42) Pennsylvania German (Palatinate Franconian, West-Germanic; Haldeman 1872: 36, 54)

a. [Dær        mann]<sub>C</sub> iss [krank]<sub>Q</sub>[-ər]<sub>D</sub> [wie]<sub>T</sub> [d'r        annər]<sub>S</sub>.  
       DEM.M.SG man    is sick-COMP        how    DET.M.SG other

---

<sup>14</sup> One can tentatively group  *eget*  with other Breton particles like  *nemet*  ‘except’,  *estreget*  ‘other than’, etc. (cf. Press 1986: 117) in a set which reminds us of the EITHER- and WHETHER-particles seen above in several Germanic varieties and is likely to pertain to CONTRAST. At any rate, this has to be left open for further research.

‘This man is sicker than the other.’

- b. *Wii kummt əs, dass dii jung-i buuwə [selli meed,*  
 how come.3SG 3N.SG that DET.PL young-PL boy.PL that.PL girl.PL  
*woo reichi, daadis hen]<sub>C</sub>, [liiw]<sub>Q</sub>[-ər]<sub>D</sub> noochschpringə [als]<sub>T</sub>[dii*  
 REL rich.PL dad.PL have.3PL dear-COMP after.jump.INF as DET.PL  
*aarmi]<sub>S</sub>?*  
 poor.PL

‘How comes it that the young men sooner run after those girls who have rich fathers than the poor ones?’

On the other hand, we also find the generalized usage of a SUB-particle which corresponds to the German general subordinator *dass* ‘that’ for encoding the TIE in the construction of equality (43a) as well as in the CCI (43b):

(43) Pennsylvania German (Palatinate Franconian, West-Germanic; Haldeman 1872: 38, 42)

- a. *des land is aw [frei]<sub>Q</sub> [for mich]<sub>C</sub> [so goet]<sub>D</sub> [das]<sub>T</sub> [for dich]<sub>S</sub>,*  
 DET land is also free for 1SG.ACC so good that for 2SG.ACC  
 ‘This country is also free for me as well as for thee.’
- b. *[wass]<sub>C</sub> is [schenn]<sub>Q</sub>[-ər]<sub>D</sub> uf dər welt [dass]<sub>T</sub> [blimlin, root un*  
 what is beautiful-COMP on DET world that flower.DIM red and  
*weiss]<sub>S</sub>?*  
 white

‘What is finer in the world than flowerets, red and white?’

The contact situation is likely to have favored the expansion of the SUB-particle – which is unprecedented in the Germanic family for encoding the TIE – to the expense of the canonical particles found in the original Rhenish varieties. Thus, far from concluding that “the PARTICLE COMPARATIVE is mostly a Romance phenomenon with the occasional parallel in several other phyla” as maintained by Stolz (2013: 87), the development of SUB-particles illustrates an important diachronic mechanism of generalization of multifunctional particles which can be recruited for serving as TIE, namely as an analytic marker for introducing complex syntactic structures employed as STANDARD. A similar account can also be suggested for the usage of the WHAT-

particles in Mòcheno and Cimbrian seen in (32) above, because they are a calque based on the Italian multifunctional particle *che* which is used – besides as a TIE, see (41) above – also as interrogative pronoun, similarly to German *was*. Thus, both Mòcheno *bos* and Cimbrian *baz* are good examples of SUB-particles. On the same track, one can also interpret the particles *fan* and *bon* found respectively in Alagna/Macugnaga Walser German and Seven Communities Cimbrian as reflecting SUB-particles calqued on the Italian multifunctional particle *di*.

Finally, this distinction leads us to our last example drawn from the Walser communities, namely the particle *ŝchu* [ʒu] for the TIE used in the Piedmontese village of Rimella, which stands alone throughout the whole Germanic family:

(44) Rimella (Alemannic, West-Germanic; SW: 97)

*Un der chalte vend het messu erchannju das [d*  
 and DET.M.SG cold.M.SG wind(M) has must.INF admit.INF that DET.F.SG  
*ŝhunna]<sub>C</sub> isch gŝchid [mis]<sub>D</sub> [ŝtarch-e]<sub>Q</sub> [ŝchu]<sub>T</sub> [ier]<sub>S</sub>.*  
 sun(F) is be.PST.PTCP more strong-F.SG so 3M.SG  
 ‘And the cold wind had to admit that the sun was stronger than itself.’

The particle *ŝchu* corresponds to the German adverb *so* ‘so’ and can be related to the onomasiological domain pertaining to SEQUENCE or SIMILARITY. Note that *ŝchu* is also employed for introducing the protasis of a conditional sentence (45a), a concessive sentence in combination with another conjunction (45b) and an interrogative sentence (45c):

(45) Rimella (Alemannic, West-Germanic; WDS: 39, 55, 97)

- a. *Ŝchu ŝchei wistet nid ŝchiéh-e, der Dŝchwànd*  
 so 3F.SG be.SUBJ.PST.3SG not sick-F.SG DET.M.SG John  
*tiéttet schpilju bet ŝchi wattà.*  
 do.SUBJ.PST.3SG play with his.F.SG sister(F)  
 ‘If she were not sick John would play with his sister.’
- b. *Tiög z wasschu d tallerà vàm dum Luis,*  
 do.IMP to wash.INF DET.PL dish.PL of DET.M.SG.DAT Luis  
*öich ŝchu hét dschà gwascht gaschter.*  
 also so have.3SG already wash.PST.PTCP yesterday

‘Let Luis wash the dishes, even if he has already washed them yesterday.’

- c. *Pì nid šheccher šchu isch en donder: matte šchi e ruvenu.*  
 be.1SG not sure so be.3SG DET thundermight be.INF DET landslide  
 ‘I am not sure whether it is a thunder: it might be a landslide.’

The usage of *so* in these three contexts is already found in Middle High German where it competes with other possible subordinators.<sup>15</sup> From this point of view, the variety of usages observed in Rimella might also speak in favor of an analysis in terms of a general subordinator, in which *šchu* is extended as a SUB-particle to the role of a TIE. This interpretation in terms of a SUB-particle might also be further supported by the parallel range of usages shared by *šchu* with the Italian multifunctional particle *se* ‘if, whether’, although they are not etymologically connected to each other nor does the Italian particle serve as TIE. Such an influence can be held to play a role on the vitality of this generalized usage of *šchu* because it is also found in other Walser German varieties as well, for instance in Gressoney, where the protasis (46a), the concessive (46b) and the interrogative value (46c) of *so* are also found:

(46) Gressoney (Alemannic, West-Germanic; DOK\_0088, DOK\_0002, DOK\_0013)

- a. *etza kammo desche ässe so eschmo*  
 now can.one this.PL eat.INF so be.3SG.one

<sup>15</sup> Besides a temporal value (i) corresponding to German *als*, we also record a modal similitive value (ii) corresponding to German *wie*, and a conditional value (iii) corresponding to German *wenn*:

Middle High German (West-Germanic; Paul 2007: 415, 425, 415)

- (i) *sô si gedâht’ an Helchen, daz tet ir inneclîche wê*  
 so 3SG.F.NOM think.PST.3SG at Helchen, this do.PST.3SG 3SG.F.DAT internal.F.SG pain  
 ‘When she thought of Helchen, this hurt her innerly.’
- (ii) *jâ huoten si ir êren, sô noch die liute tuont*  
 yes protect.3PL 3PL.NOM POSS.3 honor.ACC so still DET people do.3PL  
 ‘Yet they defend their honor, like people still do.’
- (iii) *dû kindest al der werlte fröide mêren, sô dû ez*  
 2SG.NOM can.SUBJ.PST.2SG all DET.F.SG.DAT world(F) joy increase.INF so 2SG.NOM 3N.SG  
*ze guoten dîngen woltes kêren*  
 to good.N.PL.DAT thing(N).PL.DAT want.SUBJ.PST.2SG return.INF  
 ‘You could increase the joy in the whole world, if you would turn it into a good thing.’

*enema*                      *Jägerchs-Hus*                      *engladenz*  
 DET.N.SG.DAT    hunter.GEN-house(N)    invite.PST.PTCP.N.SG

‘Nowadays one can eat these (things) when one is invited into one hunter’s house.’

b. *Fer d’oalto*                      *litté,*                      *ou*                      *sò sinn*                      *fell*                      *joar*                      *vorbi kanget,*  
 for    DET=old.PL    people    also    so be.3PL    much.PL    year.PL    over    go.PST.PTCP  
*éscht*                      *das*                      *no*                      *ni*                      *ònder*                      *d’erennròng*                      *uskanget*  
 BE.3SG    that    yet    not    under    DET=memory    go.out.PST.PTCP

‘For the old people, even if many years have passed, this has not yet gone out of their memory.’

c. *Hein*                      *éntsich*                      *gfregt*                      *so*                      *hätteber*                      *chònnò*  
 have.3PL    1PL.ACC    ask.PST.PTCP    so    have.SBJ.PST.1PL    can.INF  
*eppés*                      *séege*                      *vòn*                      *Greschòney*  
 INDEF    say.INF    of    Gressoney

‘They asked us whether we could tell something about Gressoney.’

Thus, while the usage of *schu* as particle for the TIE is only found in Rimella, its multifunctionality might also be interpreted in terms of a SUB-particle, whose generalized use found also outside Rimella is likely to have been favored by contact.

## 5.2 Analytic coding and morphological complexity

The variety of Rimella is also peculiar because it expresses the DEGREE by means of an analytic particle as shown by the example (44) above, which does not normally occur in the other Walser German varieties, although sporadic exceptions are found, for instance in the following example from Alagna:

(47) Alagna (Alemannic, West-Germanic; WDS: 127)

*Denn*                      *erfriärd*                      *lekki*                      *mich*                      [*as*                      *triku*]<sub>C</sub>                      [*mei*]<sub>D</sub>  
 because    freeze.3SG    put.1SG    1SG.ACC    DET.N.SG    shirt(N) more  
 [*schweir-s*]<sub>Q</sub>    [*van*]<sub>T</sub>                      [*d’*                      *bowolins*]<sub>S</sub>.  
 heavy-N.SG    of                      DEM    cotton.GEN

‘Because it’s cold, I will put on a heavier sweater than this of cotton.’

However, the synthetic encoding of the DEGREE generally remains quite stable throughout the Walser German islands, while only Rimella clearly testifies of a



reduction of morphological complexity in favor of analytic coding. As we have seen in Section 4.1 above, this holds partially true for the Bavarian linguistic islands found in the North-East of Italy, with remarkable exceptions found in Mòcheno (24) as well as in Seven Communities Cimbrian (25) and in Thirteen Communities Cimbrian (26).

On the other hand, the reduction of morphological complexity is also observed with regard to the inflectional properties of the adjectives used for expressing the QUALITY. In this regard, in Section 4.2 I observed that Walser German varieties – in neat contrast to most Germanic varieties – preserve the subject agreement of the adjectives in the predicative position, besides the attributive position found in German and its varieties (cf. Fleischer 2007, Gaeta 2018, 2020). This is shown by the following examples from Gressoney:

(48) Gressoney (Alemannic, West-Germanic; DOK\_0348, DOK\_0192)

a. *d'gròss-ò      lougò      ésch      gwäschn-e      kanget*  
 DET = big-F.SG    laundry(F)    be.3SG    wash.PST.PTCP-F.SG    go.PST.PTCP  
 'The big laundry has been washed.'

b. *D'schuelstòbo      éscht      gròss-e      gsid      mé      drie      fäntschtre*  
 DET = school.room(F)    be.3SG    big-F.SG    be.PST.PTCP    with    three    window.PL  
*òn    en    steinenen    ofe*  
 and    DET    stony.M.SG    oven(M)  
 'The classroom was big, with three windows and one stone stove.'

However, the adjective agreement in the comparative is only found in the attributive position (49a), while in the predicative position the uninflected form is found (49b), as in the other Walser German varieties seen in (34), (38) and (39) above:

(49) Gressoney (Alemannic, West-Germanic; DOK\_0424, DOK\_0296)

a. *Z'gèbiet    vòn    den    Éndre    hät    so    mitte    no    gròss-or-é*  
 DET = area of    DET    Éndre    have.3SG    so    still    big-COMP-F.SG  
*wéerdé    kriegt    fer    alpinismus    òn    skisport*  
 value(F)    get.PST.PTCP    for    alpinism    and    ski.sport  
 'The area of the Éndre has acquired in this way an even bigger value for alpinism and skiing.'

- b. *chant d'flammò en bétz gròss-òr si*  
can.3SG DET=flame(F) a little big-COMP be.INF  
'The flame can be a little bigger.'

Thus, when it occurs in the more complex predicative position the comparative form of the adjective follows the trend observed in German and in the other Germanic varieties, which consists in reducing the morphological complexity of inflection. Note in this case the contrast with the surrounding Romance varieties where adjective agreement is well preserved – see the Italian examples in (41) above, where the comparative form relies on an analytic particle. That the analytic construction can have an effect on the adjective agreement is shown by the cases of Rimella and Alagna in (44a) and (47) above where analytic comparatives are found which display agreement, similarly to the Italian examples. In this light, it is straightforward to conclude that the morphological complexity of the synthetic comparative militates against the occurrence of agreement in the more complex predicative position with regard to what happens in the attributive position.

### **5.3 A diachronic outlook**

The impressive variety found in the German villages of Northern Italy substantially enriches Stolz's (2013) picture and is arguably due to the complex contact situation in which any direct connection with the German-speaking home country was substantially interrupted in the last 150 years. Thus, we could identify the influence of the German standard variety only in sporadic cases – namely in Gressoney, where direct contacts with the German-speaking territory are well attested also after Italy's unification (cf. Zürrer 2009). Note that this richness also characterizes other contact-involving varieties, from Pennsylvania German to Yiddish.

In this light, it is interesting to observe that the variety found in Continental Germanic – and preserved if not further expanded in the isolated varieties – closely mirrors the manifold options which are witnessed throughout its linguistic history. While for the other two branches of Germanic the diachronic development is linear and basically testifies of the diffusion of the analytic particles already present as an alternative to dative case-marking for encoding the TIE in the older stages, namely THAN and BUT respectively for North-Sea Germanic and North-Germanic, this was not

the case for Continental Germanic. The initial Old High German stage paralleled the corresponding THAN-particle for the TIE found in the rest of the West-Germanic branch:

(50) Old High German (West-Germanic; Schrodtt 2004: 155)

[*thu*]<sub>C</sub>    *mo*            [*liab*]<sub>Q</sub>[-*ar*]<sub>D</sub>-*a*    *bist*    [*thanne*]<sub>T</sub>    [*al gifugiles*]<sub>S</sub>  
 2SG.NOM    3M.SG.DAT    dear-COMP-M.SG    be.2SG    than            all    fowl.GEN  
 ‘He likes you more than all fowl.’

Thereafter, the range of particles used for the TIE increased dramatically in dependence on grammatical (e.g., the type of STANDARD) and extra-grammatical (among others: diatopic) factors. In Middle High German, besides *danne* we record also *wan*:

(51) Middle High German (West-Germanic; DWB, s.v. *wann1*)

[*daʒ*]<sub>C</sub>    *ist*            [*beʒzer*]<sub>Q+D</sub>    [*wan*]<sub>T</sub>    [*aller creatûren werc*]<sub>S</sub>  
 this.N.SG    be.3SG    better            when            all.GEN.PL    creature.PL    work  
 ‘This is better than every creature’s work.’

Later – from the second half of the 16th century on – *als* (52a) is firstly found, subsequently *weder* (52b), and *wie* (52c), also in the combination *als wie* (52d):

(52) Early New High German (West-Germanic; Ebert et al. 1993: 480, DWB, s.v. *weder, wie*)

a. *dz*            *es*            *vnmoglich das*    [*er*]<sub>C</sub>    [*hoch*]<sub>Q</sub>[-*er*]<sub>D</sub>    *ader*    [*mehr*]<sub>D</sub>  
 this.N.SG    be.3SG    impossiblethat    3M.SG    high-COMP    or    more  
*moge*                    *geheilget*                    *werdē ...*            [*als*]<sub>T</sub>    [*er*    *gerait*  
 may.SUBJ.PRS.3SG    sanctify.PST.PTCP    become.INF    as            3M.SG    already  
*geheilget*                    *ist*]<sub>S</sub>  
 sanctify.PST.PTCP    be.3SG

‘It is impossible that he might be sanctified in a more and more elevated way than he is already sanctified.’

b. *darumb*    *das*    [*es*]<sub>C</sub>    [*wolgeschmack*]<sub>Q</sub>[-*ter*]<sub>D</sub>  
 therefore    that    3N.SG    well.taste.PST.PTCP-COMP

were [weder]<sub>T</sub> [ander fleisch]<sub>S</sub>  
 be.SBJ.PST.3SG neither other meat(N)

‘For the reason that it was more tasteful than other meat.’

c. [mer]<sub>D</sub> [daran verbrechen]<sub>C</sub> [wie]<sub>T</sub> [gutt machen]<sub>S</sub>  
 more therein break.INF how good make.INF

‘to commit a crime therein more than to do good.’

d. es kan [keiner]<sub>C</sub> [frömm]<sub>Q</sub>[-er]<sub>D</sub> seyn,  
 3N.SG can.3SG none(M) pious-COMP be.INF

[als wie]<sub>T</sub> [es jhme gott zugemessen]<sub>S</sub>  
 as how 3N.SG 3M.SG God allot.PST.PTCP

‘Nobody can be more pious than God has allotted him to be.’

This diversity appears to be only partially reflected in the actual situation found in the German-speaking territory where the AS-/HOW-divide is still observed, while it clearly strikes the observer when linguistic islands – as well as other contact-involving varieties – are considered.

## 6. Conclusion

The semasiological approach adopted in this paper has proved substantially useful in characterizing the general format of the CCI in the Germanic family and in delimiting the possible onomasiological domains filling the format. As for the semasiological side, we could pinpoint two loci of variation, namely the analytic coding of the TIE and of the DEGREE. Especially the TIE qualifies as the main point of variation with regard to the range of possible onomasiological domains providing the source morphemes. The latter are well-distributed across the main branches of the Germanic family in a rather consistent way. A remarkable exception is constituted by Continental Germanic which deviates from this neat picture because it offers a certain number of possible alternatives for the TIE, which is even larger when varieties exposed to language contact are taken into consideration. Note that in the latter case we could also identify cases of the TIE belonging to the type SUB, i.e. resulting from the generalization of multifunctional subordinators.

We can summarize the types collected in the German minorities of Italy as follows:

SemForm = ENT1 <sub>c</sub> COP PROP <sub>Q</sub> -SUFF <sub>D</sub> PART <sub>T</sub> ENT2 <sub>s</sub>			
SemElem	Dialect branch	OnomCont	
TIE	Bavarian:	Cimbrian	PART <sub>T</sub> : SOURCE/SUB ( <i>bon</i> ), SEQUENCE ( <i>mun</i> , <i>dan</i> ), SIMILARITY ( <i>alz</i> ), CONTRAST ( <i>ödar</i> ), SUB ( <i>baz</i> )
		Mòcheno	PART <sub>T</sub> : SIMILARITY ( <i>as</i> , <i>abia</i> ), SUB ( <i>bos</i> )
		Sappada	PART <sub>T</sub> : SIMILARITY ( <i>a bi</i> )
		Sauris	PART <sub>T</sub> : SIMILARITY ( <i>assbie</i> )
		Timau	PART <sub>T</sub> : SIMILARITY ( <i>a bia</i> )
	Alemannic:	Gressoney	PART <sub>T</sub> : SIMILARITY ( <i>als</i> ), SEQUENCE ( <i>wan</i> )
		Bosco Gurin	PART <sub>T</sub> : CONTRAST ( <i>widar</i> )
		Issime, Formazza	PART <sub>T</sub> : SEQUENCE ( <i>dén</i> , <i>de</i> )
		Alagna, Macugnaga	PART <sub>T</sub> : SOURCE/SUB ( <i>fan</i> )
		Rimella	PART <sub>T</sub> : SIMILARITY/SUB ( <i>šchu</i> )
DEGREE	Bavarian:	Cimbrian,	PART <sub>D</sub> : (MORE) when PROP <sub>Q</sub> is
		Mòcheno	[polysyllabic], [converted], [- native]
	Alemannic:	Rimella, Alagna	PART <sub>D</sub> : (MORE)

**Table 3:** The overall typology of CCIs in the German minorities of Italy.

This manifold picture witnesses of the high complexity of these varieties which have to be seriously taken into consideration when carrying out typological investigations and particularly areal typology. In this regard, the mixture of isolation and contact seems to enhance variation which partially exploits models occurring in the diachronic development of a language sub-family, and partially elaborates interesting new patterns calquing models present in the speakers' repertoire.

## Acknowledgements

This paper grew out from a number of research projects funded in the last years, and, in particular, the project CLiMALp (2020-23) supported by the Italian Ministry of University (MIUR) and accessible at the following site: [www.climalp.org](http://www.climalp.org). I am deeply indebted with all the collaborators in the project for help and support, as well as Adriano Murelli and two anonymous reviewers for observations and remarks. Remaining mistakes are mine.

## Abbreviations

ACC = accusative	GEN = genitive	POSS = possessive
C = COMPAREE	INDEF = indefinite	PREF = prefix
CLF = classifier	INF = infinitive	PROP = property
COMP = comparative	IMP = imperative	PS = person
COMPL = completive	LE = linking element	PTC = particle
COP = copula	LOC = locative	PTCP = participle
D = degree	M = masculine	Q = quality
DAT = dative	N = neuter	REL = relative
DEM = demonstrative	NEG = negation	S = standard
DET = determiner	NMLZ = nominalizer	SBJ = subjunctive
DIM = diminutive	NOM = nominative	SG = singular
ENT = entity	PART = particle	SUFF = suffix
EXIST = existential	PL = plural	T = TIE
F = feminine	PRS = present	
FOC = focus	PST = past	

## References

- Angster, Marco & Bellante, Marco & Cioffi, Raffaele & Gaeta, Livio. 2017. I progetti DiWaC e ArchiWals. In Livio Gaeta (ed.), *Le isole linguistiche tedescofone in Italia: situazione attuale e prospettive future (Workshop, Torino, 24 febbraio 2017)*. Special issue of *Bollettino dell'Atlante Linguistico Italiano* 41. 83-94.
- Angster, Marco & Cioffi, Raffaele & Bellante, Marco & Gaeta, Livio. 2020. Corpora e varietà minoritarie: le isole walser in Italia. *Rivista Italiana di Dialettologia* 44. 107-125.
- Burt, F. 1910. *Swahili Grammar and Vocabulary*. London: Society for Promoting Christian Knowledge.
- Bidese, Ermenegildo. 2008. *Die diachronische Syntax des Zimbrischen*. Tübingen: Narr.
- Bidese, Ermenegildo Bidese. (eds.) 2010. *Il cimbro negli studi di linguistica*. Padova: Unipress.
- Bidese, Ermenegildo & Cognola, Federica. (eds.) 2013. *Introduzione alla linguistica del mòcheno*. Torino: Rosenberg & Sellier.
- Bidese, Ermenegildo, Dow, James R. & Stolz, Thomas (eds.). 2005. *Das Zimbrische zwischen Germanisch und Romanisch*. Bochum: Brockmeyer.
- Bergmann, Gunter. 1990. Upper Saxon. In Charles V. J. Russ (ed.), *The Dialects of Modern German*, 290-312. London: Routledge.

- Bohnenberger, Karl. 1913. *Die Mundart der deutschen Walliser im Heimattal und in den Außenorten*. Frauenfeld: Huber.
- Braune, Wilhelm. 2004. *Gotische Grammatik*. 20th edn. by Frank Heidermanns. Tübingen: Niemeyer.
- Brauner, Siegmund & Bantu, Joseph K. 1964. *Lehrbuch des Swahili*. Leipzig: Enzyklopädie-Verlag.
- Croft, William. 2001. *Radical Construction Grammar*. Oxford: Oxford University Press
- Croft, William. 2003. *Typology and universals*. 2nd edn. Cambridge: Cambridge University Press.
- Dal Negro, Silvia. 2004. *The Decay of a Language*. Bern: Lang.
- Dixon, Robert M. W. 2012. *Basic Linguistic Theory 3*. Cambridge: Cambridge University Press.
- Donaldson, Bruce C. 1993. *A grammar of Afrikaans*. Berlin: Mouton de Gruyter.
- DUDEN = *Duden. Die Grammatik*, 8th edn. Mannheim: Dudenverlag.
- Durrell, Martin. 1990. North Central Westphalian. In Charles V. J. Russ (ed.), *The Dialects of Modern German*, 59-90. London: Routledge.
- Durrell, Martin & Davies, Winifred V. 1990. Hessian. In Charles V. J. Russ (ed.), *The Dialects of Modern German*, 210-240. London: Routledge.
- DWB = Grimm, Jacob & Grimm, Wilhelm. 1854-1961. *Deutsches Wörterbuch*. Leipzig: Hirzel.
- Ebert, Robert P. & Reichmann, Oskar & Solms, Hans-Joachim & Wegera, Klaus-Peter. 1993. *Frühneuhochdeutsche Grammatik*. Tübingen: Niemeyer.
- Eufe, Rembert & Mader, Anna. 2018. Das Walserdeutsche im deutschen und italienischen Sprachgebiet. In Nicole Eller-Wildfeuer & Paul Rössler & Alfred Wildfeuer (eds.), *Alpindeutsch*, 113-138. Regensburg: vulpes.
- Fleischer, Jürg. 2007. Zur Herkunft des flektierten prädikativen Adjektivs im Höchstalemannischen. *Zeitschrift für Dialektologie und Linguistik* 74. 196-240.
- Freimann, Annick & Lemal, Carine & Simon, Isabelle & Lulling, Jérôme & Sgiarovello, Alain. *Lëtzebuergesch léiere mat MP3* (<https://www.vdl.lu>).
- Gaeta, Livio. 2013. Existential constructions: a semasiological perspective. In Elly van Gelderen & Michela Cennamo & Jóhanna Barðdal (eds.), *Argument Structure in flux: the Naples/Capri Papers*, 477-510. Amsterdam: John Benjamins.
- Gaeta, Livio. 2018. Im Passiv sprechen in den Alpen. *Sprachwissenschaft* 43. 221-250.
- Gaeta, Livio. 2020. Remotivating inflectional classes: An unexpected effect of grammaticalization. In Bridget Drinka (ed.), *Historical Linguistics 2017. Selected*

*papers from the 23<sup>rd</sup> International Conference on Historical Linguistics, San Antonio, Texas, 31 July – 4 August 2017*, 205-227. Amsterdam: Benjamins.

- Gaeta, Livio. In press. The Observer's Paradox meets Corpus Linguistics: Written and oral sources for the Walser linguistic islands in Italy. In Genesin, Monica, Hempel, Gerhard & Kahl, Thede (eds.), *Endangered linguistic varieties and minorities in Italy and the Balkans*. Vienna: VLACH.
- Gaeta, Livio & Bellante, Marco & Cioffi, Raffaele & Angster, Marco. 2019. Conservazione e innovazione nelle varietà walser: i progetti DiWaC e ArchiWals. In Rosselli Del Turco, Roberto (ed.), *Dall'indoeuropeo al germanico: problemi di linguistica storica*, 141-193. Alessandria: dell'Orso.
- Goltz, Reinhard H. & Walker, Alastair G. H. 1990. North Saxon. In Charles V. J. Russ (ed.), *The Dialects of Modern German*, 31-58. London: Routledge.
- Geeraerts, Dirk. 2010. *Theories of Lexical Semantics*. Oxford: Oxford University Press.
- Glynn, Dylan. 2015. Semasiology and onomasiology. In Jocelyne Daems & Eline Zenner & Kris Heylen & Dirk Speelman & Hubert Cuyckens (eds.), *Change of Paradigms – New Paradoxes. Recontextualizing Language and Linguistics*. Berlin & Boston: De Gruyter Mouton, 47-79.
- Green, W. A. I. 1990. The Dialects of the Palatinate (Das Pfälzische). In Charles V. J. Russ (ed.), *The Dialects of Modern German*, 241-264. London: Routledge.
- Haldeman, S. S. 1872. *Pennsylvania Dutch*. London: Trübner.
- Harbert, Wayne. 2007. *The Germanic Languages*. Cambridge: Cambridge University Press.
- Heine, Bernd. 1997. *Cognitive Foundations of Grammar*. Oxford: Oxford University Press.
- Helbig, Gerhard & Buscha, Joachim. 1991. *Deutsche Grammatik*. Leipzig: Langenscheidt.
- Hopper, Paul J. 1975. *The syntax of the simple sentence in Proto-Germanic*. The Hague: Mouton.
- Jacobs, Neil G. 2005. *Yiddish*. Cambridge: Cambridge University Press.
- Kranzmayer, Eberhard. 1981. *Laut- und Flexionslehre der deutschen zimbrischen Mundart*. Ed. by Maria Hornung. Vienna: VWGÖ.
- Lehmann, Winfred P. 1975. *A Grammar of Proto-Germanic*. University of Texas at Austin: Linguistics Research Center. (<https://liberalarts.utexas.edu/lrc/resources/books/pgmc/>)



- Maitz, Péter & Lindenfelser, Siegwalt & Volker, Craig A. In press. Unserdeutsch (Rabaul Creole German), Papua New Guinea. In Hans C. Boas & Deumert, Ana Deumert & Mark L. Loudon & Péter Maitz (eds.), *Varieties of German Worldwide*. Oxford: Oxford University Press.
- Malherbe, Michel & Cheikh, Sall. 1989. *Parlons Wolof*. Paris: L'Harmattan.
- Moulton, William G. 1941. *Swiss German Dialect and Romance Patois*. (Yale University Diss. Supplement to *Language* 17). Baltimore: Linguistic Society of America.
- Orel, Vladimir. 1998. *Albanian Etymological Dictionary*. Leiden: Brill.
- PALWaM = Antonietti, Federica & Valenti, Monica & Angster, Marco (eds.). 2015. *Piccolo Atlante Linguistico dei Walser Meridionali*. Aosta: Tipografia Valdostana.
- Paul, Hermann. 2007. *Mittelhochdeutsche Grammatik*. 25th edn. by Thomas Klein & Hans-Joachim Solms & Klaus-Petera Wegera & Prell, Heinz-Peter Prell. Tübingen: Niemeyer.
- Philipp, Marthe & Bothorel-Witz, Arlette. 1990. Low Alemannic. In Charles V. J. Russ (ed.), *The Dialects of Modern German*, 313-336. London: Routledge.
- Press, Ian. 1986. *A Grammar of Modern Breton*. Berlin: Mouton de Gruyter.
- Philipp, Marthe & Bothorel-Witz, Arlette. 1990. Low Alemannic. In Charles V. J. Russ (ed.), *The Dialects of Modern German*, 313-336. London: Routledge.
- Postma, Gertjan. 2019. *A Contrastive Grammar of Brazilian Pomeranian*. Amsterdam & Philadelphia: John Benjamins.
- Reese, Johannes. 2007. *Swiss German*. München: Lincom.
- Rizzi, Enrico. 1993. *Storia dei Walser*. Anzola d'Ossola (Novara): Fondazione Arch. Enrico Monti.
- Rowley, Anthony R. 2010. *Deskriptive Grammatik des Deutsch-Fersentalerischen (Mòcheno)*. Frankfurt/Main: Universitätsbibliothek Johann Christian Senckenberg.
- Russ, Charles V. J. 1990. Swabian. High Alemannic. In Charles V. J. Russ (ed.), *The Dialects of Modern German*, 337-393. London: Routledge.
- Slaviero, Girardo. [1760] 1991. *Grammatica della lingua tedesca dei VII Comuni*. Giazza: Taucias Garëida.
- Schönfeld, Helmut. 1990. East Low German. In Charles V. J. Russ (ed.), *The Dialects of Modern German*, 91-135. London: Routledge.
- Schrodt, Richard. 2004. *Althochdeutsche Grammatik II*. Tübingen: Niemeyer.
- Spangenberg, Karl. 1990. Thuringian. In Charles V. J. Russ (ed.), *The Dialects of Modern German*, 265-289. London: Routledge.
- Stassen, Leon. 1985. *Comparison and Universal Grammar*. Oxford: Basil Blackwell.

- Šticha, František. 2011. Der kommunikative und der systembezogene Status grammatischer Phänomene mit niedriger Häufigkeit. In Marek Konopka & Jacqueline Kubczak & Christian Mair & František Šticha & Ulrich Hermann Waßner (eds.), *Grammatik und Korpora. Dritte Internationale Konferenz. Mannheim, 22.-24.9.2009*, 473-484. Tübingen: Narr.
- Stolz, Thomas. 2013. *Competing Comparative Constructions in Europe*. Berlin: Akademie-Verlag.
- SW = Antonietti, Federica (ed.). 2010. *Scrivere tra i Walser*. Borgomanero: Litopress.
- SWB = Schmeller, Johann A. 1855. *Cimbrisches Wörterbuch*. Vienna: Hof- und Staatsdruckerei. [Ed. by Guus Kroonen. Copenhagen University, 2012]
- Tiersma, Pieter. M. 1999. *Frisian Reference Grammar*. Ljouwert: Fryske Akademy.
- Tosco, Mauro. 2001. *The Dhaasanac language*. Köln: Köppe.
- Tyroller, Hans. 2003. *Grammatische Beschreibung des Zimbrischen in Lusern*. Stuttgart: Steiner.
- Wiesinger, Peter. 1990. The Central and Southern Bavarian Dialects in Bavaria and Austria. In Charles V. J. Russ (ed.), *The Dialects of Modern German*, 438-519. London: Routledge.
- WDS = Geyer, Ingeborg & Marco Angster & Marcella Benedetti. 2014. *Wortschatz aus den deutschen Sprachinseln in Italien*. Bozen: Athesia.
- WL = Ronco Hantsch, Imelda. 1993. *Willischt lesen? Kuntjini van Esopo gchierti in töitschu*. Issime.
- Zürrer, Peter. 1982. *Wörterbuch der Mundart von Gressoney: Mit einer Einführung in die Sprachsituation und einem grammatischen Abriss*. Frauenfeld: Huber.
- Zürrer, Peter. 2009. *Sprachkontakt in Walser Dialekten: Gressoney und Issime im Aostatal*. Stuttgart: Steiner.
- Zürrer, Peter. 2011. Conservazione e mutamento nei dialetti walser in Italia. In Walter Breu (ed.), *L'influsso dell'italiano sul sistema del verbo nelle lingue minoritarie*, 101-121. Bochum: Brockmeyer.

#### CONTACT

livio.gaeta@unito.it