

BOOK OF ABSTRACTS

GENERAL SESSION TALKS

58th Annual Meeting of the *Societas Linguistica Europaea*

26 – 29 August 2025

Université Bordeaux Montaigne

Anna Kisiel

General Session : Areal linguistics

Beyond evidentiality: Epistemic authority and other effects on the use of grammatical evidentials in Southern Finnic

Denys Teptiuk, Miina Norvik & Petar Kehayov
(University of Tartu, University of Tartu & University of Tartu)

Keywords: quotative mood, reported speech, reported evidentiality, epistemic stance, Estonian, Livonian

The Southern Finnic languages (Standard & South) Estonian and Livonian possess grammatical evidentials that cover two meanings: reportative with the unknown/unspecified source, and quotative with the specified source (Aikhenvald 2004; Kehayov & Skribnik 2018). The additional epistemic meaning of uncertainty in the truth-value of the report is often ascribed to Estonian (Aikhenvald 2004: 193), while this additional meaning is less typical for the Livonian grammatical evidential (Kehayov et al. 2012). Furthermore, mirative overtones arise in contexts where the reported state-of-affairs was not expected by the reporter (1). Such reports are also based on first-hand accounts of other speakers, inaccessible to the reporter (2).

(1) Estonian

Suur oli minu üllatus kui kaal
big be:PST.3SG 1SG.GEN surprise when weights
näitas, et mina kaaluvat kokku
show:PST.3SG COMP 1SG weigh:REP.EVID altogether
72,4 kilo.
NUM kilogram.PRT

‘I was surprised a lot when the weights showed that I weighed 72,4 kilos.’ (etTenTen21)

(2) Livonian

ma u'm kūlōn ku oksāka'ļdi Rīgōs
1SG be.1SG hear:APP.SG COMP three.spined.stickleback:PL.PRT Riga:INE
sāl Vēnas ve'jjjid ja Liepās ma
there Daugava:INE catch.fish:REP.EVID:PL and Liepaja:INE 1SG
u'm īž nā'nd ku kešļdōks āt ve'jjōnd
be.1SG self see:APP.SG COMP landing.net:PL.INS be.3PL catch.fish:APP.PL
oksāka'ļdi
three.spined.stickleback:PL.PRT

‘I've heard that three-spined sticklebacks have been caught in Riga, there in Daugava, and in Liepaja I've seen that three-spined sticklebacks were caught with a landing net.’ (Suhonen 1975: 26)

This study extends previous descriptions of the grammatical evidentials in these languages and scrutinizes additional connotations arising from the grammatical expression of evidentiality based on corpus material. Furthermore, we investigate how the asymmetry in epistemic authority (Grzech

2020; Bergqvist & Grzech 2023, i.a.) between the current speech participants is reflected in the use of grammatical evidentials in these languages.

Our preliminary results for Estonian show an effect of this parameter on the distribution of evidentials relative to person. While the grammatical evidential is relatively frequent in reports about the current speaker, who is also the (first person) reporter (ca. 1500 exx. in etTenTen21), it is rarely used in reports about the (second person) interlocutor (ca. 60 exx.). Furthermore, in reports about the current speaker, the grammatical evidential acquires quotative interpretation more frequently than reportative. The quotative reading often triggers additional epistemic and mirative effects (cf. First Person Effect in Aikhenvald 2004: 225), for which contextual cues and alternations with the epistemically neutral indicative mood are robust indexes. By contrast, reports about the interlocutor more frequently do not contain specification of the source. This can be viewed as a distancing strategy: to avoid potentially face-threatening situations, the speaker keeps information sources vague. A similar situation is observed in reports about the current speaker boasting about their own qualities: *ma pidavat olema ilus* [1SG must:REP.EVID beautiful] ‘People say, I am beautiful’.

Considering the rarity of epistemic overtones in the use of the Livonian evidential (cf. Kehayov et al. 2012), we expect to find differences in the use of the grammatical evidentials with the same meaning between these closely related languages, potentially stemming from different conventionalisation paths of the evidentials therein and their uses in the limited number of genres available for moribund Livonian.

References

- Aikhenvald, Alexandra Y. (2004), *Evidentiality*. Oxford: Oxford University Press.
- Bergqvist, Henrik & Karolina Grzech (2023), The role of pragmatics in the definition of evidentiality, *STUF* 76(1), 1–30.
- etTenTen21 = Estonian Web 2021. Accessed via <https://app.sketchengine.eu>.
- Grzech, Karolina (2020), Epistemic primacy, Common Ground management and epistemic perspective, in H. Bergqvist, and S. Kittilä (eds), (2020), *Evidentiality, egophoricity, and engagement*, Berlin: Language Science Press, 23–60.
- Kehayov, Petar, Helle Metslang & Karl Pajusalu (2012), Evidentiality in Livonian, *Linguistica Uralica* 48(1), 41–54.
- Kehayov, Petar & Elena Skribnik (2018), Evidentials in Uralic languages, in A. Y. Aikhenvald (ed), *The Oxford Handbook of Evidentiality*, Oxford: Oxford University Press, 525–553.
- Suhonen, Seppo (1975), *Liivin kielen näytteitä* (Castrenianumin toimitteita 5.). Helsinki: Helsingin yliopisto.

Phrasal alternation with Estonian indefinite nominal quantifiers

Maarja-Liisa Pilvik, Liina Lindström, Carl Eric Simmul & Helen Plado

(University of Tartu, University of Tartu, University of Tartu & University of Tartu)

Keywords: quantifiers, agreement, language variation, Estonian, Circum-Baltic languages

Acknowledgements: The study has been conducted within the project "Morphosyntactic variation in Estonian" (EKKD-TA2) which is funded by the Estonian Ministry of Education and Research.

Indefinite quantifiers (e.g., *few*, *some*, *several*, *many*, *much*) express a nonspecific or vague number or quantity. In Estonian, indefinite quantifiers exhibit diverse syntactic behaviors when referring to the quantity of collective (plural) entities (see Metslang 2017). Some operate exclusively within plural paradigms, while others are used in both singular and plural (e.g., SG *mõnes raamatus* 'in some books', PL *mõnedes raamatutes* 'in some books'). In both cases, the quantifiers function as modifiers within a noun phrase (NP), agreeing with the head noun in case and number. Certain quantifiers (1) alternate between agreement and government, acting as the head of a quantifier phrase (QP) in the singular paradigm. Plural quantifiers, however, can still be analyzed as modifiers within an NP.

(1) QP: <i>mitu</i>	<i>raamatu-t</i>	<i>mitme-s</i>	<i>raamatu-s</i>
several.SG.NOM	book-SG.PAR	several-SG.IN	book-SG.IN
NP: <i>mitme-d</i>	<i>raamatu-d</i>	<i>mitme-te-s</i>	<i>raamatu-te-s</i>
several-PL.NOM	book-PL.NOM	several-PL-IN	book-PL-IN
'several books'		'in several books'	

The current study focuses on the agreement patterns used with indefinite nominal quantifiers *osa* ('some; part (of)'), *enamik* ('most; majority (of)'), and *enamus* ('most; majority (of)'), which also lexically express partitivity. These quantifiers also alternate between functioning as QPs and NPs but, regardless of the number, always trigger plural marking on the noun denoting the collective whole (2). Therefore, agreement may occur in both case and number, in case alone, or in neither.

(2) QP: <i>osa</i>	<i>raamatu-i-d</i>	<i>osa-s</i>	<i>raamatu-te-s</i>
part.SG.NOM	book-PL-PAR	part-SG.IN	book-PL-IN
NP: <i>osa-d</i>	<i>raamatu-d</i>	<i>osa-de-s</i>	<i>raamatu-te-s</i>
part.PL.NOM	book-PL.NOM	part-PL-IN	book-PL-IN
'some (of the) books'		'in some (of the) books'	

The choice of NP over QP is regarded as a relatively recent development (Erelt, Metslang 1998) and has been discouraged in language planning (Kindlam 1978: 62, Erelt et al. 2020: 492). The two patterns, however, are not considered completely synonymous, as they evoke different types of construal. The use of a singular quantifier in a QP emphasizes (indefinite) quantity and part-whole relations, whereas a plural

quantifier in an NP specifies individuals, functioning similarly to non-quantifying determiners (Erelt, Metslang 1998). Nonetheless, the distinction is often ambiguous in actual usage, suggesting an ongoing shift in the language.

In this paper, we compare the extent and conditions influencing the choice between the two phrase types using data from the Estonian National Corpus 2023 (Koppel et al. 2023). Conditional inference trees and logistic regression were employed to assess whether the alternation is sensitive to factors such as genre, modification, animacy, case form, frequency, or the predictability of the collective whole within the phrase. The results indicate that the choice between QP and NP is primarily conditioned by the syntactic function of the phrase: adverbials in oblique cases emerge as a potential source of change towards number agreement, while the specifying function of the quantifier is increasingly solidified in phrases functioning as subjects. Given that the relationship between (pseudo-)partitives and indefiniteness has been attested cross-linguistically (Seržant 2021a, 2021b), and that indefinite quantifiers occur in both NPs and QPs also in other Circum-Baltic languages (e.g., Latvian *lielāka daļa grāmatu* 'most books', lit. 'biggest part of the books'), we also explore the areal typology of the investigated phenomenon.

References

- Erelt, Mati; Metslang, Helle (1998), Oma või võõras? ("Muutuv keel") [A local or a stranger? (Changing language)], *Keel ja Kirjandus* 10, 657–668.
- Erelt, Mati; Erelt, Tiiu; Ross, Kristiina (2020), *Eesti keele käsiraamat. Uuendatud väljaanne* [Handbook of Estonian. New edition], Tallinn: Eesti Keele Instituut / Eesti Keele Sihtasutus.
- Kindlam, Ester (1978), Sõltumussuhetest [On dependency relations], *Keelevoos* 76, 55–66.
- Koppel, Kristina; Kallas, Jelena; Jürviste, Madis; Kaljumäe, Helen (2023), *Eesti keele ühendkorpus 2023* [Estonian National Corpus 2023], Lexical Computing Ltd. / Eesti Keele Instituut.
- Metslang, Helle (2017), Kvantorifraas [Quantifier Phrase], in M. Erelt, and H. Metslang (eds), (2017), *Eesti keele süntaks* [Syntax of the Estonian language], Tartu: Tartu Ülikooli Kirjastus, 463–478.
- Seržant, Ilja A. (2021a), Diachronic typology of partitives, in P. Sleeman, and G. Giusti (eds), *Partitive Determiners, Partitive Pronouns and Partitive Case*, (Linguistische Arbeiten 580), Berlin, Boston: De Gruyter, 111–167. <https://doi-org/10.1515/9783110732221>
- Seržant, Ilja A. (2021b), Typology of partitives, *Linguistics* 59(4), 881–947. <https://doi-org/10.1515/ling-2020-0251>

General Session : Historical linguistics

Semantic and constructional change in Hill Mari modals

Aigul Zakirova

(University of Potsdam)

Keywords: Hill Mari, grammaticalization, modal constructions, modality.

In this talk, I discuss the diachronic development of the future/necessity verbal form *-šaš* in Hill Mari. Particular attention is paid to the semantic and constructional properties of the various uses of *-šaš* and the complex marker *-šaš-lâk* based upon it.

Data from two periods are considered: present-day and the 1890–1910s.

- The present-day data were elicited in 2016–2019 in Kuznetsovo (Gornomariyskiy district, Mari El Republic, Russia); in addition, I used corpus data from the same variety (63,522 tokens; **53 uses of *-šaš*, 69 of *-šaš-lâk***).
- The 1890–1910s are represented by oral texts in several varieties of Hill Mari (Ramstedt 1902, Beke 1951, and Wichmann 1931; overall ~31,000 tokens; **29 uses of *-šaš*, 6 of *-šaš-lâk***).

The frequencies of *-šaš* and *-šaš-lâk* in the two periods significantly differ. In the 1890–1910s, *-šaš-lâk* was used adnominally in contexts of future/ necessity (1) and was able to relativize any argument (e.g., S, P, A.). There are also two examples of *-šaš-lâk* as part of a nominal predicate, where the subject is the S/P argument of the lexical verb (2). In addition to root necessity (2), the construction could also express future.

- (1) *či-ktä-šaš-lâk* *tâgâr*
put_on-CAUS-**ŠAŠ-DEST** shirt
'shirt to be put on' (Beke 1951: 76–77)
- (2) *ti* *mešäk-šă* *kü-län* *pu-šaš-lâk?*
this sack-POSS.3SG who-DAT give-**ŠAŠ-DEST**
'Who should this sack be given to?' {according to you} (Wichmann 1931: 200)

In the present-day corpus, adnominal uses of *-šaš-lâk* are also found, however, its most frequent use (64 out of 69 examples) is in the predicate position. These examples all have S/A subjects and the meaning of root necessity (3).

- (3) *tän'-äm* *män'* *pân'imajâ-šaš-lâk* *âl-am* *vet*
you-ACC I understand-**ŠAŠ-DEST** be-NPST.1SG PTCL
'I must understand you, after all.' {according to the rules of this activity} (Kuznetsovo corpus)

I conclude that between the 1910 and 2010s the predicative use of the future/necessity form acquired a semantic restriction (only root necessity interpretations became possible) and underwent a constructional change (S/P subject > S/A subject). It also became much more frequent. I argue that this development occurred as a result of language standardization in the 1920s. The reanalyzed *-šaš-lâk* construction was needed as a translational equivalent for the Russian personal modal *dolžen*, since Hill Mari previously lacked personal modal constructions (as seen from texts in Ramstedt 1902, Beke 1951, Wichmann 1931).

In addition, I will discuss the semantic and constructional properties of the predicative uses of *-šaš* in Hill Mari, which may be characterized as preferential (synchronically already described in Zakirova 2018). These uses are absent from Meadow Mari and include intention ('I'll do P'), 1SG deliberative question ('What should I do?'), suggestion ('You, he/she better P') and optative ('If only P!'). These uses

are almost always used in dialogue or inner speech. Semantically, they all imply choosing the best course of action or outcome.

As a result, I will reconstruct the diachronic connections between all uses of *-šaš* and *-šaš-lâk* and show how the changes observed fit into the diachronic typology of modals (van der Auwera & Plungian 1998, and Narrog 2012).

Acknowledgments: The research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project ID 556302231.

References

- Beke, Ödön. (1951). *A cseremiszek (marik) népköltészete és szokásai: Volksdichtungen und Gebräuche der Tscheremissen (Maris)*. Budapest: Akadémiai.
- Kuznetsovo Hill Mari corpus. URL: <http://hillmari-exp.tilda.ws/corpus>, accessed on 10.01.2025.
- Narrog, Heiko. (2012). *Modality, subjectivity, and semantic change: a cross-linguistic perspective*. Oxford University Press.
- Ramstedt, Gustaf John. (1902). *Bergtscheremissische sprachstudien* (vol. 17). Druckerei der Finnischen Litteraturgesellschaft.
- van der Auwera, Johan, & Plungian, Vladimir. (1998). Modality's semantic map. *Linguistic Typology*. 2. 79-124.
- Wichmann, Yrjö. (1931). *Volksdichtung und Volksbräuche der Tscheremissen*. Helsinki, Suomalais-Ugrilainen Seura.
- Zakirova, Aigul. (2018). *Person constraints in modals: the case of the Hill Mari –šaš*. Talk at the 51st Annual Meeting of the Societas Linguistica Europaea. 29 August – 1st September 2018, University of Tallinn.

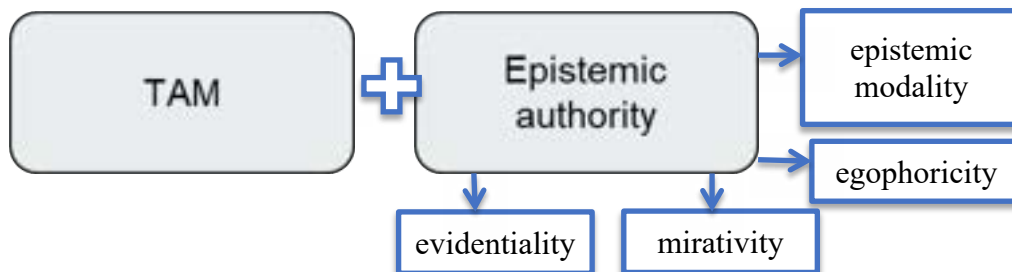
Turn out analogues in Tibetan: Diachronic and semantic aspects

Anna Kramskova

(Institute for Linguistic Studies of Saint-Petersburg)

Keywords: revelatory verbs, mirativity, Tibetan, evidentiality, epistemic authority

Modern Tibetan (MT), as well as most of its dialects, is characterized by a grammatical system (pic. 1) that encompasses not only TAM-categories, but also a functional notion of “epistemic authority” (term suggested in [Bergquist, Kittilä 2019]). Unlike in most other languages with grammaticalized evidentiality and mirativity, development of the corresponding complex analytical verbal paradigm in Tibetan can be traced through its literary tradition.



Pic. 1. Verbal grammatical categories in MT

Existence of mirativity as extension of evidentiality in Tibetan has been subject to many a debate (see [Hill 2012; DeLancey 2012; Hengeveld & Olbertz 2012]), yet some Tibetan dialects, including Lhasa Tibetan, present markers that are exclusively mirative (e.g. “revelatory” *red-sha* (*red-bzhag*) in [Tournadre 2003]).

Both semantically and etymologically the closest analogues to *turn out* verbs, which are ubiquitous throughout the European languages, can be found in the Tibetan markers for direct (1) and inferential evidentiality (2) and mirativity (3).

(1) *kho-la deb 'dug*

he-LOC the book is

‘(I see that) he has books.’/‘Oh, he has books!’ (attested by the speaker; sudden discovery) [Denwood 1999]

(2) *bod la g.yag bsdad-bzhag*

Tibet-LOC yak stay-MIR

It turns out that yaks lived in Tibet! (I see their tracks, I'm surprised)

(3) *nyon.pa re'-sha'*

crazy is.MIR

'Hey, he's crazy! (I've just realized it)' [Tournadre 2003]

Existential copulas *'dug* EXIST.DIR.EVID and *red-'dug* EXIST.MIR, as well as direct evidential *Vprs-gi-'dug* PRS.DIR.EVID and inferential perfect forms *Vpst-'dug* PRF.INFER.EVID all come back to the Old Tibetan (OT) verb *'dug* 'to sit, remain', while mirative copula *red-bzhag* EXIST.MIR (synonymous to *red-'dug*) and mirative perfective form *Vpst-bzhag* PERF.MIR go back to the OT verb *'jog* 'to put'.

The results of the comparative-historical analysis on data from available texts in OT, Classical Tibetan (CT) and MT (mainly "Old Tibetan Chronicle" (Dunhuang manuscript, before the 11th century), "Biography of Milarepa" ("Milaraspa namthar", with oral text going back presumably to the 14th century, written down in the 15th century) and own corpora of literary texts (12th-21st centuries) at www.aiire.org) show that the grammaticalization of these verbs can be traced back to the late OT period with its weakening of the semantics of several Old Tibetan verbs: *'dug* 'to sit, to remain' (4), *bzhag* 'to put.PST', *yod* 'to exist', *yin* 'to appear', *song* 'to leave.PST' and *thal* 'to pass'.

(4) *'di ni gnam-las byon-pa-'i btsan.po nyom.tshar.can zhig 'dug-pa-s*

this TOP sky-ABL come-NMZ-GEN king marvelous PART **be-NOM-ERG**

"Because this one who came from the sky **is/seems to be** a marvelous ruler" ["Nyang ral chos 'byung", 12th century]

The paper presents the results of the comparative analysis of a variety of grammatical markers with "discovery" semantics in Old, Classical and Modern Tibetan while attempting to assess the connections between the functional categories of mirativity, evidentiality and perfect aspect.

References

- Bergqvist, Henrik & Kittilä, Seppo (eds.). (2019), *Evidentiality, egophoricity and engagement*, in *Studies in Diversity Linguistics*, vol. 30, Berlin: Language Science Press.
- DeLancey, Scott (2012), *Still mirative after all these years*, in *Linguistic Typology*, vol. 16 (3), 529-564.
- Denwood, Philip (1999), *Tibetan*, Amsterdam: John Benjamins.
- Hengeveld, Kees & Olbertz, Hella (2012), *Didn't you know? Mirativity does exist!*, in *Linguistic Typology*, vol. 16 (3), 487-503.
- Hill, Nathan, (2012) "*Mirativity*" *does not exist: ḥdug in "Lhasa" Tibetan and other suspects*, in *Linguistic Typology* 16(3), 389-433.
- Tournadre, Nicolas (2003), *Manual of Standard Tibetan*. Ithaca, NY.

The fragmentation of a unified paradigm : from Indo-European $*k^w$ -expressions to French indefinite, interrogative and relative *qu*- words

Axelle Domingues
Sorbonne Université, Paris

Keywords : diachrony, corpus, Indo-European, French, TXM, indefinite, interrogative, relative

1. Background

At first glance, *qu*- words in French seem to perfectly illustrate a unified grammatical paradigm with a closed set of expressions sharing common features both in form and function. Their formal identity, specifically the initial "qu-", originates from the Indo-European roots $*k^we/o-$ and k^wei- (Meillet & Vendryès, 1968 : §§ 750-752 ; Szemerényi, 1999 : 208). However, it can partially disappear over time, as seen in the Latin words *ubi* (> *où* "where") and *unde* (> *d'où* "from where"), which have lost the "qu-" identifier. Similarly, French words like *comme* ("as") (< *quo modo*) and its derivatives *comment* ("how") and *combien* ("how much"), replacing *quant* in Middle French, require etymological insight to be recognized as "qu-" words. Functionally, these Indo-European roots express both interrogative and indefinite meanings, without a clear consensus on which value precedes the other (Haspelmath, 1997). Their function as interrogative-indefinite can be conceptualized as a variable. The $*k^w$ - expressions indicate a seemingly endless range of all the possible values of this variable within a given domain, such as time, place, quality or quantity among others (Hahn, 1942).

In Classical Latin, the notion of a unified paradigm for *qu*- words was entirely relevant. Indeed, a single word such as *qui(s)* could function as an indefinite, interrogative, or relative pronoun or adjective. However, this system is now divided into three different paradigms in French. This presentation aims to explore when, how, and why this division occurred, focusing on modeling the pivotal moment when the first signs of divergence emerged.

2. Methodology

This study uses two digital corpora, one for Late Latin and one for Old French : *PaLaFra-Lat* (Passage du Latin au Français) focuses on the Merovingian period, from the 5th to the 8th century and BFM (Base de Français Médiéval) spans from 9th to the 15th century. Through textometric analysis (TXM), we examine all instances of indefinite, interrogative and relative *qu*- words recorded during these periods.

By integrating their relative textual frequencies into Python-based executable codes within *Google Colaboratory*, we process large-scale data. Various libraries, including *Pandas* for data structure and analysis, *Matplotlib* and *Seaborn* for visualization, and *Numpy* for array processing, enable us to model linguistic change and produce graphs to illustrate our results without sampling.

3. Results

Our research identifies the end of Old French as the critical turning point in this evolution, with traces of the first cracks already evident in Late Latin. This ongoing study will present graphs to illustrate this diachronic evolution, demonstrating how the unified paradigm of indefinite, interrogative, and relative *qu*- words gradually fractured.

Few empirical studies in long-term diachrony explore this progressive restructuring. Kunstmann (1990) is a notable early work on the interrogative-relative in Old French, but it is limited

both in its diachronic scope and methodology while also dismissing indefinite words. This presentation will offer a more refined understanding of the pace of this evolution and uncover the factors that led to the divergence of these paradigms.

References

- Hahn, E. A. (1942). *The Indefinite-Relative-Interrogative Stem*. *Language*, 83-116.
- Haspelmath, M. (1997). *Indefinite pronouns*. Oxford: Oxford University Press.
- Kunstmann, P. (1990). *Le relatif-interrogatif en ancien français* (Vol. 191). Genève : Droz.
- Meillet, A. & J. Vendryes (1968). *Traité de grammaire comparée des langues classiques*. Paris : Champion.
- Szemerényi, O. (1999). *Introduction to Indo-European Linguistics*. Oxford: Oxford University Press.

New evidence for the rise of the ‘promise’ auxiliary

Bert Cornillie
(KU Leuven)

Keywords: Latin calques, grammaticalization, discourse traditions, auxiliaries

In this talk I will discuss the pace of the rise of the subjective/auxiliary uses of ‘promise’ and ‘threaten’ + infinitive in Germanic and Romance. Unlike objective uses with its illocutionary performativity, with subjective uses the speaker presents a prediction on an inferential basis. Interestingly, auxiliatization and subjectification of ‘promise’ is later than that of ‘threaten’. Verhagen (2000) accounts for this difference in terms of discourse-pragmatics.

I will show that the auxiliatization of subjective ‘threaten’ paves the way for subjective ‘promise’. Previous accounts show that subjective ‘threaten’ verbs in Romance and Germanic stem from Latin calques containing a noun and that the grammaticalized constructions with the infinitive are instances of syntactic elaboration taking place at the moment of increased prestige of the vernacular language (e.g. the 15th century for Spanish, cf. Cornillie & Octavio de Toledo 2015, and the 17th century for Dutch, cf. Cornillie 2014, and the 18th century for German, and the 19th century for English. Cf. Cornillie 2019). The question arises as to whether a similar view holds for the auxiliary uses of ‘promise’. The proposed diachronic contribution will show that, although ‘promise’ constructions arose as Latin calques, the grammaticalization of the verb into an auxiliary is patterned on the grammaticalization of the ‘threaten’ verbs.

In the Library of Latin texts (Brepols) we found that the nominal construction with *promittere* is attested in Latin authors such as Seneca (1) and Virgilius (2).

(1) *Avaritia pecuniam promittit (Seneca)*

(2) *supra quos pendens faxum ruinam promittit, nec tamen cadit (Virgilius)*

In the early Aeneis translation of the Old French *Roman d’Enéas* (12th or 13th century) we find the French verb *promettre* with a noun (3), despite the fact that *promittere* is not in the original (4), which suggests that the construction may be entrenched already in early stages.

(3) *De nule part ne veient port; Et ciels et mers lor promet mort.*

(4) *Praesentemque viris intentant omnia mortem (Virgilius, Aeneis. I, 91)*

Interestingly, in a present-day English translation of (4) we found “and all things threaten instant death for the men”. The same can be found in the 15th century Castilian translation by Villena.

(5) ... de tal guisa que todas las cosas ya a los navegantes en aquella hora les menazava la muerte
(Villena, Eneida, I, 5, 83)

In this paper I will show that, despite their presence in similar contexts in early stages, ‘promise’ constructions differ from ‘threaten’ constructions in that the former lack the shift within the same semantic field of fall. Once subjective ‘threaten’ + infinitive constructions go beyond the ‘fall’ infinitives, auxiliary uses of ‘promise’ + infinitive start to show up. Our account highlights that actualization of an exemplar, which in this case is the more entrenched ‘threaten’ auxiliary, is needed to inspire other auxiliaries such as ‘promise’. Given the shared lexical field of threat, the ‘promise’ auxiliary is an example of grammaticalization through paradigmaticization following the path of *threaten*.

References

- Cornillie, Bert (2014). "Over de subjectieve lezing van *dreigen* in het 16^e- en 17e-eeuwse Nederlands. Historische pragmatiek vs. Contact-geïnduceerde taalverandering", in: Van de Velde, F., Smessaert, H. Van Eynde, F. & S. Verbrugge, *Patroon en Argument: een dubbelfeestbundel bij het emiraat van William Van Belle en Joop van der Horst*, Leuven University Press, Leuven, 329-348.
- Cornillie, Bert & Octavio de Toledo y Huerta, Álvaro (2015). "The diachrony of subjective *amenazar* 'threaten'. On Latin-induced grammaticalization in Spanish" in Smith, Andrew D. M., Graeme Trousdale & Richard Wälchli (eds.). *New Directions in Grammaticalization Research*. Amsterdam & Philadelphia: John Benjamins, 187-208.
- Cornillie, Bert (2019). On the pace of syntactic elaboration from Latin calques. Evidence from Spanish, Dutch and English. *Belgian Journal of Linguistics*. 2019. Vol: 33: 82-107.
- Verhagen, Arie (2000). "'The girl that promised to become something': An exploration into diachronic subjectification in Dutch". In *The Berkeley Conference on Dutch Linguistics 1997: the Dutch Language at the Millennium*, Thomas F. Shannon & Johan P. Snapper, (eds.). Lanham MD: University Press of America, 197-208.

Medieval scientific writing and authorial identity: *Chaucer and The Equatorie of the Planetis* reconsidered

Bridget Drinka & Joseph Roy
(University of Texas at San Antonio)

Keywords: Authorship identification, Middle English, Chaucer, Tense & Aspect

In this paper we contrast some traditional methods for authorship identification with a more historical sociolinguistic approach in determining the likelihood of Chaucer's authorship of the *Equatorie of the Planetis* (*Equatorie*). Since the discovery of the manuscript of the *Equatorie* by Price (1954), the question of whether or not Chaucer authored the treatise has been debated by scholars from various disciplines. Schmidt (1993), examining linguistic evidence, does not find sufficient evidence to support Chaucer's authorship of the *Equatorie*. Computational literary analysts who work on authorship issues tend to focus solely on lexical distributions (e.g. Craig & Kinney 2009), but several have recently begun to include linguistic structure as well (De Langhe, De Clercq & Hoste, 2024; Grant, 2022). Specialists working on style and genre (Biber & Conrad 2019) point to constraints which make authorship identification difficult, such as the fluctuation of underlying lexical richness according to genre and the insufficiency of data (Savoy, 2020).

The data for the statistical analyses consist of 55 works in Middle English from 1200 to 1500 collected in the Penn-Helsinki Parsed Corpus of Middle English (PPCME2) (Kroch, Taylor and Santorini, 2000). We chose the present perfect and preterite as a locus of testing authorship not only because these categories represent ongoing change-in-progress in Late ME (Nevalainen & Raumolin-Brunberg 2016; Bondar 2023), but also because the individual social histories of the authors reflect greater or lesser contact with French and other Romance models (Drinka 2017). We examine the percent of the present perfect (1) versus the preterite (2) for both Chaucer and his contemporaries.

(1) tak thanne as I **haue seid** by forn the fix fot of thy compas (PPCME2/CMEQUATO/18.18)

(2) Thy blake thred whan it first **leid** thorw the pol of thyn Epicicle it shewith the verrey aux of the planete in epiciclo riht as the white thred shewith themene aux in the same epicicle. (PPCME2/CMEQUATO/38.265)

We also extract the percent of Romance-origin vocabulary versus total vocabulary in the East Midlands texts from 1380 to 1400.

Analysis of content words in the PPCME2 reveals distinct linguistic patterns in Chaucer's work. The *Treatise on the Astrolabe* and *Equatorie* show similarly high Romance vocabulary (60% and 54%), while other East Midlands texts from 1380-1400 remain below 35%. Both texts also share distinctive perfect-tense distributions, with Chaucer's works showing 29-38% present perfect usage compared to a 14% average (maximum 18%) in contemporary East Midlands texts. The *Equatorie's* 27% present perfect frequency aligns with Chaucer's pattern.

We demonstrate that utilizing a more comprehensive historical sociolinguistic framework for authorship identification yields more convincing results than methods used in previous studies. This study identifies a stable set of linguistic patterns in Chaucer's work—unique from his peers and the scientific genre in which the contested work, the *Equatorie*, occurs. It also presents a framework,

informed by corpus linguistics, historical sociolinguistics, and the intersection of the two, which can be used to determine authorship in other contested works.

References

- Biber, D., & Conrad, S. (2019). *Register, genre, and style*. Cambridge University Press.
- Bondar, V. (2023). The present perfect in past time contexts: a diachronic study of English. *Brno Studies in English*, 49(2), 5-29.
- Craig, H., & Kinney, A. F. (2009). *Shakespeare, computers, and the mystery of authorship*. Cambridge University Press.
- De Langhe, L., De Clercq, O., & Hoste, V. (2024, May). Unsupervised Authorship Attribution for Medieval Latin Using Transformer-Based Embeddings. In *Proceedings of the Third Workshop on Language Technologies for Historical and Ancient Languages (LT4HALA)@ LREC-COLING-2024* (pp. 57-64).
- Drinka, B. (2017). *Language contact in Europe: The periphrastic perfect through history*. Cambridge University Press.
- Grant, T. (2022). *The idea of progress in forensic authorship analysis*. Cambridge University Press.
- Horobin, S. (2003). *The language of the Chaucer tradition*. Boydell & Brewer Ltd.
- Nevalainen, T., & Raumolin-Brunberg, H. (2016). *Historical sociolinguistics: language change in Tudor and Stuart England*. Routledge.
- Price, D. J. (1953). The Equatorie of the Planetis. *Bulletin of the British Society for the History of Science*, 1(9), 223-226.
- Rand, K. A. (2015). The Authorship of the Equatorie of the Planetis revisited. *Studia neophilologica*, 87(1), 15-35.
- Savoy, J. (2020). *Machine Learning Methods for Stylometry*. Springer International Publishing.
- Schmidt, K. A. R. (1993). *The Authorship of the Equatorie of the Planetis* (Vol. 19). Boydell & Brewer.

Reproducibility and accessibility in Bayesian phylolinguistics

Carl Bodnaruk
(University of Sydney)

Keywords: Historical Linguistics, Bayesian. Phylogenetics, Reproduction, Research Methodologies

Statistical modelling of linguistic history through Bayesian phylogenetic methodologies borrowed from evolutionary biology has been undertaken since the mid 2000s (Bower 2018), though has become more common in recent years (e.g., Kolipakam et al 2018, Robbeets and Bouckaert 2018, Sagart et al 2019, Zhang et al 2019, Zhang et al 2020 *inter alia*). These methodologies involve collecting data, often cognate sets, from large numbers of languages within a given family and using a probabilistic analysis to generate the most likely phylogeny, or tree, through which the contemporary language situation could have come to be (Greenhill et al 2020). There are a number of potential challenges in the use of these analytical methodologies, both in terms of the methodologies themselves and their translation from biology to linguistics, as well as the presentation and publication of their results in the field of linguistics. The former of these has been discussed widely in the literature (summarised neatly by Bower 2018: 283). This presentation will focus on aspects of the latter, focusing on reproducibility and accessibility to the audience.

This presentation will compare the availability and description of the data and methods used in various Bayesian phylolinguistic studies in terms of reproducibility. It will assess the accessibility of the raw language data used, as well as its binary coding for the Bayesian analysis, and the raw output trees across a wide range of studies. It will also assess the level of detail given of other methodological decisions made by researchers to determine how possible it is, as a reader, to check the data sources and decisions made by researchers throughout their study.

The concepts of reproducibility and replication are of great importance in quantitative research. If results are to be taken as an accurate representation of the world around us, they should, with some room for natural variation, be able to be reproduced or replicated. This necessity for reproducibility extends further than the results of a given study, however. If a study is to be reproducible, its description must be sufficiently detailed that another researcher would be able to undertake it again with the same data (or data collection methods) and same analytical methods. A further benefit to the inclusion of these details and data in publications is that, even where a reader has no intention of physically reproducing the study, they can fully understand and check the full methodology used in the study.

The presentation will find that raw data and methodological details are generally available, though are not always easily accessible to a reader. Common problems include links to online resources which are no longer maintained, or have moved since publication, the availability of cognate sets but no raw data (meaning that the statistical analysis can be repeated but the judgements of cognacy cannot be checked), or the opposite. In a small number of cases, data underpinning a study is not available at all, meaning that none of the conclusions drawn can be checked against their source data. Without access to all data used and methodological decisions made in this research, it is not possible to engage with the conclusions drawn with the level of academic rigor expected in the field.

References:

Bowern, Claire (2018): Computational Phylogenetics. *Annual Review of Linguistics*. Annual Reviews. 4(1). 281–296. doi:10.1146/annurev-linguistics-011516-034142.

Greenhill, Simon J., Paul Heggarty & Russell D. Gray (2020): Bayesian Phylolinguistics. *The Handbook of Historical Linguistics*. Wiley. doi:10.1002/9781118732168.ch11.

Kolipakam, Vishnupriya, Fiona M. Jordan, Michael Dunn, Simon J. Greenhill, Remco Bouckaert, Russell D. Gray & Annemarie Verkerk (2018): A Bayesian phylogenetic study of the Dravidian language family. *Royal Society Open Science*. The Royal Society. 5(3). 171504. doi:10.1098/rsos.171504.

Robbeets, Martine & Remco Bouckaert (2018): Bayesian phylolinguistics reveals the internal structure of the Transeurasian family. *Journal of Language Evolution*. Oxford University Press (OUP). 3(2). 145–162. doi:10.1093/jole/lzy007.

Sagart, Laurent, Guillaume Jacques, Yunfan Lai, Robin J. Ryder, Valentin Thouzeau, Simon J. Greenhill & Johann-Mattis List (2019): Dated language phylogenies shed light on the ancestry of Sino-Tibetan. *Proceedings of the National Academy of Sciences* 116(21). 10317–10322. doi:10.1073/pnas.1817972116.

Zhang, Hanzhi, Ting Ji, Mark Pagel & Ruth Mace (2020): Dated phylogeny suggests early Neolithic origin of Sino-Tibetan languages. *Scientific Reports* 10(1) doi:10.1038/s41598-020-77404-4.

Zhang, Menghan, Shi Yan, Wuyun Pan & Li Jin (2019): Phylogenetic evidence for Sino-Tibetan origin in northern China in the Late Neolithic. *Nature* 569(7754). 112. doi:10.1038/s41586-019-1153-z.

Quantifying the long-term effect of sound change and analogy by forward reconstructing paradigms

Borja C. Herce & Clayton G. S. Marr
(University of Zurich; The Ohio State University)

Keywords: analogy, forward reconstruction, paradigms, verbal inflection, complexity, French

Since the dawn of historical linguistics, sound change and analogy have been recognized as two of the most important factors in the diachronic development of languages. The relationship between these two phenomena has featured in much discourse. For example, it has long been thought that the phonological force of sound change acts to disrupt paradigmatic patterns while analogy opposes it by propagating them. Sturtevant's paradox (Sturtevant 1947), that the irregular action of analogy creates morphological regularity, has long echoed. However, this and other such assertions have largely gone without large scale quantitative analysis on if their predictions are borne out by comparing the results of analogy to what would have happened if sound change alone drove paradigmatic development.

To bridge this gap, we demonstrate a method that deploys computerized forward reconstruction (or CFR; Sims-Williams 2018; Marr and Mortensen 2020) upon an inflected verbal lexicon of an ancestral language to generate the reflexes that would be predicted under the effects of sound change alone. These predicted reflexes are then compared to inflectional forms actually observed in the descendant language. While we do not assert that analogy is responsible for all differences between predicted and actual paradigmatic outcomes, it is by far the most important systematic force at work here. Thus, one can reasonably assume that patterns that emerge from large scale comparison of forward reconstructed and observed outcomes correlate strongly to the effects of analogy.

Latin and French present a rare pair of an ancestral language and a descendant language whose inflectional paradigms enjoy extensive documentation and available computational resources. We annotate the lexical and cell-level cognacy between Latin and French verbs, drawn from the inflected Latin and French verbal lexicons LatInfLexi (Pellegrini and Passarotti 2018) and Vlexique (Bonami et al. 2013) respectively. The phonological development of Latin into French likewise enjoys the availability of DiaCLEF (Marr and Mortensen 2023), recently compiled and highly accurate computerized relative chronology of regular sound changes (or cascade). Using the CFR system DiaSim (Marr and Mortensen 2020), one can predict expected French outcomes from large lists of Latin inputs (in this case verbal paradigms), which can then be compared to actual French outcomes. We then deployed Qumin (Beniamine 2018) to attain calculations of paradigmatic predictive structure and different measures of complexity of the etymologically-paired Latin and French inflectional paradigms and of the hypothetical (i.e. CFRd) expected French paradigms with only sound change.

In our results, analogy and sound change do largely exhibit the effect attributed to them in terms of paradigmatic interpredictability and conceptions of complexity. By comparing Latin to CFRd French, we observe that sound changes significantly increase allomorphy (from 62% to 85%) and the number of inflectional classes (from 107 to 239). By comparing CFR-ed French and actual French paradigms we observe that analogy significantly decreased these (85% to 60%, and 239 to 65). However, surprisingly,

neither sound change nor analogy appear to have a substantial effect upon average conditional entropy between cells (0.218 in Latin, 0.235 in CFRd French, 0.276 in French). This finding is interesting, and demands caution, given the proliferation of conditional entropy as a measure of complexity in the Paradigm Cell Filling Problem literature in recent years (Ackerman and Malouf 2013).

References

- Ackerman, F. and Malouf, R. (2013). Morphological organization: The low conditional entropy conjecture. *Language*, pages 429–464.
- Beniamine, S. (2018). *Classifications flexionnelles. Étude quantitative des structures de paradigmes*. PhD thesis, Université Sorbonne Paris Cité-Université Paris Diderot (Paris 7). <https://qumin.readthedocs.io/index.html>.
- Bonami, O., Caron, G., and Plancq, C. (2013). Flexique: an inflectional lexicon for spoken french. *Technical documentation* [<http://www.llf.cnrs.fr/flexique/documentation.pdf>]
- Marr, C. and Mortensen, D. (2023). Large-Scale Computerized Forward Reconstruction Yields New Perspectives in French Diachronic Phonology. *Diachronica*, 40(2):238–285.
- Marr, C. and Mortensen, D. R. (2020). Computerized Forward Reconstruction for Analysis in Diachronic Phonology, and Latin to French Reflex Prediction. In *Proceedings of LT4HALA 2020-1st Workshop on Language Technologies for Historical and Ancient Languages*, pages 28–36.
- Marr, C. G. S. (2024). A missed regular sound change between Latin and French: velar onset voicing. *Indogermanische Forschungen*, 129:7–43.
- Osthoff, H. and Brugmann, K. (1878). *Morphologische Untersuchungen auf dem Gebiete der indogermanischen Sprachen*, vol. 1. Hirzel.
- Pellegrini, M. and Passarotti, M. (2018). Latinflexi: an inflected lexicon of latin verbs. *Computational Linguistics CLiC-it 2018*, pages 324–329.
- Pope, M. K. (1934). *From Latin to Modern French with especial consideration of Anglo-Norman: Phonology and morphology*. Manchester University Press.
- Sims-Williams, P. (2018). Mechanising Historical Phonology. *Transactions of the Philological Society*, 116(3):555–573.
- Sturtevant, E. H. (1947). *An Introduction to Linguistic Science*. New Haven: Yale University Press.

Time alternates: On ambiguity in NP vs PP temporal expressions in the history of English

Eva Zehentner & Shimon Nakanishi
(University of Zurich & University of Mie)

Keywords: time expressions, syntactic alternations, prepositional patterns, ambiguity, history of English

This paper deals with a particular case of NP- vs PP-variation in the history of English. Specifically, we deal with instances such as (1), in which a temporal adjunct appears with or without a prepositional marker, and investigate their diachronic development in Old English (700-1100) and Middle English (1100-1500).

(1) **(On) that day**, she was born.

Despite sharing certain features with infamous phenomena like the English dative alternation (e.g. Bresnan et al. 2007) or differential argument marking (Seržant et al. 2018), this ‘time alternation’ has received little attention to date. In particular, its relation to major systemic changes in the history of English have not been addressed: Contrary to their Present-Day English counterparts, temporal expressions were still consistently marked with oblique cases in the earliest texts, as illustrated in (2). Concomitant to this decline of the morphological case system, ‘bare’ NP-uses seem to have decreased in favour of PP-patterns at the turn from Old to Middle English (Kniezsa 1986, Lundskaer-Nielsen 1993, Sato 2009). The precise role of the loss of case marking in the development of the time alternation remains unclear, however, and extant studies on the respective periods yield somewhat conflicting results (Nakanishi 2024, Zehentner 2024).

(2) **þy þridan dæge**_{INSTRUMENTAL} heo wæs eft hefigad.
‘the third day she was again distressed’
(Bede 4 B9.6.6 [0423 (21.320.27)])

The present paper tackles this issue. Taking a usage-based perspective, we hypothesise that the history of the time alternation may have been driven by cognitive pressures relating to ambiguity avoidance (e.g. Stefanowitsch 2021). That is, we assume that PP-patterns – although already present in Old English – gained ground over time, especially in contexts where the increasing loss of case marking may have heightened difficulties in determining the syntactic-semantic function of constituents (e.g. McFadden 2002, Zehentner 2024).

We test this assumption by means of logistic regression analysis of several thousand tokens of NP- and PP- time patterns from the *Dictionary of Old English Corpus* (Healey 2004) and the

Penn-Helsinki Parsed Corpus of Middle English (Kroch et al. 2000). Results show that (a) prepositional patterns are more likely when morphological case markers are lacking, and (b) PP-expression is preferred in initial position as in (1), but less strongly so in medial or final position as in (3-4). Furthermore, (c) the probability of PP-use is higher with shorter temporal expressions, while longer expressions as in (5) prefer NP-use.

(3) She was **(on) that day** born.

(4) She was born **(on) that day**.

(5) **(On) the last day that we were in town**, she was born.

(6) **That day** was when she was born.

We argue that these findings corroborate ambiguity effects: In older texts, oblique morphological case marking would have yielded robust cues regarding an expression's function. By contrast, in absence of morphological markers, there is ambiguity between adjuncts and e.g. subject interpretations as in (6), especially with initial, short constituents. In such instances, PP-marking may have been beneficial in reducing the risk of garden paths in incremental sentence processing (e.g. Schlesewsky & Bornkessel 2006).

[This work was supported by JSPS KAKENHI Grant Number JP24K22496.]

References

- Healey, Antonette diPaolo (ed) (2004), *The Complete Corpus of Old English in Electronic Form*. Dictionary of Old English Project. Centre for Medieval Studies, University of Toronto. www.doe.artsci.utoronto.ca
- Kniezsa, Veronika (1986), The progress of the expression of temporal relationships from Old English to Early Middle English, in Dieter Kastovsky, and Aleksander Szwedek (eds), (1986), *Linguistics across historical and geographical boundaries, Vol. 1.*, Berlin: De Gruyter Mouton, 423-436. <https://doi.org/10.1515/9783110856132.423>
- Kroch, Anthony, Ann Taylor and Beatrice Santorini (2000), *Penn-Helsinki Parsed Corpus of Middle English*, second edition. www.ling.upenn.edu/hist-corpora/PPCME2-RELEASE-3/index.html
- Lundskær-Nielsen, Tom (1993), *Prepositions in Old and Middle English*. Amsterdam: Benjamins.
- McFadden, Thomas (2002), The rise of the to-dative in Middle English, in D. Lightfoot (ed), (2002), *Syntactic effects of morphological change*, Oxford: OUP, 107-123. <https://doi.org/10.1093/acprof:oso/9780199250691.003.0006>
- Nakanishi, Shimon (2024), *The use and development of adverbial cases from Old to Early Middle English*. Unpublished doctoral dissertation, Kyoto University.
- Sato, Kiriko (2009), *The development from case-forms to prepositional constructions in Old English prose*. London: Peter Lang.

Schlesewsky, Matthias & Ina Bornkessel (2006), Context-sensitive neural responses to conflict resolution: Electrophysiological evidence from subject-object ambiguities in language comprehension, *Brain Research* 1098(1), 139-152.

<https://doi.org/10.1016/j.brainres.2006.04.080>

Seržant, Ilja and Alena Witzlack-Makarevich (eds) (2018), *Diachrony of differential argument marking*, Berlin: Language Science Press. <https://doi.org/10.5281/zenodo.1219168>

Stefanowitsch, Anatol (ed) (2021), Ambiguity (non-)avoidance in English, *Zeitschrift für Anglistik und Amerikanistik* 69(3). <https://doi.org/10.1515/zaa-2021-2026>

Zehentner, Eva (2024), Alternations (at) that time: NP versus PP time adjuncts in the history of English, *Linguistics Vanguard* 10(s1), 19-28. <https://doi.org/10.1515/lingvan-2023-0054>

Syntax of Coordination: Definite Article Omission in Classical Greek

Francesco Mambrini & Silvia Zampetta
(Università Cattolica del Sacro Cuore, Milan & University of Pavia)

Keywords <Definite Article, Coordination, Ancient Greek, Greek Binomials, Frame Semantics>

This study investigates the omission of the definite article in coordinated noun phrases in Classical Greek (CG), focusing on fully bare (N and N) and fully articulated (Art N and Art N) constructions. Article omission in coordination represents a striking example of syntactic alternation, offering insights into the interaction between syntactic, semantic, and pragmatic factors. While previous research (Zampetta 2023, *forth.*) observed frequent omission of the definite article in coordinated structures involving three unique nouns (sun, moon, sky), this study expands the analysis to a broader range of nouns. Adopting a frame semantics perspective (Fillmore 1975, 1976, 1979), this study examines whether fully bare constructions result from coordination reduction (Haspelmath 2007) or can be classified as idiomatic expressions.

For this study, we rely on the Glaux automatically annotated corpus (Keersmaekers 2021). We extract a random sample of 500 coordinated noun phrases from prose and poetry works of the 5th and 4th century BCE. This subsample is reviewed to identify errors in Glaux's automatic parsing and further annotated for formal, semantic, and pragmatic properties, with the aim of evaluating article omission patterns and assessing the level of idiomaticity of the constructions (Kiparsky 1976, Lambrecht 1984, Masini 2021). Formal constraints include the type of coordination link, separability, modifiability, gender and number of conjuncts, and verb agreement. Semantic constraints focus on the type of nouns (e.g., concrete, abstract, etc.) and their alignment within shared frames. Pragmatic constraints examine contextual factors, such as noun pairs linked to specific texts or literary genres.

Preliminary findings suggest that fully bare constructions (N and N) exhibit significant formal and contextual variability, undermining the hypothesis of fixed formulaicity. For instance, formal variability is evident in the use of different coordination links, such as *hēlios ē selēnē* vs. *hēlios kai selēnē* 'sun or moon' vs. 'sun and moon', (Arist. *Metaph.* 1040a 29, 1093a 12). From a semantic perspective, the meaning of fully bare constructions is entirely compositional, and they appear to be highly productive.

Instead, the data indicate that article omission aligns with coordination reduction, which is regular but governed by semantic and pragmatic constraints. Specifically, the definite article can be omitted when the coordinated nouns belong to the same semantic frame.

These frames may be either:

- Derived from general knowledge, requiring no specific context for interpretation, (e.g., *hēlios kai selēnē kinoûntai kinēseis* 'sun and moon move', Arist. *Cael.* 291b 35)
- (Con)text-dependent, reflecting contextual associations shaped by the discourse or literary conventions of Classical Greek, as *apergázesthai phúsin kai túkhēn* 'nature and fate fulfill' which is a given pair in Plato's *Leges* (889 a 5).

By contrast, fully articulated constructions (Art N and Art N) often reflect semantic frames with looser alignment, such as *hērōs tôn lóphōn kai tôn lókhōn* 'hero of the hills and of the ambushes' (Ar. *Ach.* 575).

This study clarifies how fully bare and fully articulated noun phrase constructions alternate in Classical Greek coordination, highlighting the role of semantic and pragmatic factors in shaping this phenomenon.

References

- Fillmore, Charles J. (1975), An alternative to checklist theories of meaning. In C. Cogen, H. Thompson, G. Thurgood, K. Whistler, and J. Wright (Eds.), *Proceedings of the First Annual Meeting of the Berkeley Linguistics Society*, Berkeley, CA: Berkeley Linguistics Society, 123-131.
- Fillmore, Charles J. (1976), Frame semantics and the nature of language. Origins and evolution of language and speech, *Annals of the New York Academy of Science* 280, 20-32.
- Fillmore, Charles J. (1979), *Innocence: A second idealization for linguistics*, *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 5, 63-76.
- Haspelmath, Martin (2007), Coordination. In T. Shopen (Ed.), *Language typology and syntactic description II: complex constructions*, Cambridge: Cambridge University Press, 1-51.
- Keersmaekers, Alek (2021), The GLAUx corpus: methodological issues in designing a long-term, diverse, multi-layered corpus of Ancient Greek. *Proceedings of the 2nd International Workshop on Computational Approaches to Historical Language Change 2021*, 39-50. Online: Association for Computational Linguistics. doi:10.18653/v1/2021.lchange-1.6
- Kiparsky, Paul (1976), Oral Poetry: Some Linguistics and typological considerations. In B. A. Stolz and R. S. Shannon (Eds.), *Oral literature and the formula*, Ann Arbor: University of Michigan, 73-106.
- Lambrecht, Knud (1984) Formulaicity, Frame Semantics, and Pragmatics in German Binomial Expressions, *Language* 60(4), 753-796.
- Masini, Francesca (2021), Multi-Word Expressions and Morphology. In R. Lieber (Ed.), *The Oxford Encyclopedia of Morphology*, Oxford: Oxford University Press, 1597-1618.
- Zampetta, Silvia (2023), *Sulla presenza alternante dell'articolo definito con i nomina unica in greco antico*, Bologna, Alma Mater Studiorum – Università di Bologna MA Thesis.
- Zampetta, Silvia (forthc.), Analisi corpus-based dell'articolo definito con i *nomina unica* nel greco classico. In M. Aresu, A. Giudice, R. Porqueddu (Eds.), *Dall'Anomia alla Norma: Strategie di codifica dall'antichità ai giorni odierni. Special issue of Rhesis. International Journal of Linguistics, Philology, and Literature* 16(3).

“Optional” Ergative Case Marking in Ranglong: Historical & comparative perspectives
 Hunter Brown, Institut für Sprachwissenschaft, Universität Bern
 Societas Linguistica Europea 58, Bordeaux

Keywords: Alignment, ergativity, diachronic morphosyntax, differential argument marking, language description

Ranglong is a member of the Northwestern (erstwhile “Old Kuki”) group of South-Central Trans-Himalayan (SCTH) languages spoken by 8,000–10,000 people around the Tripuri-Mizo-Assamese border in Northeast India. Like most languages of the Northwestern group, Ranglong is both endangered and highly underdescribed. This study presents a first description of the morphosyntactic alignment patterns that appear in the language, with particular attention to their diachronic development.

Ranglong exhibits two kinds of morphological alignment strategies in transitive clauses, namely nominative-accusative and ergative-absolutive. The latter is restricted to the domain of case marking, as prefixed argument indexes follow a strictly nominative-accusative pattern. A-argument nominals are differentially marked according to several factors. Animacy and definiteness appear to play a role consistent with the predictions of established reference hierarchies (e.g. Haude & Witzlack-Marakevich 2016), namely that referents “further down” in the hierarchies are less-prototypical agents, and so attract ergative marking to a greater extent than those “further up”. Verbal-semantic considerations are also important, as A-arguments of speech-act verbs are likely to take ergative marking regardless of definiteness/animacy/person value. However, none of these factors, alone or in combination, predict the presence or absence of ergative marking in a perfectly systematic way. Speakers also freely add and remove ergative marking in otherwise identical clauses in elicitation contexts. In this sense, the ergative strategy is at least partially pragmatically conditioned and so “optional” in most transitive clauses, in line with the prevailing state of affairs in Trans-Himalayan (DeLancey 2011:210).

Unlike the probabilistic factors described above, non-default syntactic position of the A-argument and/or O-argument does systematically trigger ergative marking, which here presumably serves to disambiguate core participant roles in cases where constituent order does not—another well-known phenomenon in Trans-Himalayan (Chelliah 2017). Ranglong has a default SOV order, but speakers often re-order constituents for pragmatic (information-structural or aesthetic) reasons:

- (1) *hava ha ei-mu hava sakma-te-nu ha=n*
 DIST.DEM ART 3-see DIST.DEM cucumber-DIM-F ART=ERG
 ‘Sakmatenu (lit. Little Cucumber Girl) saw that one’

Here, the A (Sakmatenu) is obligatorily marked with the ergative clitic *=in/=n* because the constituent order of the clause has been changed from SOV to OVS. The same facts hold for the other permutations, though OSV clauses with a marked A-argument are formally indistinguishable from a passive construction (on which see below).

In addition to describing the nominal alignment patterns found in contemporary Ranglong, this study will explore the origins and development of ergative morphology and the ergative construction in general. It will be argued that ergativity arose via reanalysis of a passive construction, and that in contrast to many other SCTH languages, Ranglong's ergative morphology grammaticalized from a defunct ablative marker rather than a locative. Finally, connections between the Ranglong system and the ergative/agentive marking patterns found in closely related SCTH languages (Pangkhuah, Hyow, Mizo, etc.) will be examined in order to probe for further insight into the history of ergative patterning in the family.

Sources

- Chelliah, Shobhana. 2017. Ergativity in Tibeto-Burman. In Jessica Coon, Diane Massam & Lisa Demena Travis (eds.), *The Oxford Handbook of Ergativity*, 924–947. Oxford: Oxford University Press.
- DeLancey, Scott. 2011. “Optional” “ergativity” in Tibeto-Burman languages. *Linguistics of the Tibeto-Burman Area* 34(2). 9–20.
- Haude, Katharina & Alena Witzlack-Makarevich. 2016. Referential hierarchies and alignment: An overview. *Linguistics* 54(3). 433–441.

Insights on the relation between Andoke and Urekena: Discovering linguistic history through comparative and ethnographic lenses

Jacob Menschel, Universität zu Köln

Andoke is an indigenous Amazonian language spoken by some 20-30 individuals in Southern Colombia, along the Caquetá River. With no surviving relatives, it is today considered a linguistic isolate. As pointed out by Jolkesky (2016), Urekena, a language only documented through a 40-item word list gathered by naturalist Johann Natterer in the 1830s (Adelaar & Brijnen 2014), seems to have striking similarities to Andoke. In a similar vein, French linguist Jon Landaburu considers Urekena to be a relative of Andoke (Landaburu 2023: 127), although only by virtue of impressionistic insights (Landaburu p.c.). The talk at hand provides a comparative and ethnographic approach to the distinction between Andoke and Urekena and the linguistic past of Andoke, to better understand the two lects' possible genetic relation.

Firstly, this new perspective draws on Natterer's wordlist, original fieldwork data by the author and a purportedly Andoke wordlist published in Rivet & De Wavrin (1952). By embedding the data within its cultural and ethno-historical context, I intend to achieve a more holistic view on the two language's relations. The findings of this study include an explanation for the genesis of Andoke's rather complex vowel system (Table 1), several consonant correspondences, as well as preliminary observations on the possessive paradigm of all three lects. Moreover, Urekena data can be mapped onto Andoke's word-level prosody, featuring both stress and tone.

As a second methodological approach, this new comparative perspective is amended by ethnographic research jointly conducted with Andoke speakers. Through interviews on speaker impressions, oral history and evaluating the speakers' metalinguistic attitudes towards the speech of the past, this ethnographic data extends the study's scope beyond the mere comparative method.

I thus intend to fill one gap within the web of Amazonian genetic relations. At the same time, I further the cross-disciplinary domain of language history by providing one additional case study. Especially due to the paucity of Urekena data available, this project also shows how to treat very small corpora using the comparative method. This study hence presents one piece in the puzzle of (Northwest) Amazonian language history, as well as valuable input from yet another Amazonian language 'orphan' (Van Gijn et al. 2023; Hantgan & Van Gijn 2024).

	Front		Back unrounded		Back rounded	
High	i	ĩ	u		u	
Mid	e	ẽ	ɣ	ĩ	o	õ
Low	a	ã	ʌ		ɒ	õ

Table 1. Andoke vowel inventory

References

- Adelaar, W. F. H. & H. B. Brijnen. 2014. Johann Natterer's Linguistic Heritage. *Archiv für Völkerkunde* Indigenous Heritage: Johann Natterer, Brazil, and Austria(63–64). 162–183.
- Hantgan, Abbie & Rik Van Gijn. 2024. Workshop: The Limits of the Comparative Method: Innovative Approaches to Understanding Orphan Languages. SLE 57, Helsinki.
- Jolkesky, Marcelo Pinho De Valhery. 2016. *Estudo arqueo-ecolinguístico das terras tropicais Sul-Americanas*. Brasília: Universidade de Brasília Ph.D.
<https://doi.org/10.26512/2016.02.T.21671>.
- Landaburu, Jon. 2023. Chapter 3: Andoke. In Patience Epps & Lev Michael (eds.), *Language Isolates I: Aikanã to Kandozi-Shapra* (Handbücher Zur Sprach- Und Kommunikationswissenschaft 44), vol. 1, 125–172. De Gruyter.
<https://doi.org/10.1515/9783110419405-003>.
- Rivet, Paul & Robert De Wavrin. 1952. La langue andoke. *Journal de la Société des Américanistes* 41(2). 221–233. <https://doi.org/10.3406/jsa.1952.3746>.
- Van Gijn, Rik, Sietze Norder, Leonardo Arias, Nicholas Q. Emlen, Matheus C. B. C. Azevedo, Allison Caine, Saskia Dunn, et al. 2023. The social lives of isolates (and small language families): the case of the Northwest Amazon. *Interface Focus* 13(1). 20220054.
<https://doi.org/10.1098/rsfs.2022.0054>.

Deconstructing the Cáhniks Tree-Model

This talk argues against the orthodox tree-model of the Cáhniks branch of the Caddoan language family instead proposes a dialect chain model through a synthesis of new linguistic reconstructions and the archaeo-historical record. Arikara (arik1262), Skiri Pawnee (skir1238), and South Band Pawnee (sout2966) - the Cáhniks languages as they are known today - are a group of very closely related Northern Caddoan varieties indigenous to the Great Plains of North America. All aspects of the Cáhniks languages are critically understudied, including their diachrony. However, following a forty year long documentation project represented in the works of Parks (1965/2016, 1970/2016, 1976) and Parks & Pratt (2008), it is now possible to move forward with comparative work in Cáhniks. In particular, this study builds on the nearly 400 novel cognate sets and morphological typology reconstructed in Hancock (2025) in reimagining Cáhniks historical relationships as a dialect chain.

Oral historical and archaeological work has shown that the Arikara people split from the Skiri between 1450CE and 1550CE, already after the split of the Skiri-Arikara population from other Cáhniks speaking groups presumed to be ancestral to the present-day South Band communities (Parks 1991, Van de Logt 2023). Even after these splits, all Cáhniks populations would have maintained kinship and political ties with each other in a chain stretching across the Great Plains with the Pawnee communities having the closest relationships. This historical settlement pattern is at odds with the standard model of the Cáhniks tree which has Arikara as a sibling branch to the Pari/Pawnee branch containing South Band and Skiri (Chafe 1979, Parks 1979). Parks (1979: p. 205) presents the most detailed model of this orthodox position with glottochronological dates proposed for the split of the Arikara from the Proto-Pawnee at 500 years ago and the Skiri and the South Band split at 200-300 years ago. Comparative evidence based on shared sound changes and lexicon will be presented to show that while Arikara is indeed innovative to the exclusion of Skiri and South Band, pointing to the orthodox tree model, it also shares innovations with Skiri to the exclusion of South Band, as would be predicted by the archaeo-historical evidence, and that Skiri has innovations not shared by either Arikara or South Band. In addition, it will be claimed that many of the phonological similarities shared between Skiri and South Band are retentions and so would not be used to argue for branching under traditional assumptions in the comparative method. Particular innovations to be discussed will come from changes in segmental inventories, prosodic patterns, the lexicon of nominal roots, and patterns in nominal inflection.

Based on the comparative linguistic evidence in tandem with the archaeo-historical model, it will be argued that Cáhniks is a prime example of a dialect chain in that otherwise potentially branch-defining shared innovations do not follow tree-like distributions but a geographic distribution. This finding opens new avenues for research into Cáhniks linguistic and population histories by highlighting that contact has played a significant role in the development of the contemporary languages, a conclusion which will also impact our understanding of language ecology on the Great Plains.

References

- Chafe, Wallace L. 1979. Caddoan. In Lyle Campbell & Marianne Mithun (eds.), *The Languages of Native America: Historical and Comparative Assessment*, 213–235. Austin: USA: University of Texas Press.
- Hancock, J Drew. 2025. *Reconstructing Proto-Cáhniks Nouns and Nominal Typology*. Basel: Switzerland.
- Parks, Douglas R. 1976. *A Grammar of Pawnee*. New York: USA: Garland.
- Parks, Douglas R. 1979. *The Northern Caddoan Languages: Their Subgrouping and Time Depths*. Nebraska History 60. 197–213.
- Parks, Douglas R. 1991. *Traditional narratives of the Arikara Indians*. Lincoln: USA: University of Nebraska Press.
- Parks, Douglas R. 1970/2016. *Arikara Field Notes and Lexical and Text Database*.
- Parks, Douglas R. 1965/2016. *Pawnee Field Notes and Lexical and Text Database*.
- Parks, Douglas Richard & Lula Nora Pratt. 2008. *A Dictionary of Skiri Pawnee*. Lincoln: USA: University of Nebraska Press.
- Van de Logt, Mark. 2023. *Between the Floods: A History of the Arikaras*. Norman: USA: University of Oklahoma Press.

Interactional challenges with euphemisms: the *wh-diantre/diable* 'the hell' alternation in French historical theatre plays

Jan Fliessbach (University of Potsdam) & Malte Rosemeyer (Free University of Berlin)

Keywords: interrogatives, French, euphemism, corpus, theatre, non-canonical questions

Early Modern French exhibits variation between two types of *wh-the-hell* interrogatives, formed with the nouns *diable* 'devil' (1) and *diantre* (2) (henceforth QD-interrogatives). *Diantre* is considered a euphemistic variant of *diable* (ATILF, 2002).

(1) *Le jodelet ou le maître valet*, Scarron, 1648 [French]

DON JUAN *D'où diable le sais-tu ?*
 from where devil it know.PRS.2SG-you
 'Where the hell do you know that from?'

(2) *Le dépit amoureux*, Molière, 1656

MASCARILLE *D'où diantre a-t-il pu la savoir ?*
 of where devil have.PRS.3SG-T-he can.PTCP it know.INF
 'Where the hell could he have known it from?'

While English *wh-the-hell* interrogatives are well-studied (Bednarek, 2008; Dikken & Giannakidou, 2002; Ippolito, 2024; Martin, 2020; Pesetsky, 1987), French QD-interrogatives have primarily been studied in terms of their impact on *wh*-in situ and inversion (Obenauer, 1994), leaving functional aspects and the *diable-diantre* alternation unexplored (see Celle, Jugnet, & Lansari, 2021, pp. 151–153 for functional equivalents without *diable/diantre* in 21st-century French). Studies on *wh-the-hell* often focus on surprise contexts such as those in (1–2). However, in (3), the QD-interrogative is used not to express surprise but to challenge the felicity of a previous statement by the interlocutor.

(3) *Le pédant joué*, Cyrano de Bergerac, 1654

CHASTEaufort *Que diable voulez-vous que je fasse ?*
 what hell want.PRS.2PL-YOU that I do.PRS.SBJ.1SG
 'What the hell do you want me to do?'

Perdrai-je tous les hommes pour un ?
 lose.FUT.1SG-I all the men for one
 'Shall I lose all men for one?'

We explore the discourse functions of $n = 303$ QD-interrogatives in a corpus of historical theater plays, focusing on the *diable-diantre* alternation. We define a euphemism as "an alternative to a dispreferred expression, in order to avoid possible loss of face" (Allan & Burrige, 1991, p. 11) and hypothesize that euphemistic *diantre* was typical for conflictual, face-threatening discursive contexts, particularly challenge readings. We use Latent Class (LC) Analysis (Lazarsfeld & Henry, 1969), a data-driven, non-manual method, to inductively operationalize the distinction between surprise (LC1) and rhetorical challenge (LC2) readings based on 12 parameters (e.g., grammatical person, TAM marking, syntactic function of the *wh*-form, discourse markers, turn design, clause linkage, etc.). These purely data-induced classes

then help us explain the *diable-diantre* alternation: while the use of *diantre* became more frequent from 1650 to 1750, this increase was restricted to challenge readings (cf. Figure 1). Linear mixed regression modeling confirms that, during this brief period, *diantre* emerged as a competitor to *diable* for expressing challenge readings (cf. Figure 2). We interpret *diantre* as a morphosyntactic fashion driven by audience design (Bell, 1984), with its rise and decline aligning with two broader developments in French theatrical production and reception during the late Ancien Régime: the establishment of the monopolistic Comédie-Française in 1680 and the emergence of public audiences as decisive critical agents in the 1750s (Connors, 2012).

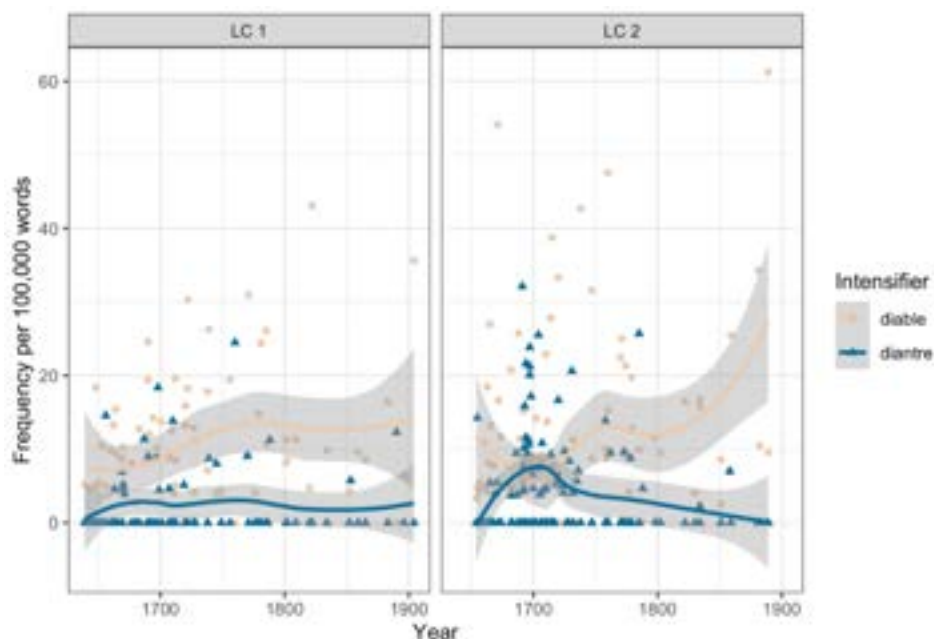


Figure 1. Usage frequency per 100,000 words of *diable*- and *diantre*-wh-interrogatives by year and discourse function (LC1: surprise, LC2: challenge). Points represent theater plays.

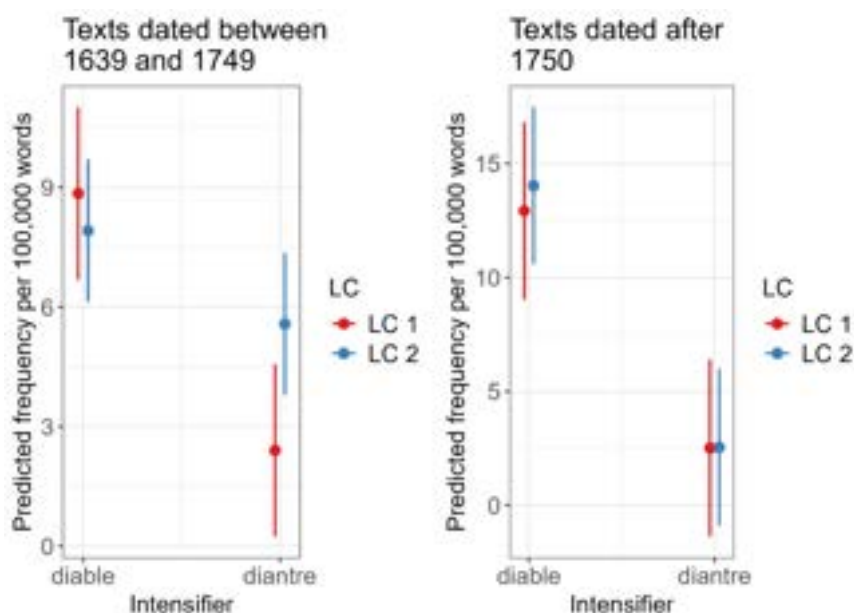


Figure 2. Predicted frequency per 100,000 words of use of *diable*- and *diantre*-wh-interrogatives by latent class (LC1: surprise, LC2: challenge).

Reference List

- Allan, K., & Burridge, K. (1991). *Euphemism and Dysphemism. Language Used as Shield and Weapon*. Oxford: Oxford University Press.
- ATILF (2002). *TLFi : Trésor de la langue Française informatisé*. Université de Lorraine, from CNRS/Analyse et Traitement Informatique de la Langue Française (ATILF): <http://atilf.atilf.fr/tlf.htm>.
- Bednarek, M. (2008). "What the hell is wrong with you?" A corpus perspective on evaluation and emotion in contemporary American pop culture. In A. Mahboob & N. Knight (Eds.), *Questioning Linguistics* (pp. 95–126). Newcastle upon Tyne: Cambridge Scholars.
- Bell, A. (1984). Language style as audience design. *Language in Society*, 13(2), 145–204.
- Celle, A., Jugnet, A., & Lansari, L. (2021). Expressive questions in English and French: *What the hell* versus *Mais qu'est-ce que*. In A. Trotzke & X. Villalba (Eds.), *Expressive Meaning Across Linguistic Levels and Frameworks* (pp. 138–166). Oxford University Press.
- Connors, L. J. (2012). *Dramatic battles in Eighteenth-century France: philosophes, anti-philosophes and polemical theatre*. Liverpool: Liverpool University Press.
- Dikken, M. d., & Giannakidou, A. (2002). From hell to polarity: "Agressively non-d-linked" wh-phrases as polarity items. *Linguistic Inquiry*, 22, 31–61.
- Ippolito, M. (2024). The Hell With Questions. *Journal of Semantics*, 41(1), 53–76.
- Lazarsfeld, P., & Henry, N. (1969). *Latent Structure Analysis*. New York: Houghton Mifflin.
- Martin, J. (2020). Wh-the-hell as a polarity insensitive, speaker-oriented domain restrictor. In J. Rhyne, K. Lamp, N. Dreier, & C. Kwon (Eds.), *Proceedings from Semantics and Linguistic Theory 30* (pp. 334–354). Linguistic Society of America.
- Obenauer, H.-G. (1994). *Aspects de la syntaxe A-barre: Effets d'intervention et mouvements des quantifieurs*, Université Paris 8, Paris.
- Pesetsky, D. (1987). Wh-In-Situ: Movement and Unselective Binding. In E. Reuland & A. t. Meulen (Eds.), *The Representation of (In)definiteness* (pp. 98–129). Cambridge, MA: MIT Press.

It would be a long businessse to discouer, step by step, how he rose vp towards perfection: Studying the diachronic development of an NPN network

Höttecke, Lara
(Osnabrück University)

Keywords: constructional networks, Corpus Linguistics, Diachronic Construction Grammar, NPNs, productivity

NPN (noun-preposition-noun) patterns such as *face to face* and *step by step* are remarkably frequent in language use and occur across many languages and in a variety of syntactic functions (Jackendoff 2008: 15, Sommerer 2022: 369). Their syntactic versatility and supposed form-function mismatch in expressing plurality despite the use of singular nouns have posed a challenge to researchers from various approaches to grammar and thus make the phenomenon a promising avenue for the diachronic study of constructional networks (Matsuyama 2004: 57, Sommerer & Baumann 2021: 102).

Many approaches propose hierarchy, similarity and sequential predictability to be main structuring principles of constructional models (cf. e.g. Langacker 1987, Goldberg 1995, Schmid 2020, and Diessel 2019, 2023). Nevertheless, the question of how to model these constructional networks when applied to specific linguistic phenomena is still controversial (Smirnova & Sommerer 2020: 3-4, Ungerer & Hartmann 2023: 38). Another recurring discussion revolves around the question of when it is possible to posit the existence of a particular construction at all (e.g. Hoffmann 2022: 43). The outstanding variability and productivity of NPN patterns provide a rich basis to address the issue of how we can explain the synchronic structure of constructional networks in general (Sommerer 2022).

Recent works have drawn attention to the diachronic development of constructions in order to better understand the structure of constructional networks (cf. e.g. Bergs & Diewald 2008, Traugott & Trousdale 2013, and Sommerer & Smirnova 2020). In adopting this approach, the talk sheds light on continuities and shifts in the NPN network over time. This explorative study is partly based on a synchronic, token-based study previously done by the author in which 3000 tokens of NPNs taken from the BNC (The British National Corpus 2001) were used to sketch a present-day network of NPN patterns in English. The talk will now extend previous findings by taking into consideration the diachronic development of the NPN network, focusing on the emergence of highly frequent patterns such as *step by step* and their influence on the formal and functional characteristics of NPN patterns. It will be examined which prepositions and nouns are used in the patterns and how the metaphorical dimension of the patterns develops. This will be done based on data retrieved from the English Historical Books Collection through Sketch Engine, which contains texts from the UK and the US between 1473 and 1820 (Sketch Engine 2025). Each of the tokens in the dataset will be analysed for syntactic and semantic features as well as cooccurrence patterns. Patterns of similarity between tokens in the data for several time intervals will be presented to show how diachronically prominent forms and functions have shaped the present-day constructional network of NPNs.

References

- Bergs, Alexander and Gabriele Diewald (eds.) (2008), *Constructions and Language Change*. Berlin, New York: Mouton de Gruyter.
- Diessel, Holger (2019), *The Grammar Network: How Linguistic Structure Is Shaped by Language Use*. Cambridge: Cambridge University Press.
- Diessel, Holger (2023), *The Constructicon: Taxonomies and Networks* (Elements in Construction Grammar). Cambridge: Cambridge University Press.
- Goldberg, Adele. E. (1995), *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago, London: The University of Chicago Press.
- Hoffmann, Thomas (2022), *Construction Grammar: The Structure of English* (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.
- Jackendoff, Ray (2008), Construction after construction and its theoretical challenges. *Language* 84(1). 8–28.
- Langacker, Ronald W. (1987), *Foundations of Cognitive Grammar: Theoretical Prerequisites* (1). Stanford: Stanford University Press.
- Matsuyama, Tetsuya (2004), The N after N construction: A constructional idiom. *English Linguistics (Journal of the English Linguistic Society of Japan)* 21(1). 55–84.
- Schmid, Hans-Jörg (2020), *The Dynamics of the Linguistic System*. Oxford: Oxford University Press.
- Sketch Engine (ed.). Historical collection of the Text Creation Partnership's (TCP). <https://www.sketchengine.eu/historical-collection-eebo-ecco-evans/>. (14 January, 2025).
- Smirnova, Elena and Lotte Sommerer (2020), Introduction: The nature of the node and the network – Open questions in Diachronic Construction Grammar. In L. Sommerer and E. Smirnova (eds.), *Nodes and Networks in Diachronic Construction Grammar* (Constructional approaches to language 27), 1–42. Amsterdam, Philadelphia: John Benjamins Publishing Company.
- Sommerer, Lotte and Andreas Baumann (2021), Of absent mothers, strong sisters and peculiar daughters: The constructional network of English NPN constructions. *Cognitive Linguistics* 32(1). 97–131.
- Sommerer, Lotte (2022), Day to day and night after night: Temporal NPN constructions in Present Day English. In L. Sommerer and E. Keizer (eds.), *English Noun Phrases from a Functional-Cognitive Perspective: Current Issues* (Studies in language companion series volume 221), 363–394. Amsterdam, Philadelphia: John Benjamins Publishing Company.
- The British National Corpus* (2001), Oxford University Computing Services on behalf of the BNC Consortium.
- Traugott, Elizabeth C. and Graeme Trousdale (2013), *Constructionalization and Constructional Changes*. Oxford: Oxford University Press.
- Ungerer, Tobias and Stefan Hartmann (2023), *Constructionist Approaches: Past, Present, Future* (Elements in Construction Grammar). Cambridge: Cambridge University Press.

*The Other Side of Unaccusativity: Old Indo-Aryan *-ti- and *-tu- Derivatives*

Since the Unaccusative Hypothesis' proposal (Perlmutter 1978), several studies have come up with a series of useful tests in order to establish the nature of an intransitive verb, both in modern and in ancient languages (Levin & Rappaport-Hovav 1995, 2005; Sorace 2000; Cennamo 2002; Lazzeroni 2002, 2004). However, with respect to ancient Indo-European languages, these tests deal only with verbal morphology (Lazzeroni 2002, 2004; Romagno 2002; Benedetti 2005).

In this contribution, we aim at showing how the selection of Old Indo-Aryan *-ti-* and *-tu-* suffixes with respect to monovalent verbal roots can be considered as a good diagnostic to establish the compatibility of a verbal root with Unergative or Unaccusative morphology (Rix 1986; Bertocci 2022).

To do so, by combining a semantic and syntactic approach (Sorace 2000; Lazzeroni 2004), we collect all the monovalent verbal roots showing *-ti-* and *-tu-* derivatives (from Whitney 1885; Mayrhofer 1992-2001) coming from three main different stages of Old Indo-Aryan (following Kümmel 2000: ṚV; AVP, AVŚ, AB, PB, JB, GB, ŚBM, TB; MBh, R). All the occurrences of these derivatives are gathered and classified in terms of types and tokens (Berg 2014; Levshina 2019, 2021). Then, we analyze: 1) all their contexts of occurrence, considering their semantic and syntactic features (Keydana 2013; Lowe 2014; Spevak 2022: 62); 2) all the occurrences of the verb they come from, verifying whether they are more likely to be or enter either telic or atelic constructions (being Unaccusative the former, Unergative the latter: Levin & Rappaport-Hovav 1995; Van Valin & Lapolla 1997; Cennamo 2002; Romagno 2002; Sorace 2011). Moreover, following the *Split Intransitivity Hierarchy* model (*SIH*, Sorace 2011), all the verbal roots are grouped into semantic classes.

Preliminary results show that: 1) suffix *-ti-* selects roots creating verbs that are either atelic or more likely to enter atelic constructions (being thus Unergative, Van Valin & LaPolla 1997: 139-141), whilst suffix *-tu-* attaches to roots creating verbs that are either telic or more likely to enter telic constructions (being thus Unaccusative, Van Valin & LaPolla 1997: 139-141); 2) the selection of these suffixes is coherent with the *SIH* model, since verbal roots capable of taking both *-ti-* and *-tu-* are located at the periphrastic semantic classes of the Unaccusative/Unergative prototype: in Early Vedic, the alternating class is *Continuation of a Pre-Existing State* (e.g., ṚV: √*vas* 'dwell' > *vasatí-* f. 'staying', *vá/ástu-* m. 'habitation'), whilst in Middle and Late Vedic it is *Change of State* (e.g., AB: √*jan* 'be born' > *jāti-* f. 'birth', *jántu-* m. 'child'), and in Epic Sanskrit suffix *-ti-* reaches the semantic class of *Change of Place* verbal roots, prototypically unaccusative (e.g., MBh: √*pad* 'go' [+ telic] > *pattí-* f. 'moving'); 3) this scenario suggests an expansion of the Unergativity at the expense of Unaccusativity already at an Early Vedic stage, similarly to what happens, with respect to the verbal domain, in Romance languages (Aranovich 2007; Legendre & Sorace 2003; Cennamo 2008).

References

Aranovich, Raúl (2007), 'Split auxiliary selection from a cross-linguistic perspective', in Aranovich, R. (ed.), *Split auxiliary systems*, Benjamins, Amsterdam.

- Benedetti, Marina (2005), 'Dispersioni formali del medio indoeuropeo', in Costamagna L., Giannini, S. (eds.), *Acquisizione e mutamento di categorie linguistiche. Atti del Convegno della Società Italiana di Glottologia (Perugia, 23-25 ottobre 2003)*, Il Calamo, Roma: 95-119.
- Berg, Thomas (2014), 'On the relation between type and token frequency', in «Journal of Quantitative Linguistics» 21.3: 199-222.
- Bertocci, Davide (2022), 'Root and stems between Indo-European and Latin', in Malzahn, Melanie; Fellner, Hannes A.; Illés, Theresa-Susanna (eds.), *Zurück zur Wurzel. Struktur, Funktion und Semantik der Wurzel im Indogermanischen. Akten der 15. Fachtagung der Indogermanischen Gesellschaft vom 13. Bis 16. September 2016 in Wien*, Verlag, Wiesbaden.
- Cennamo, Michela (2002), 'La selezione degli ausiliari perfettivi in napoletano antico: fenomeno sintattico o sintattico-semantico?', in «Archivio Glottologico Italiano» 87.1: 175-222.
- Cennamo, Michela (2008), 'The rise and development of analytic perfects in Italo-Romance', in *Grammatical change and linguistic theory: The Rosendal papers*: 115-142.
- Keydana, Götz (2013), *Infinitive im R̥gveda: Formen, Funktion, Diachronie*, Brill, Leiden/Boston.
- Kümmel, Martin Joachim (2000), *Das Perfekt im Indoiranischen*, Reichert, Wiesbaden.
- Lazzeroni, Romano (2002), 'Transitivi, causativi e incoativi nel sistema verbale vedico', in «Incontri Linguistici» 25: 105-122.
- Lazzeroni, Romano (2004), 'Inaccusatività indoeuropea e alternanza causativa vedica', in «Archivio Glottologico Italiano» 89: 139-164.
- Legendre, Géraldine & Sorace, Antonella (2003), 'Auxiliaires et intransitivité en français et dans les langues romanes', in Godard, D. (ed.), *Les langues romanes. Problèmes de la phrase simple*, Paris, Editions du CNRS (4): 185-234.
- Levin, Brett & Rappaport-Hovav, Malka (1995), *Unaccusativity. At the Syntax-Lexical Semantic Interface (Linguistic Inquiry Monograph 26)*, MA, Mit Press, Cambridge.
- Levin, Brett & Rappaport-Hovav, Malka (2005), *Argument Realization*, Cambridge University Press, Cambridge.
- Levshina, Natalia (2019), 'Token-based typology and word-order entropy: A study based on Universal Dependencies', in «Linguistic Typology» 23.3: 533-572.
- Levshina, Natalia (2021), 'Corpus-based typology: applications, challenges and some solutions', in «Linguistic Typology» 51: 611-643.
- Lowe, John (2014), 'Review of: Keydana, G. *Infinitive im R̥gveda: Formen, Funktion, Diachronie*', in «Indo-Iranian Journal» 57: 261-298.
- Mayrhofer, Manfred (1992-2001), *Etymologisches Wörterbuch des Altindoiranischen, I-III*, Winter, Heidelberg.
- Perlmutter, David M. (1978), 'Impersonal Passives and the Unaccusative Hypothesis', in *Proceedings of the 4th Annual Meeting of the Berkeley Linguistic Society*: 157-190.
- Rix, Helmut (1986), *Zur Entstehung des urindogermanischen Modusystems*, Institut für Sprachwissenschaft der Universität Innsbruck, Innsbruck.
- Romagno, Domenica (2002), 'Diatesi indoeuropea e verbi di movimento greci: alcune considerazioni sull'intransitività', in «Archivio Glottologico Italiano» 87.2: 163-174.
- Sorace, Antonella (2000), 'Gradients in auxiliary selection with intransitive verbs', in «Language» 76: 859-890.
- Sorace, Antonella (2011), 'Gradience in Split Intransitivity: the End of the Unaccusative Hypothesis?', in «Archivio Glottologico Italiano» 96: 67-86.
- Spevak, Olga (2022), *Nominalization in Latin*, Oxford Academic, Oxford.
- Van Valin, Robert D. Junior, and LaPolla, Randy J. (1997), *Syntax: structure, meaning and function*, Cambridge University Press, Cambridge.
- Whitney, William D. (1885), *The Roots, Verb-Forms and Primary Derivatives of the Sanskrit Language. A Supplement to his Sanskrit Grammar*, Breitkopf and Härtel, Leipzig.

Revaluating marginal clauses in Spanish diachronic syntax

Manuel Delicado Cantero
(Australian National University)

This paper investigates the evolution of a type of declarative finite clause in Spanish introduced by *de como* and *en como* (lit. ‘of how, in how’) (ex. 1,2). It provides an empirical and theoretical account showcasing the importance of marginal phenomena (Manetta 2020, Garachana and Sansiñena 2023).

- (1) dixéronles estas nuevas de cómo la señora de la villa era ya muerta y pesoles a todos muy de coraçón (Zifar, composition 14th c., copy 15th c.)
‘They told them the news how (=that) the lady of the town was already dead and that saddened them deeply’.
- (2) Cada día llegaban noticias de cómo los rebeldes extendiéndose como una invasión de langosta avanzaban sobre Madrid por todos lados (Barea, *La forja de un rebelde*, 1951)
‘Every day we got news about how (=the fact that) the rebels, spreading like a locust invasion, were advancing over Madrid from everywhere’.

This construction is marginal for several reasons. First, it shows *como* (‘how’), rather than the expected complementiser *que* (‘that’), introducing declarative embedded clauses (Iglesias Recuero 2002, Brucart 2009). Second, it displays overt prepositions in argumental embedded clauses much earlier than similar clauses with *que* (‘that’) (Serradilla Castaño 1997, Delicado Cantero 2013). Finally, a relative infrequency and the apparent lack of motivation of the prepositions, especially in the medieval data (Girón Alconchel 1988, Cano Aguilar 1995, Herrero Ruiz de Loizaga 2005, see Delicado Cantero under review), have seen this construction neglected in the literature.

For this study, I extracted all the tokens of *de como* and *en como*—including spelling variations—from Davies’s *Corpus del español* (n=10954 *de como* and 1691 *en como*; search period: 13th c.—20th c.) and selected those potentially declarative following the tests in the existing literature (e.g. Iglesias Recuero 2002). RAE’s *Corpus del Diccionario histórico* was consulted when necessary. To test the prepositional syntax, I explored whether the selecting predicates maintained the prepositions with other categories.

The results show attestations throughout the history of Spanish selected by similar semantic types of predicates, but with important differences. Building on current syntactic analyses of non-interrogative embedded wh-clauses in other languages, e.g. English and German (Legate 2010, Jędrzejowski and Umbach 2023), I argue for a free relative analysis with no wh-movement. Regarding the syntactic motivation of the prepositions, I show that the medieval cases illustrate *de/en* forming a constituent with *como* in a wh-PP (ex. 1), as discussed in Delicado Cantero under review, whereas the post-16th c. evidence instantiates prepositions unambiguously required by the selecting predicates, projecting outside of the wh-phrase including *como* (ex. 2); such analysis is similar to their contemporary counterparts with *que*.

This paper contributes to a better understanding of the evolution of (prepositional) clausal subordination in Spanish and to our knowledge of non-interrogative/exclamative clauses introduced by a wh-category generally. In doing so, it remarks the value of neglected data for theoretical and empirical discussions.

References

- Brucart, José M. (2009), Sobre el valor evidencial de *como* en español, in M. Veyrat Rigat, and E. Serra Alegre (eds), (2009), *La lingüística como reto epistemológico y como acción social. Estudios dedicados al profesor Ángel López García*, Madrid: Arco/Libros, 583–593.
- Cano Aguilar, Rafael (1995), *Sintaxis histórica de la comparación en español. La historia de como*. Sevilla: Universidad de Sevilla.
- Davies, Mark. *Corpus del español*, <www.corpusdelespanol.org/hist-gen/>
- Delicado Cantero, Manuel (2013), *Prepositional clauses in Spanish. A diachronic and comparative syntactic study*, Berlin: De Gruyter.
- Delicado Cantero, Manuel (under review), Reintegrating the peripheral: the syntax of medieval Spanish *en/de como* clauses.
- Garachana, Mar and María Sol Sansiñena (2023), ¿Cómo explicar el cambio lingüístico marginal? Los modelos lingüísticos actuales y el estudio de los *marginalia*, *Boletín de Filología* 58(1), 13–46.
- Girón Alconchel, José Luis (1988), *Las oraciones interrogativas indirectas en español medieval*. Madrid: Gredos.
- Herrero Ruiz de Loizaga, Francisco Javier (2005), *Sintaxis histórica de la oración compuesta en español*. Madrid: Gredos.
- Iglesias Recuero, Silvia (2002), Sobre *cómo* enunciativo en la lengua medieval y clásica., in M.T. Echenique Elizondo, and J.P. Sánchez Méndez (eds), *Actas del V Congreso Internacional de Historia de la Lengua Española*, Madrid: Gredos, 675–690.
- Jędrzejowski, Łukasz and Carla Umbach (eds) (2023), *Non-Interrogative Subordinate Wh-Clauses*. Oxford: Oxford University Press.
- Legate, Julie A. (2010), On how *how* is used instead of *that*. *Natural Language & Linguist Theory* 28: 121–134.
- Manetta, Emily (2020), *Rarely Used Structures and Lesser-Studied Languages. Insights from the Margins*. Abingdon: Routledge.
- Real Academia Española (RAE) (2013), *Corpus del Diccionario histórico de la lengua española* (CDH) <https://apps.rae.es/CNDHE>
- Serradilla Castaño, Ana (1997), *El régimen de los verbos de entendimiento y lengua en español medieval*. Madrid: Ediciones de la Universidad Autónoma de Madrid.

From Wh-element to polarity item: Reanalysis of ‘*what*’ in French and Romance

Marc Olivier
(University of Oxford)

Keywords: restructuring clauses, clitic climbing, reanalysis, Wh-elements, Romance syntax

Clitics cannot normally climb in the presence of complementizers (1), yet this pattern is violated cross-Romance with the verb ‘know’ (2).

- (1) a. * Juan **lo** quiere que Pedro compre. [Spanish]
J. it=wants that P. buys
“Juan wants Pedro to buy it.”
b. * Non **lo** so se fare. [Italian]
not it=know whether do
“I don’t know whether to do it.”
- (2) a. Non **ti** saprei che dire. [Italian]
not you=would.know what say
b. No **et** sabria què dir. [Catalan]
not you=would.know what say
“I wouldn’t know what to tell you.”

Italian *che*, Spanish *qué*, and Catalan *què* ‘what’ permit clitic climbing (CC) when the matrix verb is ‘know’. Other Wh-elements exhibit varying levels of restriction. In Italian and Catalan, *a chi/a qui* ‘to whom’ are marginally acceptable in the presence of CC (Rizzi 1982: 36, Paradís 2019: 300), while CC over *a quíen* yields ungrammatical results in Spanish (Moore 1994: fn3). Acceptability judgements are severely degraded with ‘where’ and ‘how’ (Rizzi 1982; Cardinaletti 2014), whereas ‘why’ and ‘when’ are consistently ruled out.

Cardinaletti (2014: 144) demonstrates that *know what* must be negated to be grammatical (in Italian, Catalan, and Spanish). She argues that the Wh-phrase moves from the VP-layer to the low periphery (Belletti 2004), and that it is an affective polarity item licensed by a sentential negation: (2a) has a meaning equivalent to *non ti saprei dire niente* “I could not tell you anything”, suggesting that *non ... che* operates similarly to *non ... niente* (cf. Paradís 2019: 302 for Catalan). Given that French lost CC, it does not readily show the construction in (2), yet it did so at earlier stages.

Goals:

- (i) to document the construction in (2) in the diachrony of French;
- (ii) to test Cardinaletti’s (2014) hypothesis in the context of reanalysis.

The findings were collected from the *Base de Français Médiéval* and *Frantext*. The diachrony of *know Wh* in French can be split into two:

- **Pre-1500:** the class of Wh-elements that can intervene with CC is larger than that identified in Italian, Catalan, and Spanish today and contains *que*, *à qui*, *où*, and *comment*, but also

‘(to/in/of) which X’. Moreover, and crucially, the dataset contains 14/104 instances where the matrix verb is not negated (3).

- (3) se je **le** seusce u trover. [ca. 1200]
if I it=knew where find
“if I knew where to find it.”

- **Post-1500:** the construction is much more restricted. First, the dataset counts only 1 example with *à qui*, whereas there remain 86 examples with *que*. Other Wh-elements are absent. Second, the negation is obligatory. Third, the embedded infinitive is lexically restricted to *dire* ‘say’, *faire* ‘do’, and *penser* ‘think’.

The findings show that the *know Wh* construction was initially productive and unconstrained, before *que* got reanalysed as a polarity item licenced by negation. The trigger for this reanalysis can indeed be modelled on Cardinaletti’s (2014) hypothesis, although further research is needed to understand why the construction in (3), with ‘where’, was lost from Romance.

References

- Belletti, Adriana (2004). “Aspects of the low IP area”. In: The structure of CP and IP. The cartography of syntactic structures. Ed. by Luigi Rizzi. Vol. 2. Oxford University Press, pp. 16–51.
- Cardinaletti, Anna (2014). “Locality in Restructuring: On Weak wh-Elements, and the OP-Internal ‘Left-Periphery’”. In: Locality. Ed. by Enoch Oladé Aboh, Maria Teresa Guasti, and Ian Roberts. Oxford: Oxford University Press, pp. 138–165.
- Moore, John (1994). “Romance cliticization and relativized minimality”. In: Linguistic Inquiry 25.2, pp. 335–344.
- Paradis, Anna (2019). “L’ascens dels clítics: reestructuració i control”. PhD thesis. Universitat Autònoma Barcelona.
- Rizzi, Luigi (1982). Issues in Italian Syntax. De Gruyter.

The impact of translations from Latin on (pre-)classical Spanish syntax: a case study on Absolute Constructions

Marie Molenaers
(KU Leuven)

Keywords: *translation from Latin, absolute construction, convergence, divergence, Romance syntax*

The extent to which the original Latin source text functions as a linguistic paradigm for (pre-)classical Spanish translators has been a subject of debate in the field of historical linguistics. Some academics argue that syntactic Latinisms are inherently linked to translation activities (Blatt 1957, Azofra 2006). In contrast, others contend that translations do not represent the most Latinate instances (Russell 1985, Pountain 1998, Del Rey Quesada 2019). This controversy requires a more nuanced resolution than a simplistic binary choice. A diachronic investigation of the linguistic interference between source and target texts, enriched by a comparative analysis with non-translated texts of the same authors, can reveal the relevant language-external and language-internal evolutions that shape translators' works.

In a corpus comprising both translated and non-translated texts from the 15th and 17th centuries, this research measures the level of formal and functional convergence or divergence of a typical syntactic Latinism, the Spanish absolute construction (AC), in relation to its mother construction, the Latin *ablativus absolutus*. Two types of ACs are identified: "translated ACs", Spanish ACs (1b: *lo cual dicho* 'this said') in the target texts derived from a Latin equivalent in the source text (1a), and "ex-novo ACs", Spanish newly created absolutes (2: *siendo ellos señores* 'they being masters'), appearing in both translated and non-translated works.

- (1) a. **quo inlato** *infirmatur enumeratione*. (Cicero, *De Inventione*, 86 BC)
'This concluded, the enumeration is invalidated'.
b. **lo cual dicho**, *se destruye la enumeración*. (Cartagena, *La Retórica de Tulio*, 1421/1433)
'That said, the enumeration is destroyed'.
(2) *¿de qué modo [...] puedo yo ser agradecido á los Dioses, siendo ellos señores de todas las cosas?* (Navarrete, *Los Beneficios*, 1629)
'In what way [...] can I show thanks to the Gods, they being masters of all things?'

In the 15th century, the prevalence of translated ACs exhibiting hallmarks of Latinate style (1a/b: past participial predicate, subject-predicate word order) highlights the profound impact of Latin texts. However, a strong divergence from the source texts becomes increasingly evident, as these translated ACs experience a notable decline in frequency over time (50% to 10%), while, concurrently, ex-novo ACs with pioneering features (2: gerundial predicate, predicate-subject word order) emerge in abundance (17th: 90%). In addition, in each century, noteworthy parallels are found between the translations and the original compositions of the same author. Firstly, a striking correspondence is observed between the total number of ACs (15th: 123-117; 17th: 330-323), as well as between their respective percentage of gerundial ex-novo ACs with syntactically and semantically elaborated characteristics (15th: 47%-47%; 17th: 94%-92%). The growing divergence from the source text, in combination with the persistent congruence between translated and non-translated works, is not unexpected, given that the author is subject to the same time-specific interplay of cultural-historical and linguistic dynamics (Molenaers 2024: priming and overlap mechanisms, prevailing conceptualisation styles, the presence of V2).

References

- Azofra, Elena (2006), Problemas de sintaxis y traducción. El caso del participio de presente en español, *Epos* 22, 67–80.
- Blatt, Franz (1957), Latin Influence on European Syntax, *Classica et mediaevalia* 18, 133–178.
- Del Rey Quesada, Santiago (2019), Participial and gerundial clauses in sixteenth-century Spanish prose, *Belgian Journal of Linguistics* 33, 43–81.
- Molenaers, Marie (2024), *The absolute construction in 15th- to 18th- century Spanish and Dutch translations from Latin. On Latin influence and natural language change*, KU Leuven & Universidad de Salamanca dissertation.
- Pountain, Christopher (1998), Learned syntax and the Romance Languages: The ‘Accusative and Infinitive’ construction with declarative verbs in Castilian, *Transactions of the Philological Society* 96(2), 159–201.
- Russell, Peter (1985), *Traducciones y traductores en la península Ibérica (1400-1550)*, Barcelona: Bellaterra.

Mood alternation with epistemic adverbials of doubt and probability in Spanish: A multivariate, cross-dialectal, and diachronic corpus approach

Matti Marttinen Larsson
(University of Gothenburg)

Keywords: mood, language change, usage-based linguistics, corpus linguistics, quantitative linguistics

Despite long-standing efforts towards a comprehensive theory of mood that can capture language-internal and cross-linguistic variation, there is still little consensus regarding the workings of mood. One controversial issue pertains to its determinants: whereas mood is commonly argued to be a meaning-bearing device (Lenz 1920 and Terrell & Hooper 1974, among many others) and whose use "does not depend on other structures except in a loose statistical way" (Bolinger 1974: 465), syntactic and usage-based accounts (e.g., Poplack 1991, 2001, Poplack et al. 2013, 2018 and Torres et al. 2017) emphasize that mood variability can largely be ascribed to frequency effects by which a limited set of matrices are entrenched in their combination with either mood form which, therefore, bears limited semantic meaning.

This study explores whether semantic or non-semantic effects best predict mood selection in Spanish (indicative vs subjunctive). Focusing on epistemic adverbials of doubt and probability (e.g., *tal vez* 'maybe')—a context exhibiting free mood selection, presumably based on speaker intent—the analysis examines two macroregional varieties undergoing change: Mexican and Argentinian Spanish. Using data from the *Corpus diacrónico del español* (-1974) and the *Corpus de Referencia del Español Actual* (1975-2001), over 7,000 occurrences from the 1850s to 2001 were analyzed using Bayesian mixed-effects logistic regression (Bürkner 2021).

The predictors, analyzed in interaction with real time and variety, included semantic factors (epistemic adverbial, grammatical aspect, the presence of modal auxiliaries, polarity) and non-semantic factors (distance between matrix and verb, conditional probability of verb infinitive and verb forms, morphological similarity between indicative and subjunctive counterparts).

Results indicate that imperfective aspect favors the subjunctive mood, suggesting that subjunctive denotes *irrealis* (Fleischmann 1995). However, the semantic effect of mood is constrained by numerous cognitive and usage-determined factors. Non-recency of epistemic adverbials triggers indicative use; subjunctive forms morphologically most similar to their indicative counterparts favor the subjunctive, arguably due to high accessibility; and many subjunctive uses result from entrenched patterns of specific form instantiations learned through repetition (e.g., *quizá sea*). These findings suggest that the semantic contribution of mood is limited and heavily constrained by non-semantic predictors.

Diachronically, there is stable variability in most contexts, with the analyzed varieties largely converging in conditioning factors. However, diachronic change also occurs: first, subjunctive mood usage increases in frequency in both varieties; second, the epistemic adverbial system changes, with some adverbials increasing in frequency while others increasingly fall into disuse, altering the overall distribution of mood-conditioning contexts.

These developments are argued to be interrelated: change in mood selection and increased subjunctive usage result chiefly from concomitant developments in the epistemic adverbial systems of Argentinian and Mexican Spanish. Moreover, these developments differ between varieties, creating different and shifting probabilities of variable contexts across dialects. This leads to the emergence of dialectally diverging substructures steering change and, ultimately, community-specific courses of actualization (cf. Marttinen Larsson 2025).

Acknowledgements: This work was supported by funding from the Swedish Research Council (grant number 2022-00303).

References

- Bolinger, Dwight (1974), One Subjunctive or Two? *Hispania* 57(3). 462. <https://doi.org/10.2307/339180>
- Fleischman, Suzanne (1995), Imperfective and Irrealis, in J. L. Bybee and S. Fleischman (eds.), (1996), *Modality in Grammar and Discourse*, 519–552. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.32.23fle>
- Lenz, Rodolfo (1920), *La oración y sus partes*. Madrid: Centro de estudios históricos.
- Marttinen Larsson, Matti (2025), Pathways of actualization across regional varieties and the real-time dynamics of syntactic change, *Language Variation and Change* aop, 1-27. <https://doi.org/10.1017/S0954394525000055>
- Poplack, Shana (1991), The inherent variability of the French subjunctive, in C. Laeufer and T. A. Morgan (eds.), (1991), *Theoretical Analyses in Romance Linguistics: Selected papers from the Linguistic Symposium on Romance Languages XIX, Ohio State University, April 21–23, 1989*, Amsterdam/Philadelphia: John Benjamins, 235–263.
- Poplack, Shana (2001), Variability, frequency, and productivity in the irrealis domain of French, in J. L. Bybee and P. J. Hopper (eds.), (2001), *Frequency and the Emergence of Linguistic Structure*, Amsterdam: John Benjamins, 405–428. <https://doi.org/10.1075/tsl.45.20pop>
- Poplack, Shana, Allison Lealess & Nathalie Dion (2013), The evolving grammar of the French subjunctive, *Probus* 25(1), 139–195. <https://doi.org/10.1515/probus-2013-0005>
- Poplack, Shana, Rena Torres Cacoullos, Nathalie Dion, Rosane De Andrade Berlinck, Salvatore Digesto, Dora Lacasse & Jonathan Steuck (2018), Variation and grammaticalization in Romance: a cross-linguistic study of the subjunctive, in W. Ayres-Bennett & J. Carruthers (eds.), (2018), *Manual of Romance Sociolinguistics*, Berlin: De Gruyter, 217–252. <https://doi.org/10.1515/9783110365955-009>
- Terrell, Tracy & Joan Hooper (1974), A Semantically Based Analysis of Mood in Spanish, *Hispania* 57(3), 484. <https://doi.org/10.2307/339187>
- Torres, Rena, Dora Lacasse & Michael Johns (2017), El subjuntivo: hacia la rutinización, *Moenia* 23.

When *eat* means *take*:

Accounting for an apparent relationship between applicatives, causatives, and middles

Muhammad Zakaria, University of Bern; David Peterson, Dartmouth College

In this talk, we present a discussion on the source of cognate morphemes that mark applicative and causative in the Southwestern and middle and applicative in the Southeastern group of the South Central Tibeto-Burman (SC, also known as Kuki-Chin) branch.

In SC, an element of approximately the form *-(h)aj* developed from a verbal source into an adposition, *haj* ‘with’; this element also grammaticalized into a marker of applicatives, including comitative and portative ones. In the Southwestern portion of SC (e.g., Khumi), while retaining applicative functions (1a), the element developed causative ones (1b). In the Southeastern part of SC (e.g., Hyow), on the other hand, it developed into a highly productive middle marker ((2a), (2b)) and retained applicative (especially portative) uses (2c).

- (1) a. la^1 a^{1ti^3} $nay^{11b^1}l\ddot{o}^1$ $t^1ko^{11}-hay^{11}=bo^2$
 take etc.AUX then leave-COM.APPL=REAL
 ‘...they took her and so on, and then **left with** her.’ (24.66)
- b. $kh\ddot{a}y^1tewng^{11}=te^1$ $ne^1-hay^{11}-ma^{11}=n\ddot{o}^2$ $nang^1=p\ddot{o}^1$
 pn.LOC=EVID drink-CAUS-FIRST=NMLZ 2SG=also
 ‘They had Kh\ddot{a}y¹tewng drink first, you know...’ (42.93)
- (2) a. $n\acute{a}ng$ $l\acute{a}$ $k\acute{e}y$ $in\acute{i}-hm\acute{u}-\acute{e}y-\acute{a}^?y$ $\acute{o}l\acute{o}$
 2SG and 1SG 1INCLS-see_B-MID-IRR again
 ‘You and I will see each other again.’ [ZM_SK_THP_082015_HYOW_0024_0068]
- b. $\acute{e}yh\ddot{n}\acute{o}^?l\acute{a}ts\acute{a}^?$ $ts\acute{u}\acute{a}=n\acute{i}$ $w\acute{a}t=kh\acute{o}l$ $hy\acute{a}ng-\acute{e}y-hy\acute{o}$
 after.that DIST=FOC cloth=SPL strew_B-MID-VCF
 ‘After that, (he saw) the clothes and such were strewn about there.’
 [ZM_SMTB_SPW_082007_HYOW_0002_0126]
- c. $y\acute{o}yt\acute{a}n=\acute{a}ng$ $\acute{i}-hn\acute{i}^?-\acute{o}k\acute{e}y-\acute{a}l=n\acute{i}$ $\acute{e}y$ $\acute{o}ng\acute{o}=l\acute{a}$
 clothes.line=INE 3S-DU-get.stuck-DEP=TEMP ANAPH.DEM crow=ERG
 $hm\acute{u}^?=n\acute{i}$ $\acute{e}y=\acute{a}ng=k\acute{o}^?$ $hn\acute{g}\acute{a}t$ $gr\acute{i}g\acute{a}k$ $n\acute{u}-\acute{a}k=n\acute{i}$ $ts\acute{e}n-\acute{e}y$
 see_D=TEMP ANAPH.DEM=INE =GEN one octopus CLF-one=FOC flee-PORT.APPL
 ‘When they two (octopi) got stuck in the clothesline, and that crow saw them, he **fled with** one of the octopi.’ [ZM_VSO_HP_122013_HYOW_0041_028]

We have argued elsewhere (Peterson and Zakaria, in press) that the development of middle semantics from the comitative applicative use proceeded via a *portative* applicative (Zúñiga and Creissels 2024) bridging context, as exemplified in (2c). Portative applicatives frequently have autobenefactive nuances: the item, e.g., the octopus in (2c), taken by the agent, e.g., the crow in (2c), benefits the agent. Middle functions thus arise from such autobenefactive nuances.

The main point of this talk, however, is to argue further that the source for this element was the Proto-Tibeto-Burman root **kay*, reconstructed with the semantics ‘eat’ (Matisoff

2015:225). While initial PTB **k* is generally retained at the Proto-SC level, there are a number of SC roots (e.g., PSC **uj* ‘dog’ < PTB **d-k^wəy-n*, PSC **im* ‘house’ < PTB **d-k-y(i/u)m*), where root-initial **k* is either lost or weakened at the PSC level. We suggest that the root VanBik 2009 reconstructs as PSC **aj* ‘eat’ is one such root.

We propose that the original meaning for this root must have actually been ‘take’, a much more likely source than the verbal element ‘eat’ for a comitative applicative marker/adposition. However, it seems reasonable that ‘take’ can shift in meaning to ‘eat’ (cf. English *take medicine*, for a simple example). This shift has happened widely in Tibeto-Burman with this root, including in SC itself, leading to the impression that its original meaning must have been ‘eat’.

The development of middle marking from *eat* is a well-known areal grammaticalization path (see especially Creissels 2010 and Coupe 2018). We suggest that, although there were independent reasons for the marker to drift towards a middle in function in Southeastern SC, perhaps this evolution was further reinforced by the fact that the independent verbal element became the basic verb ‘eat’ in Southeastern languages, much as it did in many other SC languages which preserve it.

Abbreviations

1	first person	DIST	distal	NMLZ	nominalizer
2	second person	DU	dual	PORT	portative
ANAPH	anaphoric	ERG	ergative	REAL	realis
AUX	auxiliary	EVID	evidential	S	single argument
APPL	applicative	FIRST	first		of canonical
B	base stem	FOC	focus		intransitive verb
CLF	classifier	GEN	genitive	SG	singular
COM	comitative	INCL	inclusive	SPL	simulative plural
CAUS	causative	INE	inessive	TEMP	temporal
D	derived stem	IRR	irrealis	VCF	verb complex
DEM	demonstrative	LOC	locative		final
DEP	departative	MID	middle		

References

- Coupe, Alexander R. 2018. Grammaticalization processes in the languages of South Asia. 189-218 in Heike Narrog and Bernd Heine, eds. *Grammaticalization from a typological perspective*. Oxford: Oxford University Press.
- Creissels, Denis. 2010. Benefactive applicative paraphrases: A typological approach. 29-69 in Zúñiga, Fernando and Seppo Kittilä, eds. *Benefactives and malefactives: Typological perspectives and case studies*. Amsterdam: John Benjamins.
- Matisoff, James A. 2015. *The Sino-Tibetan etymological dictionary and Thesaurus*. University of California.
- Peterson, David A and Muhammad Zakaria. In press. Evolutionary pathways between applicative, causative, and middle. *Studies in Language*.
- Van Bik, Kenneth. 2009. *Proto-Kuki-Chin: A reconstructed ancestor of the Kuki-Chin languages*. Berkeley: University of California [STEDT monograph series 8.]
- Zúñiga, Fernando and Denis Creissels. 2024. Applicative constructions: An introductory overview. 3-56 in Fernando Zúñiga and Denis Creissels, eds. *Applicative constructions in the world's languages*. Berlin/Boston: De Gruyter Mouton.

Between code-switching and borrowing: Terminological considerations in Latin technical texts

Quinten Goethals
(Ghent University)

Keywords: <code-switching>, <borrowing>, <bilingualism>, <Latin-Greek contact>, <corpus research>

Code-switching, i.e. the alternation of two or more languages within a single sentence by bilinguals, has emerged as a popular subject in the fields of living languages (see Myers-Scotton 2001) and, more recently, Latin (see Dunkel 2000; Swain 2002; Adams 2003; Elder & Mullen 2019). One key issue in these fields, however, is the persistent lack of consensus on the specific criteria that differentiate code-switching from related phenomena, e.g. borrowing and interference (Bullock & Toribio 2009: 2). Whereas studies on living languages often deem pronunciation to be a decisive factor, in closed-corpus languages, such as Latin and Greek, a different model is highly suggestive (cf. Dickey 2023: 7–16).

To address this issue, this paper will propose a novel framework for distinguishing code-switching and borrowings more accurately in Latin texts. Expanding on Matras (2020: 115–119), it will incorporate a multifactorial approach for classifying words along a continuum from full-code-switching to full-borrowing. These factors include the frequency of a foreign (i.e. Greek) word in Latin (cf. Treffers-Daller 2023), the presence of either foreign or Latin morphological endings (see ex. (1)), the use of Greek or Latin alphabet (see ex. (2)) and spelling adaptation (see ex. (4), from Greek *Sikéloi*). For instance, a word can, despite its Greek ending, still be considered as a borrowing, provided it appears with significant frequency in Latin texts (see ex. (3); cf. Kossmann 2010). By contrast, nonce borrowings are infrequent in the receiving language, despite showing morphological adaptation (Poplack et al. 1988). Contrary to previous models, this framework offers a systematic alternative for a rigid dichotomy between code-switching and borrowing.

(1)a. *contra* ***analogian*** *atque Aristarchum* *est+nixus*
against analogy-ACC.SG and Aristarchus-ACC.SG strive-PRF.3SG.MPASS
“he strove against analogy and Aristarchus” (Varro *ling.* 9,1,1)

b. *etiam* ***syllabam*** *auget*
even syllable-ACC.SG strengthen-PRS.3SG.ACT
“it even strengthens the syllable” (Cassiod. *gramm.* VII 150,6)

(2)a. *nihil* ***ἐτυμολόγον*** *Latinum*
nothing-NOM.SG etymologist-NOM.SG Latin-NOM.SG
“not a Latin etymologist at all” (Varro *ling.* 5,4,29)

b. *si* ***etymologus*** *principia* *uerborum*
if etymologist-NOM.SG elements-ACC.PL word-GEN.PL
postulet *mille*
postulate-SBJ.PRS.3SG a-thousand
“if an etymologist postulates a thousand elements of words” (Varro *ling.* 6,5,39)

(3)a. *trahit* ***Hectora***
pull-PRS.3SG Hector-ACC.SG

“he pulls Hector” (Serv. gramm. IV 461,20)

b. ut	quaestorem	praetorem,	sic	Hectorem
as	quaestor-ACC.SG	praetor-ACC.SG	thus	Hector-ACC.SG

Nestorem

Nestor-ACC.SG

“such as *quaestorem* and *praetorem*, thus *Hectorem* and *Nestorem*” (Varro *ling.* 10,3,70)

(4) <i>Italus</i>	<i>enim rex</i>	<i>Siculorum</i>
Italus-NOM.SG	indeed king-NOM.SG	Sicilian-GEN.PL

“Italus, indeed, the king of the Sicilians” (Serv. 1,1,2)

As a case-study, this framework will be applied to 3 Latin grammatical texts, composed by Servius (*Commentarii in Vergilium*; 4th century CE), Consentius (*De barbarismis et metaplasms*; 5th century CE) and Cassiodorus (*De orthographia*; 6th century CE). This analysis will illustrate how the factors outlined above can be effectively utilised in diachronic research, by connecting these to existing sociolinguistic models of bilingualism (cf. Berruto 2007). The genre of grammatical texts offers several advantages for such an investigation. First, their long and continuous tradition allows for the study of diachronic changes in the use of Greek. Second, the earliest grammatical texts were likely written in bilingual settings, since they relied on Greek sources (cf. Fögen 2011). Indeed, it is widely accepted that bilingualism provides the ideal circumstances for code-switching.

References

- Adams, James (2003), *Bilingualism and the Latin Language*, Cambridge: Cambridge University Press.
- Berruto, Gaetano (2007), Situazioni sociolinguistiche e tutela delle lingue minoritarie. Considerazioni alla luce della Survey Ladins, *Mondo ladino* 31, 37–63.
- Bullock, Barbara, and Toribio, Almeida Jacqueline (2009). Themes in the study of code-switching, in B. E. Bullock, and A. J. Toribio (eds.), *The Cambridge Handbook of Linguistic Code-switching*, Cambridge: Cambridge University Press, 1–18.
- Dickey, Eleanor (2023), *Latin loanwords in ancient Greek: a lexicon and analysis*, Cambridge: Cambridge University Press.
- Dunkel, George (2000), Remarks on code-switching in Cicero’s letters to Atticus, *Museum Helveticum* 57(2), 122–129.
- Elder, Olivia, and Mullen, Alex (2019), *The Language of Roman Letters: Bilingual Epistolography from Cicero to Fronto*, Cambridge: Cambridge University Press.
- Fögen, Thorsten (2011), Latin as a technical and scientific language, in J. Clackson (ed.), *A Companion to the Latin Language*, Hoboken: John Wiley & Sons, Incorporated, 445–463.
- Kossmann, Maarten (2010), Parallel system borrowing: Parallel morphological systems due to the borrowing of paradigms, *Diachronica* 27(3), 459–488.
- Matras, Yaron (2020), *Language contact*, Cambridge: Cambridge University Press.
- Myers-Scotton, Carol, and Bolonyai, Agnes (2001), Calculating speakers: Codeswitching in a rational choice model, *Language in society* 30(1), 1–28.
- Poplack, Sara, Sankoff, David, and Miller, Christopher (1988), The social correlates and linguistic processes of lexical borrowing and assimilation, *Linguistics* 26(1), 47–104.

- Swain, Simon (2002), Bilingualism in Cicero? The evidence of code-switching, in J. N. Adams, M. Janse, and S. Swain (eds), *Bilingualism in Ancient Society. Language Contact and the Written Text*. Oxford: Oxford University Press, 128–167.
- Treffers-Daller, Jeanine (2022), The simple view of borrowing and code-switching, *International Journal of Bilingualism* 29(2), 347–370.

Quantitative Perspectives on Morphological Productivity in Ancient Greek: The Case of Deverbal Nouns

Silvia Zampetta
(University of Pavia)

Keywords <Derivational Morphology, Deverbal Nouns, Ancient Greek, Productivity, Quantitative Analysis>

This study investigates the productivity of six Ancient Greek (AG) deverbal suffixes in a diachronic perspective. The theoretical framework is the corpus-based quantitative approach to morphological productivity proposed by Baayen (1989, 1992, 1993, 2001, 2009; Baayen and Lieber 1991; Baayen and Renouf 1996), which crucially links morphological productivity to the number of *hapax legomena*, i.e., words with a frequency of 1, occurring in a sufficiently large corpus.

Deverbal suffixes have been much studied from an Indo-European perspective (Debrunner 1916, Chantraine 1933, Benveniste 1948, Risch 1974), with a focus on morpho-phonological aspects of word formation and comparison with other ancient Indo-European languages. However, a diachronic and quantitative investigation into the morphological productivity of deverbal AG suffixes is still lacking. This study aims to fill this gap by providing a comprehensive analysis of six AG deverbal suffixes, i.e., *-eía*, *-mos/-mós*, *-sia*, *-sis*, *-tis* and *-tus* (and their allomorphs).

For this study, I created a corpus of AG texts containing approximately four million tokens, divided into four chronological sub-corpora: Archaic, Classical, Hellenistic, and Imperial Greek, which are comparable in terms of token count and literary genres. I then automatically extracted all lemmas containing the target suffixes from the online *Liddell Scott Jones* lexicon provided by Perseus Library, and manually checked the extracted lemmas to ensure that the final list included only the deverbal nouns relevant to this study. The final dataset consists of 1,905 types and 50,637 tokens.

Three productivity measures were employed: (1) the *P* measure, calculating the ratio of *hapax legomena* to tokens for each suffix (Baayen 1989), (2) the *P** measure, assessing the proportion of hapaxes formed with a suffix relative to total corpus hapaxes (Baayen 1993), and (3) the LNRE (*Large Number of Rare Events*) models of word frequency distribution (Baayen 2001, Evert 2004, Evert and Baroni 2006).

The results highlight distinct patterns of suffix productivity across periods and genres. For instance, the suffix *-sis* exhibits consistently high productivity, especially in technical and philosophical texts during the Classical Period, peaking at 75.9%. In contrast, *-mos/-mós* shows continuously increases in productivity from the Archaic Period onward, particularly in comedy and historiography. The suffix *-tus* declines significantly after the Archaic Period and disappears entirely by the Imperial period, remaining primarily associated with epic texts. Moderate productivity is observed for *-eía* and *-sia* across periods and genres.

Finally, Kendall's Tau correlation measure was applied to evaluate Chantraine's (1933) claim that *-mos/-mós* and *-sia* compete with *-sis*. The analysis revealed that only *-mos/-mós* shows a significant negative correlation with *-sis*, supporting the hypothesis of competition between these suffixes. Conversely, no significant correlation was found between *-sis* and *-sia*, while a low degree of correlation emerged between *-sia* and *-eía*. These observations will be further assessed qualitatively to deepen the analysis.

This study provides empirical validation for previous qualitative observations on AG derivational morphology (e.g., Chantraine 1933) and uncovers new patterns of suffix productivity across historical periods and literary genres, contributing to a deeper understanding of the diachronic dynamics of AG morphological processes.

References

- Baayen, Harald (1989), *A corpus-based approach to morphological productivity. Statistical analysis and psycholinguistic interpretation*. Unpublished doctoral dissertation, Vrije Universiteit, Amsterdam.
- (1992), Quantitative aspects of morphological productivity. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 1991*, 109-149. Dordrecht: Kluwer.
- (1993), On frequency, transparency and productivity. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 1992*, 181-208. Dordrecht: Kluwer.
- (2001), *Word-Frequency Distributions*. Dordrecht: Kluwer.
- (2009), Corpus linguistics in morphology: Morphological productivity, In Lüdeling, Anke & Kytö, Merja (eds.), *Corpus Linguistics. An International Handbook*, Vol. 2, 899-919. Berlin: Mouton de Gruyter.
- Baayen, Harald and Lieber, Rochelle (1991), Productivity and English word-formations: a corpus- based study. *Linguistics*, 29, 801-843.
- Baayen, Harald and Renouf, Antoinette (1996), Chronicling the Times: Productive lexical innovations in an English newspaper. *Language*, 72, 69–96.
- Benveniste, Émile (1948), *Noms d'agent et noms d'action en indo-européen*. Paris: Adrien- Maisonneuve.
- Chantraine, Pierre (1933), *La formation des noms en grec ancien*. Paris: Libraire C. Klincksieck.
- Debrunner, Albert (1917). *Griechische Wortbildungslehre*. Heidelberg: Carl Winters Universitätsbuchhandlung.
- Evert, Stephanie (2004), A simple LNRE model for random character sequences. *Proceedings of JADT 2004*, 411-422.
- Evert, Stephanie and Marco Baroni (2006), Testing the extrapolation quality of word frequency models. In *Proceedings of Corpus Linguistics 2005*.
- Risch, Ernst (1974), *Wortbildung der homerischen Sprache*. Berlin/New York: de Gruyter.

The evolutionary dynamics of ergative case

Tai Hong, Alena Witzlack-Makarevich, Balthasar Bickel
(University of Zurich, The Hebrew University of Jerusalem & University of Zurich)

Keywords: ergative case, language evolution, stationary bias, phylogenetic model, Silverstein's hierarchy

A preference towards agents as the causes or initiators of events has been identified as one of the most outstanding characteristics of event cognition in hominids including humans (Brocard et al. 2024 and V. Wilson et al. 2022). Evidence from language comprehension (Egurtzegi et al. 2022, Huber et al. 2024, Isasi-Isasmendi et al. 2024 and Sauppe et al. 2021), gist apprehension (Gerwien and Flecken 2016, Isasi-Isasmendi et al. 2023, Webb, Knott, and MacAskill 2010 and F. Wilson et al. 2011), and referential decisions (Brennan 1995, Du Bois 2003, Haig and Schnell 2016 and Haig, Schnell, and Schiborr 2022) converges in predicting a preference for agents in language. While this prediction is arguably supported by the much higher frequency of agent-first than patient-first languages (Dryer 2013), the evidence remains debated because patient-first orders readily emerge in the absence of language contact, for example in sign languages without literacy (Meir et al. 2017). Here we focus on a different prediction, the “anti-ergative” hypothesis: the agent preference would lead to a stationary bias (Maslova 2000) during linguistic evolution for unmarked coding of agents as ‘nominatives’, i.e., the case used for naming and for the single argument of intransitives (“S”).

While one evolutionary study confirmed this bias (Bickel et al. 2015), more explicit phylogenetic modeling on Pama-Nyungan challenges these findings (Phillips and Bower, 2022). We re-assess the evidence for the preference towards nominatives by more richly articulated phylogenetic models applied to 11 language families including Pama-Nyungan. We compare models that assume the same rates for all families with models that allow for different rates (Jäger and Wahle 2021), and furthermore allow for different regimes in different parts of the trees. Our data coding allows for mixed states for languages with both ergative and non-ergative alignments. The result shows that, across all families, the best-fitting model requires two regimes (log Bayes Factor > 2). The stationary probabilities of the model suggest that in the long run, nominative case marking is expected to be more frequent than ergative case marking. When language evolution reaches stationarity, on average 75% of human languages will be non-ergative, and the proportion of languages in ergative and mixed states is expected to significantly decrease. Thus, the stationary distribution confirms the anti-ergative hypothesis.

Given the relevance of languages with mixed states, we furthermore assess whether transitions between these states follow a hierarchy from inanimate nouns to first and second person pronouns, which was claimed in Silverstein (1976). We translate Silverstein's argument into quantitative models to examine whether a model that places intermediate mixed states in the order of the Silverstein's hierarchy outperforms any other model, where the intermediate states are permuted (e.g. ranking first person between inanimate and animate states). Comparing models with different pathways and orders, we find models that are consistent with the hierarchy fail to show significant advantages over models that violate the hierarchy (log Bayes Factor within [-2,2]). This confirms earlier studies (Bickel, Witzlack-Makarevich, and Zakharko 2014) challenging the relevance of Silverstein's hierarchy for ergative case marking.

References

- Bickel, Balthasar, Alena Witzlack-Makarevich, Kamal K. Choudhary, et al. (2015), The neurophysiology of language processing shapes the evolution of grammar: Evidence from case marking, *PLoS One* 10(8), e0132819.
- Bickel, Balthasar, Alena Witzlack-Makarevich, and Taras Zakharko (2014), Typological evidence against universal effects of referential scales on case alignment, in I. Bornkessel-Schlesewsky, A. Malchukov, and M. Richards (eds), (2014), *Scales: a cross-disciplinary perspective on referential hierarchies*, Berlin: De Gruyter Mouton, 7-43.
- Brennan, Susan E (1995), Centering attention in discourse, *Language and Cognitive processes* 10(2), 137-167.
- Brocard, Sarah et al. (2024), A universal preference for animate agents in hominids, *iScience* 27(6).
- Dryer, Matthew S. (2013), Order of Subject, Object and Verb (v2020.3), in M. S. Dryer and M. Haspelmath (eds), (2013), *The World Atlas of Language Structures Online*, Zenodo, DOI: 10.5281/zenodo.7385533.
- Du Bois, John W (2003), Argument structure, *Preferred argument structure*, 11-60.
- Egurtzegi, Aitor et al. (2022), Cross-linguistic differences in case marking shape neural power dynamics and gaze behavior during sentence planning, *Brain and Language* 230, 105127.
- Gerwien, Johannes and Monique Flecken (2016), First things first? Top-down influences on event apprehension, *38th Annual Meeting of the Cognitive Science Society (CogSci 2016)*, Cognitive Science Society, 2633-2638.
- Haig, Geoffrey and Stefan Schnell (2016), The discourse basis of ergativity revisited, *Language*, 591-618.
- Haig, Geoffrey, Stefan Schnell, and Nils Norman Schiborr (2022), Universals of reference in discourse and grammar: Evidence from the Multi-CAST collection of spoken corpora, in G. Haig, S. Schnell, and F. Seifart (eds), (2022), *Doing corpus-based typology with spoken language corpora: State of the art* 25, University of Hawai'i Press, 141-177.
- Huber, Eva et al. (2024), Surprisal from language models can predict ERPs in processing predicate-argument structures only if enriched by an Agent Preference principle, *Neurobiology of Language* 5(1), 167-200.
- Isasi-Isasmendi, Arrate, Caroline Andrews, et al. (2023), The agent preference in visual event apprehension, *Open Mind* 7, 240-282.
- Isasi-Isasmendi, Arrate, Sebastian Sauppe, et al. (2024), Incremental sentence processing is guided by a preference for agents: EEG evidence from Basque, *Language, Cognition and Neuroscience* 39(1), 76-97.
- Jäger, Gerhard and Johannes Wahle (2021), Phylogenetic typology, *Frontiers in Psychology* 12, e682132.
- Maslova, Elena (2000), A dynamic approach to the verification of distributional universals, *Linguistic Typology* 4, 307-333.
- Meir, Irit et al. (2017), The effect of being human and the basis of grammatical word order: Insights from novel communication systems and young sign languages, *Cognition* 158, 189-207.
- Phillips, Joshua and Claire Bower (2022), Bayesian methods for ancestral state reconstruction in morphosyntax: Exploring the history of argument marking strategies in a large language family, *Journal of Language Evolution* 7(1), 1-15.

- Sauppe, Sebastian et al. (2021), Neural signatures of syntactic variation in speech planning, *PLoS biology* 19(1), e3001038.
- Silverstein, Michael (1976), Hierarchy of features and ergativity, in R.M.W. Dixon (eds), (1976), *Grammatical categories in Australian languages*, Australian Institute of Aboriginal Studies, 112-171
- Webb, Andrew, Alistair Knott, and Michael R MacAskill (2010), Eye movements during transitive action observation have sequential structure, *Acta Psychologica* 133(1), 51-56.
- Wilson, Frances et al. (2011), Rapid extraction of event participants in caused motion events, *Proceedings of the annual meeting of the cognitive science society* 33(33).
- Wilson, Vanessa, Klaus Zuberbühler, and Balthasar Bickel (2022), The evolutionary origins of syntax: Event cognition in nonhuman primates, *Science Advances* 8(25), eabn8464.

General Session : Language acquisition

IconicITA: Iconicity ratings within the Italian lexicon and across L1 and L2 speakers

Andrea Amelio Ravelli, Tommaso Lamarra, Andrea Gregor de Varda, Chiara Saponaro, Beatrice Giustolisi & Marianna Bolognesi

(University of Bologna, University of Bologna, University of Milano-Bicocca, University of Milano-Bicocca, University of Milano-Bicocca, University of Bologna)

Keywords: Iconicity, Ratings, Dataset, Italian, Second Language Acquisition

The concept of iconicity in language has long been considered a marginal topic in Linguistics, that prioritized the idea of arbitrariness of the linking between form and meaning (signifier and signified). However, it is now widely recognized that iconicity is a core component of natural languages (Winter et al. 2023). Lexical resources for operationalizing and measuring word iconicity have been created for many languages, such as English (Perry et al. 2015 and Winter et al. 2023), Spanish (Perry et al. 2015 and Hinojosa et al. 2021), and Japanese (Thompson et al. 2020).

In this study, we propose the first iconicity resource for the Italian language. We present iconicITA, a dataset of iconicity ratings for more than 1 thousand Italian words, collected on a 7-points Likert scale and hereby analyzed and discussed. Ratings have been elicited from two populations: native Italian speakers (L1) and fluent, non-native Italian speakers (L2).

The dataset consists of the words included in ANEW-IT (Montefinese et al. 2014), an existing psycholinguistic database that translate the original stimuli from the English ANEW (Bradley et al. 1999). On this same dataset, ratings measuring various other psycholinguistic variables are currently available, enabling us to perform comprehensive statistical analyses that include now also iconicity. The list of words was expanded with ten control words - five onomatopoeic and five phonosymbolic words - which are expected to elicit high iconicity ratings due to their inherent sound-symbolic relationship with their meanings. Both L1 and L2 participants were recruited online, with L2 participants' linguistic skills verified through the lexITA test (Amenta et al. 2021).

Research questions addressed by this study can be summarized as follows:

1. What is the relationship between speakers' dominant language (L1) and their perceived iconicity in Italian? We hypothesize that Italian L1 speakers will on average provide higher iconicity ratings, given their deeper linguistic experience.
2. Does the perception of iconicity transfer from L1 to L2? We expect a positive finding, as we expect ratings of iconicity from Italian L2 speakers to correlate more strongly with preexisting L1 English ratings (Winter et al. 2023).
3. Can we reproduce in Italian the relationships between iconicity ratings and various psycholinguistic variables previously reported for English? We aim to uncover similar patterns.
4. Does iconicity influence lexical decision times in Italian? Based on Vergallito et al. (2020) and Sidhu et al. (2020), we expect a facilitatory effect of iconicity ratings by L1 speakers on lexical decision data, consistent with triangle models of reading that posit bidirectional paths between orthography, phonology, and semantics (Harm & Seidenberg 2004).

Our primary objective is to offer a comprehensive analysis of iconicity in Italian and its impact on second language acquisition, shedding light on the interplay between language experience, psycholinguistic factors, and lexical processing.

Acknowledgments

AAR, TL, and MMB were funded by ABSTRACTION European Union (GRANT AGREEMENT: ERC-2021-STG-101039777). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

References

Amenta, Simona, Badan, Linda, & Brysbaert, Marc (2021). LexITA: A quick and reliable assessment tool for Italian L2 receptive vocabulary size, *Applied Linguistics*, 42(2), 292–314.

Bradley, Margaret M., & Lang, Peter J. (1999). Affective norms for English words (ANEW): Instruction manual and affective ratings (Vol. 30, No. 1, pp. 25–36). Technical report C-1, the center for research in psychophysiology, University of Florida.

Harm, Michael W., & Seidenberg, Mark S. (2004). Computing the Meanings of Words in Reading: Cooperative Division of Labor Between Visual and Phonological Processes, *Psychological Review*, 111(3), 662–720.

Hinojosa, José A., Haro, Juan, Magallares, S. et al. Iconicity ratings for 10,995 Spanish words and their relationship with psycholinguistic variables, *Behav Res* 53, 1262–1275 (2021). <https://doi.org/10.3758/s13428-020-01496-z>

Montefinese, Maria, Ambrosini, Ettore, Fairfield, Beth, & Mammarella, Nicola (2014). The adaptation of the Affective Norms for English Words (ANEW) for Italian, *Behavior research methods*, 46(3), 887–903.

Perry, Lynn K., Perlman, Marcus, & Lupyan, Gary (2015). Iconicity in English and Spanish and Its Relation to Lexical Category and Age of Acquisition, *PLoS ONE* 10(9), e0137147. doi:10.1371/journal.pone.0137147

Sidhu, David, Vigliocco, Gabriella, & Pexman, Penny M. (2020). Effects of iconicity in lexical decision, *Language and Cognition*, 12(1), 164–181.

Thompson, Arthur L., Akita, Kimi, & Do, Youngah (2020). Iconicity ratings across the Japanese lexicon: A comparative study with English, *Linguistics Vanguard*, 6(1), 20190088.

Vergallito, Alessandra, Petilli, Marco A., & Marelli, Marco (2020). Perceptual modality norms for 1,121 Italian words: A comparison with concreteness and imageability scores and an analysis of their impact in word processing tasks, *Behavior Research Methods*, 52(4), 1599–1616.

Winter, Bodo, Lupyan, Gary, Perry, Lynn K., Dingemanse, Mark, & Perlman, Marcus (2023). Iconicity ratings for 14,000+ English words, *Behavior research methods*, 10.3758/s13428-023-02112-6. Advance online publication.

Linguistics in the language class: On null, overt and postverbal subjects in L2 Italian

Elisa Di Domenico, Federico Piersigilli & Maria Laura Restivo
(Università per Stranieri di Perugia)

Keywords: null pronouns, overt pronouns, Italian, L2 acquisition, negative evidence.

Parameters setting in a second language does not always proceed in a natural fashion, especially when properties at the syntax-discourse interface are involved (Sorace & Filiaci 2006). Near-native L2ers of Italian do not resort to the postverbal subject option (a well-known consequence of the positive setting of the null subject parameter) as native speakers, in answers requiring the identification of the subject, although they produce null subjects (Belletti & Leonini 2004). Moreover, as shown by many studies, these speakers overuse overt subject pronouns and interpret them also as co-referent with subject antecedents, differently from native speakers (Calabrese 1986, Carminati 2002). In this work, we investigate whether a specific kind of negative evidence, namely explicit mention of relevant properties grounded on linguistic theory, influences L2ers choices in anaphora resolution. Additional questions are whether this influence leads to native-like choices and whether it is long-lasting. 22 adult advanced late L2ers of Italian and 26 Controls (adult native speakers of Italian) were administered two forced-choice tasks (Task 1: null pronouns [Np], overt pronouns [Op] or lexical DPs [LEX] for subject/object resumption; Task 2: SV, VS, clefts or reduced clefts [RC] in answers requiring the identification of the subject). A sub-group of 11 L2ers was then taught on the null subject option and its consequences, and then participated in an immediate post-test session. A third step of this ongoing research (not yet accomplished) includes a follow-up test 5 weeks later.

In line with previous literature, pre-test L2ers differ from Controls: they choose less Np ($p \leq .005$) and more Op ($p \leq .05$) to resume subject antecedents; they choose more LEX ($p \leq .05$) and less Op ($p \leq .05$) to resume object antecedents (Figure 1); they choose more SV ($p \leq .005$) and RC ($p \leq .005$) and less VS ($p \leq .005$) (Figure 4).

The pre-test/post-test comparison reveals significant changes. In the post-test, Op choices decrease to resume subject antecedents ($p \leq .05$), and increase to resume object antecedents ($p \leq .005$), while LEX choices decrease ($p \leq .005$) (Figure 2). SV and RC choices decrease ($p \leq .005$, $p \leq .001$), while VS choices increase ($p \leq .005$) (Figure 4).

Finally, in the post-test, L2ers and Controls are alike in most cases, with some differences: L2ers chose more Op than Controls ($p \leq .001$) to resume object antecedents (Figure 3); SV choices are fewer in L2ers than Controls ($p \leq .05$) and RC are still more chosen by L2ers ($p \leq .005$) (Figure 6). To conclude, in the short term, negative evidence as explicit mention of the relevant properties, has a strong effect on L2ers choices in anaphora resolution. This effect does not always lead to native-like choices.

Figure 1. Task 1: L2ers pre-test/ Controls

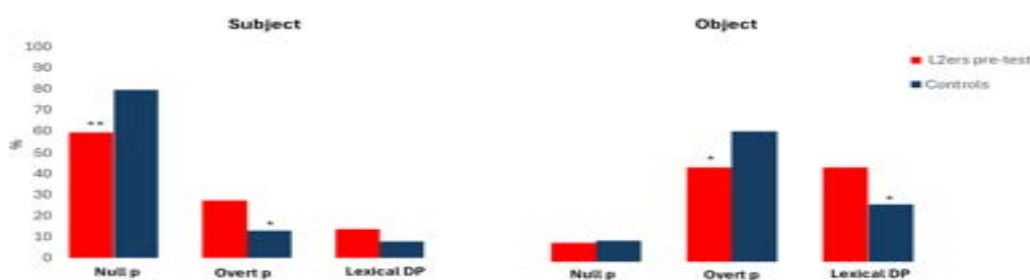


Figure 2. Task 1: L2ers pre-test/ L2ers post-test

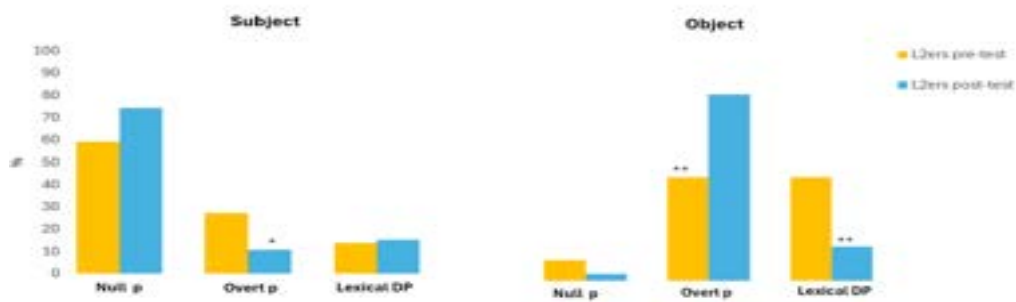


Figure 3. Task 1: L2ers post-test/ Controls

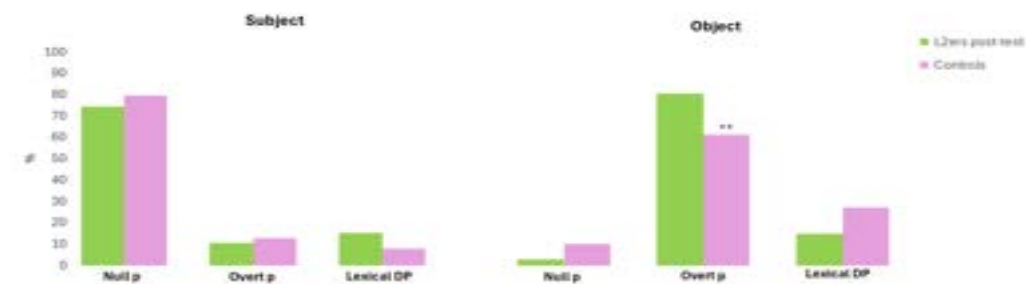


Figure 4. Task 2: L2ers pre-test/ Controls

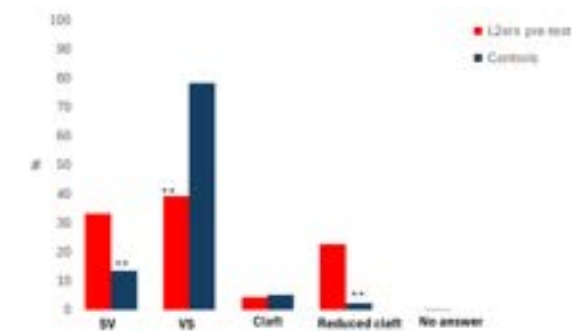


Figure 5. Task 2: L2ers pre-test/ L2ers post-test

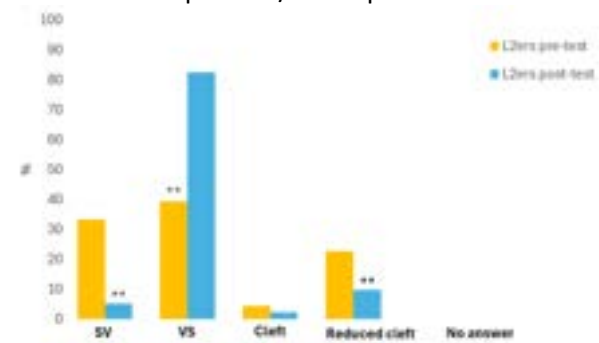
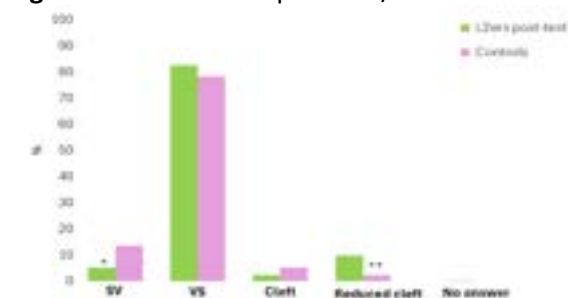


Figure 6. Task 2: L2ers post-test/ Controls



References

- Belletti, Adriana & Leonini, Chiara (2004), Subject inversion in L2 Italian, in S. Forster Cohen *et al.* (eds), (2004), *Eurosla Yearbook 4*, Amsterdam: Benjamins, 95-118.
- Calabrese, Andrea (1986), Pronomina. In N. Fukui *et al.* (eds), (1986), *Papers in Theoretical Linguistics*, Cambridge MA: MIT Working Papers in Linguistics, 1-46.
- Carminati, Maria Nella (2002), *The Processing of Italian Subject Pronouns*. Ph.D. thesis, University of Massachusetts Amherst.
- Sorace, Antonella & Filiaci, Francesca (2006), Anaphora resolution in near-native speakers of Italian, *Second Language Research* 22: 339-368.

The development of probabilistic grammars in spoken EFL: The case of the genitive alternation

Tanguy Dubois
(UCLouvain and KU Leuven)

Whereas most previous research on alternation phenomena in learner language focuses on the relation between learners' L1 and the choice of variant, this study investigates how proficiency level modulates the probabilistic grammar of learners of English as a Foreign Language (EFL) regarding the use of the genitive variants (e.g., *the dog's tail* vs. *the tail of the dog*). First, I collected genitive observations from the Trinity Lancaster Corpus (Gablasova, Brezina & McEnery 2019), which consists of recordings from an official language exam between a native speaker of British English and low-intermediate (B1) to advanced (C2) learners of English from several L1 backgrounds. Non-interchangeable genitive structures such as partitive genitives (e.g., *the cup of wine*) and classifying genitives (e.g., *a children's book*) were excluded. The remaining 2,302 genitive observations were annotated for various probabilistic constraints, such as the length, animacy and definiteness of the constituents (see Rosenbach 2014). The data was then analyzed via mixed-effects logistic regression, where the constraints were allowed to interact with the speaker's proficiency level so as to reveal how proficiency modulates the constraints' influence on the choice of variant. Other sources of variation, such as learners' mother tongue, were controlled for through random effects. Results show that although native speakers and learners are similar, low-proficiency learners are less sensitive to possessor definiteness and possessor animacy. This corpus-based analysis was complemented with a rating experiment, where 25 native speakers and 101 learners from various L1 backgrounds rated the naturalness of either variant in corpus excerpts (see Engel et al. 2022). The ratings were found to correlate with the predictions from the corpus-based model, indicating that learners' intuitions reflect their language production in the corpus. This correlation is slightly weaker for low-proficiency learners, which can partially be explained by low-proficiency learners again being less sensitive to the animacy constraint, which is a very strong predictor of genitive choice in the corpus model. I argue that learners struggle mainly with possessor animacy because it is encoded by the semantics of the noun phrase in a given context rather than by any formal cues, which makes it more difficult to learn its statistical association with the choice of genitive variant.

Keywords: genitive alternation; proficiency level; SLA; probabilistic grammar; rating task experiment

References

- Engel, A., Grafmiller, J., Rosseel, L., & Szmrecsanyi, B. (2022). Assessing the complexity of lectal competence: The register-specificity of the dative alternation after *give*. *Cognitive Linguistics*, 33(4), 727–766.
- Gablasova, D., Brezina, V., & McEnery, T. (2019). The Trinity Lancaster Corpus: Development, description and application. *International Journal of Learner Corpus Research*, 5(2), 126–158.
- Rosenbach, A. (2014). English genitive variation—The state of the art. *English Language and Linguistics*, 18(2), 215–262.

Assessing the statistical validity of multinoun alternation metrics as features of L2-English proficiency

Thomas Gaillat, Cyriel Mallart, Andrew Simpkin, Nicolas Ballier, Rémi Venant, Bernardo Stearns & Jen-Yu Li

(Université Rennes 2, Université Rennes 2, University of Galway, Université Paris Cité, Université Le Mans, University of Galway & Université Rennes 2)

Keywords: multinoun constructions, syntactic alternations, L2 English, L2 microsystems, proficiency

The study of the development of proficiency in writings has been approached with the use of the Complexity, Accuracy, Fluency framework including holistic complexity metrics (Bulté and Housen 2012; Norris and Ortega 2009) and the use of syntagmatic forms mapped to specific functions (Biber, Gray, and Poonpon 2011; Biber et al. 2020; Biber, Larsson, and Hancock 2023). In both cases, lexico-grammatical patterns are counted along the syntagmatic axis and analyzed in terms of correlations with proficiency/developmental stages. For all their benefits, these studies have left paradigmatic production out of the equation. Interestingly, some studies have focused on analysing the internal structure of form-function mappings identifying cases of paradigmatic instability within mappings. This instability has been evidenced in the context of syntactic alternations with the case of the dative alternation (Bresnan et al. 2007) and that of the genitive (Gries and Wulff 2013). However, little work has been conducted on the relationship between syntactic alternations and proficiency (Dubois, Paquot, and Szmrecsanyi 2023; Gaillat et al. 2022). Our proposal is to focus on a microsystem of alternations which poses a problem to learners, i.e., multinoun constructions.

The objective is to model proficiency as a function of how learners of English alternate between genitive, NOUN of NOUN and NOUN-NOUN constructions. To do so, we apply NLP methods to compute probability scores of occurrence of constructions in context. After, pre-processing the EFCAMDAT corpus (Shatz 2020) with UDPipe (Straka, Hajič, and Straková 2016), we extract multinoun constructions with GrewMatch (Guillaume 2021). We train a probabilistic model (multinomial logistic regression with Nnet library in R) on a randomized training subset (80% - $N = 141,528$) based on contextual linguistic features including n-gram tokens and POS, dependencies and their heads. We apply the same extraction strategy, evaluate it against human annotation (F1-Score = 0.72), and test the model's prediction performance on the EFCAMDAT test set as well as the CELVA.Sp (Mallart et al. 2023), an external test set. We use ordinal logistic regression to evaluate associations between probabilities and CEFR-based proficiency (Council of Europe 2018).

Results show that the odds of the constructions' probabilities per text are associated with proficiency (Kruskal-Wallis rank sum test across CEFR: $p < .001$). Table 1 shows the results of the ordinal regression with odds ratios indicating the likelihood of proficiency for every unit increase of the construction probability. Results are consistent across the datasets. The ordinal regression indicates that N OF N probabilities are associated with higher levels, whilst genitives and Noun-Noun probabilities tend to be associated with low levels, in spite of discrepancies between A1 and A2 levels. These new metrics provide measures of paradigmatic dimension in the assessment of L2 writing. Further work is underway to create more syntactic

alternation metrics. The purpose is to exploit them in a general CEFR prediction model.

Alternation microsytem	Structures	EF-CAMDAT		CELVA.Sp	
		Odds ratio	95\% CI	Odds ratio	95\% CI
Multinoun	N2 N1	.993	.993, .993	.989	.976, 1.001
	N2 S N1	.999	.999, 1	.995	.984, 1.005
	N1 OF N2	1.006	1.006, 1.007	1.004	.996, 1.013

Table 1. Results of the ordinal regression for Multinoun constructions as a function of CEFR in two corpora.

Grants

This research was funded by the French *Agence Nationale de la Recherche* as part of the *Analytics for Language Learning* project (A4LL). Grant number: ANR-22-CE38-0015-01

References

- Biber, Douglas, Bethany Gray, and Kornwipa Poonpon. 2011. "Should We Use Characteristics of Conversation to Measure Grammatical Complexity in L2 Writing Development?" *TESOL Quarterly* 45 (1): 5–35. <https://doi.org/10.5054/tq.2011.244483>.
- Biber, Douglas, Bethany Gray, Shelley Staples, and Jesse Egbert. 2020. "Investigating Grammatical Complexity in L2 English Writing Research: Linguistic Description versus Predictive Measurement." *Journal of English for Academic Purposes* 46:100869.
- Biber, Douglas, Tove Larsson, and Gregory R. Hancock. 2023. "The Linguistic Organization of Grammatical Text Complexity: Comparing the Empirical Adequacy of Theory-Based Models." *Corpus Linguistics and Linguistic Theory*, June. <https://doi.org/10.1515/cllt-2023-0016>.
- Bresnan, Joan, Anna Cueni, Tatiana Nikitina, and R. Harald Baayen. 2007. "Predicting the Dative Alternation." In *Cognitive Foundations of Interpretation*, edited by Bouma Gerlof, Irene Kramer, and Joost Swarts, 69–94. Amsterdam: Royal Netherlands Academy of Arts and Sciences.
- Bulté, Bram, and Alex Housen. 2012. *Defining and Operationalising L2 Complexity*. John Benjamins Publishing Company.
- Council of Europe. 2018. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment: Companion Volume with New Descriptors*. Strasbourg: Council of Europe. http://www.coe.int/t/dg4/linguistic/Source/Framework_FR.pdf.
- Dubois, Tanguy, Magali Paquot, and Benedikt Szmrecsanyi. 2023. "Alternation phenomena and language proficiency: the genitive alternation in the spoken language of EFL learners." *Corpus Linguistics and Linguistic Theory* 19 (3). <https://dial.uclouvain.be/pr/boreal/object/boreal:266084>.

- Gaillat, Thomas, Andrew Simpkin, Nicolas Ballier, Bernardo Stearns, Annanda Sousa, Manon Bouyé, and Manel Zarrouk. 2022. "Predicting CEFR Levels in Learners of English: The Use of Microsystem Criterial Features in a Machine Learning Approach." *ReCALL* 34 (2): 130–46. <https://doi.org/10.1017/S095834402100029X>.
- Gries, Stefan Th, and Stefanie Wulff. 2013. "The Genitive Alternation in Chinese and German ESL Learners: Towards a Multifactorial Notion of Context in Learner Corpus Research." *International Journal of Corpus Linguistics* 18 (3): 327–56. <https://doi.org/10.1075/ijcl.18.3.04gri>.
- Guillaume, Bruno. 2021. "Graph Matching and Graph Rewriting: GREW Tools for Corpus Exploration, Maintenance and Conversion." In . <https://hal.inria.fr/hal-03177701>.
- Mallart, Cyriel, Andrew Simpkin, Rémi Venant, Nicolas Ballier, Bernardo Stearns, Jen Yu Li, and Thomas Gaillat. 2023. "A New Learner Language Data Set for the Study of English for Specific Purposes at University Level." In *Proceedings of the 4th Conference on Language, Data and Knowledge - LDK 2023*, 1:281–87. Vienna, Austria. <https://hal.science/hal-04247635>.
- Norris, John M., and Lourdes Ortega. 2009. "Towards an Organic Approach to Investigating CAF in Instructed SLA: The Case of Complexity." *Applied Linguistics* 30 (4): 555–78. <https://doi.org/10.1093/applin/amp044>.
- Shatz, Itamar. 2020. "Refining and Modifying the EFCAMDAT: Lessons from Creating a New Corpus from an Existing Large-Scale English Learner Language Database." *International Journal of Learner Corpus Research* 6 (2): 220–36. <https://doi.org/10.1075/ijlcr.20009.sha>.
- Straka, Milan, Jan Hajič, and Jana Straková. 2016. "UDPipe: Trainable Pipeline for Processing CoNLL-U Files Performing Tokenization, Morphological Analysis, POS Tagging and Parsing." In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, 4290–97. Portorož, Slovenia: European Language Resources Association (ELRA). <https://aclanthology.org/L16-1680>.

Parallelism between the Diachrony and Ontogeny of Formal Nouns: An Analysis based on UG and Third-Factor-Principles

Yoshiki Ogawa
(Tohoku University)

Keywords: <formal nouns, grammaticalization, acquisition, aspect/modality, Japanese>

Formal nouns (FN) are distinguished from lexical nouns in that they are functional categories and bound morphemes. While there are few examples of FNs in English, as illustrated by *one/thing/body of someone/something/somebody*, there are more than 60 FNs in Japanese. Most of them are grammaticalized and can take a verbal or clausal complement, and some of them can be combined with an external verbal copula to make a composite predicate, as in (1) to (3), where *koto* ‘fact’, *tokoro* ‘place’, and *hazu* ‘the two edges of a bow’ are semantically bleached and have gained a grammaticalized aspectual/modal usage:

- (1) a. Taro-wa Tokyo-ni itta **koto**-ga aru. <aspectual *koto*>
Taro-Top Tokyo-to went KOTO-Nom is (Kato 2010)
‘Taro has been to Tokyo.’
b. Taro-wa Hanako-no **koto**-ga suki-da. <modal *koto*>
Taro-Top Hanako-Gen-KOTO-Nom fond-is (Sasaguri 1998)
‘Taro likes Hanako.’
- (2) a. Taro-wa ima kaette-kita **tokoro**-da. <aspectual *tokoro*>
Taro-Top now got.back.home TOKORO-is (Ohori 2001)
‘Taro has just got back home.’
b. Boku-wa ima syokuji-o si-teiru **dokoro**-de-wa-nai. <modal *tokoro*>
I-Top now meal-Acc do-Asp TOKORO-at-Top-Neg
‘I am too busy now to take a meal.’
- (3) a. Taro-wa soko-ni iru **hazu**-da. <modal *hazu*(A)>
Taro-Top there-at is HAZU-is
‘I believe that Taro is there.’
b. Taro-wa soko-ni iru **hazu**-ga-nai. <modal *hazu*(B)>
Taro-Top there-at is HAZU-Nom-Neg
‘There is no way that Taro is there.’

In this presentation, I will demonstrate, on the basis of the Corpus of Historical Japanese, the

Balanced Corpus of Contemporary Written Japanese and the Child Language Data Exchange System, that the grammaticalization processes of these FNs in the history of Japanese and the acquisition processes of these FNs by nine children learning Japanese as their native language are in complete harmony. More specifically, their modality usages always appear(ed) in a later stage than their aspectual usages in both diachrony and ontogeny. In addition, the acquisition of more basic functional elements such as demonstratives and case-particles takes place earlier than that of the grammaticalized usages of these FNs.

Diessel (2011:130-138) argues based on various data across languages that similarities between diachrony and ontogeny play an important role in certain theories of language change, and that there are intriguing parallels in the development of grammatical markers, though the parallels are restricted to semantic and pragmatic features and the developments of morphosyntactic and phonological features are different. On the other hand, Ziegeler (1997) argues that ontogenetic and diachronic routes of grammaticalization correlates in a number of morphosyntactic cases too, such as the *ba* particle in Chinese and grammatical subjects.

My discovery supports Ziegeler's (1997) view (cf. also Slobin 2002, López-Couso 2011). I will argue that the morphosyntactic parallelism is observed for the following three reasons related to the human endowment of Universal Grammar (UG) and/or universal learning mechanism.

- (4) a. Cinque's (2006) cartography (universal functional hierarchy) as part of UG is valid.
- b. Grammaticalization is upward reanalysis along the universal functional hierarchy.
(Roberts 2012)
- c. Third-factor principles and our genetic endowment lie behind the common course in the semantic and grammatical development of our internal grammar.
(cf. Chomsky 2005)

Acknowledgments

This work was financially supported by the Grant-in-Aid for Scientific Research (C) from JSPS (Grant Number: 24K03939) and by the Grant-in-Aid for the Frontier Research Duo (FRiD) of Tohoku University, titled 'Understanding the brain mechanism of language computation from analyzing the communication signals of birds' for the second author.

References

- Chomsky, Noam (2005), Three factors in Language Design, *Linguistic Inquiry* 36, 1-22.
- Cinque, Guglielmo (2006), *Restructuring and Functional Heads: The Cartography of Syntactic Structures, Volume 4*, New York: Oxford University Press.
- Diessel, Holger (2011), Grammaticalization and Language Acquisition, in Heiko Narrog and Bernd

- Heine (eds.), (2011), *The Oxford Handbook of Grammaticalization*, Oxford: Oxford University Press, 130-141.
- Kato, Shigehiro (2010), Nihongo-niokeru Bunpooka to Setsu-Gensyoo [Grammaticalization and Clause Reduction in Japanese], *Asian and African Languages and Linguistics* 5, 35-57.
- Kishimoto, Hideki (2000), Indefinite Pronouns and Overt N-raising, *Linguistic Inquiry* 31, 557-566.
- López-Couso, Mariá José (2011) Developmental Parallels in Diachronic and Ontogenetic Grammaticalization: Existential *There* as a Test Case, *Folia Linguistica* 45, 81-102.
- Ohuri, Toshio (2001), Clause Integration as Grammaticalization: A Case from Japanese *Tokoro*-Complements, in Kaoru Horie and Shigeru Sato (eds.), (2001), *Cognitive-Functional Linguistics in an Easy Asian Context*, Tokyo: Kurosio Publishers, 279-301.
- Roberts, Ian G. (2012), Diachrony and Cartography: Paths of Grammaticalization and the Clause Hierarchy, in Laura Brugè, Anna Cardinaletti, Giuliana Giusti, Nicola Munro, and Cecilia Poletto (eds.), (2012), *Functional Heads: The Cartography of Syntactic Structures, Volume 7*, New York: Oxford University Press, 351-367.
- Sasaguri, Junko (1998), Meishi-ku-no Modaritii-tosite-no 'Koto': 'N-no-koto' to Zyutubu-no Sokan-kara ['Koto' as the Modality of NP: A Perspective from Its Interealtion between 'N-no-koto' and the Predicates], in Yukiko Sasaki Alam (ed.), (1998), *Gengogaku to Nihongo Kyooiku: Jituyooteiki-na Gengo-Kyoiku-no Kotiku-o Mezasite* [Linguistics and Japanese Education: Toward a Construction of Practical Language Education], Tokyo: Kurosio Publishers, 161-176.
- Slobin, Dan I. (2002), Language Evolution, Acquisition and Diachrony: Probing the Parallels, in Talmy Givón & Bertram F. Malle, (eds.), (2002), *The Evolution of Language out of Pre-language*, Amsterdam/Philadelphia: John Benjamins, 375-392.
- Ziegeler, Debra (1997), Retention in Ontogenetic and Diachronic Grammaticalization, *Cognitive Linguistics* 8, 207-241.

Used Corpora:

Corpus of Historical Japanese (CHJ), National Institute for Japanese Language and Linguistics.

<https://clrd.ninjal.ac.jp/chj/chj-reference-en.html>

Balanced Corpus of Contemporary Written Japanese (BCCWJ), National Institute for Japanese Language and Linguistics.

<https://clrd.ninjal.ac.jp/bccwj/>

MacWhinney, Brian (2000), *The CHILDES Project: Tools for Analyzing Talk*, 3rd ed., Lawrence Erlbaum Associates, Mahwah, NJ.

Miyata, Susanne (2012), *CHILDES Japanese Version*.

<https://childes.talkbank.org>

General Session : Language contact

Adverbial clauses in language contact: A study of Vaupés linguistic area

Alexandra Nogina
(The Hebrew University of Jerusalem)

Keywords: adverbial clauses, language contact, areal typology, Amazonian languages, nominalizations

Adverbial clause is a clause that modifies a verb phrase or a main clause, stating the circumstances under which the event of the main clause takes place (Cristofaro 2003: 155, Thompson et al. 2007: 238). The goal of the current research is to establish whether in close language contact the behaviour of adverbial clauses, namely, the strategies according to which the adverbial clauses are formed, may be contact-induced.

The research was carried out on the basis of the Vaupés linguistic area, located in the northwest Amazonia in South America. It hosts speakers of languages belonging to four different families: Tucanoan, Naduhup, Kakua-Nukak, and Arawakan, and is famous for being an area with obligatory multilingualism and established long-term language contact (see Sorensen 1967, Aikhenvald 1999, 2012, Epps 2007, 2009, Stenzel and Gomez-Imbert 2009, and others).

The study was based on a sample of 22 languages, 12 of which are found in the Vaupés area, and 10 are located outside of it, but belong to the same language families as those in the Vaupés. For each language all possible strategies of formation of adverbial clauses were registered, resulting in four major groups: those formed with the use of subordinators (adverbial linkers), cases or adpositions, bare (unmodified) nominalizations, and switch-reference markers.

Comparing the data on languages of the area to that of closely related languages outside of it, I have singled out common features that cannot be explained by genealogical inheritance. Those are the features that I believe to be a result of areal diffusion. For example, adverbial subordination marked with bare nominalizations, while being a genealogical feature of Tucanoan languages, was also attested in the languages of Vaupés belonging to other families, and was absent in languages of the same families outside of the linguistic area. Example (1) features a bare-nominalization-based adverbial clause in Cubeo (← Tucanoan): here it is a v-classifier type of nominalization, a strategy of nominalizing a verb by modifying it with a nominal classifier. Example (2) shows a similar construction in Hup (← Naduhup), both languages belonging to Vaupés linguistic area:

- (1) *põe-wã* *upa-i=wĩ* *ihowe* *da-ḏa-ma*
man-AN.PL dance-ST=CLS:AN.COL sweat come-CAUS-TAEM.I.3PL
'The men are sweating as they dance'. (Chacon and Genetti 2019: 424)
- (2) *yúp* *hũyʔah j'ób* *næn-n'ĩh* *wædó* *nɔh-cud-yé-ay-n'ĩh=mah*
that.ITG after night come-NMZ sun fall-be.inside-enter-INCH-NMZ=REP
tĩh-añ *baʔtĩb'* *ye-yíʔ-ay-áh*
3SG-OBJ spirit enter-TEL-INCH-DECL
'So after this, at nightfall, when the sun was going down, they say, an evil spirit came to her (house)'. (Epps 2008: 852)

In this talk, I will describe the four main strategies of forming an adverbial clause in the languages of Vaupés, and will focus on those instances which are believed to be a result of language contact within the linguistic area. Special attention will be given to nominalizations, whose ability to form adverbial clauses is, on the one hand, characteristic of the Tucanoan languages, and seems to have spread from them to other languages of Vaupés area. On the other hand, this ability has already been described as a universal tendency (see Cristofaro 2003, Schmidtke-Bode and Diessel 2019, and others). I will propose an analysis that shows that the notion of a universal tendency in this instance is, in fact, tightly connected to areal diffusion and language contact.

References:

- Aikhenvald, Alexandra Y. (1999), Areal diffusion and language contact in the Içana-Vaupés basin, north-west Amazonia, in R. M. W. Dixon and A. Y. Aikhenvald (eds.), *The Amazonian languages*, Cambridge / New York: Cambridge University Press, 384–416.
- Aikhenvald, Alexandra Y. (2012), *The languages of the Amazon*, Oxford / New York: Oxford University Press.
- Chacon, Thiago Costa and Genetti, Carol (2019), Grammatical relations in Kubeo, in A. Witzlack-Makarevich and B. Bickel (eds.), *Argument selectors: A new perspective on grammatical relations*, Amsterdam / Philadelphia: John Benjamins, 399–432.
- Cristofaro, Sonia (2003), *Subordination*, Oxford / New York: Oxford University Press.
- Epps, Patience (2007), The Vaupés melting pot: Tukanoan influence on Hup, in A. Y. Aikhenvald and R. M. W. Dixon (eds.), *Grammars in contact: A cross-linguistic typology*, Oxford / New York: Oxford University Press, 267–289.
- Epps, Patience (2008), *A grammar of Hup*, Berlin / New York: Mouton de Gruyter.
- Epps, Patience (2009), Language classification, language contact, and Amazonian prehistory, in *Language and linguistics compass*, 3(2), 581–606.
- Schmidtke-Bode, Karsten and Diessel, Holger (2019), *The typology of non-argument clauses*, manuscript.
- Sorensen, Arthur (1967), Multilingualism in the Northwest Amazon, in *American Anthropologist*, 69(6), 670–684.
- Stenzel, Kristine and Gomez-Imbert, Elsa (2009), Contato linguístico e mudança linguística no noroeste Amazônico: O caso do Kotiria (Wanano), in *Revista da Abralin*, 8(2), 71-100.
- Thompson, Sandra A., Longacre, Robert E. and Hwang, Shin Ja J. (2007), Adverbial clauses, in T. Shopen (ed.), *Language typology and syntactic description, vol.2: Complex constructions*, Cambridge / New York, Cambridge University Press, 237–300.

Exploring the global impact of isolation and contact on linguistic structures through population genetics

Anna Graff, Chiara Barbieri & Balthasar Bickel

(University of Zurich, University of Cagliari & University of Zurich, University of Zurich)

Keywords: language contact, borrowing, diversification, population genetics,

The dynamics of contact vs. isolation has been known to impact the distributions of linguistic traits since the beginning of modern historical linguistics. Contact typically leads to borrowing (Weinreich, 1953; Thomason & Kaufman, 1988; Ranacher et al., 2021), but it can also lead to divergence (Bateson, 1935; François, 2011; Evans, 2019; Epps, 2020; Mansfield et al., 2023). Isolation typically leads to diversification through regular processes of language change (Nichols, 1992).

Yet, despite extensive research our understanding of these effects remains fragmented and unsystematic, as most evidence comes from individual case studies. It is therefore unclear how contact and isolation of human populations impact languages across the world and whether the same linguistic features are subject to the similar effects globally and across different social and demographic conditions. This uncertainty raises broader theoretical questions: are contact effects so widespread that language requires entirely different models of evolution than biology (Gould, 1988; Fracchia & Lewontin, 1999; Gray et al., 2007; Gray & Watts, 2017)? Are they driven by general principles of second language learning (Lupyan & Dale, 2010) or more accidental? To address such questions, we need systematic measures of contact that are independent of language, so that the net effects of contact on language can be systematically studied (Bickel & Nichols, 2006; di Garbo & de Souza, 2023; Sinnemäki et al., 2024).

Here, we propose a novel approach that leverages insights from population genetics as a proxy for (language) contact and isolation. In our two global-scale studies we employ multilevel Bayesian models to integrate genetic and linguistic data from the *Genes and Languages Together (GeLaTo)* database (Barbieri et al., 2022) and two new aggregations of typological databases that minimize inter-feature dependencies (Graff et al., 2025). The first study uses global patterns of genetic admixture as a proxy for contact, and then quantifies the impact of contact between unrelated languages on linguistic structures across different demographic conditions. The second study uses excess genetic homozygosity to infer the degree of local contact vs. isolation at different geographical scales as part of a geostatistical analysis of global structural linguistic diversity.

Both studies robustly confirm that contact favors structural convergence. Isolation further favors structural diversification. Beyond these overall trends, however, we find substantial variation in how individual features respond to contact and isolation. The differences only partly align with expectations from case studies or from second language acquisition studies, in line with other recent challenges of received scholarship (Widmer et al., 2020; Shcherbakova et al., 2023; Koplenig, 2024). Some features also yield divergence under contact, likely as an effect of schismogenesis, i.e. cumulative processes of differentiation.

Our analyses provide a nuanced perspective on how linguistic features interact with the dynamics of contact and isolation, revealing broad global patterns but challenging previous generalizations.

References

- Barbieri, C. et al. (2022). A global analysis of matches and mismatches between human genetic and linguistic histories. *Proceedings of the National Academy of Sciences*, 119(47), e2122084119. <https://doi.org/10.1073/pnas.2122084119>
- Bateson, G. (1935). Culture contact and schismogenesis. *Man*, 35, 178–183.
- Bickel, B., & Nichols, J. (2006). Oceania, the Pacific Rim, and the Theory of Linguistic Areas. *Annual Meeting of the Berkeley Linguistics Society*, 3–15. <https://doi.org/10.3765/bls.v32i2.3488>
- Di Garbo, F., & De Souza, R. N. (2023). A sampling technique for worldwide comparisons of language contact scenarios. *Linguistic Typology*, 27(3), 553–589. <https://doi.org/10.1515/lingty-2022-0005>
- Epps, P. (2020). Amazonian linguistic diversity and its sociocultural correlates. In P. Epps, *Language Dispersal, Diversification, and Contact* (pp. 275–290). Oxford University Press. <https://doi.org/10.1093/oso/9780198723813.003.0016>
- Evans, N. (2019). Linguistic divergence under contact. In M. Cennamo & C. Fabrizio (Eds.), *Historical Linguistics 2015: Selected papers from the 22nd International Conference on Historical Linguistics, Naples, 27-31 July 2015* (pp. 563–592). John Benjamins Publishing Company. <https://doi.org/10.1075/cilt.348.26eva>
- Fracchia, J., & Lewontin, R. C. (1999). Does Culture Evolve? *History and Theory*, 38(4), 52–78. <https://doi.org/10.1111/0018-2656.00104>
- François, A. (2011). Social ecology and language history in the northern Vanuatu linkage: A tale of divergence and convergence. *Journal of Historical Linguistics*, 1(2), 175–246. <https://doi.org/10.1075/jhl.1.2.03fra>
- Gould, S. J. (1988). *An Urchin in the Storm: Essays about Books and Ideas*. Collins Harvill.
- Graff, A., Chousou-Polydouri, N., Inman, D., Skirgård, H., Lischka, M., Zakharko, T., Barbieri, C., & Bickel, B. (2025). Curating global datasets of structural linguistic features for independence. *Scientific Data*, 12(1), 106. <https://doi.org/10.1038/s41597-024-04319-4>
- Gray, R. D., Greenhill, S. J., & Ross, R. M. (2007). The Pleasures and Perils of Darwinizing Culture (with Phylogenies). *Biological Theory*, 2(4), 360–375. <https://doi.org/10.1162/biot.2007.2.4.360>
- Gray, R. D., & Watts, J. (2017). Cultural macroevolution matters. *Proceedings of the National Academy of Sciences*, 114(30), 7846–7852. <https://doi.org/10.1073/pnas.1620746114>
- Koplenig, A. (2024). Still No Evidence for an Effect of the Proportion of Non-Native Speakers on Natural Language Complexity. *Entropy*, 26(11), Article 11. <https://doi.org/10.3390/e26110993>
- Lupyan, G., & Dale, R. (2010). Language Structure Is Partly Determined by Social Structure. *PLoS ONE*, 5(1), e8559. <https://doi.org/10.1371/journal.pone.0008559>
- Mansfield, J., Leslie-O'Neill, H., & Li, H. (2023). Dialect differences and linguistic divergence: A crosslinguistic survey of grammatical variation. *Language Dynamics and Change*, 13(2), 232–276. <https://doi.org/10.1163/22105832-bja10026>
- Nichols, J. (1992). *Linguistic diversity in space and time*. University of Chicago Press.
- Ranacher, P. et al. (2021). Contact-tracing in cultural evolution: A Bayesian mixture model to detect geographic areas of language contact. *Journal of The Royal Society Interface*, 18(181), 20201031. <https://doi.org/10.1098/rsif.2020.1031>
- Shcherbakova, O., Michaelis, S. M., Haynie, H. J., Passmore, S., Gast, V., Gray, R. D., Greenhill, S. J., Blasi, D. E., & Skirgård, H. (2023). Societies of strangers do not speak less complex languages. *Science Advances*, 9(33), eadf7704. <https://doi.org/10.1126/sciadv.adf7704>

- Sinnemäki, K., Garbo, F. D., Souza, R. N. de, & Ellison, T. M. (2024). A typological approach to language change in contact situations. *Diachronica*, 41(3), 379–413. <https://doi.org/10.1075/dia.23029.sin>
- Thomason, S., & Kaufman, T. (1988). *Language Contact, Creolization, and Genetic Linguistics*. University of California Press.
- Weinreich, U. (1953). *Languages in Contact: Findings and Problems*. Mouton Publishers.
- Widmer, M., Jenny, M., Behr, W., & Bickel, B. (2020). Morphological structure can escape reduction effects from mass admixture of second language speakers: Evidence from Sino-Tibetan. *Studies in Language*, 45, 707–752. <https://doi.org/10.1075/sl.19059.wid>

Semantic change and conceptual re-organization within the Kalaallisut demonstrative paradigm

Hilary McMahan
(Independent scholar)

Keywords: Inuit, deixis, typology, cognitive semantics, language change, language contact

This paper investigates the complexities of ongoing semantic change occurring across a grammatical paradigm, involving a combination of internal change, contact-induced change, and sociocultural processes. Language variation and change are motivated by both internal and external factors, and to some degree language contact is always present (Thomason 2010). Thus, challenges occur in understanding the mechanisms and motivations for language change especially as it unfolds. The research presented here sheds light on the multifaceted dynamics of language change through an examination of the demonstrative paradigm of Kalaallisut (iso 639-3: kal; Unangan-Yupik-Inuit). Demonstrative semantics arise from the paradigmatic oppositions made by terms of a system in the localization of referents with respect to an origo (Levinson et al. 2018). However, there exists a lack of in-depth linguistic studies of demonstrative semantics, and even more so of their changes over time.

The Kalaallisut paradigm consists of 11-12 stems given in Table 1, inflected to form nominals, adverbs, or predicatives. A majority encode a typologically-unusual directional semantics with which they indicate a direction or vector from the deictic origo to the referent (Burenhult 2008's 'spatial-coordinate demonstratives'). This directional semantics anchors the system to the geophysical environment surrounding its traditional usage, the west coast of Greenland. However, the system is currently undergoing rapid change, made evident by considerable synchronic variation in meanings and usage with an overall trend toward a reduced system. The focus of the current study is thus on understanding precisely how the system is changing and which factors are motivating these changes. We collected data on demonstrative usage from 33 Kalaallisut speakers in Greenland and Denmark using a combination of structured elicitation tasks.

	Proximal/Medial	Distal
Near speaker	<i>ma-</i> <i>uv-</i>	<i>(im-</i> 'non-visible')
Away from speaker, same level	<i>ik-</i>	<i>av-</i> 'north' <i>qav-</i> 'south'
Down from speaker	<i>kan-</i>	<i>sam-</i> 'seaward, west'
Up from speaker	<i>pik-</i>	<i>pav-</i> 'inland, east'
Inside/outside	<i>qam-</i> 'inside/outside' <i>kig-</i> 'just outside', 'south'	

Table 1: Kalaallisut demonstrative paradigm (McMahan 2022)

Broadly, our findings depict a grammatical system that is undergoing change and reduction through the influence of broader sociocultural transformations in Greenland, including urbanization, multilingualism, and globalization. On a smaller scale, we find evidence of semantic changes such as reinterpretation, metaphorical extension, bleaching and loss. Put together, these individual processes display more systematic changes occurring across the paradigm which illustrate a shift in its conceptual organization. Such a shift mirrors wide-scale sociocultural shifts for Greenlanders, as both linguistic and social factors are clearly shown to drive change in this domain. Further, contact with Danish and English arises as a likely motivation for change. This research therefore helps to illuminate the potential contact effects between demonstrative and other deictic systems and the possibilities for the convergence of conceptual structures across grammatical paradigms (see Ross 1985, 1987, 1996). Finally, this study increases our understanding of change-in-progress, particularly with respect to wide-scale social and economic transformations such as those occurring across the Arctic.

Acknowledgments: This research was supported by the National Science Foundation under grant BCS-2004015 *Doctoral Dissertation Research: Documenting contact and shift in a demonstrative system*.

References

- Burenhult, N. (2008). Spatial coordinate systems in demonstrative meaning. *Linguistic Typology*, 12: 99–142.
- Levinson, S. C., S. Cutfield, M. J. Dunn, N. J. Enfield, and S. Meira. (2018). *Demonstratives in Cross-Linguistic Perspective*. Language, Culture & Cognition 14. Cambridge: Cambridge University Press.
- McMahan, H. (2022). *Spatial Deixis and the Demonstrative System of Kalaallisut*. [Phd thesis, the University of Chicago].
- Ross, M. D. (1985). Current use and expansion of Tok Pisin: Effects of Tok Pisin on some vernacular languages. In Wurm, S. A. & Mühlhäusler, P. (eds.), *Handbook of Tok Pisin (New Guinea Pidgin)*, 539-556. Pacific Linguistics C-70. Australian National University, Canberra.
- Ross, M. D. (1987). A contact-induced morphosyntactic change in the Bel languages of Papua New Guinea. In Laycock, D. C. & Winter, W. (eds.), *A world of language: Papers presented to Professor S. A. Wurm on his 65th birthday*, 583-601. Pacific Linguistics C-100. Australian National University, Canberra.
- Ross, M. D. (1996). Contact-induced change and the comparative method: Cases from Papua New Guinea. In Ross, M. & Durie, M. (eds.), *The Comparative Method Reviewed: Regularity and Irregularity in Language Contact*, 180-217. Oxford University Press.
- Thomason, S. (2010). Contact explanations in linguistics. In Hickey, R. (ed.), *The Handbook of Language Contact*, 31-47. West Sussex: Wiley-Blackwell.

A New Dimension of Tat Redoubled Non-Finite Verb Forms: Affective Concessives?

Murad Suleymanov & Ağacamal Soltanov
(EPHE-PSL / ILARA & independent researcher, Baku)

Keywords: converb, reduplication, concessive clauses, Tat, Iranian

Converbs is not an inherited formation in Tat. As in many Iranian (Soper 1996, Perry 1979, Johanson 2002) and some non-Iranian (Haspelmath 1995) languages, their development is influenced by contact with Turkic. The converb's recent origin is evident from its great morphological variation across Tat, which has seen limited description (Miller 1907, Authier 2012, Mammadova 2018, Suleymanov 2020).

Tat converbs occur in adverbial clauses, expressing an event taking place before (sequence) or at the same time (simultaneity) as that expressed by the main verb. The aim of this study is to examine the development of adverbial clauses termed here “affective concessives” and their areal and typological significance.

A redoubled converb typically indicating simultaneity is, depending on the variety, derived from an infinitive (e.g. *xard-än* <eat-INF> → *xard-än-xard-än* ‘while eating’) or a participle (e.g. *xard-ä* <eat-PTCP> → *xard-ä-xard-ä* ‘while eating’), with the latter more common in varieties where participles have merged with infinitives due to dropping word-final *n*.

Şirvan Tat (Suleymanov 2020: 324)

- (1) *bă magistratura xund-an-xund-an*
LOC graduate_school read₂-INF-read₂-INF
häm=iş bă kor boštamiş bir-um.
also=ADD LOC work starting be₂:PST-1SG
‘While attending graduate school, I started working.’

Qonaqkənd Tat (Grjunberg 1963: 99)

- (2) *ägär män in şir düşir-ä-düşir-ä*
if I this milk boil₂-PTCP-boil₂-PTCP
tiyan=ä sä rah bă xari bül-üm vögür-üm
cauldron=OBL three road LOC ground MOD.leave₁-1SG MOD.take₁-1SG
yəqin duxtär=ä bəstän bi-yar-üm.
maybe girl=OBL PROSP MOD-bring₁-1SG
‘If I set down and lift the cauldron three times while this milk is boiling, I will probably have the girl appear.’

Ərüs-küş–Dağ Quşçu (field data) is the only Tat variety with converbs derived from both redoubled participles and infinitives:

- (3) *tü=riş unjā yaşāmiş bir-än-bir-än*
you=OBL.ADD there living be₂-INF-be₂-INF
üşun=ä zuhun=şun=ä xub amuxt-äni.
they=OBL tongue=POSS:3PL=OBL good learn₂-PROSP:2SG
‘You too will learn their language well while living there.’

- (4) *i injä yä sal-lä yašämiš bir-ä-bir-ä*
 3SG here one year-DIM living be₂-PTCP-be₂-PTCP
injä-ha=rä rä-riz=i=rä äz=tü xub-tä mü-šunaxt-än.
 here-PL=OBL road-trail=POSS:3=OBL from=you good-COMP IPFV-recognise₂-PRS:3
 ‘He, having only lived here for a year, knows every trail here better than you.’

The variety is spoken in an area where *n*-dropping and non-*n*-dropping varieties converge, which initially suggests that the two forms are in complementary distribution. A comparative analysis of indirectly elicited examples reveals that infinitive-derived converbs express events simultaneous with the main verb, making them imperfective by default. Meanwhile, participle-derived converbs are aspect-neutral and convey grounded information implying causality between the clauses. Semantically patterning concessives (cf. English *while*), they constitute a specific type, termed here “affective concessives” or “concessives of reproach and regret”.

The redoubled-participle constructions fulfil the definition of a concessive clause in that they involve a presupposed contrast and counter-expectancy (Givón 2001: 336). Their specificity compared to other concessives is that their apodoses can be isolated and reformulated as “reproachative” clauses (per Van Olmen 2017). With the adverbial clause featuring a non-finite verb, the construction contrasts with previous findings, which suggest that concessives are cross-linguistically situated among adverbial clauses the least prone to morphological downgrading (Hetterle 2015).

Abbreviations

ADD = additive, COMP = comparative, DIM = diminutive, INF = infinitive, LOC = locative, MOD = modal, OBL = oblique, POSS = possessive, PROSP = prospective, PRS = present, PST = past, PTCP = participle, SG = singular

Bibliography

- Givón, Talmy. 2001. *Syntax: An Introduction*, vol. 2. Amsterdam / Philadelphia: Benjamins.
- Grjunberg, Aleksandr. 1963. *Jazyk severoazerbajdžanskix tatov* [The Language of the Tats of Northern Azerbaijan]. Leningrad: Izdatel'stvo Akademii nauk SSSR.
- Haspelmath, Martin. 1995. “Contextual and specialized converbs in Lezgian”. In: Haspelmath, Martin and König, Ekkehard, eds. *Converbs in Cross-Linguistic Perspective: Structure and Meaning of Adverbial Verb Forms - Adverbial Participles, Gerunds*. Berlin / Boston: De Gruyter Mouton, pp. 415–440.
- Hetterle, Katja. 2015. *Adverbial Clauses in Cross-Linguistic Perspective*. Trends in Linguistics Studies and Monographs 289. Berlin / Boston: De Gruyter Mouton.
- Johanson, Lars. 2002. *Structural Factors in Turkic Language Contacts*. London: Curzon.
- Mammadova, Nayiba. 2018. “Eléments de description et documentation du tat de l’Apshéron, langue iranienne d’Azerbaïdjan”. Ph.D. dissertation, Inalco, Paris.
- Miller, Vsevolod. 1907. *Tatskie étyudy II: Opyt grammatiki tatskogo jazyka* [Tat Studies II: Towards a Grammar of Tat]. Trudy po vostokovedeniju 26 [Works on Oriental Studies 26]. Moscow: Tipografija Vjač. Al. Gatsuk.
- Perry, John R. 1979. “Uzbek Influence on Tajik Syntax: The Converb Constructions”. In: Clyne, Paul R. & Hanks, William F. & Hofbauer, Carol L., eds. *The Elements: A Parasession on Linguistic Units and Levels*. Chicago: University of Chicago, pp. 448–461.
- Soper, John. 1996. *Loan syntax in Turkic and Iranian*. Bloomington: Eurolingua.
- Suleymanov, Murad. 2020. *A Grammar of Şirvan Tat*. Beiträge zur Iranistik 46. Wiesbaden: Reichert.
- Van Olmen, Daniël. 2018. “Reproachatives and imperatives”. *Linguistics* 56/1, pp. 115–162.

Coca, jaguar, and beyond: The dynamics of shared words in the languages of Northwest Amazonia

This study investigates the diffusion of culturally significant lexical items across the languages spoken in the Caquetá-Putumayo (CP) region of the Amazon. The CP region is home to unrelated languages (Witotoan, Boran, Arawak families, and the Andoke language isolate), spoken by communities historically connected through trade networks and intermarriage. These connections have led to the formation of culturally homogeneous groups distinct from neighboring peoples to the north, such as those in the Vaupés region, and to the south (e.g., Epps, 2020; Aikhenvald, 2022). The inhabitants of the CP region have long shared a common cultural fabric, setting them apart from more diverse groups found in adjacent areas (e.g., Echeverri, 1997; Eriksen, 2011; Wojtylak, 2020).

Our focus is on basic CP vocabulary related to fauna, flora, and cultural terms, drawing on data from the Languages of Hunter-Gatherers and Their Neighbors Database (Bowern et al., n.d.), complemented by firsthand fieldwork conducted by the authors over 20 years. This study builds on the work of Haynie et al. (2014) and highlights the phenomenon of *Wanderwörter*—lexical items widely shared across Amazonian languages—and the processes through which they spread.

While there is evidence of grammatical convergence among the languages of the CP region, lexical convergence remains limited, consistent with trends in Northwest Amazonia (e.g., Aikhenvald, 2002). Bowern et al. (2011) note that lexical borrowing rates in the region are low, with no language exhibiting more than 4% loans in its basic vocabulary, and most ranging between 1–2%. However, certain cultural terms have spread across the region, albeit not uniformly. For example, the word for the culturally significant plant coca is widely shared across CP languages, appearing as *ípií* (Bora, Boran), *hiibiro* (Ocaina, Witotoan), *hiibi?é* (Resígaro, Arawak), and *hi?píe* (Andoke). The related form in Yagua (Peba-Yagua), south of the CP region, is *xapatij* (Bowern et al., n.d.; cf. Haynie et al., 2014), suggesting that the word coca spread from the south, likely through contact with other language groups. In contrast, tobacco—another culturally important plant—does not exhibit the same widespread borrowing pattern, as each CP language has its own term for it. Similarly, the term *kumu* for "signal drum" is shared across all CP languages and extends beyond the region (Wojtylak, 2019). While the term for maize appears only in Murui-Muina (Witotoan) and Muinane (Boran), suggesting close-knit contact, other shared terms include hummingbird and jaguar. The latter appears in both Witotoan and Boran languages but has distinct forms in Arawak and Andoke languages. Zamponi (2020) discusses similar widespread lexical forms, such as the term for jaguar, showing how these forms spread across different linguistic lineages in Greater Amazonia.

This paper offers new insights into lexical diffusion in the CP region, emphasizing cultural diffusion's role in the spread of terms. By incorporating interviews with speakers, this study provides an emic perspective on how these communities perceive shared words, enriching our understanding of the sociolinguistic dynamics of the Amazon.

References

- Aikhenvald, A. Y. (2002). *Language contact in Amazonia*. Oxford: Oxford University Press.
Aikhenvald, A. Y. (2022). The Amazon Basin: Linguistic Areas and Language Contact. In S. Mufwene & A. M. Escobar (Eds.), *The Cambridge Handbook of Language Contact: Volume 1*:

- Population Movement and Language Change* (pp. 232–260). Cambridge: CUP.
- Bowern, C., Epps, P., Hill, J., & Hunley, K. (n.d.). *Languages of hunter-gatherers and their neighbors: A collection of lexical, grammatical, and other information about languages spoken by hunter-gatherers and their neighbors*.
- Bowern, C., Epps, P., Gray, R., Hill, J., Hunley, K., McConvell, P., & Zentz, J. (2011). Does lateral transmission obscure inheritance in hunter-gatherer languages? *PLoS ONE*, 6(9), e25195.
- Echeverri, J. A. (1997). *The people of the center of the world: A study in culture, history, and orality in the Colombian Amazon* (Ph.D. dissertation). New York: New School for Social Research.
- Epps, P. (2020). Amazonian linguistic diversity and its sociocultural correlates. In M. Crevels & C. P. Muysken (Eds.), *Language dispersal, diversification, and contact* (pp. 275–290). Oxford: Oxford University Press.
- Eriksen, L. (2011). *Nature and culture in prehistoric Amazonia: Using G.I.S. to reconstruct ancient ethnogenetic processes from archaeology, linguistics, geography, and ethnohistory*. Lund: Lund University.
- Haynie, H., Bowern, C., Epps, P., Hill, J., & McConvell, P. (2014). Wanderwörter in languages of the Americas and Australia. *Ampersand*, 1, 1-18.
- Zamponi, R. (2020). Some precontact widespread lexical forms in the languages of Greater Amazonia. *International Journal of American Linguistics*, 84, 527-573.
- Wojtylak, K. I. (2019). Traversing language barriers: Murui signal drums from Northwest Amazonia. *International Journal of Language and Culture*, 6(1), 195–216.
- Wojtylak, K. I. (2020). *A Grammar of Murui (Bue): A Witotoan Language from Northwest Amazonia*. Leiden: Brill.

Northern Akhvakh fused verbal forms

Vladimir de Haldat du Lys

(Ecole Pratique des Hautes Etudes, Paris)

Keywords: morphophonology, univerbation, sandhi, aspect marking, serial verbs

Akhvakh belongs to the Nakh-Daghestani language family, also known as Northeast Caucasian. In Akhvakh, as in other languages of the western group of this family, aspect is not marked by the verbal root, but by suffixes and periphrastic forms. As indicated by Magomedbekova (1967), these analytic forms can morphophonologically merge, which will be the topic of this presentation. The phonological processes are explained by Creissels (2010 and 2016).

More specifically, the presentation will focus on the study of external sandhi phenomena – which are very flexible, as shown by the provided examples – affecting the verb, i.e. the fusion and univerbation of two or more words, including a verbal lexeme. The forms thus created can be divided into four categories: underlying periphrastic forms with aspectual value including an auxiliary, fused form including the quotative, fused forms resulting in preverbation and verbal locutions with a light verb. Particular attention will be paid to aspect marking and lexical creation. These sandhi periphrases significantly complicate the verbal morphology of Akhvakh, creating new forms, some of which being considerable as serial verbs as defined by Aikhenvald (2018), and expanding the TAME paradigm.

Aspect can be marked by periphrastic forms, optionally univerbated. Akhvakh thus creates new aspecto-temporal categories that cannot be expressed as such by a specific suffix different from the imperfective and perfective suffixes. In example (1), the form *b-eq'erēk^m-āwudi* consists of the lexical verb *b-eq'-uruḷa* “to know” in the progressive converbial form *-ere* and the auxiliary *b-ik'-uruḷa* in the perfect tense. This combination is used to express the imperfect, and thus constitutes a paradigm augmentation, since synthetic forms cannot express imperfect in Akhvakh. It is also possible to augment the evidential paradigm with the auxiliary *m-ič-unuḷa* “to find (oneself)” in the conditional *m-ič-ala* (2). The auxiliary expresses an event that has not been witnessed. It is therefore an evidential auxiliary univerbated with the short infinitive (marked with *-u* instead of *-uruḷa*), creating a synthetic counterfactual form.

The univerbation of two elements can allow one of them to change part of speech. This is witnessed in Akhvakh with the verb “to pass” *b-oḷ'-uruḷa*, which can, when combined with a noun, form a new denominative verbal lexeme (3). This new verbal lexeme is incorporated into a perfect form. Nonverbal lexemes can also, when merged with verbal lexemes, give rise to preverbal forms (4).

Finally, there is the case of light verb construction (5). In this case, the compound coverb fuses with the light verb, creating a new verbal lexeme expressing perfect.

Studying merged verb forms is a necessary step in understanding how aspect marking works in Akhvakh. The forms analysed in this presentation might help clarify the typology of aspect marking in this language.

Ex. (1)

<i>b-eq'erēk^w-āwudi</i>	<	<i>b-eq'-ere</i>	<i>b-ik^w-āwudi</i>
N-know.CV:PROG+(N)BE-PERF		N-know-CV:PROG	N-be-PERF

<i>hugu-ṣu-la</i>	<i>b-eq'erēk^w-āwudi</i>	<i>inṣu-be</i>	<i>koša</i>	<i>eḡ-u</i>
DIST-M-DAT	N-know.CV:PROG+(N)be-PERF	REFL.M-N(GEN)	bad	look-INF
<i>šo-t-ika</i>	<i>k'eha</i>	<i>b-ik^w-e.</i>		
good-VBLZ-IPF.NEG	eye	N-be-SEQ		

‘He knew he had an evil eye and could not look without bringing bad luck.’

Ex. (2)

<i>b-ik^weč-ala</i>	<	<i>b-ik'-u</i>	<i>m-ič-ala</i>
N-be.INF+(N)find-COND		N-be-INF	N-find-COND

<i>t'ēki</i>	<i>Zamaḡo</i>	<i>kurak'e</i>	<i>k^wīle</i>	<i>b-ik^weč-ala</i>
throw.IPF.NEG	Zamako	abricot	wish.MSD	N-be.INF+(N)find-COND
<i>mede-dala</i>	<i>b-ił-u</i>	<i>b-uḡ-e</i>	<i>b-ik'-uwa</i>	<i>ruša.</i>
2SG.ERG-at.least	N-put-INF	N-fall-SEQ	N-be-PERF?/FUT?	tree

‘I won’t throw any at you, Zamaḡo, if you wanted an apricot all you had to do was topple the tree.’

Ex. (3)

<i>waranoḡ^w-ehe</i>	<	<i>warani b-oḡ^w-ehe</i>
camel+(N)go-SEQ.N		camel-N-go-SEQ.N

<i>dene</i>	<i>waranoḡ^w-ehe</i>	<i>gwede.</i>
1SG.NOM	camel+(N)go-SEQ.N	COP.N

‘I turned into a camel.’

Ex. (4)

<i>kadabuḡ-alaq'o</i>	<	<i>kadiga</i>	<i>b-uḡ-alaq'o</i>
ground.LAT+N.fall-CV:ANT		ground.LAT	N-fall-CV:ANT

<i>kadabuḡ-alaq'o</i>	<i>b-adaḡ-ewudi</i>	<i>šari</i>
ground.LAT+N.fall-CV:ANT	N-laugh-PERF	fox

‘The fox laughed his head off.’

Ex. (5)

<i>bakala-kirōḫ-ēwudi</i>	<	<i>bakala-kiri</i>	<i>oḫ-ēwudi</i>
thanks-benediction+give-perf		thanks-benediction	give-perf

<i>qe</i>	<i>hu</i>	<i>minarēk^wa-ṣ^w-e</i>	<i>īṣu-da</i>	<i>ḫ'e-ṣu-ḫa</i>
then	dist	host-OBL.M-ERG	REFL.M(GEN)-ATTR	guest-OBL.M-DAT

bakala-kirōḫ-ēwudi
thanks-benediction+give-perf

‘The host then thanked his guest and blessed him’

References

- AIKHENVALD, Alexandra Y., *Serial Verbs*, Oxford Studies in Typology and Linguistic Theory (Oxford, 2018; online edn, Oxford Academic, 20 Dec. 2018), <https://doi-org.proxy.rubens.ens.fr/10.1093/oso/9780198791263.001.0001>, accessed 18 May 2025.
- CREISSELS, Denis, 2010. “Liaison and grammaticalization in Northern Akhvakh”, [online] <https://disk.yandex.ru/d/fqtPmzQMOPDEw/Andic/Axvax>.
- CREISSELS, Denis, 2016. “Univerbation via liaison and the evolution of lexicon and grammar in Northern Akhvakh” [online] http://www.deniscreissels.fr/public/Creissels-liaison_Akhvakh.pdf
- MAGOMEDBEKOVA, Zagidat M., 1967. *Axvaxskij jazyk (grammatičeskij analiz, teksty, slovar')* [Akhvakh language (grammatical analysis, texts, lexicon)], Tbilissi: Mecniereba.

General Session : Language variation

Es hät vs. *es git*: Investigating the lexical split of Swiss German existential clauses

Eyal Liron Dolev
(University of Zurich)

Keywords: Swiss German, existential clauses, syntax, pragmatics, semantics

Swiss German, the group of Alemannic dialects spoken in German-speaking Switzerland, displays a case of a lexical split in its existential clauses, resulting in two competing, semantically distinct types. Formally, the two types are constructed using the expletive pronoun *es* 'it' and the verbs *hät* or *git*, the 3.SG forms of *haa* 'have' and *gää* 'give', respectively. Due to their non-canonicity (expletive subject, bleached verb) and their function to assert the existence or presence of objects at a location, cf. McNally (2011) and Sarda & Lena (2023), I interpret them as existential constructions.

I investigate this semantic contrast based on data collected from a corpus ("What's Up, Switzerland", Ueberwasser & Stark, 2017) and an online survey conducted with native speakers of Swiss German (N=93).

Competing types of existentials within one linguistic system also exist in other languages. Koch (2012) offers a typological review of a split between existential and locational sentences. Czinglar (2002) compares *es hot* (the Austrian-Alemannic variant of *es hät*) with the Standard German *es gibt*, but not as competing constructions within one linguistic system. However, the data shows that the case of Swiss German existentials cannot be understood strictly in terms of existential vs. locational vs. presentational.

The contrast involves several linguistic categories. *Es hät* existentials have an imperfective modality, they require a specific, instantiated referent identifiable from within the speech act and being within deictic reach. They state the specific existence/presence of an extant object. *Es git* existentials, on the other hand, are unable to access instantiated objects within the hearer-speaker deixis. They rather refer to entities outside the deixis with two possible readings: (1) generic and kind reference; (2) reference to the instantiation of an object coming into existence – a resultative-transformative *Aktionsart*-reading with a prospective aspect, which I call the EXISTENTIVE.

- (1) a. *uf-em Wysshorn hät-s hüt schnee*
 on-the PN has-it today snow
 "There's snow on the Wysshorn today."
 b. *uf-em Wysshorn git-s hüt schnee*
 on-the PN gives-it today snow
 "There is going to be snow on the Wysshorn today."

Sentence (1a) refers to concrete, instantiated snow currently lying on the ground. In sentence (1b), there is necessarily still no snow, but the future presence of snow is stated, i.e., the process, at the end of which there will be snow – the EXISTENTIVE reading – hence the future-time reference in the translation.

The contrast thus lies at the conjunction of several categories: modality and aspect, genericity, and referentiality. It is an example of the interaction between categories, mutually affecting each other across domains.

Finally, attempting a holistic perspective, I suggest that the contrast between *es hät* and *es git* has a strong pragmatic component and has to do with a contrast between the ‘pragmatic’ plane and the ‘objective/real’ plane, cf. Bühler (1934) and Benveniste (1971): *Es hät* is a predication embedded in the act of speech and requires knowledge of the speaker’s origo-deixis; *Es git* is a predication that can be detached from the speech act.

References

- Benveniste, Émile (1971), The Nature of Pronouns. In *Problems in General Linguistics*. University of Miami Press.
- Bühler, Karl (1934), *Sprachtheorie: Die Darstellungsfunktion der Sprache* (Ungekürzter Nachdruck d. Ausgabe v. 1934). Gustav Fischer Verlag.
- Cztinglar, Christine (2002), *Decomposing Existence: Evidence from Germanic*. in: Abraham, Werner & C. Jan Wouter Zwart (eds.) (2002/in print): *Issues in Formal German(ic)Typology*. Amsterdam, Benjamins (Linguistik Aktuell/Linguistics Today 45). 336pp. <https://doi.org/10.1075/la.45.06czi>
- Koch, Peter (2012), Location, existence, and possession: A constructional-typological exploration. *Linguistics*, 50(3), 533–603. <https://doi.org/10.1515/ling-2012-0018>
- McNally, Louise (2011), Existential sentences. In K. V. Heusinger, C. Maienborn, & P. Portner (Eds.), *Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science* (pp. 1829–1848). De Gruyter. <https://doi.org/10.1515/9783110255072.1829>
- Sarda, Laure & Lena, Ludovica (2023), Existential constructions: In search of a definition. In L. Sarda & L. Lena (Eds.), *Human Cognitive Processing* (Vol. 76, pp. 1–32). John Benjamins Publishing Company. <https://doi.org/10.1075/hcp.76.01sar>
- Ueberwasser, Simone & Stark, Elisabeth (2017), What’s up, Switzerland? A corpus-based research project in a multilingual country. *Linguistik Online*, 84(5). <https://doi.org/10.13092/lo.84.3849>

Alternation between marking strategies for contrastive subjects: A corpus-based analysis in formal and informal written French

Jorina Brysbaert

Fund for Scientific Research (F.R.S.-FNRS) and UCLouvain

Keywords: alternation, contrastive subjects, register, corpus analysis, French

GOAL AND BACKGROUND. This talk examines the probabilistic constraints governing the alternation between four strategies for marking contrastive subjects in French: contrastive adverbs (C-Advs, e.g. *par contre* ‘on the other hand’); emphatic pronouns (E-Pros, e.g. *lui* ‘him’); emphatic pronouns introduced by *quant à* (*quant à* E-Pros, e.g. *quant à lui* ‘as for him’); and possessive-based adverbial phrases (PB-Advs, e.g. *pour sa part* ‘for his part’). The analyzed constraints include the syntactic type, definiteness, animacy and post-modification of the subject, alongside the register (formal vs. informal).

- (1) – *Que mangent les oiseaux dans ton jardin ?* ‘What do the birds in your garden eat?’
– *Le moineau adore les graines de tournesol.* ‘The sparrow loves sunflower seeds.’
Le rouge-gorge, [par contre / lui / quant à lui / pour sa part] préfère les vers de farine.
‘The robin, [on the other hand / him / as for him / for his part] prefers mealworms.’

Previous research on this alternation is sparse, except for Brysbaert and Lahousse (2022). Some linguists mention that C-Advs, (*quant à*) E-Pros and PB-Advs have a similar information-structural function, i.e. to focus on and single out the subject (Caddéo 2004; Cappeau 1999; Nølke 1997), but they do not elaborate on the underlying conditions of the alternation.

METHODS. We present a **quantitative corpus analysis** of C-Advs (323 occurrences), E-Pros (3119 occurrences), *quant à* E-Pros (1479 occurrences) and PB-Advs (1023 occurrences) in formal written (newspaper; 74 million words) and informal written (online discussion platform; 6,1 million words) French. We use multinomial logistic regression, conditional inference trees and conditional random forests to model the alternation.

RESULTS. Our data show that the alternation is affected by:

1. **Register.** The chance of a *quant à* E-Pro and a PB-Adv is significantly lower in informal written French, probably because these are more ‘complex’ markers, i.e. they contain more lexical material (Table 1).
2. **Semantico-pragmatic constraints.** (*Quant à*) E-Pros and PB-Advs mainly modify subjects with a definite, human referent – which are considered to be ‘good’ preverbal subjects (e.g. Cappeau and Deulofeu 2006; Karssenberg 2017). C-Advs combine significantly more often with ‘bad’ preverbal subjects, i.e. subjects with an (i) indefinite (14-15% of the cases, Table 2) and (ii) inanimate referent (36-51% of the cases, Table 3). In line with the idea that contrast functions as a ‘survival strategy’ (e.g. Leonetti 2013), we argue that C-Advs can improve the acceptability of otherwise ‘bad’ preverbal subjects by highlighting their contrastive nature. The other three markers do not possess this ‘saving potential’, probably because (i) they do not have an inherently contrastive meaning and (ii) they are morpho-syntactically dependent on their subject.
3. **Formal complexity.** Compared to E-Pros, *quant à* E-Pros and PB-Advs are more often used in linguistic environments with a high formal complexity: (i) they combine more often with lexical NPs than with simple relative pronouns (Table 2), and (ii) they are favored by complex post-modified NPs

(Table 4). This is in line with the Complexity Principle (e.g. Rohdenburg 1996), according to which more explicit items are favored in more complex environments.

Table 1. Frequency of marker per register

	Formal	Informal
C-Advs	5% (276)	12% (47)
E-Pros	51% (2819)	76% (300)
Quant à E-Pros	26% (1446)	8% (33)
PB-Advs	18% (1009)	4% (14)
Total	100% (5550)	100% (394)

Table 2. Syntactic type and (in)definiteness of subject per marker per register

		C-Advs		E-Pros		Quant à E-Pros		PB-Advs	
		Formal	Informal	Formal	Informal	Formal	Informal	Formal	Informal
Noun	Definite	70% (194)	68% (32)	87% (2459)	53% (160)	97% (1403)	97% (32)	93% (939)	71% (10)
	Indefinite	4% (12)	4% (2)	1% (39)	2% (5)	1% (19)	0% (0)	2% (19)	0% (0)
Pronoun	Relative	12% (32)	4% (2)	11% (310)	38% (114)	1% (19)	0% (0)	3% (34)	21% (3)
	Indefinite	10% (28)	11% (5)	0% (0)	0% (0)	0,1% (2)	0% (0)	0% (0)	0% (0)
	Other	3% (9)	6% (3)	0,4% (11)	7% (21)	0,2% (3)	3% (1)	2% (17)	7% (1)
Verb phrase		0,4% (1)	6% (3)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)
Total		100% (276)	100% (47)	100% (2819)	100% (300)	100% (1446)	100% (33)	100% (1009)	100% (14)

Table 3. Animacy of subject per marker per register

		C-Advs		E-Pros		Quant à E-Pros		PB-Advs	
		Formal	Informal	Formal	Informal	Formal	Informal	Formal	Informal
Human		34% (94)	38% (18)	69% (1949)	62% (187)	68% (982)	48% (16)	70% (707)	64% (9)
Animate		30% (83)	11% (5)	15% (420)	14% (42)	15% (220)	12% (4)	26% (258)	14% (2)
Inanimate		36% (99)	51% (24)	16% (450)	24% (71)	17% (244)	39% (13)	4% (44)	21% (3)
Total		100% (276)	100% (47)	100% (2819)	100% (300)	100% (1446)	100% (33)	100% (1009)	100% (14)

Table 4. Post-modification of lexical NP subjects per marker per register

	E-Pros		Quant à E-Pros		PB-Advs	
	Formal	Informal	Formal	Informal	Formal	Informal
Post-modification	30% (744)	15% (25)	45% (637)	19% (6)	39% (372)	20% (2)
No post-modification	70% (1754)	85% (140)	55% (785)	81% (26)	61% (586)	80% (8)
Total	100% (2498)	100% (165)	100% (1422)	100% (32)	100% (958)	100% (10)

REFERENCES

- Brysbaert, Jorina and Lahousse, Karen (2022), Marking contrastive topics in a topic shift context: Contrastive adverbs versus emphatic pronouns, *Discours: Revue de linguistique, psycholinguistique et informatique* 31, 1-27.
- Caddéo, Sandrine (2004), Lui, le propriétaire, le propriétaire, lui : deux constructions bien distinctes, *Recherches sur le Français Parlé* 18, 145-161.
- Cappeau, Paul (1999), Sujets éloignés. Esquisse d'une caractérisation des sujets lexicaux séparés de leur verbe, *Recherches sur le Français Parlé* 15, 199-231.
- Cappeau, Paul and Deulofeu, José (2006), Les "indéfinis" en relation avec la position sujet dans trois types de constructions prédicatives, in F. Corblin, S. Ferrando, and L. Kupferman (eds), (2006), *Indéfini et prédication*, Paris: Presses Universitaires de Paris Sorbonne, 125-138.
- Karssenbergh, Lena (2017), *Ya les oiseaux qui chantent. A corpus analysis of French il y a clefts*, Leuven: KU Leuven PhD dissertation.
- Leonetti, Manuel (2013), On contrastive readings in the interpretation of NPs/DPs, in S. Chiriacescu (ed), (2013), *Proceedings of the VI Nereus International Workshop: Theoretical implications at the syntax-semantics interface in Romance*, Konstanz: Fachbereich Sprachwissenschaft der Universität Konstanz, 99-116.
- Nølke, Henning (1997), Anaphoricité et focalisation: Le cas du pronom personnel disjoint, in W. De Mulder, L. Tasmowski-de Ryck, and C. Vetter (eds), (1997), *Relations anaphoriques et (in)cohérence*, Amsterdam: Brill/Rodopi, 55-67.
- Rohdenburg, Günter (1996), Cognitive complexity and increased grammatical explicitness in English, *Cognitive Linguistics* 7(2), 149-182.

Head to Tail: Bridging Constructions in the Linguistic Landscape of Amazonia

This paper investigates bridging linkage constructions across the Witotoan language family, comprising Murui-Muina (MM), Nonuya (NO), and Ocaina (OC), spoken by indigenous communities in Northwest Amazonia (Colombia and Peru). Bridging constructions (Guérin & Aiton, 2019)—also referred to in the literature by various terms, such as tail-head linkage (de Vries, 2005), head-tail linkage (Fabian et al., 1998), tail-head recapitulation (Farr, 1999), and recapitulation clauses (Genetti, 2007)—function as a discourse strategy to enhance cohesion and structure by connecting sentences or paragraphs, typically through the repetition of elements from the preceding clause (Thurman, 1975).

This study identifies two primary types of bridging constructions in Witotoan languages: **recapitulative linkage**, characterized by verbatim repetition of the reference clause in the bridging clause, ensuring textual continuity, as in (1) from MM (Wojtylak 2020: 476), and **summary linkage**, which uses generic verbs such as *fino*- ‘to do’ in MM, alongside demonstratives to encapsulate prior events without direct repetition. These strategies are particularly inherent to procedural discourse of the Witotoan languages.

- (1)

huidzi-hi _{OBJ}	ti-iti-kue _{PRED}	kore _{OBJ} o-dze-na
yuca-CLF:TUBER	grate-FUT.LK-1SG	starch get-FUT.ACT.NMLZ-NSBJ
kore _{OBJ}	o-a-no	(...) [kue φarie] _{OBJ} beei-iti-kue _{PRED}
starch	get-ACT.NMLZ-SEQ	1SG farina.Sp toast-FUT.LK-1SG

‘I will grate the yuca, to get the starch. After getting the starch, (...) I will toast my *farina* (cereal meal).’

A distinctive feature of Witotoan languages is the use of paragraph-initial connectives, which serve anaphoric and cohesive functions, as in (2) from OC (Fagua Rincón, 2013: 68). This type of construction parallels summary bridging constructions, with the connective referring anaphorically to the preceding event and marking thematic continuity and cohesion in the text. These connectives are integral to narrative discourse, performing a role similar to the neighboring Bora (Boran) connective described by Seifart (2010).

- (2)

nú?ũ	tiamóó?	ha-naa-βí-na _{PRED}	bu?	[haṇáá-hũ-?
DISC.CONN	there	3SG-lay-ANTIC-PST	maybe	two-CLF:G-two
dʰoʔoo-hĩ] _{NP}	nú?ũ	[dʰaa-φo	háɖia] _{POSTP}	ha-taruuu-na-ʔi-na _{PRED}
warm-CLF:G	DISC.CONN	hole-CLF:CAVITY	inside	3AN.S-cast.spell-VBZ-HAB-PST

‘So/thus, (he) was there for about two years (lit. hot seasons), and so/thus, inside the hole, (he) was casting spells.’

This study draws on first-hand fieldwork conducted with Nonuya and Murui-Muina communities since 2013, complemented by second-hand data on Ocaina and discussions with linguists specializing in Witotoan languages. By situating Witotoan bridging constructions within the broader Amazonian typological framework, the paper identifies parallels and distinctions with Tariana (Arawak; Aikhenvald 2019) and Tucanoan languages (Tucano and Wanano; Ramirez, 1997; Waltz & Waltz 1997 in Aikhenvald 2002: 169-171) languages. It also examines the potential for areal diffusion, particularly in the Vaupés region, where multilingual contact has shaped linguistic structures (Aikhenvald, 2002). Because the Caquetá-Putumayo people that the Witotoan people form part of have been in contact with other regional groups, and other morphosyntactic features have been shown to be shared in the region, bridging constructions are

likely to have been influenced by contact-induced processes. This analysis provides a detailed typological comparison and addresses how bridging constructions in Witotoan languages complement and expand our understanding of these structures in discourse typology.

References

Aikhenvald, A. Y. (2002). *Language contact in Amazonia*. New York: Oxford University Press.

Aikhenvald, A. Y. (2019). Bridging linkage in Tariana, an Arawak language from Northwest Amazonia. *International Journal of American Linguistics*, 85(1), 455–496.

de Vries, L. (2005). Towards a typology of tail-head linkage in Papuan languages. *Studies in Language*, 29(2), 363–384.

Fabian, G., Fabian, E., & Waters, B. (1998). *Morphology, syntax and cohesion in Nabak, Papua New Guinea* (Pacific Linguistics 148). Canberra: Australian National University.

Fagua, D. (2013). *Aspects morphosyntaxiques de l'ocaina: Autour des classes lexicales* (Doctoral dissertation). Université Paris Diderot.

Farr, C. J. M. (1999). *The interface between syntax and discourse in Korafe, a Papuan language of Papua New Guinea* (Pacific Linguistics C144). Canberra: Australian National University.

Genetti, C. (2007). *A grammar of Dolakha Newar*. Berlin: Mouton de Gruyter.

Guérin, V., & Aiton, G. (2019). Bridging constructions in typological perspective. In V. Guérin (Ed.), *Bridging constructions* (Studies in Diversity Linguistics, Vol. 24, pp. 1–44). Berlin: Language Science Press.

Seifart, F. (2010). The Bora connector pronoun and tail-head linkage: A study in language-specific grammaticalization. *Linguistics*, 48(4), 893–918.

Ramirez, H. (1997). *A fala Tukano dos Yepâ-masa. Tomo I: Gramática. Tomo II: Dicionário. Tomo III: Método de aprendizagem*. Manaus: Inspetoria Salesiana.

Thurman, R. C. (1975). Chuave medial verbs. *Anthropological Linguistics*, 17(7), 342–352.

Waltz, N., & Waltz, C. (1997). *El agua, la roca y el humo: Estudios sobre la cultura wanana de Vaupés*. Santafé de Bogotá: Instituto Lingüístico de Verano.

Wojtylak, K. I. (2020). *A grammar of Murui (Bue), a Witotoan language from Northwest Amazonia*. Leiden: Brill.

Linking universality to language complexity: A fuzzy model

M. Dolores Jiménez López, Adrià Torrens Urrutia, Antoni Brosa Rodríguez,
Susana M. Campillo Muñoz & Carlos Martín Vide
(Universitat Rovira i Virgili)

Keywords: complexity of language, relative complexity, absolute complexity, fuzzy models, language universals

Recent studies in theoretical and applied linguistics highlight a growing interest in measuring linguistic complexity and addressing whether all languages are equally complex (Di Domenico 2017, Kortman and Szmrecsanyi 2012, La Mantia et al. 2017, McWhorter 2012, Newmeyer and Preston 2014 and Ortega and Han 2017). However, calculating these differences remains challenging. This paper contributes to these discussions by proposing a novel approach to assess linguistic complexity through its relationship with universality. The central premise is that languages with more universal traits tend to be less complex and easier to learn, while those with rarer features are more challenging and complex. By focusing on universality, the study aims to offer a new framework that avoids exhaustive pairwise comparisons and instead provides an efficient, objective measure of linguistic complexity.

The research addresses key questions: How can linguistic universality be quantified and linked to complexity? Can fuzzy modeling provide a more nuanced analysis than traditional binary classifications?

Our model aims to jointly approach and interrelate absolute (McWhorter 2001 and Dahl 2004) and relative complexity (Kusters 2003), integrating the objective properties of the system with user-related factors. Given the inherent challenges of calculating the global complexity of a language, our initial focus is on local complexity (Miestamo 2008), which allows for a more manageable analysis by examining particular subdomains of languages.

The study employs a mixed-method approach combining mathematical modeling and empirical data analysis. A fuzzy model is developed to treat universality and complexity as gradual, rather than discrete, categories. To validate this model, the study utilizes Greenberg's universals (Greenberg 1963), formalized through the Grew-Match (Guillaume 2021), alongside the Universal Dependencies dataset (2.11), which includes 241 corpora from 143 languages spanning all major linguistic macro-areas (except Australia). The dataset features both living and dead languages, as well as Creoles, sign languages, and isolated languages, ensuring broad diversity.

Using this dataset, the study validates language universals using a fuzzy model to assess their universality and the complexity of languages. It involves calculating theoretical and relative universality and complexity and interpreting the results qualitatively with fuzzy evaluative expressions. The process includes three steps: computing the theoretical relationship of Greenberg's universals to complexity, comparing their relative universality and complexity

across languages, and applying fuzzy logic (Novák 2008) to translate quantitative findings into qualitative insights.

The results confirm an inverse relationship between universality and complexity. Languages with more universal features exhibit lower complexity, validating the hypothesis that linguistic structures shared across languages are generally simpler. The fuzzy model successfully establishes a scalable and adaptable measure of complexity, enabling cross-linguistic comparisons. The analysis also reinforces the robustness of Greenberg's universals, extending their applicability to a broader and more diverse dataset than prior studies.

Despite limitations, such as focusing on a subset of universals and potential biases in language distribution, the findings provide valuable insights into linguistic complexity. Future research could expand this work by including additional universals, refining the fuzzy model, and exploring its application to other linguistic phenomena, further advancing the study of linguistic complexity.

Acknowledgments: This paper was supported by the project PID2020-120158GB-I00 funded by MCIN/AEI/10.13039/501100011033.

References

Dahl, Östen (2004), *The growth and maintenance of linguistic complexity*, Amsterdam: John Benjamins.

Di Domenico, Elisa (2017), *Syntactic Complexity from a Language Acquisition Perspective*, Cambridge: Cambridge Scholars Publishing.

Greenberg, Joseph H. (1963), *Universals of language*, Cambridge, Mass.: MIT Press.

Guillaume, Bruno (2021), Graph matching and graph rewriting: GREW tools for corpus exploration, maintenance and conversion, in D. Gkatzia, and D. Seddah (eds), (2021), *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations*, Association for Computational Linguistics, 168-175.

Kortmann, Bernd and Szmrecsanyi, Benedikt, (2012), *Linguistic Complexity: Second Language Acquisition, Indigenization, Contact*, Berlin & Boston: Walter de Gruyter

Kusters, Wouter (2003), *Linguistic Complexity: The Influence of Social Change on Verbal Inflection*, Amsterdam: LOT.

La Mantia, Francesco, Licata, Ignazio, Perconti, Pietro (eds), (2017), *Language in Complexity. The Emerging Meaning*, Berlin: Springer.

McWhorter, John (2001), The world's simplest grammars are creole grammars, *Linguistic Typology* 6, 125–166.

McWhorter, John (2012), *Linguistic Simplicity and Complexity: Why do Languages Undress?*, Berlin & Boston: Mouton de Gruyter.

Miestamo, Matti (2008), Grammatical complexity in a cross-linguistic perspective, in M. Miestamo, K. Sinnemäki and F. Karlsson (eds), (2008), *Language Complexity: Typology, Contact, Change*, Amsterdam: John Benjamins, 23–42.

Newmeyer, Frederick J. and Preston, Laurel B. (eds.) (2014), *Measuring Grammatical Complexity*, Oxford: Oxford University Press.

Nivre, J., De Marneffe, M.-C., Ginter, F., Hajič, J., Manning, C. D., Pyysalo, S.,... Zeman, D. (2023), Universal Dependencies. Retrieved from <https://universaldependencies.org/>

Novák, Vilem (2008), Mathematical fuzzy logic: From vagueness to commonsense reasoning, in G. Kreuzbauer, N. Gratzl and E. Hielb (eds), (2008), *Retorische Wissenschaft: Rede und Argumentation in Theorie und Praxis*, Berlin: LIT-Verlag, 191–223.

Ortega, Lourdes and Han, ZhaoHong (2017), *Complexity Theory and Language Development*, Amsterdam: John Benjamins.

Morphological borrowing: the case of augmentative suffixes in Italiot

Michail I. Marinis

CHS; Harvard University – GLOS; The Ohio State University

Keywords: language contact, griko, greko, derivational suffixes, morphological borrowing

Italiot is a Greek dialect spoken in two enclaves in southern Italy: Greko in Calabria and Griko in Puglia. As the language system has undergone long-term and intense contact with Romance dialects (Ralli 2012; Ledgeway 2012; Marinis 2020; to print), borrowing phenomena have occurred not only at the lexical level (Marinis & Ralli 2015) but across all levels of linguistic analysis.

This study examines the borrowing of Southern Italian Greek (Italiot) from Romance dialects at the level of derivational morphology, focusing specifically on the case of augmentative derivational suffixes. The data come from both written sources (i.e., Karanastasis 1984) and fieldwork data I personally collected in the Greek-speaking area during the summers of 2016 and 2018.

The research reveals that in both varieties, the derivation of lexemes with the feature [+augmentative] is typically realized using the Romance-origin derivational suffix *-une*, combined with Italiot bases. These bases may be either native to the Greek system or of Romance origin. For *-une* to integrate into the derivational morphology of Italiot, it loses its final segment *-e* and adopts native inflectional suffixes, enabling it to align with a specific inflectional class and acquire gender. As a result, it manifests as either *uni-N* (i.e., *li 'kuni.M* ‘big wolf’ < Gr. *'lik(os)* ‘wolf’) or *una-F* (i.e., *ane 'muna.F* ‘strong wind’ < Gr. *'anemos* ‘wind’).

At the same time, the more productive Modern Greek derivational suffixes, namely *-ara-F* and *-aros-M*, while present in Medieval Greek (Holton et al. 2019) and still in use in Standard Modern Greek (i.e., *tur'tara.F* ‘big cake’ < *'turta.F* ‘cake’), are entirely absent from the morphological system of both Griko and Greko (**tur'tara.F* ‘big cake’ < *'turta.F* ‘cake’). This indicates that *-une* entered the Greek linguistic system as a borrowed element and not only became fully integrated as a morpheme but also displaced the previously available native derivational suffixes. It is noteworthy that while Romance-origin derivational suffixes can also be found in other Greek dialects that were in intense contact with Romance languages (i.e., Heptanesian; Krimpas 2018), Italiot is the only dialect where the Romance-origin element prevailed, effectively replacing the native derivational suffixes used to denote augmentation.

Finally, research confirms that in Italiot, borrowing involves not only lexical elements but also sub-lexical elements (Ralli 2012). Beyond the lexical stems discussed by Marinis & Ralli (2015), derivational suffixes are also readily borrowed. However, Marinis (2020; to print) asserts that in the case of Southern Italian Greek, inflectional morphology resists borrowing.

References

- Holton, D., Horrocks, G., Janssen, M., Lendari, T., Manolassou, I., & Toufexis, N. (2019). *The Cambridge grammar of medieval and early modern Greek*. Cambridge: Cambridge University Press.
- Karanastasis, A. 1984. *Ιστορικό λεξικόν των ελληνικών ιδιωμάτων της Κάτω Ιταλίας – Volume 1*. Academy of Athens.
- Ledgeway, A. 2012. Greek disguised as Romance? The case of southern Italy. *Modern Greek Dialects and Linguistic Theory*, 5.1: 184-227.

- Marinis, M. I. & Ralli, A. 2015. Περιστάσεις έντονης επαφής και δανεισιμότητα λεξικών στοιχείων: η περίπτωση της Γκρίκο και της Καπαδοκικής. *Patras Working Papers in Linguistics* 4, 63-74. <https://doi.org/10.26220/pwpl.v4i0.2297>
- Marinis, M. I. 2015. Κλίμακες δανεισιμότητας στην γκρίκο και την κρητική: μία πρώτη προσέγγιση. [Borrowability scales in Griko and Cretan dialects]. *Studies in Greek Linguistics* 35, 351–363. <https://doi.org/10.5281/zenodo.14211644>
- Marinis, M. I. 2020. Κλίση και οργάνωση κλιτικών παραδειγμάτων. Doctoral dissertation, University of Patras. <https://doi.org/10.12681/eadd/46842>
- Marinis, M. I. to print. Greek enclaves in Southern Italy: an introduction to contemporary language and its history. *Greek around the world*, ed. by Brian D. Joseph. Publications of the Laboratory for the Study of the Greek Language. Volume 1.
- Ralli, A. 2012. Verbal loanblends in Griko and Heptanesian: A case study of contact morphology. *L'Italia Dialettale. Rivista de dialettologia italiana* 73: 111-132.

Growing nominal number marking: the case of South American languages

Olga Krasnoukhova & An Van linden

(University of Liège/ Liège & University of Leiden/Leiden; & University of Liège/ Liège)

Keywords: verbal number marking; nominal number marking; typology; diachrony; South American Indigenous languages

In this paper we present an account of language change through which number marking can spread from the verbal to the nominal domain, with a focus on South American (SA) languages.

The category of number is usually associated with nominals. For instance, in English examples like *I see clouds in the sky*, the noun *cloud* obligatorily carries number marking if the referent of the noun is plural. In the Peruvian language Harakmbut, by contrast, nouns go unmarked for number. That is, plurality of a nominal referent is encoded on the verb by the verbal plural prefix *mba-~ma-~mã-*, marking plurality of the object (on transitives) or subject (on intransitive verbs) (Van linden 2022:464). Verbal number marking, in which plurality of participants is not marked on nouns but on verbs, is found throughout the globe (Corbett 2000:245; Mattioli 2009:17), but seems particularly prominent in SA languages (Krasnoukhova 2022:631). Our dataset of ±50 SA languages show that some SA languages use formally identical plural markers on their nouns and verbs. Number marking on nouns is optional in most such cases. Chibchan languages are a case in point. For example, in Buglere, the plural marker *-dre* on the verb signals plurality of the subject participant (1a), and the same marker can mark plural on nouns (1b). Proto-Chibchan did not have number marking on nouns at all; inflectional morphology reconstructed for this family is all verbal (Constenla Umaña 2012:404).

(1) Buglere (Chibchan)

(a) *muire* *isi* *ngarribega-ble-dre*

woman snake kill-REM.PST-PL

‘The women killed a snake.’ (Quesada 2012:86)

(b) *muire-dre*

woman-PL

‘women’ (Quesada 2012:74)

We hence suggest that number marking can ‘spread’ from the verbal to the nominal domain, and propose at least two scenarios for this process. The first one involves the use of verbal plural marking on deverbal nouns and nominalizations as an intermediary stage. For example in Nivacle (Matacoan), the marker *=sha7ne* is used on verbs to indicate plurality of the subject (on intransitive verbs) or the object participant (on transitives), as in (2). This plural marker also occurs on adjectives (which are nouny in Nivacle) and some nouns. Importantly, all nouns that take this plural marker are either deverbal nouns or nominalizations (Campbell et al. 2020:101).

(2) Nivacle (Matacoan)

ja7-p’ôw-ja-yan=sham=sha7ne

1ACT-notch-LIG-VBLZ=THROUGH/INSIDE =PL

‘I make notches.’ (Campbell et al. 2020:260, p.c.)

A second, overall rarer scenario can be observed in Teribe (Chibchan), where the plural marker for nominal subjects presumably originates in a positional verb. In present-day Teribe, the element *lok* ‘plural’ predominantly encodes plurality of subject participants (Quesada 2000:66), shown in (3).

Diachronically, however, *lok* is likely to be a positional verb ‘be plural in a place/state(?)’ (*ibidem*). Syntactically, its behavior matches its typical occurrence as part of serial verb constructions (Quesada 2000:65, 131).

(3) Teribe (Chibchan)

<i>dbong</i>	<i>ī-na</i>	<i>lok</i>
tiger	see-PERF.3	PL

‘They saw a tiger.’ (Quesada 2000:66)

In short, this study contributes to our knowledge of pathways to nominal number marking, revealing phenomena which have so far remained under the radar.

Abbreviation:

1= 1st person; 3 = 3rd person; ACT = actor; LIG = ligature; PERF = perfective; PL = plural; REM.PST = remote past; VBLZ = verbalizer; VPL = verbal plural marker.

References:

- Campbell, Lyle, Luis Díaz & Fernando Ángel (2020), *Nivacle grammar*. Salt Lake City: University of Utah Press.
- Corbett, Greville (2000), *Number*. Cambridge: Cambridge University Press.
- Krasnoukhova, Olga (2022), Number in the languages of South America. In Paolo Aquaviva & Michael Daniel (eds.) *Number in the World's Languages: A Comparative Handbook*. Comparative Handbooks of Linguistics nr. 5: De Gruyter Mouton.
- Mattiola, Simone (2019), *Typology of Pluractional Constructions in the Languages of the World*. (Typological studies in language 125). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Quesada, Juan Diego (2012), *Gramática del buglere*. Quito, Ecuador: Abya-Yala.
- Van linden, An (2022), Harakmbut. In Epps, P. & L. Michael (eds.), *Amazonian Languages, An International Handbook*. Berlin: de Gruyter Mouton.

General Session : Morphology

Differentiating between derivation and compounding: prefixes vs. compounding constituents in complex Latvian nouns

Andra Kalnača, Tatjana Pakalne & Ilze Lokmane
(University of Latvia, Riga)

Keywords: prefixes, prepositions, compounding, derivation, word-formation

In Latvian, almost all prefixes are homonymous with prepositions, from which they have historically originated, e.g., *aiz(-)*, *ap(-)*, *no(-)*, *pa(-)*, *pār(-)*, *pie(-)*, *uz(-)*. This means that there simultaneously exist ‘pairs’ of corresponding homonymous bound and free morphemes (Nītiņa & Grigorjevs 2013). While morphemes functioning in verb and deverbal derivation, as well as, to a lesser extent, in adjective derivation, are prototypical prefixes, morphemes that appear in non-deverbal noun formation often do not exhibit features characteristic of grammaticalization, i.e. of having become an affix, such as semantic bleaching and dissociation from the corresponding free form by developing a different semantics in a specific fixed position (Amiot 2004, Trips 2009, Van Goethem 2009, 2020). The classification of this second group of morphemes has, in fact, always been controversial in Latvian, probably because of the simultaneous co-existence of homonymous free and bound morphemes, which may create a bias towards classifying these morphemes as prefixes when occurring as part of complex words and as prepositions when used as separate words (typologically, see e.g., Bauer 2005, Booij 2010). E.g., word-initial morphemes in (1) have been interpreted, at different points in time, as compounding constituents (Endzelīns 1951) and prefixes (Nītiņa, Grigorjevs 2013).

- (1) *ap-kakle* ‘collar’
aiz-durve ‘space behind doors’
no-male ‘edge, outskirts’
pa-galde ‘space under a table’
pie-kraste ‘coastal area’

The objective of the proposed study, therefore, has been to find out whether it is possible to formulate consistent criteria for differentiating between prefixes and compounding constituents, i.e. prefixation and $[P + N]_{N/A}$ compounding in Latvian as distinct types of word-formation processes. The study is based on a list of complex non-deverbal nouns with a word-initial element corresponding simultaneously to a prefix and a preposition extracted from the Balanced Corpus of Modern Latvian LVK2018. A comparative analysis of the semantic and grammatical properties of word-initial elements against the ‘prototypical’ prefixes, on the one hand, and prepositions, on the other hand (Bauer 2017, Olsen 2017, Stump 2017) has resulted in singling out three groups of complex nouns: 1) prefixal derivatives expressing a quantitative meaning (excessiveness), *pārgalvība* ‘recklessness’, 2) a small group of possible exocentric subordinate compounds, probably historically motivated by a PP with the prepositions *no* and *pie* as their left-hand constituent, *novakare* ‘early evening’, *piekrāsa* ‘a touch of, a nuance’, 3) a large open group of exocentric subordinate compounds with various, mostly, spatial prepositions as their head element, synchronically expressing the semantics of the underlying PP, the referent of a compound being that of which the PP is a complement, *pārupe* ‘territory across a river’, *aizvēsture* ‘pre-history, time before history’. The properties of word-initial elements in the third group

vs. those in other groups vs. ‘prototypical’ prefixes and prepositions have allowed to formulate the following criteria for differentiating [P + N]_{N/A} compounds from derivatives: 1) absence of dissociation from the corresponding preposition by developing a different prefix-like semantics, 2) the grammatical relationship of subordination inherited from a PP, 3) sensitivity to preposition argument structure/ complementation relation between compound constituents (e.g., Melloni 2020), 4) a transparent synchronic semantic motivation by a PP.

This study was carried out in the framework of the Latvian Council of Science research project “Database of Latvian Morphemes and Derivational Models” (No. lzp-2022/1-0013).

References

- Amiot, Dany (2004), Préfixes ou prépositions? Le cas de *sur(-)*, *sans(-)*, *contre(-)* et les autres, *Lexique* 16, 67–83.
- Bauer, Laurie (2005), The borderline between derivation and compounding, in W. Dressler, D. Kastovsky, O. Pfeiffer, & F. Rainer (eds), *Morphology and Its Demarcations*, Amsterdam: John Benjamins, 97–108.
- Bauer, Laurie (2017), *Compounds and Compounding*, Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108235679>
- Booij, Geert (2010), *Construction morphology*, Oxford: Oxford University Press.
- Endzelīns, Jānis (1951), *Latviešu valodas gramatika*, [Latvian Grammar.] Rīga: Latvijas Valsts izdevniecība.
- Melloni, Chiara (2020), Subordinate and Synthetic Compounds in Morphology, *Oxford Research Encyclopedia of Linguistics*, retrieved 15.01.2025, <https://oxfordre-com.datubazes.lanet.lv/linguistics/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-562>
- Nītiņa, Daina & Juris Grigorjevs (eds), (2013), *Latviešu valodas gramatika*, [Latvian Grammar.] Rīga: Latvijas Universitātes Latviešu valodas institūts.
- Olsen, Susan (2017), Delineating derivation and compounding, in R. Lieber, and P. Štekauer (eds), *The Oxford Handbook of Derivational Morphology*, Oxford: Oxford University Press, 16–49. <https://doi.org/10.1093/oxfordhb/9780199641642.001.0001>
- Stump, Gregory (2017), Derivation and function words, in R. Lieber, and P. Štekauer (eds.), *The Oxford Handbook of Derivational Morphology*, Oxford: Oxford University Press, 317–337. <https://doi.org/10.1093/oxfordhb/9780199641642.001.0001>
- Trips, Carola (2009), *Lexical semantics and diachronic morphology. The development of -hood, -dom and -ship in the history of English*, Tübingen: Niemeyer. <https://doi.org/10.1515/9783484971318>
- The Balanced Corpus of Modern Latvian LVK2018*, available at: <https://repository.clarin.lv/repository/xmlui/handle/20.500.12574/11>
- Ungerer, Friedrich & Hans-Jörg Schmid (2006), *An Introduction to Cognitive Linguistics*, London / New York: Pearson & Longman.
- Van Goethem, Kristel (2009), *L’emploi préverbal des prépositions en français. Typologie et grammaticalisation*, Louvain-la-Neuve: De Boeck-Duculot.
- Van Goethem, Kristel (2020), Affixation in Morphology. *Oxford Research Encyclopedia of Linguistics*, retrieved 15.01.2025, <https://oxfordre-com.datubazes.lanet.lv/linguistics/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-678>

Contact verbs in East Caucasian

Gilles Authier (EPHE, PSL, Paris), Ayten Babaliyeva (UFAZ, Baku), Gasangusen Sulaibonov (EPHE, PSL, Paris)

English ‘hit verbs’ () and ‘touch verbs’ show particular syntactic properties (Fillmore 1970, Levin 1993), which can be interpreted in terms of reduced transitivity: the Target does not undergo a change of state, and the two participants are less “semantically distinct in terms of their roles in the event described by the clause” (Næss 2007).

In indigenous languages of the Caucasus, all characterized by low transitivity prominence (Haspelmath 2015), contact verbs display either intransitive bivalent frames as in (1), found in North-West Caucasian and East Caucasian, including the anti-impersonal type (Lazard 1985) typical of Kartvelian as in (2), also found in East Caucasian, or ditransitive verbs with a bodypart or an instrument cast as a Patient as in (3), as well as ‘coverb constructions’ as in (4), found only in East Caucasian. The common denominator of these constructions is the encoding of the Surface/Target argument, as Indirect Object, using either ‘oblique’ cases (Dative or Spatial cases) and/or Indirect Object indexing; the Direct Object slot is either absent or reserved for instruments or a body part of the Agent. This “pan-Caucasian valence pattern” (Nichols 1984:188), involving an apparent reversal in the expression of Target and Instrument arguments as Direct and Indirect objects, suffers few exceptions, such as (5), including ‘locative alternations’ as in (6).

Focusing on East Caucasian, we will study the valence frames of contact verbs found in the Bivalentyp database (Say 2020), in search of possible micro-classes, phylogenetic signals or explanations by contact, lexical shift or language-internal structural parameters for the diversity present in this family.

(1) Bivalent intransitives

a. (North-West Caucasian) Abaza (=Adyghe)

<i>ala</i>	<i>sara</i>	<i>j-ša-sə-cha-d</i>
DEF.dog	I	3SG.N.ABS-CISL-1SG.IO-bite(AOR)-DCL
‘The dog bit me.’		

b. East Caucasian

i. Kadar Dargwa

<i>mehamad-i-ke</i>	<i>χ^ʷa</i>	<i>q’ac’-b-ik-ib</i>
Mahammad-OBL-SUPER	dog(N)	bite-N-fall.PFV-PRET
‘A dog bit Muhammad.’		

ii. Tsakhur

<i>murad-i-k</i>	<i>xuvaa</i>	<i>ats’abk’in</i>
Murad-OBL-CONT	dog(A)	A.bite.PFV
‘A dog bit Murad.’		

(2) Anti-impersonal intransitives

a. (South Caucasian=Kartvelian) Svan (= Georgian, Megrelian, Laz)

<i>žey-d</i>	<i>maizer-s</i>	<i>ä-x-qäp</i>
dog-ERG	Maizer-DAT	PV:DIST-IO3-bite(AOR)
‘A dog bit Maizer.’		

b. East Caucasian

i. Avar

<i>šali-da</i>	<i>hanč’-ana</i>	<i>ho-jał</i>
Ali-LOC	bite-AOR	dog-ERG
‘A dog bit Ali.’		

ii. Zilo Andi

<i>χ^ʷedo-di</i>	<i>q’ammij</i>	<i>pat’imati-?a</i>
dog.OBL-ERG	bite.AOR	Patimat-SUPER
‘The dog bit Patimat.’		

- iii. Khwarshi
ɤ^wa-ji haj-na šamil-es
 dog-ERG bite-AOR Shamil-GEN1
 ‘A dog bit Shamil.’
- (3) Ditransitives: East Caucasian
 Mukhad Rutul
murad-i-s tɬije-re sis viʔi-ri
 Murad-OBL-DAT dog.OBL-ERG bite(A) A.do.PFV-AOR
 ‘A dog bit Murad.’
- (4) Coverb constructions: East Caucasian
- i. Chechen
žala-s islam-ana ka toex-na
 dog-ERG Islam-DAT bite hit.PFV-AOR
 ‘A dog bit Islam.’ (cf. *cerg* ‘tooth’)
- ii. Kandyk Tabasaran
χu-ji ma^ha^hmad.a-z q’ac’ ɤ-ap’-nu.
 dog-ERG Mahamad-DAT bite PFV-do-PST
 ‘A dog bit Mahamad.’ (cf. *silib* ‘tooth’)
- iii. Chirag Dargwa
χ:uru-d ʔa li-l-i q’ac’ ib.
 dog-ERG Ali-OBL-SUPER bite say:PFV.AOR(3)
 ‘The dog bit Ali.’ (cf. *cula* ‘tooth’)
- (5) Transitive: East Caucasian
- i. Archi
paša g^wači-li w-eq:’u.
 Pasha(M) dog-ERG M-bite.PFV
 ‘The dog bit Pasha.’
- ii. Sagada Tsez
ğway-e xan-i pat’imat
 dog-ERG bite-AOR Patimat
 The dog bit Patimat.’
- (6) Locative alternation: (East Caucasian) Budugh (=Northern Akhvakh)
- i. transitive
vida’di-ri yeç k’ıpsa-ci
 Vidadi-ERG apple(A) A.bite.PFV-AOR
 ‘Vidadi bit in the apple.’
- ii. anti-impersonal
xor-ur aslan-a’ k’ısa-ci.
 dog-ERG Aslan-LOC bite.PFV-AOR
 ‘A dog bit Aslan.’

References

- Fillmore, Charles. J. 1970. The Grammar of *hitting* and *breaking*, in R. Jacobs & P. Rosenbaum, eds., *Readings in English Transformational Grammar*, Ginn, Waltham, MA, 120-133.
- Haspelmath, Martin. 2015. Transitivity prominence. In: Malchukov, Andrej L. & Comrie, Bernard (eds.) *Valency classes in the world’s languages*. Berlin: De Gruyter Mouton
- Lazard, Gilbert. 1985. Anti-impersonal verbs, transitivity continuum and the notion of transitivity. In Hansjakob Seiler & Gunter Brettschneider, eds. *Language invariants and mental operations*. Tübingen: Gunter Narr Verlag, 115-123.
- Levin, Beth. 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago, IL:University of Chicago Press.
- Næss, Åsschild. 2007. *Prototypical transitivity*. Amsterdam / Philadelphia: John Benjamins.
- Nichols, Johanna. 1984. Direct and Oblique Objects in Chechen-Ingush and Russian, in F. Plank, ed., *Objects*, Academic Press, New York, 183-209.
- Say, Sergey (ed.). 2020-. BivalTyp: Typological database of bivalent verbs and their encoding frames. (Available online at <https://www.bivaltyp.info>, Accessed on 25 December 2024.)

Table 2: Derivational patterns across template patterns

	SIMPLE	INTENSIVE	CAUSATIVE
√wqr	‘become rare’	‘make rare’	‘cause to become rare’
√lmd	‘learn’	‘inform’	‘cause to learn’
√lk	‘go’	/	‘cause to go’
√prs	‘separate, cut off’	‘chop off’	‘cause to separate’

Kouwenberg (1997, 2010) suggests that unaccusatives and EXPERIENCER-transitives form the causative-alternation through the INTENSIVE pattern, while unergatives and AGENT-transitives causativise through the CAUSATIVE. As shown in Redacted (Accepted), however, certain classes of unaccusatives exclusively form the causative-alternation through the CAUSATIVE (e.g., verbs of existence, following Levin 1993), while classes of unergatives demonstrate mixed behaviour with some verbs transitivity through the INTENSIVE, the CAUSATIVE, or both (e.g., verbs of emission).

To account for these alternations in causative formation, I argue that INTENSIVE verbs denote direct causation through the Merge of an AGENT at VoiceP (following Kastner 2020), while CAUSATIVE verbs denote indirect causation with a CAUSER. Whether or not a root then causativises through the INTENSIVE or CAUSATIVE, relies on a. aspectual properties of roots, and b. their inner predisposition for internally-caused change (ICC, Bentley 2023): more durative roots predisposed for ICC take INTENSIVE causatives, while punctual roots non-predisposed for ICC require the introduction of external, indirect causers through the CAUSATIVE. The heightened syntactic agentivity of INTENSIVES further accounts for its ‘intensive’, semantically increased-agentive usages.

References:

- Bentley, Delia. 2023. Internally caused change as change by inner predisposition: Comparative evidence from Romance. *Journal of Linguistics*. 60(3) 483–525.
- Kastner, Itamar. 2020. *Voice at the interfaces: The syntax, semantics and morphology of the Hebrew verb*. Berlin: Language Science Press.
- Kouwenberg, N. J. C. 1997. *Gemination in the Akkadian verb*. Uitgeverij Van Gorcum.
- Kouwenberg, N. J. C. 2010. *The Akkadian Verb and its Semitic Background* (Languages of the Ancient Near East 2). Winona Lake: Eisenbrauns.
- Levin, Beth. 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago: University of Chicago Press.
- Redacted. Accepted. The causative-inchoative alternation in Akkadian – Inherited and innovated strategies. In Guglielmo Inglese, Giulia Mazzola & Eugenio Gorla (eds.), *Title TBD* (Typological Studies in Language). Amsterdam & Philadelphia: John Benjamins.

Understanding and out-organizing:
Identifying taxonomic relations between morphological constructions
through collostructional analysis and clustering methods

Droste, Kim-Kristin
Osnabrück University

Keywords: Construction Morphology, locative prefixoids, collostructional analysis, semantic vector spaces, linguistic productivity

This paper employs a range of corpus methods to examine antonymous pairs of locative prefixoids (*up-down*, *in-out*, *over-under*, *on-off*) in complex words such as *downfall*, *outsource*, *undergraduate*, and *offshore*. Being situated within the framework of Construction Morphology (Booij 2010), it posits constructional idioms with specified parts and open slots and sketches a network of constructions that includes meaning extensions of the individual prefixoids as well as taxonomic relations between them.

Bauer et al. (2013: 345) provide an overview of opposite spatial axes of these antonymous prefixoids and Lieber (2004: 107) posits antonymous semantics for the corresponding prepositions. However, the relations between directional antonyms of locative prefixoids have not yet been systematically researched. Although the prefixoids can be attached to the same bases to create antonyms (*upstairs-downstairs*, *ingrown-outgrown*, *overstate-understate*, *online-offline*), semantic equivalence is not consistently maintained across all constructions. Specifically, some meanings expressed by one construction, e.g. the ‘defeat’ meaning for *out-* in words like *outdance* (Kotowski 2023: 134), do not have a corresponding antonym, here in the *in-* construction.

Van Goethem & Norde (2020) have shown that a combination of multiple distinctive collexeme analysis, semantic-distributional profiles by way of semantic vector spaces, and productivity measures is a profitable method of determining similarities among functionally similar constructions, in their case of extravagant “fake” morphemes in Dutch. Research has also shown that clustering methods are able to identify antonyms (Gries & Otani 2010). This paper applies the methodology employed by Van Goethem & Norde (2020) to pairs of directional antonyms to determine similarities and differences among them.

First, samples of the eight constructions are extracted from the ELEXIS English Web 2020 corpus (Jakubíček et al. 2022). Second, a multiple distinctive collexeme analysis (Stefanowitsch 2013) is conducted. This method “is concerned with collexemes that are significantly associated with a (particular slot in a) construction as compared to a semantically or functionally similar construction” (Stefanowitsch & Gries 2005: 8). In the present analysis, this method can determine whether specific bases are significantly attracted by one or more of these constructions. In a third step, these results are submitted to a clustering procedure, clustering the constructions by their bases (Levshina 2015: 323–332). Fourth, the distinctive collexeme ratio of each construction is calculated, i.e., “its number of distinctive collexemes divided by its number of types in the corpus sample” (Van Goethem & Norde 2020: 443). Multiple productivity measures (Baayen & Lieber 1991) serve to further compare the constructions and locate them on Barðdal’s (2008) productivity cline.

The aim of this study is to determine to which extent complex words with locative prefixoids have similar semantic-distributional profiles and productivity measures and whether there are differences within and among the four antonymous pairs. For example, antonyms where one member

exhibits more extended meanings (e.g. *over-*, *out-*) are expected to show larger differences in their semantic-distributional profile, productivity, and distinctive collexeme ratio. This study also investigates whether locatives making reference to the same end of the spatial axis share certain similarities (e.g. *up-*, *over-* and *down-*, *under-*).

References

- Baayen, Harald & Rochelle Lieber (1991), Productivity and English derivation: A corpus-based study, *Linguistics* 29, 801–843.
- Barðdal, Jóhanna (2008), *Productivity: Evidence from case and argument structure in Icelandic*, Amsterdam: John Benjamins.
- Bauer, Laurie, Rochelle Lieber & Ingo Plag (2013), *The Oxford reference guide to English morphology*, Oxford: Oxford University Press.
- Booij, Geert (2010), *Construction morphology*, Oxford: Oxford University Press.
- Gries, Stefan Th. & Naoki Otani (2010), Behavioral profiles: A corpus-based perspective on synonymy and antonymy, *ICAME Journal* 34, 121–150.
- Jakubíček, Miloš, Vít Suchomel, Frederico Martelli & Roberto Navigli (2022), Semantically annotated corpora: ELEXIS - European Lexicographic Infrastructure, https://elex.is/wp-content/uploads/ELEXIS_D4_6_Semantically_annotated_corpora.pdf.
- Kotowski, Sven (2023), Modeling locative prefix semantics. A formal account of the English verbal prefix *out-*, *Morphology* 33(2), 115–152.
- Levshina, Natalia (2015), *How to do linguistics with R: Data exploration and statistical analysis*, Amsterdam: John Benjamins.
- Lieber, Rochelle (2004), *Morphology and lexical semantics*, Cambridge: Cambridge University Press.
- Stefanowitsch, Anatol (2013), Collostructional analysis, in Th. Hoffmann, and G. Trousdale (eds), (2013), *The Oxford Handbook of Construction Grammar*, Oxford: Oxford University Press, 290–306.
- Stefanowitsch, Anatol & Stefan Th. Gries (2005), Covarying collexemes, *Corpus Linguistics and Linguistic Theory* 1(1), 1–43.
- Van Goethem, Kristel & Muriel Norde (2020), Extravagant “fake” morphemes in Dutch: Morphological productivity, semantic profiles and categorical flexibility, *Corpus Linguistics and Linguistic Theory* 16(3), 425–458.

A pilot study on neoclassical compounds with *-skopas* ‘-scope’ in contemporary Lithuanian

Lina Inčiuraitė-Noreikienė & Erika Rimkutė
(Vytautas Magnus University)

Keywords: neoclassical compounds, free stems, bound stems, word-formation, contemporary Lithuanian

Over the past ten years, a growing interest in neoclassical compounds has been observed in linguistics worldwide (cf. Díaz-Negrillo 2020: 213–261; Bauer 2017: 148–156; Panocová 2016: 192–207, etc.). In Lithuanian, only hybrid compounds and their integration into the native word-formation system have received more attention (cf. Inčiuraitė-Noreikienė 2017; Inčiuraitė-Noreikienė 2015; Inčiuraitė-Noreikienė & Stundžia 2015).

A characteristic feature of Lithuanian compounding is the combination of two stems, mostly simplex and functioning as independent words, e.g., the native compound *saulegrąža* ‘sunflower’ combines two free stems, *saule* ‘sun’ and *grąža* ‘return’, both of which are independent words (for more on Lithuanian compounds, see Stundžia, 2016: 3091–3093). In contrast, neoclassical compounds differ significantly as they are usually formed from Greek or Latin elements, many of which do not function as independent words but as bound stems, e.g., *mikroskopas* ‘microscope’ is formed from two bound stems, *mikro-* ‘micro’ and *-skopas* ‘-scope’, which cannot function independently in Lithuanian. This fundamental difference in stem status justifies analysing neoclassical compounds based on the relationship between free and bound stems.

This pilot study employs a synchronic approach to investigate neoclassical compounds featuring *-skopas* ‘-scope’ as the second constituent in contemporary Lithuanian from a word-formation perspective. We address two research questions: 1. To what extent are neoclassical compounds prevalent compared to the prefixed derivatives in the analysed corpora? 2. How can neoclassical compounds be categorized based on their internal morphological structure? The comparative method focuses on identifying similarities and differences among neoclassical compounds. The descriptive method offers insights into their combinability properties. The analytical method is used to examine the internal morphological structure of neoclassical compounds and the groups they may form.

The investigation draws upon data sourced from the *Corpus of Contemporary Lithuanian Language* (CCLL) and the *Joint Corpus of Lithuanian* (JCL). A total of 186 nouns with *-skopas* ‘-scope’ were extracted from these sources. Of the analysed formations, 96% are neoclassical compounds (e.g., *megaskopas* ‘megascope’ ← *mega* ‘mega’ + *skopas* ‘scope’) and 4% are prefixed derivatives (e.g., *ultramikroskopas* ‘ultramicroscope’ ← *ultra* ‘ultra’ + *mikroskopas* ‘microscope’). Based on their internal morphological structure, we categorized the neoclassical compounds into four groups: 1) compounds consisting of two bound stems (59%), e.g., *galvanoskopas* ‘galvanoscope’ ← *galvano* ‘galvano’ + *skopas* ‘scope’, 2) a free and a bound stem (19%), e.g., *kapiliaroskopas* ‘capillaroscope’ ← *kapiliaras* + *skopas* ‘scope’, 3) a bound and a free stem (17%), e.g., *biomikroskopas* ‘biomicroscope’ ← *bio* ‘bio’ + *mikroskopas* ‘microscope’ and 4) two free stems (5%), e.g., *rentgenodefektoskopas* ‘X-ray defectoscope’ ← *rentgenas* ‘X-ray’ + *defektoskopas* ‘defectoscope’. The predominance of compounds with two bound stems highlights the typical morphological pattern of neoclassical compounds.

Acknowledgements: This project has received funding from the Research Council of Lithuania (LMTLT), agreement No S-PD-24-170.

We would like to thank Virginijus Dadurkevičius, a data analyst at Vytautas Magnus University's Institute of Digital Resources and Interdisciplinary Research (SITTI), for providing the data.

References

- Bauer, Laurie (2017), *Compounds and Compounding*, (Cambridge Studies in Linguistics 155), Cambridge: Cambridge University Press.
- Díaz-Negrillo, Ana (2020), Neoclassical word-formation in English: A paradigm-based account of -scope formations, in J. Fernández-Domínguez, A. Bagasheva, and C. Lara-Clares (eds), (2020), *Paradigmatic relations in word-formation*, Leiden: Brill, 213–261.
- Inčiuraitė-Noreikienė, Lina, and Bonifacas Stundžia (2015), On the patterns of word-formation in hybrid neoclassical nouns and adjectives in Lithuanian, in S. Kessler, and A. Judžentis (eds), (2015), *Contributions to Syntax and Morphology*, Berlin: Logos, 27–50.
- Inčiuraitė-Noreikienė, Lina (2017), *Morphemic structure and derivation of Lithuanian nominal formations with (neo)classical elements*, Doctoral dissertation, Vilnius: Vilniaus universiteto leidykla.
- Inčiuraitė-Noreikienė, Lina (2015), Lietuvių kalbos dūriniai su neoklasikiniais dėmenimis [Lithuanian compounds with neoclassical constituents], *Baltistica* 50(2), 245–259.
- Panocová, Renáta (2016), Neoclassical compounds in the onomasiological approach, in Pius ten Hacken, and Renáta Panocová (eds), (2016), *The Semantics of Compounding*, Cambridge: Cambridge University Press, 192–207.
- Stundžia, Bonifacas (2016), Word-formation in Lithuanian, in P. O. Müller, I. Ohnheiser, S. Olsen, and F. Rainer (eds), *Word-Formation: An International Handbook of the Languages of Europe*, vol. 5, Berlin: Mouton de Gruyter, 3089–3106.

Sources

- DLKT – Dabartinės lietuvių kalbos tekstynas [Corpus of Contemporary Lithuanian Language], Kaunas: Vytauto Didžiojo universitetas (<http://tekstynas.vdu.lt/tekstynas/>), 2011, (Accessed 2025-06-01).
- JCL – Virginijus Dadurkevičius, Wordlist of Lemmas from the Joint Corpus of Lithuanian, CLARIN-LT digital library in the Republic of Lithuania (<http://hdl.handle.net/20.500.11821/41>), 2020, (Accessed 2025-01-06).

Iconicity in verbs of motion in English – where do we go from here?

Merlin Kim Gray

(Université Of Caen Normandie – CRISCO)

Keywords: Iconicity, diachronic change, manner of motion, verbs of motion

This paper, within the context of my PhD, investigates the impact of iconicity on semantic changes in manner-of-motion verbs in English — more specifically analyzing the development of figurative meanings from a diachronic perspective using a combined method (see Smith (2016, 2018, 2020)). I propose to investigate a subset of motion verbs with iconic formations — imitative verbs such as *plash*, *plod*, and *plop*. Per Dingemanse (2012, p. 663), iconic verbs are widely used in describing conceptual categories of sound, movement, visual impressions, inner feelings, and cognitive states, making them critical in the study of semantic change. By linking the study of semantic change in motion verbs to their iconic origins, this research seeks to illustrate how iconicity influences semantic evolution (see Flaksman’s treadmill hypothesis (2017)).

To construct a dataset and imitative subset, several steps were taken following Smith (2016,2018,2020). Using the Oxford English Dictionary (OED3) the keywords ‘move’ and ‘motion’ were searched within the ‘definition’ sections, yielding 1438 lexemes from Old English to contemporary usage. Keywords such as ‘echoic,’ ‘onomatopoeic,’ ‘imitative,’ and ‘expressive’ were then used to identify iconic verbs. Further searches within individual word pages, as with *seal* v1 (Oxford English Dictionary, 2024c), were conducted to capture instances of imitative usage not listed under etymology.

The resulting subset consists of 136 verbs characterized under etymological keywords ‘imitative’: 69, ‘echoic’: 17, ‘expressive’: 12 or ‘onomatopoeic’: 41. Of these 136 iconic verbs, 78 verbs exhibit figurative usages as defined by the dictionary, underscoring the significance of figurative meanings in this study. Morphologically echoic verbs with figurative senses exemplify Traugott and Dasher’s (2002, p. 62) assertion that new meanings expand words’ semantic scopes, enabling broader usage beyond their original morphological constructions. Early alphabetical examples include *birl*, *bob* & *bump* with figurative meanings occurring respectively in meanings 1, 2.c. & 1.5.a.

Semantic and chronological analyses of the dataset are carried out to determine the types of semantic change (e.g. metaphor, metonymy, broadening) and the chronology of these changes. For example, *bump* is first attested in 1588, with its figurative sense (‘1.5.a’) emerging 200 years later in 1788. In contrast, *birl* displays figurative usage synchronically with its first attestation in 1789, while its oldest meaning (‘3.a.’) dates to 1724.

These findings will provide useful insight into the processes of semantic change relating to verbs expressing manner-of-motion and the role of iconicity in effecting this development.

Bibliography:

- Dingemanse, M. (2012). Advances in the Cross-Linguistic Study of Ideophones. *Language and Linguistics Compass*, 6(10), 654–672. <https://doi.org/10.1002/lnc3.361>
- Flaksman, M. (2017). Iconic treadmill hypothesis: The reasons behind continuous onomatopoeic coinage. In M. Bauer, A. Zirker, O. Fischer, & C. Ljungberg (Eds.), *Iconicity in Language and Literature* (Vol. 15, pp. 15–38). John Benjamins Publishing Company. <https://doi.org/10.1075/ill.15.02fla>
- Oxford English Dictionary. (2023). *Birl*, v.². Oxford University Press; Oxford English Dictionary. <https://doi.org/10.1093/OED/1451075119>
- Oxford English Dictionary. (2024a). *Bob*, v.³. Oxford University Press; Oxford English Dictionary. <https://doi.org/10.1093/OED/1031359048>
- Oxford English Dictionary. (2024b). *Bump*, v.¹. Oxford University Press; Oxford English Dictionary. <https://doi.org/10.1093/OED/8299687169>
- Oxford English Dictionary. (2024c). *Seal*, v.¹. Oxford University Press; Oxford English Dictionary. <https://doi.org/10.1093/OED/2709161265>
- Smith, C. A. (2016). Tracking semantic change in fl – Monomorphemes in the Oxford English Dictionary. *Journal of Historical Linguistics*, 6(2), 165–200. <https://doi.org/10.1075/jhl.6.2.02smi>
- Smith, C. A. (2018). Where do new words like boobage, flamage, ownage come from? Tracking the history of -age words from 1100 to 2000 in the OED3. *Lexis. Journal in English Lexicology*, 12, Article 12. <https://doi.org/10.4000/lexis.2167>
- Smith, C. A. (2020). A Case Study of -some and – able Derivatives in the OED3: Examining the Diachronic Output and Productivity of Two Competing Adjectival Suffixes. *Lexis. Journal in English Lexicology*, 16, Article 16. <https://doi.org/10.4000/lexis.4793>
- Traugott, E. C., & Dasher, R. B. (2002). *Regularity in semantic change*. Cambridge University Press.

Mru Numeral Classifiers: Synchronic and Diachronic Perspectives

Mithun Banerjee, PhD researcher, University of Helsinki

Erika Sandman, Docent in General Linguistics and grant-funded researcher, University of Helsinki

This presentation will explore the meanings and functions of Mru numeral classifiers, as well as their position within a noun phrase from both synchronic and diachronic perspectives. We will also explore the word order change in the Mru noun phrase in a broader Sino-Tibetan context and contrast it with languages undergoing similar changes, such as Northwest Mandarin. Mru is an extremely understudied and unclassified Sino-Tibetan language spoken by approximately 30,000 (BSB, 2011) speakers in the southeastern region of Bangladesh. There are four classifiers in Mru: one generic classifier *-tɕɔ*, one human classifier *-iya*, and two non-human classifiers *-buk* and *-klun*. In addition to the semantic parameter of humanness, the choice of a classifier in Mru depends on the number of items being counted. The generic classifier *-tɕɔ* is used when counting only one item, and it is used for both human and non-human nouns, while the human classifier *-iya* is used exclusively for humans in other contexts. For non-human nouns, the non-human classifier *-buk* is used with numbers two and three, while *-klun* is used with numbers higher than three. Nominal classifiers in Mru are used with numbers, demonstratives, and sometimes with possessive constructions. In terms of their position, numerals and classifiers show a mixed pattern of pre-nominal and post-nominal structures, which is uncommon in Sino-Tibetan languages. Tibeto-Burman languages typically have classifiers after the noun (77 out of 105 languages, see Dryer 2008), whereas in Sinitic languages, such as Standard Mandarin, numeral classifiers are used with numerals, demonstratives, and quantifiers placed before the noun. The examples (1) and (2) exhibit the mixed pattern in Mru noun phrases (Banerjee, forthcoming).

1. *ʔaŋko pre-buk kui woi*
I two-CLF.NON.HUMAN dog have
“I have two dogs”
2. *ʔapa ran ɕiŋ hamut-klun*
Father buy tree ten-CLF.NON.HUMAN
“father buys ten trees”

In example (1), the non-human classifier *-buk* occupies the pre-verbal position, while the non-human classifier *-klun* is used in the post-verbal position. Yurayong et al. (2024) argue that Mruic languages are shifting from a head-final to a head-initial directionality in their noun phrases, which is possibly due to the combination of internal drift and language contact with predominantly head-initial languages (e.g. Karenic). We will compare the Mru with data on other Sino-Tibetan languages that have undergone similar changes, notably two Northwest Mandarin contact varieties, Wutun (Sandman 2016) and Zhoutun (Zhou 2022), spoken in Qinghai Province, P.R. China. Both languages have developed mixed word order pattern for numerals and classifiers due to the language contact with Amdo Tibetan. The mixed pattern is illustrated by two Wutun examples (Sandman 2016: 44, 47), in which (3) shows the post-nominal order and (4) the pre-nominal order.

3. *Qhichai Liang-ge*
car two-CLF.GENERIC
“two cars”
4. *liang-ge yai*
two-CLF.GENERIC month
“two months”

The Mru data presented in this research is primarily derived from the on-site fieldwork (October 2023 to December 2023, Bandarban, CHT in Bangladesh) and distributed fieldwork (2023-present).

References

- Banerjee, Mithun. Forthcoming. *Descriptive Grammar of Mru*. PhD dissertation. University of Helsinki.
- BSB. 2011. Bangladesh Statistical Bureau, People's Republic of Bangladesh.
- Dryer, Matthew S. 2008. Word order in Tibeto-Burman languages" *Linguistics of the Tibeto-Burman Area* 31: 1–88.
- Sandman, Erika. 2016. *A grammar of Wutun*. Ph.D. dissertation, University of Helsinki.
- Yurayong Chingduang, Sandman Erika, Kerbs Richard, Banerjee Mithun & Pui Yiu Szeto (in press). Head directionality in Northern Sinitic varieties: An areal-typological and comparative Sino-Tibetan perspective. In Redouane Djamouri, Christine Lamarre, Julie Lefort (eds.) *Language Contacts in Northern China*. Leiden: Brill.
- Zhou, Chenlei 2022. *Zhoutun*. London: Routledge.

The typology and morphosyntax of discontinuous numerals

Xavier Bach

This presentation is concerned with the syntax and morphology of cardinal numerals, specifically larger numerals, above the numeral base of the language system, when they are used attributively. In a number of languages, such numerals can be discontinuous, i.e. they do not appear to form a coherent constituent but can instead be interrupted, generally by the noun they quantify over. Only a handful of languages have been reported to present such numerals in the literature, including Guajiro, Kikongo, and Breton (Hammarström 2008), Welsh and Obolo (Dryer 2013), to which Grambank adds a handful of examples. In Obolo for example (Benue-Congo, Nigeria, cited after Dryer 2013), the discontinuity is due to the normal differential placement of some numerals: numeral words for 10, 20 and 400 always precede the noun they quantify over, while all other numerals consisting of only one word follow it, as shown in example 1:

- 1) étíp úwù mè gò
 twenty house and five
 ‘twenty-five houses’

In his study about the typological distribution of the order of the noun and the numeral in NPs, Dryer (2013) indicates that a minority of languages can present two different orders: NNum or NumN; in such cases, it is usual for lower numerals to appear on one side of the noun, while all others appear to the other side. This may well be the origin of discontinuous numerals in such languages, as shown for Obolo.

The appearance in the literature of two Celtic languages is no surprise: discontinuous numerals are in fact a feature of this language family, including those not reported in the typological literature. Thus in Cornish, as is the case in Welsh, the noun follows the first numeral element in a complex numeral, although in Cornish the order of elements does not seem to be as fixed (example 2 and 3, from Jenner 2010):

- 2) trŷ igans bledhan ha deg
 sixty year and ten
3) deg bledhan war trŷ igans
 ten year on sixty
 ‘seventy years’

In Celtic languages, the discontinuity of numerals follows one of two patterns: either the first numeral element is directly followed by the noun (Cornish, Welsh), or the noun immediately follows the unit element (up to 19 in the vigesimal system of counting, as is the case in Breton, and in Scottish Gaelic for complex numbers below 30), although some varieties can allow for more flexibility as to where the noun will appear in larger, complex numerals (as in Scottish Gaelic, see Lamb 2024: 203).

I describe the various patterns attested in the languages of the world, with particular emphasis on the Celtic languages. The data obtained question the syntactic status of complex numerals. If lower numerals are often considered more adjectival and higher numerals more nominal (see Veselinova 2020) when not classified as a specific word class, they also exhibit some of the properties of phrases: discontinuous constituents are widely attested in the languages of the world, but all available tests for wordhood emphasize that words should not be interruptible.

References

Dryer, Matthew. 2013. Order of numeral and noun. In Matthew Dryer & Martin Haspelmath (eds.) *WALS Online* (v. 2020.4), (Available online at <http://wals.info/chapter/89>, Accessed on 2024-10-26.)

Hammarström, Harald. 2008. Complexity in numeral systems with an investigation into pidgins and creoles. In Matti Miestamo, Kaius Sinnemäki & Fred Karlsson (eds.), *Language Complexity: Typology, contact, change*. Amsterdam: Benjamins. 287-304.

Jenner, Henry. 2010. *Handbook of the Cornish Language*. Revised by Michael Everson. Cnoc Sceichín: Everttype. [original edition London: David Nutt, 1904].

Lamb, William. *Scottish Gaelic: A Comprehensive Grammar*. Abingdon, Oxon / New York: Routledge.

Veselinova, Ljuba. 2020. Numerals in Morphology. *Oxford Research Encyclopedia of Linguistics* Retrieved 26 Oct. 2024, from <https://oxfordre.com/linguistics/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-559>.

Word Formation Processes in Proper Names and Common Nouns

Rafael Antonio Gutiérrez Martínez
(Universitat Pompeu Fabra)

Keywords: common nouns, predicativism, proper names, referentialism, word formation processes

The differences in word formation processes between proper names (PNs) and common nouns (CNs) are understudied. In practice, morphologists often assume that the processes affecting these two types of words are the same and operate in similar ways (e.g., Fradin, 2017, Bonami & Guzmán, 2023, Mauri & Masini, 2024). Moreover, while studies exploring the morphosyntactic distinctions between PNs and CNs have examined e.g. modification, case marking, and definiteness (Caro Reina & Helmbrecht, 2022), they have overlooked potential differences in derivational processes. However, given that the two types of words belong to distinct semantic classes—one referring to specific individuals and the other to classes of individuals—it is reasonable to expect some morphological distinctions, specifically in (a) the precise morphological processes employed, (b) the productivity of the morphemes, and (c) the semantic effect of the morphemes on each type of nominal.

Here I report on the results of a large-scale comparison between words derived from PNs and those derived from CNs, based on Spanish corpus data. I focus specifically on prefixation and issues (a) and (b). The data, from CORPES XXI and the Corpus del Español (Davies, 2016), were extracted by first defining a set of PNs (belonging to various categories, e.g., philosophers, cities) and CNs (broadly of the same ontological types as the PNs: humans, places) and then manually looking for all the prefixed words of these categories. This yielded 574 words based on PNs and 841 based on CNs.

The results show: (1) some prefixes are found only in words derived from PNs (e.g., *vetero-*, *tardo-*), (2) some prefixes are found only with those derived from CNs (e.g., *tri-*, *exo-*, *pluri-*), and (3) although some prefixes are equally productive with both PNs and CNs (e.g., *anti-*, *pre-*), others are very productive only with PNs (e.g., *neo-*, *pro-*) or only with CNs (e.g., *ex-*, *extra-*).

While these results should be interpreted with caution—particularly since the absence of certain forms in the data does not necessarily imply their impossibility (a Spanish speaker, for instance, can easily conceive of CNs with *tardo-*)—I argue that some of the observed differences are tied to the specific semantic contributions of the bases. For example, the tendency of PNs to resist quantifier prefixes (Fabregas, 2024), such as *pluri-*, *tri-*, and *bi-*, compared to their presence in CNs, can be linked to the contrasting denotational properties of these word types (unique individual vs. class of individuals).

I conclude by proposing that these results raise questions for both predicativist and referentialist theories of PNs. If PNs function as predicates in the same way as CNs, it remains unclear how predicativist accounts can explain the observed differences in the morphological profiles of PNs and CNs: All things being equal, we might expect the same morphological processes to apply equally to both. Conversely, if PNs contribute only their referent, how words formed from PNs come to have the specific meanings they do is challenging to explain.

References

Bonami, Olivier and Guzmán Naranjo, Matías (2023), Distributional evidence for derivational paradigms in S. Kotowski and I. Plag (eds), (2023), *The Semantics of Derivational Morphology: Theory, Methods and Evidence*. Berlin: De Gruyter. 219-258.

- Caro Reina, Javier and Helmbrecht, Johannes (2022), *Proper Names versus Common Nouns. Morphosyntactic Contrasts in the Languages of the World*. Berlin: De Gruyter Mouton.
- Davies, Mark (2016), *Corpus del Español: Web/Dialects*. Available online at <http://www.corpusdelespanol.org/web-dial/>.
- Fábregas, Antonio (2024), Deconstructing Spanish prefixation, *Glossa: a journal of general linguistics* 9(1), 1–35.
- Fradin, Bernard (2017), The multifaceted nature of denominal adjectives, *Word Structure* 10(1), 27–53.
- Mauri, Caterina and Masini, Francesca (2024), Multi-layered indexicality: When proper names become categories, in P. J. Nielsen and M. S. Sansiñena (eds), (2024), *Indexicality: The Role of Indexing in Language Structure and Language Change*. Berlin: De Gruyter, 171–196.
- RAE: Banco de datos (CORPES XXI) [online]. *Corpus del Español del Siglo XXI (CORPES)*: <http://www.rae.es> [Consulted online November, 2024].

General Session : Morphosyntax

The light verb *eman* ‘give’ as a causative verb in Basque psych complex predicates

Ane Berro (University of Deusto) and Beatriz Fernández (UPV/EHU)

In this talk we aim to analyze Basque psych complex predicates that involve the light verb *eman* ‘give’ (1). These complex predicates consist of a psych noun, such as *beldur* ‘fear’, along with the light verb *eman* ‘give’. The psych nouns attested in this constructions include *amorr* ‘rage, anger’, *ardura* ‘concern, worry’ *atsegin* ‘pleasure, pleasant’ among many others. Most of these nouns also occur -as denominal verbs– in verbs belonging to Class II psych verbs of Belletti and Rizzi’s (1988) (2b). Additionally, in Basque, these verbs alternate between the intransitive and transitive constructions – (2a) and (2b), respectively–, depending on whether the experiencer surfaces as the subject (2a) or the object (2b). Psych predicates with *eman* exhibit a third pattern, where the experiencer appears in a ditransitive-like construction, marked by dative case and agreement (1).

In this talk we will argue that in psych causative constructions, the light verb *eman* does not function as a transfer verb, as it habitually does, but as a causative verb that introduces a causer –not an agent– and relates it to the psychological state.

We base this hypothesis on the following evidence. First, thematically most psych nouns that combine with *eman* in these predicates resemble Class II psych verbs rather than transfer verbs. Secondly, with regards to derivational morphology, *eman* complex predicates pattern with change-of-state verbs where the subject is a causer rather than an agent. For instance, when these complex predicates take the suffix *-garri* (similar to the suffix *-ble* in English or Catalan, Oltra-Massuet 2013) to derive an adjective, they give rise to adjectives that have a causative interpretation: e.g. from *lotsa eman* [shame give] ‘shame’ to *lotsa eman-garri* ‘embarrassing’. This is also the case of change-of-state verbs such as *hil* ‘kill/die’, *hil-garri* ‘lethal’. This distribution contrast to that of agentive verbs like *eraman* ‘carry’, which under the suffix *-garri*, they become dispositional adjectives: *eraman-garri* ‘portable, sth that can be carried’. Other phenomena in derivational morphology point in the same direction. In fact, leaving aside some rare exceptions, there are no agentive *-le* nouns derived from the causative *eman*, whereas *-le* nouns derived from the transfer (light) verb *eman* are common. For example, we do not find **beldur-emaile* [fear-giver] ‘intended: one who frightens’ but we do find *aholku emaile* [advice-giver] ‘one who gives counsel, counselor’. Thirdly, psych complex predicates are incompatible with the morphological causative suffix *-arazi*. All verbs in Basque can be selected by this causative morpheme, e.g. from *eraman* ‘carry’, *eraman-arazi* ‘cause to carry’. But from *beldur eman* [fear give] ‘frighten’, we cannot get **beldur eman-arazi* ‘intended: cause to frighten’. We will argue that all these contrasts derive from the fact that the light verb *eman* in psych complex predicates is a causative verb. As such, it is incompatible with another causative morpheme, and its subject behaves as a causer, rather than as an agent.

- (1) Mamu-a-k Jon-i beldur(-ra) ema-ten
ghost-DET-ERG Jon-DAT fear-DET(ABS) give-IPFV
dio.
(3ABS).have.3SGDAT.3ERG
‘The ghost frightens Jon.’ (lit. ‘The ghost gives (a) fear to Jon’)

- (2) a. Jon beldur-tzen da mamu-a-rekin.

- Jon(ABS) frighten-IPFV (3ABS).be ghost-DET-COMIT
 ‘Jon becomes frightened with the ghost.’
- b. Mamu-a-k Jon beldur-tzen du.
 ghost-DET-ERG Jon(ABS) frighten-IPFV (3ABS).have.3ERG
 ‘The ghost frightens Jon.’

References

- Alexiadou, Artemis and Gianina Iordăchoaia. 2014. The psych causative alternation. *Lingua* 148: 53–79.
- Belletti, Adriana and Luigi Rizzi. 1988. Psych-verbs and θ -theory. *Natural Language and Linguistic Theory* 6: 291–352.
- Oltra-Massuet, Isabel. 2013. *Deverbal adjectives at the interface. A crosslinguistic investigation into the morphology, syntax and semantics of -BLE*. De Gruyter Mouton.

Encoding future in Bangime (Mali)

Montébran Aurore

(LLACAN - projet ERC Bang; CNRS/INALCO/EPHE)

Keywords: Future, Bangime, TAM auxiliaries, imperfective, morphosyntax

This paper presents a descriptive study on the morphosyntactic encoding of future in Bangime. It analyzes the synchronic tense-aspect-mood (TAM) picture, with a specific focus on future, with regards to the diachronic development of the TAM markers (in particular the future one). This language (spoken in Mali), is considered as a mysterious isolate surrounded by the Dogon languages, Fulfulde (Atlantic) and the Mande languages.

In the West African languages, aspect is a fundamental element of the TAM system. The main aspectual opposition is between perfective and imperfective, and future is mostly related to the imperfective aspect. In Bangime future has an overt morphological marking, in an analytic construction. One of the specificities of this language is the use of word order variations as a feature distinguishing perfective and imperfective, and the extensive use of complex predicates involving auxiliaries. In Bangime, the imperfective system gathers the imperfective and future constructions (terminology of Heath & Hantgan 2018). Synchronically, the habitual and progressive meanings are encoded by the imperfective paradigm, whose dedicated marker is the auxiliary *dà*. The future is marked by the auxiliary *nàw*, which has a minority variant *ndàw* (composed of a nasal proclitic subject morpheme [n=] and an auxiliary *dàw*). However, “[t]he distinction between imperfective and future constructions is neutralized in favor of the imperfective when a constituent is focalized” (Heath and Hantgan 2018: 283). Looking at the auxiliaries *dà* and *nàw* more closely, Heath and Hantgan suggest that in a diachronic perspective *nàw* may be an allomorph of *dà* occurring in specific grammatical contexts (Heath and Hantgan 2018: 283). They note that “[t]he variants with *nd* [...] point to an origin as *[ŋ dàw], with a nasal [proclitic subject] morpheme and an auxiliary *dàw*. The latter suspiciously resembles imperfective *dà* [...]. Imperfective *dà*, unlike future *nàw*, is always nonfinal in the clause, so it too could reflect an older **dàw* whose *w* has been lost” (Heath and Hantgan 2018: 283). Bybee et al. (1994) showed the diachronic path from progressive (itself mostly originating from a locative construction) to general imperfective up to habitual. Considering the range of meanings encoded by the *dà/nàw/n-dàw* auxiliary in Bangime, the hypothesis is that the development of these morphemes may be an example of this grammaticalisation path, including future.

Based on the data presented in Heath and Hantgan (2018), the study will firstly explore and synchronic uses of future; then describe the morphological and tonal features of the future paradigm; and lastly look at the diachronic evolution of the auxiliary marking future.

References

Bybee, J. L., Perkins, R & Pagliuca W. 1994. *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. The University of Chicago Press.

Heath, J. & Hantgan, A. 2018. Grammar of Bangime: Langue isolate of Mali. Mouton Grammar Library. Berlin: De Gruyter Mouton.

Polish Conditionals between Auxiliary Verb Constructions and Simple Verb Constructions

Beata Trawiński
(Leibniz Institute for the German Language, Mannheim)

Keywords: Conditional, Auxiliary Verb Construction, Simple Verb, Polish, Corpus

As in other Slavic languages, the conditional mood in Polish is typically indicated by the element *by*, which historically can be traced back to the aorist form of the auxiliary *być* ‘be’ (Urbańczyk 1978, Długosz-Kurczabowa & Dubisz 2006) and is now usually classified as a modal particle. It is phonologically weak (showing properties similar to those of pronominal clitics; Dłuska 1974, Rappaport 1988) and requires personal marking: 1SG *by-m*, 2SG *by-ś*, 1PL *by-śmy*, 2PL *by-ście*, 3SG/PL *by-Ø*. The interesting feature of the inflected forms of the conditional *by* is that they can appear to the left of the verb (more precisely, to the *l*-participle form of the verb) as auxiliary verbs (1), or immediately to the right of the verb, attached to the *l*-participle form of the verb (2).

(1) (Ja) rower **by-m** sobie kupić-a.
I bike by-1SG POSS.REFL buy.l-PART-FEM
‘I would buy a bike for myself.’

(2) (Ja) kupić-a-**by-m** sobie rower.
I buy.l-PART-FEM-by-1SG POSS.REFL bike
‘I would buy a bike for myself.’

Kupść & Tseng (2005) claim that the Polish conditional is in a transitional state between an auxiliary verb construction (AVC) and a simple verb construction (SVC). In this paper, we test this hypothesis empirically using corpus data. The research questions we ask are: how similar or different is the distribution of AVCs versus SVCs in a large corpus of Polish, and whether any usage patterns of AVCs versus SVCs can be revealed in the corpus data. In our study, we used the Polish subcorpus of the multilingual parallel corpus InterCorp v16ud with over 300 million tokens (Bańczyk et al. 2020, Čermák & Rosen 2012). Using dependency relations provided by the corpus annotation according to the Universal Dependencies framework (Nivre et al. 2017), AVCs and SVCs were identified and analyzed in terms of their frequency, position (preverbal versus postverbal), and the occurrence with specific lemmas. The results of the study show that SVCs account for 78% (669,496 tokens) of all conditionals, while AVCs account for only 22% (188,850 tokens). This distribution seems to support the transition hypothesis (although more in-depth studies using historical corpora would be needed to better understand this change). Our results also revealed an interesting pattern: SVCs occur significantly more often with modal verbs and the verb *być* ‘be’ than with main verbs. The distribution of modal verbs, the verb *być* ‘be’ and main verbs is exactly reversed for AVCs (see Figure 1). Thus, SVCs and AVCs clearly show complementary usage preferences. The results of our study thus demonstrate that the hypothetical transition from AVC to SVC in Polish mainly affects modal verbs, which in turn reflects a strong link between modality and grammatical mood, sometimes even considered as expressions of the same category (Thieroff 2010, van der Auwera & Aguilar 2016).

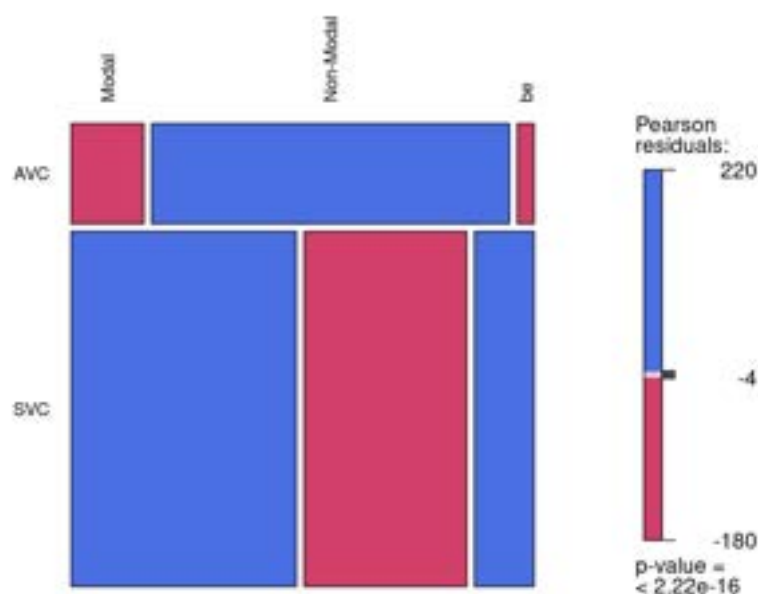


Figure 1 Association plot illustrating the over-representation of AVCs with non-modal verbs and over-representation of SCVs with modal verbs and the verb *be*

References

- Bańczyk, Ł., Dybalska, R., Vavřín, M. & Zasina, A. J. 2024. *Korpus InterCorp – polština, verze 16ud z 13*. 3. 2024. Ústav Českého národního korpusu FF UK, Praha. Accessible at <http://www.korpus.cz>.
- Čermák, F. & Rosen, A. 2012. The case of InterCorp, a multilingual parallel corpus. In *International Journal of Corpus Linguistics*, 17(3). Pp. 411–427.
- Długosz-Kurczabowa, K. & Dubisz, S.. 2006. *Gramatyka Historyczna Języka Polskiego*. Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.
- Dłuska, M. 1974. *Prozodia Języka Polskiego*. Warsaw: PWN.
- Kupść, A. & Tseng, J. 2005. A new HPSG approach to Polish auxiliary constructions. *Proceedings of the 12th International Conference on Head-Driven Phrase Structure Grammar*. Pp. 253–273.
- Nivre, J., Zeman, D., Ginter, F. & Tyers, F. 2017. Universal Dependencies. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics*. Association for Computational Linguistics.
- Rappaport, G. C. 1988. On the Relationship between Prosodic and Syntactic Properties of Pronouns in the Slavic Languages. In Schenker, A. M. (ed.), *American Contribution to the Tenth International Congress of Slavists*. Columbus, OH: Slavica. Pp. 301-327.
- Thieroff, R. (2010). Moods, Moods, Moods. In Rothstein, B. & Thieroff, R. (eds.), *Mood in the Languages of Europe* Amsterdam: John Benjamins. Pp. 1-29.
- Urbańczyk, S. 1978. *Encyklopedia wiedzy o języku polskim*. Wrocław: Zakład Narodowy im. Ossolińskich
- van der Auwera, J. & Aguilar, A. Z. (2015). The history of mood and modality. In Nuyts, J. & van der Auwera, J. (eds.), *The Oxford handbook of Modality and Mood*. (pp.9-27) Oxford: Oxford University Press.

Predication focus, progressive and evidentiality - grammaticalization of form and function of auxiliary *-tenda* in Southern Tanzania

Eva-Marie Bloom Ström, University of Gothenburg

Rasmus Bernander, University of Gothenburg

Aron Zahran, CNRS-Llacan, Paris and University of Ghent, Belgium

In the endangered Tanzanian language Ndengeleko, the speaker always needs to use one of two possible verb forms when expressing a proposition in (at least) four of its tense-aspect (TA) categories, e.g. in the present. This choice depends mainly on syntax, i.e. one form is used when the verb is followed directly by an argument, the other when this is not the case. The latter form is therefore more independent and is the only form that can end a sentence.

This more ‘independent’ verb form shows similarities across the different TA verb forms, with a prefix of the form *and(o)-* and varieties. Bloom Ström and Zahran (2024) argue that this form has additional functions in the language, and hypothesize that it started out as a predication focus strategy, much in line with Güldemann (2003). Apart from its function in polarity focus (no other types of predication focus), it also takes on progressive readings in certain contexts (1), and functions as an evidential strategy. The sentence in example (2) is not accepted by the speaker in the context when the person is inside the house and hearing noise, but only when the person is sitting outside, seeing (and hearing) the boys singing:

1. Andóandíka balúa.
a-**ando**-andik-a balua
SM1SG-PRS.DJ-write-FV letter
‘He is writing letters.’ (‘What is your brother doing right now?’) (Ström 2013: 233)
2. BaandóyIimba amíséembe.
ba-**ando**-yimb-a amisembe
SM2-PRS.DJ-sing FV boys
‘Boys are singing.’ (Scott 2017)

Bloom Ström and Zahran (2024) hint to the possibility of this polyfunctionality being related to the origins of the prefix in an auxiliary *-tenda* ‘to do’. In this study, we expand on this by situating the Ndengeleko data within a broader(micro-)comparative perspective, examining cognate constructions and their formal and functional diversity. These constructions occur within a cluster of related and/or geographically proximate Bantu languages – many of which are understudied and endangered – that are renowned for other significant areal features (see, e.g., Nurse (2003), Bernander et al. (2022)).

Drawing on both secondary and primary field work data, we demonstrate that the disjoint form in Ndengeleko has its counterpart in auxiliary constructions with *-tenda* ‘to do’ in other Bantu languages in the region, expressing various functions. An example is another Southern Tanzanian language; Ngindo:

3. Tu-teenda ku-lim-a.

Furthermore, we reconstruct its grammaticalization trajectory, addressing both its formal maturation — from auxiliary to prefix — and its functional expansion – from marking focus to progressive and, ultimately, to evidential(-epistemic) notions. In this endeavor, we provide a comprehensive description and analysis of this construction, underscoring its Bantu family-internal and typologically significant development (cf. Jäger 2007), and its connection to the historically complex patterns of language affinity witnessed in this area (cf. Nurse 1999).

- Bernander, Rasmus, Eva-Marie Bloom Ström and Hannah Gibson. 2022. Post-verbal negative particles in Southern Tanzania: Form, distribution and historical development. *Africana Linguistica* 28:55-113.
- Bloom Ström, Eva-Marie and Aron Zahran. 2024. Syntactic control and progressive-evidential polyfunctionality: the conjoint/disjoint in Ndengeleko. *Linguistique et Langues Africaine* 10:1-27.
- Güldemann, Tom. 2003. Present progressive vis-à-vis predication focus in Bantu. A verbal category between semantics and pragmatics. *Studies in Language* 27:323–360.
- Jäger, Andreas. 2007. Grammaticalization paths of periphrastic ‘do’- constructions. *Studies van de Belgische Kring voor Linguïstiek* 2:1-18.
- Nurse, Derek. 1999. Towards a historical classification of East African Bantu languages. In Jean-Marie Hombert and Larry M. Hyman (eds): *Bantu historical linguistics: theoretical and empirical perspectives*, 1-41. Stanford: CLSI.
- Nurse, Derek. 2003. Aspect and Tense in Bantu languages. In Derek Nurse and Gérard Philippson (eds): *The Bantu Languages*, 90-102. London: Routledge.
- Odden, David A. 2003. Rufiji-Ruvuma (N10, P10-20). In Derek Nurse and Gérard Phillipson (eds): *The Bantu Languages*, 529-545. London: Routledge.
- Scott, Tessa. 2017. Ndengeleko fieldwork collection: University of California, Berkeley, <http://cla.berkeley.edu/item/25049>.
- Ström, Eva-Marie. 2013. The Ndengeleko language of Tanzania. Doctoral dissertation. Göteborg: University of Gothenburg.

Argument sharing and semantic unification in verbal complex predicates: the case of Jaminjung-Ngaliwurru

Eva Schultze-Berndt
(University of Manchester)

Keywords: complex predicates, generic verbs, monosemy, semantic decomposition, Australian languages

This paper proposes an analysis of complex predicates (CPs) in Jaminjung-Ngaliwurru, an Australian language of the Mirndi family, where CPs are composed of members of two distinct parts of speech: a semantically generic inflecting verb from a closed class, and an open class of uninflecting verbs (Schultze-Berndt 2017). The analysis is based on the principles of compositionality and semantic unification in a Construction Grammar framework. In addition, the principles of monosemy and monotonicity will be adopted – the latter stipulating that compositional operations do not remove meaning components (Rappaport Hovav & Levin 1998; Koontz-Garboden 2008). The constraints on unification of nuclear complex predicates state that neither of the components can introduce a participant that would duplicate a grammatical function already contributed by the other. Thus, while the two components of a CP can have different valencies, syntactic arguments have to be unified. Specifically, unification is licensed in the case of one of three relationships between the predicates involved:

- (i) In the case of inclusion, one of the elements of the CP makes no contribution to the overall interpretation, i.e. it is semantically included in the other (including its argument structure), but is required to form a finite clause. Such semantically generic elements are reminiscent of a grammatical classifier (cf. McGregor 2002). In (1), the generic verb *-angga* 'get/handle', more appropriately translated as 'affect by extended contact', is analysed as semantically included in the uninflecting verb *yurr* 'rub'.
- (ii) Partial semantic overlap, including overlapping argument structure, is permitted for certain *aktionsart* combinations (e.g. processes and states can overlap, but also processes and degree achievements).
- (iii) No semantic overlap (apart from overlapping argument structure) is only permitted if the relationship between the predicate components is construed as one of cause and effect (or motion-endpoint), since the CP construction itself requires that the combination is construed as a single event. This is illustrated in (2), where the generic verb *-angga* 'get/handle' encodes the causing event and the uninflecting verb *bag* 'break' the result.

I will also discuss the limits of existing models of semantic decomposition for representing such unification, in particular Baker & Harvey's (2010) proposed constraints on complex predicate formation. These limitations are particularly obvious for the extremely polyfunctional verb *-Junggu* shown in (3), variously translatable as 'say', 'do', 'become', 'start', and 'internally cause'. Baker & Harvey alternatively represent it with the primitive predicates MOVE and BECOME. I propose instead a monosemic analysis of this generic verb as encoding the initiation of an event or state.

Examples

- (1) **Yurr** **ba-ngu** guri-j-ni jarlig, gunbarr-wurru.
rub IMP-get/handle fat-INST child sore-PROPR
'Rub the child with fat, he has a sore.'
- (2) Jajaman-ni **bag** **gan-angga-m** langiny.
wind-ERG break 3SG>3SG-get/handle-PRS wood
'The wind breaks off a branch.'
- (3) Gurrany **yugung** **yanthi-yu!**
NEG run IRR:2SG[>3SG]-initiate
'Don't run off!'

References

- Baker, Brett & Mark Harvey (2010), Complex predicate formation. In Mengistu Amberber, Brett Baker and Mark Harvey (eds.), *Complex predicates. Cross-linguistic perspectives on event structure*. Cambridge: Cambridge University Press, 13-47.
- Koontz-Garboden, Andrew (2008), Monotonicity at the lexical semantics-morphosyntax interface. In E. Elfner & M. Walkow (eds.), *Proceedings of the 37th annual meeting of the NE Linguistic Society*.
- McGregor, William B. (2002), *Verb classification in Australian languages*. Berlin: Mouton.
- Rappaport Hovav, Malka & Beth Levin (1998), Building verb meanings. In M. Butt & W. Geuder (eds.), *The projection of arguments*. Stanford: CSLI, 97-134.
- Schultze-Berndt, Eva (2017), Two classes of verbs in Northern Australian languages: implications for the typology of polycategoriality. In V. Vapnarsky & E. Veneziano (eds.), *Lexical Polycategoriality: Cross-linguistic, cross-theoretical and language acquisition approaches*. Amsterdam: Benjamins, 243–271.

Personal and deictic pronouns in Dogon

Anonymous

Proto-Dogon is the hypothesised ancestral languages from which all of the modern Dogon languages emerged. Today the twenty-some distinct Dogon languages spoken across the Bandiagara region vary immensely with respect to number and animacy marking. However, there is converging research (Heath 2022 & *various*; Zuk 2024; Prokhorov 2024) that Proto-Dogon made a distinction between animate and inanimate nouns. There is continuing discussion as to which modern Dogon languages best-preserve the archaic structures of Proto-Dogon: Yanda Dom (Prokhorov 2024), Najamba-Kindige (Hantgan, *Personal Communication*), or Jamsay and Donno So (Calame-Griaule 1956). In this talk we present our results of the attempted reconstruction of the Dogon pronominal system.

By employing internal-reconstruction based upon the grammars of Heath (*various*), Prokhorov (2012), McPherson (2013), Dyachkov (*Under Review*), and the CLDF transformation of the Comparative Dogon Wordlist (Heath et al. 2023) we present several pan-Dogon reflexes and the manner in which they diverged in function and form from inherited etyma throughout the language family (see table 1 below).

Table 1: Reconstruction of PD Grammatical Morphemes & Reflexes in 3 daughter languages

Proto-Dogon	Najamba	Tebul Ure	Tommo So
*mí <i>1sg pro</i>	mí <i>1sg pro</i>	mí <i>1sg pro</i>	mí <i>1sg pro</i>
*ú <i>2sg pro</i>	ó <i>2sg pro</i>	ú <i>2sg pro</i>	ú <i>2sg pro</i>
*ná <i>anim 3sg pro</i>	ná <i>anim 3sg pro</i>	íné <i>anim 3sg pro</i>	...wó... <i>anim 3sg pro</i>
*bV <i>anim 3pl pro.</i>	bé <i>3pl animate anaphoric pro.</i>	bú <i>3pl animate anaphoric pro.</i>	bé <i>3pl anaphoric pro.</i>
*mú <i>an. deictic pro.</i>	mó <i>anim. near-distal dem. pro.</i>	-m <i>deictic dem. pro.</i>	-m <i>human plural marker</i>
*kú <i>inanim. deictic pro</i>	kó <i>inanim. dem. pro.</i>	kú <i>inanim pro: 'it'</i>	kó <i>inanim. dem. pro.</i>
*gV <i>gram. deict. pro</i>	-	-gè <i>def. art. inanim. pl.</i>	-ge/-go <i>def. art.</i>

While the first- and second-person singular are effortlessly reconstructable for Proto-Dogon, the third person, both singular and plural, show additional difficulties as they also mark animacy distinctions.

Furthermore, there is no cohesive series of plural personal pronouns, contributing to our hypothesis is that Proto-Dogon did not overtly mark number on the noun. This is line with a larger tendency for Niger-Congo identified by Creissels (2024). Rather, in Dogon, as elsewhere in Niger-Congo we see the appearance of various singulative and plurative marking strategies across the system, generally emerging from the hitherto proposed, previously independent personal and deictic pronouns.

We conclude our talk with a presentation of sound correspondences, partial cascades of phonological and morphological change between the proto-language and the daughter languages.



Map 1: Vowel of the 2.sg pronoun in Dogon and Bangime

References

Author 2024.

Calame-Griaule, Geneviève. 1956. Les Dialectes Dogon. *Africa: Journal of the International African Institute*. [Cambridge University Press, International African Institute] 26(1). 62–72.
<https://doi.org/10.2307/1156770>.

Creissels, Denis. 2024. Noun inflection and gender in Atlantic languages. In Friederike Lübke (ed.), *The Oxford Guide to the Atlantic Languages of West Africa*, 463–482. Oxford: Oxford.

Dyachkov, Vadim. *Under Review. A Grammar of Tomo Kan Dogon* (Institute of Linguistics RAS).
UCL Press.

- Heath, Jeffrey. 2022. Origins of Dogon NP tonosyntax. *Diachronica*. John Benjamins Publishing Company 39(5). 707–741. <https://doi.org/10.1075/dia.20071.hea>.
- Heath, Jeffrey, Steven Moran, Robert Forkel, Kirill Prokhorov, Abbie Hantgan & Johann-Mattis List. 2023. CLDF dataset derived from Heath et al’s “Dogon Comparative Wordlist” from 2016. CLDF. Jena: Max Planck Institute for the Science of Human History. <https://doi.org/10.5281/zenodo.8238983>.
- McPherson, Laura. 2013. *A Grammar of Tommo So*. Berlin: De Gruyter Mouton. <https://doi.org/10.1515/9783110301076>.
- Prokhorov, Kirill. 2024. Development of Dogon number marking. Paris.
- Prokhorov, Kirill, Jeffrey Heath & Steven Moran. 2012. Dogon Classification. In. Paris.

On the relationship between light verbs and voice constructions in Odia, an Indo-Aryan language

Foong Ha Yap & Anindita Sahoo

(Chinese University of Hong Kong, Shenzhen & Indian Institute of Technology Madras, Chennai)

Keywords: grammatical voice, light verbs, transitivization, Indo-Aryan languages, Odia

Previous studies have shown that some languages make use of light verbs not only to form particular types of voice constructions but also to modulate the illocutionary force. For example, Chinese ‘give’ light verbs are used not only to form causative, passive and ‘unaccusative’ (or middle voice) constructions but also to express speaker stance (e.g. Matthews, Xu & Yip 2005; Huang 2013). This paper extends this line of investigation beyond the Sinitic language family to the Indo-Aryan language family. More specifically, we will focus on the role of the light verbs *delaa* (< ‘give’) and *galaa* (< ‘go’) in the formation of causative and passive voice constructions in Odia, an Indo-Aryan language spoken in the Indian state of Odisha. Our data for analysis come from 8 hours of natural conversations as well as elicited examples from native speakers of Odia. We extracted causative and passive uses of *delaa* and *galaa*, then analyzed their functions across different verb types: unaccusative intransitives, unergative intransitives, labile verbs and transitive verbs (we leave aside ditransitives for the moment). Our analysis reveals that both light verbs—*delaa* and *galaa*—yield assertive readings, with *delaa* contributing a stronger illocutionary force. As seen in (1a), (2a), (3a) and (4a), in the absence of light verbs, the simplex constructions are non-assertive, while the more complex constructions involving light verbs (except in the case of (4f)) yield assertive readings. Among intransitive constructions, unaccusative verbs favor light verb *galaa* (< monovalent ‘go’) over *delaa* (< trivalent and transitivizing ‘give’), as illustrated in (1b), whereas unergative verbs favor *delaa* over *galaa*, as seen in (2b), arguably because of the semantic affinity between the dynamicity of unergative verbs and the transitivizing potential of *delaa* (< ‘give’). Further evidence of semantic affinity can be seen in (1c) and (2c). When an intransitive construction (be it unaccusative or unergative) is causativized and hence transitivized by suffix *ei*, it is *delaa* instead of *galaa* that is used to express assertive mood.

The bias toward semantic affinity is also evident for labile and transitive verbs. As illustrated in (3c), causativized labile verbs (e.g. *nachei* ‘dance.CAUS’) favor the use of light verb *delaa* to help convey assertive mood, whereas as seen in (3d), passivized labile verbs (e.g. *nachaa* ‘dance.PASS’) favor (and in fact, require) light verb *galaa*. Transitive verbs further allow for mixed combinations —e.g. causative-passive constructions, using either causative suffix *-aa* + light verb *galaa* or causative suffix *-ei* + light verb *diala galaa*, as seen in (4e) and (4f) respectively. Whereas the presence of *delaa* in causative constructions is optional, the presence of *galaa* in passive constructions is obligatory, and this has facilitated its extended use as a passive marker alongside its assertive mood function in Odia. Findings from this study show that the inherent semantics of light verbs—e.g. whether transitivizing or non-transitivizing—can significantly influence their collocation with particular types of voice constructions.

With unaccusative verb *paD* ‘fall’

- (1) a. *mili baranda-ru paDilaa*
Mili veranda-from fall.PST.3SG.NH
‘Mili fell from the veranda.’
- b. *mili baranda-ru paDi galaa/*delaa*
Mili veranda-from fall go.PST.3SG.NH/*give.PST.3SG.NH
‘Mili did fall down.’ (assertive)
- c. *mini mili-ku baranda-ru pakei delaa/*galaa*
Mini Mili-DAT veranda-from fall.CAUS give.PST.3SG.NH/*go.PST.3SG.NH
‘Mini caused Mili to fall down.’ (causative, assertive)

With unergative verb *has* ‘laugh’

- (2) a. *mili hasilaa*
Mili laugh.PST.3SG.NH
‘Mili laughed.’
- b. *mili hasi delaa/*galaa*
Mili laugh give.PST.3SG.NH/*go.PST.3SG.NH
‘Mili did laugh.’ (assertive)
- c. *mini mili-ku hasei delaa/*galaa*
Mini Mili-DAT laugh.CAUS give.PST.3SG.NH/*go.PST.3SG.NH
‘Mini caused Mili to laugh.’ (causative, assertive)

With labile verb *naacha* ‘dance’

- (3) a. *mili raati saaara naachilaa*
Mili night all dance.PST.3SG.NH
‘Mili danced all night.’
- b. *mili raati saara flamenco naacha naachilaa*
Mili night all flamenco dance dance.PST.3SG.NH
‘Mili danced the flamenco all night.’
- c. *mini mili-ku raati saara flamenco naacha nachei delaa/*galaa*
Mini Mili-DAT night all flamenco dance dance.CAUS give.PST.3SG.NH
‘Mini made Mili dance the flamenco all night.’ (causative & assertive)
- d. *kali jaatraa-re bhala naacha-Tie nachaa *(galaa)/*delaa*
yesterday theater-in good dance-DEF dance.PASS go.PASS.PST.3SG.NH
Lit. ‘A good dance was danced in the theatre yesterday.’
‘There was a good dance in the theatre yesterday.’ (impersonal passive, assertive)

With transitive verb *khaa* ‘eat’

- (4) a. *mili bhaata khaailaa*
Mili rice eat.PST.3SG.NH
‘Mili ate rice.’
- b. *mili bhaata khaai delaa/*galaa*
Mili rice eat.PFV give.PST.3SG.NH/*go.PST.3SG.NH
‘Mili did eat rice.’ (assertive)
- c. *mili sabu bhaata khaai delaa/galaa*
Mili all rice eat.PFV give.PST.3SG.NH/go.PST.3SG.NH
‘Mili ate up all the rice.’ (strongly assertive)

- d. *mini mili-ku bhaata khueilaa*
 Mini Mili-DAT rice eat.CAUS.PST.3SG.NH
 'Mini had (someone make) Mili eat rice.' / 'Mini fed Mili rice.' (causative)
- e. *mini dwaaraa mili-ku bhaata khuua *(galaa)/*delaa*
 Mini by Mili-DAT rice eat.CAUS go.PST.3SG.NH
 'Mili was made to eat rice by Mini.' (causative-passive & non-assertive)
- f. *mini dwaaraa mili-ku bhaata khuei *(diaa galaa)/*delaa*
 Mini by Mili-DAT rice eat.CAUS give go.PST.3SG.NH
 'Mili was made to eat rice by Mini.' (causative-passive & assertive)
- g. *kaali bahut bhaata khia hei galaa*
 yesterday much food eat.PASS be-PFV go.PASS.PST.3SG.NH
 'A lot of food was eaten yesterday.' (impersonal passive)

References

- Huang, C-T. James (2013), Variations in non-canonical passives, in A. Alexiadou, and F. Schäfer (eds), (2013), *Non-canonical passives*. Amsterdam / Philadelphia: John Benjamins Publishing Company, 95-114.
- Matthews, Stephen, Huiling Xu, and Virginia Yip (2005), Passive and unaccusative in the Jieyang dialect of Chaozhou, *Journal of East Asian Linguistics* 14, 267-298.

Irina Nikolaeva (SOAS and University of Münster)

Dejan Matic (University of Münster)

Differential possessive morphology in obviation

Obviation is a grammatical system of reference tracking which regulates the co-occurrence of 3rd person nominals in a given syntactic domain. Obviation in possessive constructions was discussed e.g. for Tzotzil (Aissen 1997), Passamaquoddy (Bruening 2001) and Plains Cree (Dahlstrom 1986; Aissen 1997; Oshima 2007). In some languages it is only identified by syntactic effects and there is no explicit coding of the Proximate status on possessors, while in other languages Proximate possessors are encoded in the same way as Proximate arguments. We will discuss two more languages where obviation has a morphological reflex in possessive constructions but is only limited to them, Tundra Nenets (TN; Uralic) and Tundra Yukaghir (TY; Yukaghiric).

In both languages, SAP possessors are always encoded identically, but there are two synonymous constructions with 3rd person possessors. In TN, they optionally trigger agreement on the possessed noun, cf. *ñū-h tí* [child-GEN reindeer] vs. *ñū-h te-da* [child-GEN reindeer-3SG.POSS] both meaning ‘child’s reindeer’. In TY, pronominal genitive possessors do not agree but nominative possessors must be cross-referenced, cf. *tude eñe*: [3SG.GEN mother] vs. *tudel eñe:-gi* [3SG.NOM mother-3POSS] both meaning ‘his/her mother’. We propose that agreeing possessors in TN and genitive possessors in TY are specified as Proximate, while non-agreeing possessors in TN and nominative possessors in TY are underspecified for obviation.

Arguments for this claim come from distributional restrictions. Crucially, only a single referent in the relevant domain can be Proximate (“Proximate Uniqueness”). In TN, clauses with agreeing possessors cannot contain another non-coreferential 3rd person Proximate nominal, e.g. a 3rd person subject or a 3rd person pronoun (Nikolaeva & Bárány 2019). In TY, the nominative possessor signals disjoint reference, while the genitive indicates co-reference with the 3rd person subject. This follows because subject is a default Proximate, and when two phrases are both Proximate, they have to co-refer.

- (1) Tude ugurčə pul-l'əl-mələ / Tudel ugurčə-gi pul-l'əl-mələ.
3SG.GEN leg kill-EV-3SG / 3SG.NOM leg-3POSS kill-EV-3SG
‘He_i hurt his_{i/*j} leg / He_i hurt his_{j/*i} leg.’

When the subject is a SAP, obviation is not an issue and both options are available.

- (2) Tude čoyoje nu:-məŋ / Tudel čoyoje-gi nu:-məŋ.
 3SG.GEN knife find-1/2SG / 3SG.NOM knife-3POSS find-1/2SG
 'I found his knife.'

Binding principles, which also impose restrictions on the co-occurrence of 3rd person nominals, cannot account for these distributions. They only come into play when expressions co-refer, but Proximate possessors are ruled out by other 3rd person nominals even when these do not have the same referent. Moreover, Proximate possessors are available in non-reflexive contexts (2).

TN and TY appear to be unique in grammaticalizing the Proximate/Obviative distinction by differential possessive morphology. Another interesting property of these languages is the functional load associated with Proximate possessors. Proximates are likely topics, but topicality and Proximate status do not necessarily come together: it can be defined on such factors as empathy, agency or point of view (e.g. Oshima 2007; Dahlstrom 2017). This is also true of TN and TY, where the referent of Proximate possessors is not always topical but is a highly ranked discourse participant.

References

- Aissen, Judith. 1997. On the syntax of obviation. *Language* 73: 705–750.
- Bruening, Benjamin. 2001. Syntax at the edge: Cross-clausal phenomena and the syntax of Passamaquoddy. Doctoral dissertation, MIT, Cambridge, MA.
- Dahlstrom, Amy. 1986. Plains Cree morphosyntax. Doctoral dissertation, University of California, Berkeley.
- Dahlstrom, Amy. 2017. Obviation and information structure in Meskwaki. *Papers of the Forty-sixth Algonquian Conference 2014*, ed. by Monica Macaulay and Margaret Noodin, 39–54. East Lansing: Michigan State University Press.
- Nikolaeva, Irina and András Bárány. 2019. Proximate possessors. In *Prominent internal possessors*, ed. by András Bárány, Oliver Bond, and Irina Nikolaeva, 228–258. Oxford: Oxford University Press.
- Oshima, David Y. 2007. Syntactic direction and obviation as empathy-based phenomena: A typological approach, *Linguistics: An Interdisciplinary Journal of the Language Sciences* 45(4): 727–764.

Investigating Basic Valency Orientation in Hindi

Lucrezia Carnesale
(University of Pavia)

Keywords: Valency orientation, Hindi, anticausative, causative, complex predicates

This study examines basic valency orientation in Modern Standard Hindi from a typological perspective (Nichols et al. 2004, Plank & Lahiri 2015, Zúñiga & Kittilä 2019). While the extensive productivity of the (anti)causative paradigm is a well-known feature of South Asian languages (Masica 1976, Emeneau 1980), and existing studies on Hindi valency-changing operations are abundant (Saksena 1982, Haspelmath 1993, Montaut 2011, Butt 2002), a language-specific analysis on Hindi basic valency orientation remains unexplored. Typological studies suggest that Hindi is a transitivizing language, where non-causative verbs are morphologically basic, and causatives are derived (Masica 1976, Saksena 1982, Montaut 2004). However, previous research has predominantly focused on simple verbs and has overlooked the role of complex predicates (CPs) in the verbal lexicon. As I will argue in this paper, CPs in Hindi constitute a predominant part of the verbal lexicon (Mohan 1994, Kulkarni 2011) and have become extremely productive in the language due to a long-lasting contact with Persian (Montaut 2016). Consequently, when analyzing any phenomenon related to Hindi verbs, it is essential to consider the dual structure of the verbal lexicon.

The methodology of the present study follows previous studies on basic valency orientation (Inglese 2021; Luraghi 2012) and starts with the selection of English pairs of verbs consisting of a basic meaning and its semantic causative. For each pair the Hindi equivalents are selected in order to assess whether the language preferably lexicalize spontaneous events or externally caused ones as morphologically basic verbs (see also Haspelmath 1993). The list of verbs analyzed consists of the 24 meanings based on the guidelines in Nichols (2017), extended with 15 additional pairs. Examples are extracted from a corpus of 20th-century Hindi literary texts and the HiTenTen corpus available on SketchEngine (<https://www.sketchengine.eu/hitenten-hindi-corpus/>).

The analysis reveals that Hindi employs two distinct strategies: (i) simple verbs predominantly use a transitivizing strategy (see 1 and 2), while (ii) complex predicates exhibit a neutral orientation, with alternations marked by auxiliary variation rather than morphological complexity (see 3 and 4). No intransitivizing strategies are used. These findings highlight the productivity of the neutral strategy with Hindi CPs, which has never been properly recognized in previous studies. This paper also points towards a declining productivity of the anticausative paradigm of simple verbs, accompanied by the increasing role of non-oriented strategies within complex predicates. This shift suggests that Hindi may be evolving toward a non-oriented basic valency profile, driven by the widespread use of complex predicates in MSH.

1. *kyā tum sacmuc use mār-t-e?*
what 2PL.NOM really 3SG.ACC kill-IPRF-M.PL
'Would you really have killed him?'
2. *thāyas mar rah-ī hai?*
thayas die PRGR-F.SG be.PRS.3SG
'Thayas is dying?'
3. *mair̥=ne saphar=kī taiyārī śurū k-ī.*
1SG=ERG journey=GEN preparation(F)SG.NOM start do-PRF.F.SG
'I started the preparation for the journey.'

4. *dūsre* *din=se* *boāī* *śurū* *huī*.
 second day=from sowing(F)SG.NOM start be.PRF-F.SG
 'The sowing started from the second day.'

Acknowledgments

Research for this paper has been supported by European Union funding – NextGenerationEU – Missione 4 Istruzione e ricerca - componente 2, investimento 1.1" Fondo per il Programma Nazionale della Ricerca (PNR) e Progetti di Ricerca di Rilevante Interesse Nazionale (PRIN)" progetto 20223XH5XM "Verbs' constructional patterns across languages: a multi-dimensional investigation" CUP F53D23004570006.

References

- Butt, Miriam. 2002. *The morpheme that wouldn't go away*, talk presented at the Workshop on Pertinacity, Shloss Freudental.
- Emeneau, Murray B. 1980. *Language and Linguistic Area: Essays*. Amsterdam: John Benjamins Publishing Company.
- Haspelmath, Martin. 1993. More on the typology of inchoative/causative verb alternations. In: *Causatives and transitivity*, ed. by Bernard Comrie and Maria Polinsky, 87–120. Amsterdam: John Benjamins.
- Inglese, Guglielmo. 2021. Anticausativization and basic valency orientation in Latin. In *Valency over Time*, Silvia Luraghi & Elisa Roma (eds.), 133-168. Berlin & New York: de Gruyter.
- Kulkarni, Amba. 2011. Agreement in Hindi Conjunct Verbs. In *Proceedings of ICON-2011: 9th International Conference on Natural Language Processing*. India: Macmillan Publishers.
- Luraghi, Silvia. 2012. Basic valency orientation and the middle voice in Hittite. *Studies in Language* 36 (1). 1–32.
- Masica, C.P. 1976. *Defining a Linguistic Area: South Asia*. University of Chicago Press.
- Mohanan, Tara. 1994. *Argument Structure in Hindi*, Stanford, CSLI Publications.
- Montaut, Annie. 2004. *A Hindi Grammar*. LINCOM EUROPA.
- Montaut, Annie. 2011. "Basic Intransitivity: A Typologically Relevant Feature for Indo-Aryan". in Koul O.N (ed.) *Indo-Aryan Linguistics*, Central Institute of Indian Languages, pp.29-46.
- Montaut, Annie. 2016. "Noun-Verb Complex Predicates in Hindi and the rise of non-canonical subjects", In *Approaches to Complex Predicates*, Pollet Samvelian & Lea Nash (eds) Benjamins pp 142-174.
- Nichols, Johanna. 2017. Realization of the causative alternation: Revised wordlist and examples. Available at [https://www.academia.edu/34318209/ Realization of the](https://www.academia.edu/34318209/Realization_of_the_causative_alternation) causative alternation Revised wordlist and examples (accessed 1 October 2020)
- Nichols, Johanna, David Peterson & Jonathan Barnes. 2004. Transitivity and detransitivising languages. *Linguistic Typology* 8 (2). 149–211.
- Plank, Frans & Aditi Lahiri. 2015. Microscopic and macroscopic typology: Basic valence orientation, more pertinacious than meets the naked eye. *Linguistic Typology* 19 (1). 1–54.
- Saksena, Anuradha. 1982. *Topics in the Analysis of Causatives, with an Account of Hindi Paradigms*. Los Angeles: University of California Press.
- Zúñiga, F., and Kittilä, S. (2019). Changing Semantic Valency: Causatives, Applicatives, and Related Constructions. In *Grammatical Voice (Cambridge Textbooks in Linguistics)*, pp. 12-81. Cambridge: Cambridge University Press.

What contexts favor the use of personal possessive pronouns over the reflexive in Czech?

Michal Lázníčka & Prokop Hanžl
(Charles University, Prague)

Keywords: Adnominal possession; corpus analysis; Czech morphosyntax; possessive pronouns; usage-based linguistics

This paper explores the variation in the use of the Czech reflexive possessive pronoun *svůj* and its personal counterparts *můj* ‘my’ and *tvůj* ‘your’. While *svůj* is only permitted in contexts where the possessor is coreferential with the subject of the clause, personal possessive pronouns also appear in such contexts, e.g. *Miluju svůj/můj život* ‘I love REFL/my life.’ Traditionally, the use of the personal possessives has been tied to contexts where a non-finite predication with a different subject is present to avoid ambiguity (Daneš and Hausenblas 1962), e.g. *Nechal mě sníst svůj koláč* ‘He_i let me_j eat REFL_{i/j} cake.’ However, other factors have been argued to influence pronoun choice, such as discourse genre or alienability of the possessum (Čmejrková 1998), but such proposals have been based on introspection, or informal corpus queries.

This paper is a follow-up study on the first systematic quantitative analysis of the phenomenon conducted by Perevozchikova (2023). Our aim was to use a larger sample and extend the analysis to *tvůj* ‘your’. Furthermore, we wanted to confront the findings with speakers’ intuitions (cf. Klavan & Veismann 2017). We analyzed a total of 2382 sentences with 1SG and 2SG possessors, sampled from a corpus of posts on forums and social media and annotated for a suite of structural and semantic variables. Separate samples of approximately 600 occurrences were drawn for each person and pronoun variant.

We found the personal possessives (both 1SG and 2SG) to be associated with the presence of a competitor for possessor identity, higher structural distance from the predicate, and preverbal position of the possessive phrase (cf. a representative example in A). These results are in line with Perevozchikova’s findings for 1SG. However, no effects of possessum semantics were found on the level of the whole sample, although pragmatic and semantic factors seem to play some role in contexts that otherwise favour the reflexive. The effect of “competing” referents may point, albeit indirectly, to an awareness of shallow, “good-enough” processing (Ferreira and Patson 2007).

A) *...když mu hned ve druhém řádku **tvého komentáře** nadáváš* ‘...when, right in the second line of **your comment**, you start calling him names’

We then selected 20 sentences from our sample, most strongly associated with the reflexive or the personal possessive, respectively, based on the results. We used these to confront speakers’ intuitions with our results, using two tasks, in which participants either guessed the original form of the sentence (N = 110), or indicated on a 5-point scale whether the sentence sounds more natural with the reflexive or the personal possessive (N = 139). The results show the reflexive to be the default, but speakers do show a sensitivity towards contexts favouring the use of personal pronouns, i.e. contexts of potential ambiguity. A closer look at the items with a personal pronoun in the original that were guessed and rated as more natural with the reflexive points to other variables that might influence the variation, such as, e.g., the presence of any “pointers” to subject identity.

References

- Čmejrková, Světlá (1998), Syntactic and discourse aspects of reflexivization in Czech: The case of the reflexive pronoun *svůj*, in E. Hajičová (ed.), (1998), *Issues of Valency and Meaning. Studies in Honour of Jarmila Panevová*, Karolinum, 75-87.
- Daneš, František and Karel Hausenblas (1962), Přivlastňovací zájmena osobní a zvrtná ve spisovné češtině, *Slavica Pragensia* IV, 191–202.
- Ferreira, Fernanda and Nikole D. Patson (2007), The Good Enough Approach to Language Comprehension, *Language and Linguistics Compass* 1(1-2), 71-83.
- Klavan, Jane and Ann Veismann (2017), Are corpus-based predictions mirrored in the preferential choices and ratings of native speakers? Predicting the alternation between the Estonian adessive case and the adposition *peal* 'on', *Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics* 8(2), 59–91.
- Perevozchikova, Tatiana (2023), Reflexive or Not? Choosing a Possessive in Bulgarian, Czech, and Russian. *Scando-Slavica*, 69(2), 244–263.

The entropy of syntax: exploring word order variation in Persian

Mohammad Tavakoli *¹ and Kateryna Krykoniuk²

* Corresponding Author: mahtav@amu.edu.pl

¹Adam Mickiewicz University, Poznań, Poland

² School of English, Communication and Philosophy, Cardiff University, Cardiff, UK

Understanding word order is crucial for both typological consideration and insights into language processing and generation. Recent studies suggest that basic SVO word order, compared to SOV, is associated with higher predictive processing (Engelhardt et al. 2024) and lower working memory capacity (Amici et al. 2019). These cognitive differences may help explain the variation between SVO and SOV orders in languages such as Persian, where both variants occur in different modes of language use.

Persian syntax predominantly follows the SOV word order (Mirdehghan & Mehmanchian 2012: 374), with verb inflection indicating person and number, facilitating pro-drop. However, what makes a typological description of Persian both interesting and difficult is the presence of two distinct variants: written and spoken. Whereas written Persian adheres to the SOV word order, spoken Persian varies between the SOV and SVO word order (ibid: 372). Pro-drop situations, which mainly occur in spoken language, also allow for VO and OV patterns (ibid: 378). This word order variation calls for further study to better understand Persian's syntactic structure and the discrepancies between written and spoken modes of language use, given that today's formal written Persian does not accurately reflect everyday spoken usage (Haig et al. 2024: 36).

The aim of this study is twofold: to empirically establish word order variation in Persian and explain it. To achieve these objectives, two random samples of Persian were drawn from written and spoken registers. The written register is represented by a BBC website news corpus, compiled specifically for this study using a constructed week sampling method (Luke et al. 2011). The spoken register is drawn from the *Corpus of Conversational Persian* (Mohammadi 2019), comprising transcripts of informal conversations in the Tehrani dialect. Both samples were further randomly reduced to 500 transitive sentences each. Following a two-step reliability test (Cohen's Kappa = 0.986, $p = 0$), the samples were manually annotated for word order by two human raters using a carefully developed, formalised scheme with rigorously

defined criteria for transitivity. Finally, the encoded sequences were analysed and compared through approximate entropy (ApEn; Pincus 1991; Rukhin 2000), a method used to quantify the (un)predictability of a time series or sequence of data based on its local patterns (in the case of this study, the local patterns are SOV and SVO). A key advantage of ApEn is its ability to capture the degree of (ir)regularity in a dataset, even when derived from relatively small sample sizes (Pincus 1991: 2297). Additionally, ApEn's sequential nature may make it more suitable for studying word order than other entropy measures, such as Shannon entropy.

In our presentation, we share the findings of our research, which demonstrate that written Persian (primarily SOV) shows lower approximate entropy compared to spoken Persian, which features a mix of SVO and SOV structures. This variation in spoken Persian supports the view that SOV structures impose greater cognitive demands on working memory (Amici et al. 2019), prompting speakers to alleviate this load by resorting to SVO order during the real-time production of language. We explore this hypothesis in depth, providing new insights into the cognitive dynamics of word order variation in Persian.

REFERENCES

- Amici, F., Sánchez-Amaro, A., Sebastián-Enesco, C., Cacchione, T., Allritz, M., Salazar-Bonet, J., & Rossano, F. (2019). The word order of languages predicts native speakers' working memory. *Scientific Reports*, 9(1), p. 1124.
- Engelhardt, P. E., Filipović, L., & Hawkins, J. A. (2024). Prediction in SVO and SOV languages: processing and typological considerations. *Linguistics*, 62(2), pp. 349–383.
- Haig, G., Rasekh-Mahand, M., Stilo, D., Schreiber, L., & Schiborr, N. (2024). *Post-predicate elements in the Western Asian Transition Zone: Data, theory, and methods*. In (Haig et al. eds): *Post predicate elements in the Western Asian Transition Zone: A corpus-based approach to areal typology*, pp. 3–54.
- Luke, D. A., Caburnay, C. A., & Cohen, E. L. (2011). How much is enough? New recommendations for using constructed week sampling in newspaper content analysis of health stories. *Communication Methods and Measures*, 5(1), pp. 76–91.
- Mirdehghan, M., Mehmanchian, Sh. (2012). The word order in Saravi dialect and spoken Persian language: A Typological Comparative Study. *Iranian Studies*, 45:3, pp. 371–394.

- Mohammadi A. N. (2019). Corpus of Conversational Persian Transcripts. Available at: <https://catalog.ldc.upenn.edu/LDC2019T11?>
- Pincus, S. M. (1991). Approximate entropy as a measure of system complexity. *Proceedings of the national academy of sciences*, 88(6), pp. 2297–2301.
- Pincus, S., and Singer, B. H. (1996). Randomness and degrees of irregularity. *Proceedings of the National Academy of Science of the USA*, 93, pp. 2083–2088.
- Rukhin, A. L. (2000). Approximate entropy for testing randomness. *Journal of Applied Probability*, 37(1), pp. 88–100.

Three generalizations of space: Focusing on axial locations

Shogo Mizuno (Leipzig University/Kyoto University); efforts.0213@gmail.com

In this presentation, I report three generalizations from typological exploratory research on axial locations and suggest that these generalizations are best explained by frequency of use.

Axial location is a cover term for intrinsic and relative frames of reference and is defined as a region that is spatially separated in an axial manner relative to a ground, such as ‘in front of’, ‘behind’, ‘above’, ‘under’, and ‘beside’. Axial locations have been investigated in various approaches, including semantic typology (Levinson & Wilkins 2006), generative grammar (Svenonius 2006; Cinque 2010), FDG (Mackenzie 2013), and grammaticalization theory (Lehmann 2015; Svorou 1994). However, they have not been examined from a morphosyntactic typological perspective. The present study aims to fill that gap.

This study conducts exploratory research based on a 40-language convenience sample and reports the following three generalizations:

- (1) Strategies used for axial locations are more complex (or longer) than those for general locations.
- (2) Strategies used for source relationships in axial locations tend to be more complex than those used for static location and goal relationships in axial locations.
- (3) When strategies for axial locations exhibit variations, those used for UNDER relationships tend to be simpler than those used for the other relationships.

Concerning (1), I have identified several types of strategies used for axial locations, and all of them are more complex (or longer) than those used for general locations. This is illustrated in (4) and (5). This finding can be situated within the coding asymmetries (Greenberg 1966; Haspelmath 2021).

(4) English

- a. *in Bordeaux*
- b. *in front of the building*

(5) Ulwa (Keram; Barlow 2023)

- | | |
|-----------------------|------------------|
| a. <i>ma=in</i> | b. <i>ma=wan</i> |
| 3SG.OBJ=in | 3SG.OBJ=above |
| ‘in it (=the garden)’ | ‘over it’ |

As for (2), in general locations, source relationships are encoded more complexly than goal relationships (Haspelmath 2008; Haspelmath 2006; Haspelmath 2021). The present study confirms that this generalization holds true for axial locations as well. This is illustrated in (6) and (7).

(6) English

- a. *My cat goes **under** the table.*
- b. *My cat comes **from under** the table.*

(7) Japanese

- | | |
|-----------------------------|-------------------------|
| <i>tukue=no</i> | <i>sita=e/sita=kara</i> |
| table=GEN | under=ALL/under=ABL |
| ‘(to)/from under the table’ | |

Regarding (3), I find that the strategies used for axial locations are highly homogenous. However, when languages exhibit variation in strategies, those used for UNDER relationships tend to be the simpler. As illustrated in (8), UNDER relationships in Koromfe can be expressed with a shorter form than IN FRONT OF relationships.

(8) Koromfe (Atlantic-Congo; Rennison 1997)

- | | |
|---|---|
| a. <i>dãŋ</i> <i>koŋ</i> <i>jika</i> <i>nɛ</i>
house.SG DET face.SG at
'(to) in front of the house' | b. <i>fɛkɔ</i> <i>koŋ</i> <i>hogo</i>
tree.SG DET under
'from under the tree' |
|---|---|

As an explanation for these three generalizations, I compare several competing frameworks, such as markedness, iconicity, and frequency, and suggest that they can be most effectively explained in terms of frequency, as Haspelmath (2006; 2008; 2021) suggests. While markedness and iconicity fail to account for cases where general and axial locations are coexpressed using the same marker, the frequency-based explanation can account for these instances.

References

- Barlow, Russel. 2023. *A grammar of Ulwa*. Berlin: Language Science Press.
- Cinque, Guglielmo. 2010. Mapping Spatial PPs: An Introduction. In Guglielmo Cinque & Luigi Rizzi (eds.), *Mapping Spatial PPs: The Cartography of Syntactic Structures, Volume 6*, 3–25. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195393675.003.0001>.
- Greenberg, Joseph H. 1966. *Language universals, with special reference to feature hierarchies*. The Hague: Mouton. <https://doi.org/10.1515/9783110899771>.
- Haspelmath, Martin. 2006. Against markedness (and what to replace it with). *Journal of linguistics*. Cambridge University Press 42(1). 25–70.
- Haspelmath, Martin. 2008. Frequency vs. iconicity in explaining grammatical asymmetries. *Cognitive Linguistics* 19(1). 1–33. <https://doi.org/10.1515/COG.2008.001>.
- Haspelmath, Martin. 2019. Differential place marking and differential object marking. *STUF - Language Typology and Universals* 72(3). 313–334. <https://doi.org/doi:10.1515/stuf-2019-0013>.
- Haspelmath, Martin. 2021. Explaining grammatical coding asymmetries: Form–frequency correspondences and predictability. *Journal of Linguistics* 57(3). 605–633. <https://doi.org/10.1017/S0022226720000535>.
- Lehmann, Christian. 2015. *Thoughts on grammaticalization*. 3rd edn. Berlin: Language Science Press. <https://doi.org/10.17169/langsci.b88.99>.
- Levinson, Stephen C. & David P. Wilkins (eds.). 2006. *Grammars of Space: Explorations in Cognitive Diversity* (Language Culture and Cognition). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511486753>.
- Mackenzie, J. Lachlan. 2013. Spatial adpositions between lexicon and grammar. In J. Lachlan Mackenzie & Hella Olbertz (eds.), *Casebook in Functional Discourse Grammar* (Studies in Language Companion Series), 67–94. Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/slcs.137.04mac>.
- Rennison, John R. 1997. *Koromfe* (Descriptive Grammars Series). London & New York:

- London & New York: Routledge.
- Stolz, Thomas, Sander Lestrade & Christel Stolz. 2014. *The Crosslinguistics of Zero-Marking of Spatial Relations*. München: De Gruyter (A).
<https://doi.org/doi:10.1524/9783050065304>.
- Svenonius, Peter. 2006. The Emergence of Axial Parts. *Nordlyd* 33(1). 49–77.
<https://doi.org/10.7557/12.85>.
- Svorou, Soteria. 1994. *The Grammar of Space*. Amsterdam: John Benjamins.
<https://www.jbe-platform.com/content/books/9789027276575>.

Voice Syncretism in Takituduh Bunun

Victor Bogren Svensson
(Lund University; Academia Sinica)

Keywords: Takituduh Bunun, Austronesian Voice System, Voice Syncretism, Transitivity, Verb Class

This paper presents evidence from original fieldwork for voice syncretism in Takituduh Bunun (Austronesian: Taiwan), where the Locative Voice (LV) marker has taken over the function typically associated Patient Voice (PV) for a subset of the verbal lexicon, while still retaining its originally function. The pattern is illustrated in (1a-b) & (2a-b), where the verb *kulut* ‘cut’ exhibits both Patient and Locative Voice forms, with Patient and Location as clause Pivot, respectively, and the verb *tabal* ‘chop’ exhibits the same alternation in terms of semantic role of the pivot (Patient and Location, respectively), but fulfills both functions with the LV form.

- | | | | | |
|-----|--|---------------|----------------|-----------------------|
| 1a. | <i>kulut-un</i> | <i>uva’az</i> | <i>ca</i> | <i>sanglav</i> |
| | cut-PV | child | PIVOT | vegetables |
| | ‘The child cut the vegetables.’ | | | |
| 1a. | <i>kulut-an</i> | <i>uva’az</i> | <i>sanglav</i> | <i>ca pitdidai’an</i> |
| | cut-LV | child | vegetables | PIVOT kitchen |
| | ‘The child cut vegetables in the kitchen.’ | | | |
| 2b. | <i>tabal-an</i> | <i>uva’az</i> | <i>ca</i> | <i>lukic</i> |
| | chop-LV | child | PIVOT | wood |
| | ‘The child chopped the wood.’ | | | |
| 2b. | <i>tabal-an</i> | <i>uva’az</i> | <i>lukic</i> | <i>ca quma</i> |
| | chop-LV | child | wood | PIVOT field |
| | ‘The child chopped wood in the field.’ | | | |

In a preliminary list of 200 transitive verbs, the same pattern is attested across different verb classes (Tsukida & Zeitoun 2023 and Ross 2015), semantic categories (Levin 1993), degrees of transitivity (Hopper and Thompson 1980) and phonological contexts (Li 1988), illustrated using six verbs in (3a-c) across different semantic and transitivity categories, where the verbs to the right lack an expected PV form.

- | | | | |
|-----|------------------|--------------------------------|----------------------|
| 3a. | Perception: | <i>tan’aun/tan’aan</i> ‘hear’ | ---/sadu’an ‘see’ |
| 3b. | Physical Impact: | <i>ludaqun/ludaqan</i> ‘hit’ | ---/lintauzan ‘push’ |
| 3c. | Movement: | <i>sinapun/sinapan</i> ‘chase’ | ---/pavazan ‘leave’ |

Since this pattern is attested across different parts of the lexicon, the emergent pattern appears to be that of two inflectional classes, here tentatively named PV/LV-verbs and LV-verbs (approximately 30% of the current verb list). This conclusion is further corroborated by intra- and interspeaker variation; some speakers use the forms *painukun*(PV)/*painukan*(LV) ‘to wear’ interchangeably, while others prefer *painukan*.

	PV/LV-Verbs	LV-Verbs
Agent Voice	M-	M-
‘Patient’ Voice	-un	-an
Locative Voice	-an	-an
Circumstantial Voice	is-	is-

Proposals for the development of this pattern remain speculative. A pattern found among languages with the Austronesian Voice System is that transitive verbs can alternate between Patient and Locative Voice marking on the verb to signal a lower degree of transitivity for the latter (cf. Puyuma, Teng 2008, 111). The latter use could then have been reanalyzed as the default voice marking, leading to a loss of the PV form. Another possible origin is via analogy; verbs of perception often use LV for object pivot clauses (cf. Tsou, Huang & Huang 2007, 434), and this usage could have spread via analogy to other categories of verbs as well, replacing the PV form.

This paper documents an a-typical Austronesian Voice System, making it a relevant contribution to Austronesian linguistics. It also provides evidence for emerging inflectional classes and how a symmetrical voice system and transitivity marking can change over time, while also constituting the first attempt at a systematic description of voice marking in the Taktiuduh Bunun lexicon, highlighting its broader empirical and theoretical contributions.

References

- Hopper, Paul. J., & Thompson, Sandra. A. (1980). Transitivity in grammar and discourse. *Language*, 251-299.
- Huang, Huei-ju, & Huang, Shuan-fan. (2007). Lexical perspectives on voice constructions in Tsou. *Oceanic Linguistics*, 424-455.
- Levin, Beth. (1993). *English verb classes and alternations: a preliminary investigation*. Chicago: University of Chicago Press.
- Li, Paul Jen-keui. (1988). A comparative study of Bunun dialects. *Bulletin of the Institute of History and Philology*, 479-508.
- Ross, Malcom. (2015). Reconstructing proto-Austronesian verb classes. *Language and Linguistics*, 16(3), 279-345.
- Teng, Stacy. (2008). *A reference grammar of Puyuma, an Austronesian language of Taiwan*. Canberra, A.C.T.: Pacific Linguistics.
- Tsukida, Naomi., & Zeitoun, Elizabeth. (2023). Verbal Morphology of Formosan Languages. In P. J. Li, E. Zeitoun and R. De Busser (eds.), *Handbook of Formosan Languages Online: The Indigenous Languages of Taiwan*. Brill.

Conceptualization of event packaging in Chinese Action-Motion Serial Verb Construction

Yuqi YIN

(The Graduate University for Advanced Studies ,SOKENDAI)

Keywords: event conceptualization, Serial Verb Construction, monoclausality, eventhood, spatio-temporal continuity, resultant connection

Acknowledgment: This research is part of the research conducted under the NINJAL collaborative research project ‘Evidence-based Theoretical and Typological Linguistics’ (project leader: Masayuki ASAHARA) subproject ‘Data-oriented Typological Study of the Semantics and Grammar of Predicates’ (subproject leader: Yo Matsumoto). This work was supported by The Graduate University for Advanced Studies, SOKENDAI.

This study experimentally investigates the cognitive constraints on event conceptualization in Chinese Serial Verb Constructions (hereafter CSVs) that describe consecutive events within a single clause. Serial Verb Constructions are often said to represent a single event (Aikhenvald 2006, 2018), a characteristic closely tied to their monoclausality (Haspelmath 2016). To support this perspective, this study examines how factors such as spatio-temporal disruptions and causal disconnections influence event continuity in speakers’ conceptualization of CSVs.

We compare the conceptualization differences between “Action-Motion” CSVs, such as (1), and their biclausal counterparts, such as (2).

- (1) *Zhāngsān ná shū qù jiàoshì le.*
Zhangsan take book go classroom PRT
‘Zhangsan takes a book/books to the classroom’.

- (2) *Zhāngsān ná le shū, zài nà zhī hòu qù jiàoshì le.*
Zhangsan take PFV book after.that go classroom PRT
‘Zhangsan takes a book/books and then goes to the classroom’.

Syntactic tests, such as the scope of manner adverbials and temporal adverbials, and negation scope, demonstrate that Action-Motion CSVs exhibit monoclausal characteristics.

We report the results of an experiment, in which participants evaluated the appropriateness of sentences for pictures presented. The stimuli consisted of 6 combinations of texts and pictures: 2 types of text stimuli (monoclausal CSVs and their biclausal counterparts) and 3 types of pictorial stimuli (Set A, B, and C). Set A includes stimuli representing a sequence of events (action and motion) without any disruptions of the continuity of the two events. Set B features illustrations depicting spatio-temporal disruptions of the two events (e.g., going somewhere else to perform an unrelated action). Set C contains illustrations in which the resultative state caused by the action is disrupted (e.g., the result of picking up a book is canceled by subsequently placing the book down), representing causal disconnection.

20 Chinese native speakers rated 82 stimuli, presented in random order, on the appropriateness of using the text to describe the illustrations on a scale of 1 to 5 (1 being unacceptable, and 5 being perfectly acceptable).

Results were analyzed using a cumulative link mixed model (R, version 4.4.1, function “clmm()”) reveal significant effects for Sets A and C compared to the reference category (Set B), but only in monoclausal Action-Motion CSVc sentences (Table 1, Significant level of $p < 0.001$ is denoted as ‘***’, $0.001 \leq p < 0.01$ as ‘**’, $0.01 \leq p < 0.05$ as ‘*’.).

Table 1 Cumulative link mixed model results of the appropriateness ratings

Ref = Set B	Monoclausal CSVcs	Biclausal counterparts
Set A	0.021*	0.051
Set C	<0.001***	0.502

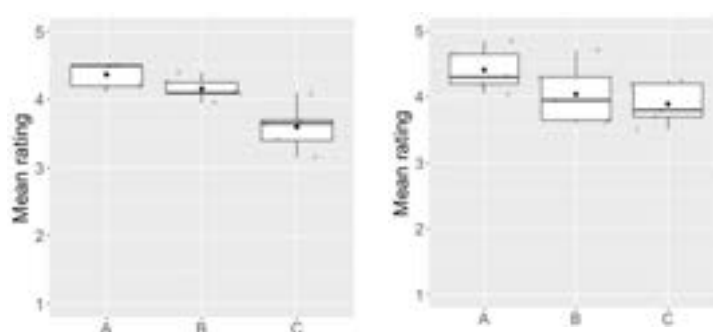


Figure 2 Mean of the appropriateness ratings of monoclausal (left) and biclausal (right) sentences

Our findings indicate that disruptions to spatio-temporal continuity and cancellations of the result of the first event segment are disfavored in the use of Action-Motion CSVcs. In contrast, such disruptions are acceptable in their biclausal counterparts.

This study has thus revealed that monoclausal CSVcs exhibit cognitive constraints on the spatio-temporal continuity and causal connection of events. This substantiates the view that monoclausal sentences represent a single event characterized by spatio-temporal and causal continuity.

Glossary

PRT = sentence-final particle expressing the change of a new situation/state

References

- Aikhenvald, Alexandra Y. (2006) Serial verb constructions in typological perspective, in Alexandra Y. Aikhenvald and Robert M. W. Dixon (eds.), (2006), *Serial verb constructions: a cross-linguistic typology*, Oxford: Oxford University Press, 1–68.
- Aikhenvald, Alexandra Y. (2018) *Serial Verbs*, Oxford: Oxford University Press.
- Haspelmath, Martin (2016) The Serial Verb Construction: Comparative Concept and Cross-Linguistic Generalizations. *Language and Linguistics* 17(3): 291–319.

General Session : Phonetics and phonology

Prosodic influence on syntactic alternation at the phrase level: Evidence from French attributive adjective placement

Anna Pressler
(Goethe-University Frankfurt)

Keywords: French, prosody, syntactic alternation, attributive adjectives

Phonological factors induce syntactic alternation at the sentence-level in a variety of typologically diverse languages (e.g. Bennett et al., 2016 for Irish; Faghiri, 2016 for Persian; Thuilier, 2012 for French; Yao, 2018 for Mandarin). Their influence on the linearization within phrases is less studied (e.g. Schlüter, 2005 for English; Shih & Zuraw, 2017 for Tagalog; Thuilier, 2014 for French).

This study addresses prosodic effects on the linearization at the phrase-level. Hence, the placement of a set of French attributive adjectives that can occur in two syntactically possible positions is investigated. Their position does not systematically provoke distinct interpretations, as illustrated by the example *un charmant garçon* and *un garçon charmant* ('a charming boy') (Thuilier, 2014, p. 289).

I explore the effects of two prosodic factors. First, the impact of relative length (here the number of syllables) is examined. Second, the effect of the rhythmic make-up of the noun phrase is investigated. Previous research found constituents to be preferably ordered in terms of increasing length in French (Thuilier, 2012). Consequently, I predict that this preference applies likewise to noun phrases consisting of a single adjective and a noun, see (1). The tendency to alternate metrically strong and weak syllables has been found to affect syntactic ordering (Schlüter, 2005; Shih, 2017). Thus, I predict that French adjective-noun pairs are ordered with alternating phrasal stress. In French, each lexical item can be marked prosodically (Post, 2000) and marking can be done using phrase-final stress (Dell, 1984; Hirst & Di Cristo, 1996; Jun & Fougeron, 2002). Hence, some sequences of a polysyllabic adjective and a monosyllabic noun present a potential environment for juxtaposed strong syllables, see (2a.). This unfavoured configuration can be avoided by placing the adjective in postnominal position (2b.).

- | | | | | | | | | | |
|-----|----|-----------|----------|-----------|----|-----------|-----------|----------|---------------|
| (1) | a. | une | énorme | perte | b. | une | perte | énorme | |
| | | INDF.F.SG | huge | loss | | INDF.F.SG | loss | huge | 'a huge loss' |
| (2) | a. | un | charmant | prof | b. | un | prof | charmant | 'a charming |
| | | INDF.M.SG | charming | professor | | INDF.M.SG | professor | charming | professor' |

Two studies were conducted: a forced-choice task using written material (89 participants) and an elicited production task involving spoken material (59 participants). In the production task, participants combined two sentences that elicited adjectives and nouns separately. The material manipulates the relative length (shorter, longer, or equally long adjective), the rhythmic make-up of the adjective-noun pair (potentially adjacent strong syllables vs. alternating strong and weak syllables) and the position of the adjectives (prenominal or postnominal). The preferred positioning of the adjectives is investigated by means of generalized mixed models and an acoustic analysis is performed.

Results confirm that speakers of French prefer an ordering in terms of increasing length, both with respect to the number of syllables and the actual duration of the adjectives and nouns. As such, adjectives that are shorter than nouns are preferred prenominally, while longer adjectives are preferred postnominally. Furthermore, native speakers are sensitive to the rhythmic make-up of the noun phrases. Thus, they tend to avoid juxtaposed phrasal stress. This illustrates that, in addition to influencing the ordering of constituents at the sentence-level, prosodic factors affect the linearisation of words at the phrase-level in French.

Acknowledgments

This research was financially supported by the German Research Foundation (DFG, GRK 2016), Goethe-University and the Embassy of France in Germany. I thank all colleagues within and outside of Goethe-University for fruitful discussions about this work.

References

- Bennett, R., Elfner, E., & McCloskey, J. (2016). Lightest to the Right: An Apparently Anomalous Displacement in Irish. *Linguistic Inquiry*, 47(2), 169–234. https://doi.org/10.1162/LING_a_00209
- Dell, F. (1984). L'accentuation dans les phrases en français. In F. Dell, D. J. Hirst, & J.-R. Vergnaud (Eds.), *Forme sonore du langage: Structure des représentations en phonologie* (pp. 65–122). Hermann.
- Faghiri, P. (2016). *La variation de l'ordre des constituants dans le domaine préverbal en persan: Approche empirique*. Université Sorbonne Paris Cité.
- Hirst, D., & Di Cristo, A. (1996). Y at-il des unités tonales en français. *Proceedings of XXIèmes Journées d'étude Sur La Parole, Avignon*, 223–226.
- Jun, S.-A., & Fougeron, C. (2002). Realizations of accentual phrase in French intonation. *Probus*, 14(1). <https://doi.org/10.1515/prbs.2002.002>
- Post, B. (2000). *Tonal and phrasal structures in french intonation*. Thesus.
- Schlüter, J. (2005). *Rhythmic grammar: The influence of rhythm on grammatical variation and change in English*. Mouton de Gruyter.
- Shih, S. S. (2017). Phonological Influences in Syntactic Alternations. In V. Gribanova & S. S. Shih (Eds.), *The Morphosyntax-Phonology Connection* (pp. 223–252). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780190210304.003.0009>
- Shih, S. S., & Zuraw, K. (2017). Phonological conditions on variable adjective and noun word order in Tagalog. *Language*, 93(4), e317–e352. <https://doi.org/10.1353/lan.2017.0075>
- Thuilier, J. (2012). *Contraintes préférentielles et ordre des mots en français*. Université Paris-Diderot - Paris VII.
- Thuilier, J. (2014). An Experimental Approach to French Attributive Adjective Syntax. *Empirical Issues in Syntax and Semantics*, 10(ed. Christopher Piñón), 287–304.
- Yao, Y. (2018). NP weight effects in word order variation in Mandarin Chinese. *Lingua Sinica*, 4(1), 5. <https://doi.org/10.1186/s40655-018-0037-8>

Re-evaluating a non-contrast: *t₁, *t₂, *c₁ and *c₂ in Proto-Inuit-Yupik-Unangan

B. W. E. Bethlehem
University of Cambridge

Keywords: Inuit-Yupik-Unangan, Eskimo-Aleut, historical phonology, reconstruction, sound change

Although the link between the Unangan (U, also called Aleut) and Inuit-Yupik (IY, also called Eskimo) clades has been recognised since the 19th century, the first rigorous reconstruction for Proto-Inuit-Yupik-Unangan (p-IYU) was only proposed by Knut Bergsland in 1986. Since then, re-evaluations of the material such as in Berge (2018) and Fortescue, Jacobson and Kaplan's *Comparative Eskimo Dictionary* (CED) have considerably advanced the field, but many questions still remain unanswered.

One key area of debate in the historical phonology of IYU is the reconstruction of the coronal obstruents, as correspondences are not straightforward between the two branches. Initially /č/ in U corresponds to PIY *t as well as *č, whereas U /t/ corresponds to PIY *t as well as initial *n before back vowels; U /š/ (usually cited as /s/ and nowadays so pronounced, but formerly a palato-alveolar sibilant (Bergsland 1997)) only corresponds to PIY *č. Fortescue et al. and Bergsland (1986: 71) both reconstructed these as four separate phonemes, as shown below. Medially, the reflexes are more coelex. Note that for PIY data, I have used forms from the CED adjusted to fit Bergsland (1986)'s notation, i.e. *c *đ are replaced with *č *ř.

Unangan	Gloss	PIY	Gloss	CED	Bergsland
<i>chidag-lix</i>	'to stretch'	*čiray-	'to spread out'	*c ₁	*č
<i>chnga-ř</i>	'body hair'	*taŋa	'pubic hair'	*t ₂	*c
<i>sngag-lix</i>	'to dream'	*čaŋay-	'to sleep'	*c ₂	*š
<i>tař-six</i>	'to get dark'	*tarar-	'to be dark'	*t ₁	*t

Using lexical data chiefly from the CED and Bergsland's 1994 *Aleut Dictionary*, I examine roughly 150 comparanda, i.e. cognate sets of varying levels of confidence. I propose that the reflexes are determined by (a) the following vowel and (b) position, i.e. initial/intervocalic/other medial/final. Under this schema, there is no difference between *t₁ and *t₂, nor between *c₁ and *c₂, allowing the phonological system to be simplified. The sound shifts I posit are:

- U: *t to /č/ before *a; *č to /š/ before all vowels except for *i
- U: intervocalic lenition of *t *č *š (of whatever provenance) to /ð š ȳ/.
- IY: *a and *i merge to *a after *t and to *i after *č, leading to obfuscation.

These are to be taken in conjunction with the long-recognised U shifts of initial *n to /t/ (before *a, *a, *u) or its loss before *i (Bergsland 1986: 79).

References

Berge, A., 2018. Re-evaluating the reconstruction of Proto-Eskimo-Aleut. In *Journal of Historical Linguistics*, 8:2 (2018), pp. 230–272.

- Bergsland, K., 1986. Comparative Eskimo-Aleut phonology and lexicon. In *Journal de la Société Finno-Ougrienne*, 82.7–80, pp. 63–138.
- Bergsland, K., 1994. *Aleut Dictionary / Unangam Tunudgusii*. Alaska Native Language Center, University of Alaska Fairbanks.
- Bergsland, K., 1997. *Aleut Grammar / Unangam Tunuganaan Achixaasiġ*. Alaska Native Language Center, University of Alaska Fairbanks.
- Fortescue, M., Jacobson, S. and Kaplan, L., 2010. *Comparative Eskimo Dictionary with Aleut Cognates*. Alaska Native Language Center, University of Alaska Fairbanks.

Phonotactics and morphonotactics of word-initial clusters in Belarusian: A corpus-based study

Dziyana Sabaleuskaya, Paulina Zydorowicz, Alona Kononenko-Szoszkiewicz & Katarzyna Dziubalska-Kořaczyk

(Adam Mickiewicz University)

Keywords: phonotactics of consonant clusters, morphonotactics, Belarusian, auditory distance, corpus frequency

This study provides the first quantitative overview of Belarusian phonotactics based on a written corpus, complemented by a qualitative analysis focusing on the emergence of consonant clusters and the distinction between phonotactic and morphonotactic types. Morphonotactics, as defined by Dressler and Dziubalska-Kořaczyk (2006), explores the interface between phonotactics and morphotactics, identifying consonant clusters that arise due to morphological intervention.

We claim that a distinction should be made between phonotactic clusters, which are phonologically motivated and occur within a single morpheme, e.g.:

(1)

/skn-/ скнара ['sknara] 'miser'

and morphonotactic clusters, which arise due to concatenative and nonconcatenative morphology. The former refers to contexts in which a morpheme is added to a root, triggering a (longer) cluster, e.g.:

(2)

/pt-/ плыць [ptʰitsʲ] 'to swim' (imperfective)

/spʰ-/ сплыць [spʰtʰitsʲ] 'to swim away' (perfective)

The non-concatenative morphology encompasses those sequences of consonants which emerge as a result of, e.g. deletion of a root vowel:

(3)

лён /'lʲɔn/ 'flax' (NOM.SG.)

льну /'lʲnu/ 'flax' (GEN.SG.)

or in plural formation, as in:

(4)

сон /'sɔn/ 'dream' (NOM.SG.)

сны /'snɨ/ 'dreams' (NOM.PL.).

The analysis adopts the Beats-&-Binding Phonology model (Dziubalska-Kořaczyk 2002), which evaluates clusters using the Net Auditory Distance (NAD) principle, defined by manner and place of articulation (MOA and POA), as well as the obstruent-sonorant distinction. This model, free from traditional syllabification constraints, formulates universal preferences for optimal clustering based on cluster length and position within the word (initial, medial, final). Large-scale calculations are performed by a phonotactic calculator available for nine languages (Dziubalska-Kořaczyk et al. 2014).

The study draws data from a Standard Belarusian written corpus comprising over 1.5 million words (Mazzitelli 2021). Analyses are conducted across cluster types, lemmas, word types, and word tokens, incorporating frequency information. While Russian and Ukrainian have been previously studied within the Beats-&-Binding framework, this research marks Belarusian as the last East Slavic language to undergo such analysis. Building on previous research on other Slavic languages, such as Polish (Zydorowicz et al. 2016) and Slovak (Dressler et al. 2015), we expect that double consonant clusters represent the most numerous group in terms of cluster types and are predominantly phonotactic. In contrast, triple consonant clusters are expected to be less frequent and primarily morphonotactic in nature. According to Maddieson (2013), all Slavic languages can be classified as predominantly consonantal. However, the degree of consonantism and the prevalence of consonant clusters within each language reveal gradual typological differences (Sawicka 2001). Therefore, cluster preferences and their distribution are highly language-specific, reflecting the unique phonological characteristics of each language.

References

- Dressler, Wolfgang U. and Katarzyna Dziubalska-Kořaczyk. 2006. "Proposing morphonotactics", *Journal of Italian Linguistics/Rivista di Linguistica* 18, 2: 249-266.
- Dressler, Wolfgang U., Miroslava Hliničanová, Matej řurčo, Karlheinz Mörth and Katharina Korecky-Kröll. 2015. "Phonotactic vs. morphonotactic obstruent clusters in Slovak and German", *Italian Journal of Linguistics* 27: 45–60.
- Dziubalska-Kořaczyk, Katarzyna, Dawid Pietrala and Grzegorz Aperliński. 2014. The NAD Phonotactic Calculator – an online tool to calculate cluster preference in English, Polish and other languages. (<https://wa.amu.edu.pl/nadcalc/>) (date of access: 13 Jan. 2025).
- Dziubalska-Kořaczyk, Katarzyna. 2002. *Beats-and-Binding Phonology*. Frankfurt am Main: Peter Lang.
- Dziubalska-Kořaczyk, Katarzyna. 2014. "Explaining phonotactics using NAD", *Language Sciences* 46: 6-17. (doi: 10.1016/j.langsci.2014.06.003).

- Maddieson, Ian. 2013. "Consonant Inventories", in: Matthew S. Dryer and Martin Haspelmath (eds.) *WALS Online*. v2020.4. [Data set]. Zenodo. (<https://doi.org/10.5281/zenodo.13950591>) (<http://wals.info/chapter/1>) (date of access: 14 Jan. 2025).
- Mazzitelli, Lidia Federica. 2021. *Belacorus*. (<https://github.com/Belarusian-Corpus>) (date of access: 12 Jan. 2025).
- Sawicka, Irena. 2001. *An outline of the phonetic typology of the Slavic languages*. Toruń: Wydawnictwo Uniwersytetu M. Kopernika.
- Zydorowicz, Paulina, Paula Orzechowska, Katarzyna Dziubalska-Kołaczyk, Michał Jankowski, Piotr Wierzchoń and Dawid Pietrala. 2016. *Phonotactics and morphonotactics of Polish and English: theory, description, tools and applications*. Poznań: Adam Mickiewicz University Press.

Complex Sound sequences as (un)ambiguous signals of word structure: Experimental and corpus approaches

Irene Böhm & Nikolaus Ritt
(University of Vienna)

Keywords: phonotactics, artificial language learning experiment, sound change, corpus, English

Our paper discusses sound patterns as cues to morphological structure and presents the findings of two studies, an experiment on artificial language learning and a corpus study on language change, that explore the impact of ambiguity in this relationship.

Speakers pay attention to where and how often sounds (co-)occur in word forms, and use these distributional frequencies in the learning and processing of words (e.g., Storkel 2001, Vitevitch and Luce 2016, Kelley and Tucker 2017). Some phonotactic patterns act as cues to word structure because they strongly correlate with either morphologically simple or complex forms. For instance, in English, /-mp/ (as in *lamp* or *jump*) is associated with morphologically simple forms, while word-final /-gz/ (as in *egg+s*, *leg+s*, or *hug+s*) always results from suffixation and therefore indicates morphological complexity. Such patterns facilitate processing: complex forms whose sound patterns are reliable signals of complexity are processed both more quickly and accurately (Korecky-Kröll et al. 2014).

However, not all patterns are predictable signals of word structure. Some, like /-nd/ in *hand* and *sinn+ed*, occur in both simple stems and morphologically complex forms, making them morphotactically ambiguous. This ambiguity introduces processing challenges and slows it down instead of making it easier (Post et al. 2008, Celata et al. 2015). Consequently, such ambiguous patterns should be harder for speakers to learn and pass on and, as a result, be selected against in language change.

We explore this hypothesis from two different angles. On the one hand, we corroborate the plausibility of this hypothesis by means of a corpus study of sporadic past tense /-d/ devoicing in Early Modern English (Lahiri 2009, Wełna 2009). We argue that this somewhat puzzling sound change and the continued stability of the irregular forms, such as *spoilt* or *burnt*, can be understood better if one assumes that sound changes tend to prevent or reduce this type of ambiguity (Böhm et al. [accepted]). On the other hand, we explore underlying causalities in two artificial language learning experiments of coda consonant clusters. We demonstrate that physiologically difficult consonant sequences are learnt more easily when they act as cues to morphological complexity, and that this advantage in learning is lost when their signalling function is compromised by ambiguity.

Our talk presents the findings from both studies in greater detail, discusses the limitations and complementary strengths of the two methods, and connects our insights to research on probable sound patterns as attractors in language change (Wedel 2006, Blevins 2009) and consonant clusters within the field of morphonotactics (Dressler and Dziubalska-Kołaczyk 2006, Dressler et al. 2015, Baumann, Prömer and Ritt 2019).

References

- Baumann, Andreas, Christina Prömer, and Nikolaus Ritt (2019), Word form shapes are selected to be morphotactically indicative, *Folia Linguistica* 40(1), 129–151.
- Blevins, Juliette (2009), Structure-preserving sound change: A look at unstressed vowel syncope in Austronesian, in A. Adelaar, and A. Pawley (eds), *Austronesian Historical Linguistics and Culture History: A Festschrift for Bob Blust*, Canberra: Pacific Linguistics, 33–49.

- Böhm, Irene, Nikolaus Ritt, and Theresa Matzinger, Sound changes are selected by a bias against morphotactic ambiguity, *Journal of Historical Linguistics*, manuscript accepted for publication.
- Celata, Chiara, Katharina Korecky-Kröll, Irene Ricci, and Wolfgang U. Dressler (2015), Phonotactic processing and morpheme boundaries: Word-final/Cst/clusters in German, *Italian Journal of Linguistics* 27, 185-110.
- Dressler, Wolfgang U., and Katarzyna Dziubalska-Kořaczyk (2006), Proposing morphonotactics, *Wiener Linguistische Gazette* 73(1-19), 108-9.
- Dressler, Wolfgang U., Miroslava Hliničanová, Matej Ďurčo, Karlheinz Möřth, and Katharina Korecky-Kröll (2015) Phonotactic vs. morphonotactic obstruent clusters in Slovak and German, *Italian Journal of Linguistics* 27(1), 45-59.
- Kelley, Matthew C., and Benjamin V. Tucker (2017), The effects of phonotactic probability on auditory recognition of pseudo-words, *The Journal of the Acoustical Society of America* 141(5), 4038.
- Korecky-Kröll, Katharina, Wolfgang U. Dressler, Eva Maria Freiberger, Eva Reinisch, Karlheinz Möřth, and Gary Libben (2014), Morphonotactic and phonotactic processing in German-speaking adults, *Language Sciences* 46, 48-58.
- Lahiri, Aditi (2009), The dental preterites in the History of English, in K. Hanson, and S. Inkelas (eds), *The Nature of the Word: Studies in Honor of Paul Kiparsky*, Cambridge, MA: Mit Press, 507-525.
- Post, Brechtje, William D. Marslen-Wilson, Billi Randall, and Lorraine K. Tyler (2008), The processing of English regular inflections: Phonological cues to morphological structure, *Cognition* 109, 1-17.
- Storkel, Holly L. (2001) Learning new words: phonotactic probability in language development, *Journal of Speech, Language, and Hearing Research* 44(6), 1321-1337.
- Vitevitch, Michael S., and Paul A. Luce (2016), Phonological neighborhood effects in spoken word perception and production, *Annual Review of Linguistics* 2(1), 75-94.
- Wedel, Andrew (2006), Exemplar models, evolution and language change, *The Linguistic Review* 23(3), 247-74.
- Weřna, Jerzy (2009), The post-sonorant devoting of [d] in the past/past participle forms of weak verbs (sent, spend, etc.), in M. Krygier, and L. Sikorska (eds), *Pe laurer of oure Englische tonge*, Frankfurt: Peter Lang, 21-34.

Evolutionary dynamics of maximal syllable complexity

Laura C. Dees, Shelece M. Easterday, Paul Widmer, Balthasar Bickel & Chundra A. Cathcart
(University of Zurich, University of Hawai'i, University of Zurich, University of Zurich & University of Zurich)

Keywords: linguistic evolution, phylogenetic modeling, syllable structure, phonological complexity

Linguistic theory has since long proposed that the consonant-vowel (CV) syllable, or at least an asymmetry between strong onsets and weak codas, is a universally preferred (or even optimal) structure of language (Blevins, 2006; Gordon, 2016; Greenberg, 1978; Mailhammer et al., 2015 building on Jakobson, 1941; Zec, 2007; cf. Easterday, 2019). This claim is partially grounded in findings from acquisition (Davis & MacNeilage, 1995; Levelt et al., 2000), incremental processing (Content et al., 2001; Sun & Poeppel, 2023; Wedel et al., 2019) and the biomechanics of the articulation apparatus (Byrd, 1996; MacNeilage, 1998). However, large-scale surveys do not readily support a universal preference for CV syllables (Joo & Hsu, 2024; Maddieson, 2013) and it remains an open question whether linguistic evolution indeed favors the predicted onset-coda asymmetry.

The current study aims at shedding new light on this question by means of a quantitative phylogenetic approach. Since we expect a universal bias towards a CV template to constrain the evolution of grammars above and beyond synchronic frequencies of syllables, the focus is on maximal templates. Modeling change in syllable structure according to a Continuous Time Markov process, we estimate transition rates for maximal onset and coda sizes (in number of consonants). Our manually collected data come from reference grammars of 401 languages spread out over 5 families (Austronesian, Indo-European, Pama-Nyungan, Sino-Tibetan and Uto-Aztecan). Two competing models are compared: one that fits separate CTM processes on individual families, and one with a shared CTM process based on all attested states in each of the 5 families. Comparison is achieved by means of Bayes Factors based on the marginal likelihoods.

Results show that the lineage-specific model clearly outperforms the global model in both onset and coda (both Bayes Factors > 100 (Jeffreys, 1998)). We further explore lineage-specific stationary probabilities (informative of long-term preferences (Bickel, 2015; Maslova, 2000)), gain and loss rates and speed of change, which together help to characterize the evolutionary dynamics that variably contribute to diversification in this domain of language. These results imply that lineage-specific linguistic and/or socio-geographic contingencies impact syllable complexity more than any universal cognitive biases or biomechanical constraints on the articulators do – counter to the widespread belief that the CV template is a strong constraint on optimal syllable sizes in grammars (cf. Gordon, 2016).

References

- Bickel, B. (2015). Distributional Typology. In B. Heine & H. Narrog (Eds.), *The Oxford Handbook of Linguistic Analysis*. Oxford University Press.
- Blevins, J. (2006). Syllable: Typology. In K. Brown (Ed.), *Encyclopedia of language & linguistics* (Second edition, pp. 333–337). Elsevier.
- Byrd, D. (1996). Influences on articulatory timing in consonant sequences. *Journal of Phonetics*, 24(2), 209–244.

- Content, A., Kearns, R. K., & Frauenfelder, U. H. (2001). Boundaries versus onsets in syllabic segmentation. *Journal of Memory and Language*, 45(2), 177–199.
- Davis, B. L., & MacNeilage, P. F. (1995). The articulatory basis of babbling. *Journal of Speech and Hearing Research*, 38(6), 1199–1211.
- Easterday, S. (2019). *Highly complex syllable structure: A typological and diachronic study*. Language Science Press.
- Gordon, M. K. (2016). *Phonological typology*. Oxford University Press.
- Greenberg, J. H. (1978). Some generalizations concerning initial and final consonant sequences. In J. H. Greenberg, C. A. Ferguson, & E. A. Moravcsik (Eds.), *Universals of human language* (Vol. 2, pp. 243–279). Stanford University Press.
- Jakobson, R. (1941). *Kindersprache, Aphasie und allgemeine Lautgesetze*. Uppsala Universitets Arsskrift.
- Jeffreys, H. (1998). *The theory of probability*. Oxford University Press.
- Joo, I., & Hsu, Y.-Y. (2024). Phonotacticon: A cross-linguistic phonotactic database. *Linguistic Typology*.
- Levelt, C. C., Schiller, N. O., & Levelt, W. J. (2000). The acquisition of syllable types. *Language Acquisition*, 8(3), 237–264.
- MacNeilage, P. F. (1998). The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences*, 21(4), 499–511.
- Maddieson, I. (2013). Syllable Structure. In M. S. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Zenodo.
- Mailhammer, R., Restle, D., & Vennemann, T. (2015). Preference laws in phonological change. In P. Honeybone & J. Salmons (Eds.), *The Oxford handbook of historical phonology* (pp. 450–466). Oxford University Press.
- Maslova, E. (2000). A dynamic approach to the verification of distributional universals. *Linguistic Typology*, 4(3).
- Sun, Y., & Poeppel, D. (2023). Syllables and their beginnings have a special role in the mental lexicon. *Proceedings of the National Academy of Sciences*, 120(36), e2215710120.
- Wedel, A., Ussishkin, A., & King, A. (2019). Incremental word processing influences the evolution of phonotactic patterns. *Folia Linguistica*, 53(40–1), 231–248.
- Zec, D. (2007). The syllable. In P. de Lacy (Ed.), *The Cambridge handbook of phonology* (pp. 161–194). Cambridge University Press.

Are sound changes selected by their effects on the identifiability of words and their morphological structure?

Nikolaus Ritt and Irene Amparo Böhm
(University of Vienna)

Keywords: sound change, lexicostatistics, evolutionary phonology, morphonotactics, boundary signals

We explore the hypothesis that the actuation and implementation of sound changes are constrained by a preference for word forms with shapes that are (a) easy to identify as exponents of words and that are (b) indicative of their morphological structure.

For example, the voicing (or lenition) of the English *-es* suffixes (3SGPRES, PL, GEN) to /z/ (as in *stones* /EME /stɔ:nəs/ > ModE /stəʊnz/ had the effect of establishing final /z/ as a good signal not only of morphological complexity (nearly 90% of items with word final /z/ in the CUBE *Dictionary of Current English Pronunciation* are complex), but also of a word final boundary (more than 75% of all items containing /z/ have it in final position). – Similarly, the loss of word final /g/ and the phonemicization of the velar nasal /ŋ/ (as in *thing* ME /θɪŋg/ > ModE /θɪŋ/) established the phoneme as a signal of word (or stem) finality, while the retention of /h/ in foot initial onsets (but in no other positions) turned it into a signal of word initial boundaries.

Such cases suggest that phonological evolution may reflect a bias in favour of recognizable and morphotactically transparent word forms. This hypothesis is in line with research on lexical phonotactics, which demonstrates that speakers are sensitive to the occurrence frequencies of phonotactic patterns in the lexicon (e.g., Blevins 2009; Wedel 2006; Divjak 2019) and use them for identifying words in the speech stream (Saffran, Aslin, Newport 1996; Hay, Saffran 2012). Likewise, there is also evidence – from experiments (Hay 2004) – and from language acquisition (Kamandulytė-Merfeldienė 2006; Zydorowicz 2007; Zydorowicz 2010; e.g., Dressler, Hliničanová, Ďurčo, Mörrth, Korecky-Kröll 2015; Kelić, Dressler 2019) that speakers rely on phonotactic probabilities specifically for the recognition and decomposition of morphologically complex word forms (see also (Dressler, Dziubalska-Kořaczyk 2006; Dziubalska-Kořaczyk 2019).

For these reasons, phonological word form variants that are more easily identifiable as words and/or indicate their morphotactic structure better than competing variants should also be transmitted more faithfully and be historically more stable. This predicts a general bias in favour of sound changes that increase (or at least do not reduce) the identifiability of words as words and their morphotactic transparency.

We consider a set of well-known sound changes from this perspective. For example, it provides a plausible explanation of final devoicing in German, which started after the German Consonant Shift removed voiceless stops from word-initial and word medial position and thereby turned them into highly suitable signals of word endings. However, the focus of our presentation is on the question how the hypothesis that a preference for identifiable and morphotactically transparent word forms is a universal selection pressure in phonological evolution could at all be tested, falsified, or corroborated.

References

- Ambridge, Ben, Evan Kidd, Caroline F. Rowland, Anna L. Theakston (2015), The ubiquity of frequency effects in first language acquisition, *Journal of Child Language* 42 (2), 239-273.
- Blevins, Juliette, Structure-preserving sound change: A look at unstressed vowel syncope in Austronesian, in: K. Alexander Adelaar, Andrew Pawley, Robert A. Blust (eds), *Austronesian historical linguistics and culture history: A festschrift for Robert Blust*, Canberra: Pacific Linguistics, 38-55.
- Divjak, Dagmar (2019), *Frequency in language: Memory, attention and learning*, New York, NY: Cambridge University Press (CUP).

- Dressler, Wolfgang U., Katarzyna Dziubalska-Kołaczyk (2006), Proposing morphonotactis, *Wiener Linguistische Gazette* 73, 69-87.
- Dressler, Wolfgang U., Miroslava Hliničanová, Matej Ďurčo, Karlheinz Mörrth, Katharina Korecky-Kröll (2015), Phonotactic vs. morphonotactic obstruent clusters in Slovak and German, *Italian Journal of Linguistics* 27 (1), 45-59.
- Dziubalska-Kołaczyk, Katarzyna (2019), On the structure, survival and change of consonant clusters, *Folia Linguistica* 53 (s40-s1), 107-127.
- Goldrick, Matthew, Meredith Larson (2008), Phonotactic probability influences speech production, *Cognition* 107 (3), 1155-1164.
- Hay, Jennifer (2004), *Causes and Consequences of Word Structure*, New York, NY: Routledge.
- Hay, Jessica F., Jenny R. Saffran (2012), Rhythmic grouping biases constrain infant statistical learning, *Official Journal of the International Society on Infant Studies* 17 (6), 610-641.
- Kamandulytė-Merfeldienė, Laura (2006), The acquisition of morphonotactics in Lithuanian, *Wiener Linguistische Gazette* 73, 88-96.
- Kelić, Maja, Wolfgang U. Dressler (2019), The development of morphonotactic and phonotactic word-initial consonant clusters in Croatian first-language acquisition, *Suvremena Lingvistika* 45 (88), 179-200.
- Kelley, Matthew C., Benjamin V. Tucker (2017), The effects of phonotactic probability on auditory recognition of pseudo-words, *The Journal of the Acoustical Society of America* 141 (5_Supplement), p. 4038.
- Saffran, Jenny R., Richard N. Aslin, Elissa L. Newport (1996), Statistical learning by 8-month-old infants, *Science* 274 (5294), 1926-1928.
- Storkel, Holly L. (2001), Learning new words: phonotactic probability in language development, *Journal of Speech, Language, and Hearing Research* 44 (6), 1321-1337.
- Vitevitch, Michael S., Faisal M. Aljasser, Phonotactics in Spoken-Word Recognition, in: Jennifer S. Pardo, Lynne C. Nygaard, Robert E. Remez, David B. Pisoni (eds), *The Handbook of Speech Perception*, Hoboken, NJ: Wiley-Blackwell, 286-308.
- Vitevitch, Michael S., Paul A. Luce (1998), When Words Compete: Levels of Processing in Perception of Spoken Words, *Psychological Science* 9 (4), 325-329.
- Wedel, Andrew B. (2006), Exemplar models, evolution and language change, *The Linguistic Review* 23 (3), 247-274.
- Zydorowicz, Paulina (2007), Polish morphonotactics in first language acquisition, *Wiener Linguistische Gazette* 74, 24-44.
- Zydorowicz, Paulina (2010), Consonant Clusters Across Morpheme Boundaries: Polish Morphonotactic Inventory and its Acquisition, *Poznań Studies in Contemporary Linguistics* 46 (4), 565-588.

How languages echo nature: an acoustic analysis of natural sounds and their human imitations in English and Slovak

Renáta Gregová

(Pavol Jozef Šafárik University in Košice, Slovakia)

Keywords: onomatopoeia, natural sound, imitation, acoustic analysis, Slovak, English

Onomatopoeias – items representing a non-arbitrary relationship between form and meaning in language – allow various insights into sound symbolism and iconicity and offer a unique window into how languages capture extralinguistic reality through sound. These "words that sound like their referent" (Sidhu & Pexman 2017: 1623) demonstrate variation intralingually, with some languages having multiple onomatopoeic expressions for the same sound (*crack* – *crake* – *crash* in English), and extralingually, where one sound has different onomatopoeic versions across languages (*cock-a-doodle-doo* in English vs. *cocorico* in French for rooster sounds) (Krupa 1992, and Perniss et al. 2010). This variation depends on each language's phonemic inventory and limitations imposed by human perception and imitative abilities (Bredin 1996). Previous research suggests that humans echo natural sounds by following their acoustic properties (Tsur 2001, and Gregová & Hrytsu 2023). However, systematic experimental evidence comparing natural sounds with their linguistic representations across languages still remains limited.

To fill the gap, this study analyses the acoustic properties of nine natural sounds categorized into three frequency bands (high, middle, low) and compares them with their onomatopoeic representations in Slovak and English. The analysis examines: (a) complete sound profiles, (b) initial phases, and (c) final phases of both natural sounds and their verbal imitations. Acoustic measurements were conducted using spectral analysis, with particular attention to duration, fundamental frequency, formant structures, and intensity patterns. Preliminary results indicate that both languages show consistent adaptation patterns where low-frequency sounds generate longer imitations, mid-frequency sounds are typically shortened, and high-frequency sounds demonstrate reduction in both duration and frequency. Individual segments within verbal versions of natural sounds maintain remarkable acoustic stability, with minimal formant variation and fundamental frequency remaining consistent. The relationship between natural sounds and linguistic imitations reveals that speech sounds participating in onomatopoeias are selected based on optimal acoustic fit with perceived qualities of imitated sounds, supporting findings that "the speech sounds that participate in the linguistic imitations of natural sounds are those whose acoustic qualities best fit the perceived acoustic qualities of the imitated sounds" (Gregová & Hrytsu 2023: 15). This is in accordance with Tsur's hypothesis about acoustic property matching in sound imitation and aligns with Taub's (2001) findings that in the English word *ding*, the abrupt onset and fading offset reflect corresponding features in actual bell sounds.

This research indicates that onomatopoeia creation follows systematic rather than arbitrary patterns. Speakers employ consistent strategies across different sound types while conforming to language-specific phonological constraints. The findings contribute to understanding sound symbolism in human language, cognitive processing of natural sounds, and the interface between acoustic perception and linguistic production, providing evidence about which acoustic characteristics of

sounds shape their human imitation and determining the degree of agreement between natural sounds and their verbal representations in different language systems.

Acknowledgments

This work was supported by the Slovak Research and Development Agency (APVV), grant no. APVV-23-0027.

References

- Bredin, Hugh (1996), Onomatopoeia as a figure and a linguistic principle, *New Literary History* 27(3), 555–569.
- Gregová, Renáta and Hrytsu, Dmytro (2023), An acoustic analysis of a sample of English and Ukrainian onomatopoeias: A pilot study, *Govor* 40 (1), 3 – 26.
- Krupa, Viktor (1992), Ikonické akustické prvky v slovnej zásobe, *Jazykovedný časopis* 43, 26 – 33.
- Perniss, Pamela, Thompson, Robin L., and Vigliocco, Gabriella (2010), Iconicity as a general property of language: evidence from spoken and signed languages, *Frontiers in Psychology* 1, 227.
- Sidhu, David M. and Pexman, Penny M. (2018), Five mechanisms of sound symbolic association, *Psychon Bull* 25, 1619–1643.
- Taub, Sarah. F. (2001), *Language from the body: Iconicity and metaphor in American Sign Language*, Cambridge, UK: Cambridge University Press.
- Tsur, Reuven (2001), *Onomatopoeia: cuckoo-language and tick-tocking. The constraints of semiotic systems*, https://www.tau.ac.il/~tsurxx/Cuckoo_onomatopoeia.html.

Presentation of the OTPS project:

Onomatopoeia-Translation-Plurilingual-Submorphology

Sophie Saffi & Maruszka E-M. Meinard

(University of Aix-Marseille & University of Lille)

Keywords: onomatopoeia ; translation; comics ; submorphology ; classification

Within the LICOLAR team (Linguistique COMparée des LANGues Romanes) from the CAER laboratory (Centre Aixois d'Etudes Romanes) of Aix-Marseille University, Stéphane Pagès (Hispanic linguistics professor) and Sophie Saffi (Italian linguistics professor) lead a research project to study and translate onomatopoeia since 2023: the OTPS project (Onomatopoeia-Translation-Plurilingual-Submorphology), developed as an extension of a previous Comic Book Translation project (2019-2021, <https://traductionbd.hypotheses.org>). The Wordpress platform for the OTPS project is currently under development (<https://otps.univ-amu.fr>).

The OTPS project aims at creating an international network of researchers, translators and linguists who study onomatopoeia. The objectives of this project currently under development are the following ones: to set up a multilingual corpus of onomatopoeia in several language families (Romance, Germanic, Slavonic languages, etc....) as well as super-families (Indo-European, Sino-Tibetan, Afroasiatic languages, etc...); to set up a typology of onomatopoeia classified firstly according to the types of sounds they imitate and secondly according to the word formation process they stem from (thanks to a diachronic investigation); propose a submorphological analysis of their structure; translate them, produce descriptive sheets that include the results from the above-mentioned analyses. All this data will constitute an online resource for translators and cartoonists.

We will present our methodological questions, notably on how to classify onomatopoeia and how to identify their etymology, and we will also display examples of descriptive sheets that are already available on the online database.

Phonological rarities in areal perspective: consonants in Europe

Thomas Stolz & Nataliya Levkovych

(University of Bremen)

Keywords: areal phonology, consonants, rare phonemes

Advocating the distributional perspective for phonological typology, Hyman (2018: 17) puts areal matters on the agenda because “[p]honologists can and should be involved in [] identifying the geographical [] distributions of the phenomena”. This task is relevant for the study of phonological rarities most of which have to be located on linguistic maps yet. Maddieson (2005a–b, 2023: 251–256) shows that it is feasible to capture the global distribution of (selected) rarities cartographically. We assume that what is possible globally must also be possible areally (Guzmán Naranjo/Mertner 2023). We identify uncommon consonants and determine their distribution in Europe. Previous studies (e.g. Ternes 2010) on the areal phonology of Europe have largely ignored the issue of rarities.

The sample comprises 210 languages (doculects) that were alive in Europe in the 20th century. The geographical boundaries of Europe correspond to those introduced by König/Haspelmath (1999) for EUROTYPE. The analysis yields 311 different consonantal types which are unevenly distributed over the 210 languages. We evaluate the data synchronically. The phenomena are analysed both quantitatively and qualitatively. Diagram 1 reveals that there is a minority of widely distributed consonants which contrasts with a majority of consonants with a very limited distribution over the languages of Europe. Only 22 consonants are attested in $n \geq 50\%$ of the sample. With 289 consonants the vast majority of the types fails to reach this threshold.

What is important for the topic of rarities is the high number of consonants that are attested only sporadically. If we classify as rare any consonant that is reported in 5% (~ 16 languages) or less of the sample, the result is striking insofar as 246 consonants (= 79% of all consonants) in Europe have the status of areal *rara*. With 169 languages and 54% of the sample, the turnout and share for areal *rarissima* – i.e. for consonants which occur in 1% (~ 3 languages) or less of the sample – are remarkably big. This means that the inventory of European consonants is predominantly composed of areal rarities.

European *rarissima* tend to be phonologically complex in the sense that they involve secondary articulations (frequently combined with other marked features) such as the aspirated voiceless labialized denti-alveolar plosive /t^{wh}/ whose existence is stated for Lezgian (Haspelmath 1993: 33). Many but not all of the *rarissima* cluster in the Caucasian region. We address the geographical (and genetic) distribution of the rarities within the European macro-area. Many but not all European rarities are also cross-linguistically uncommon, e.g. /t^{wh}/ is absent from Maddieson (1984) but attested in eight of 2,186 languages in Phoible2.0 (Moran/McCloy 2019). We compare the European patterns with those emerging from global accounts of consonant distribution. The results impact not only on the hitherto phonologically underinformed EUROTYPE (Haspelmath 2001) but also add the much needed

areal component to both phonological typology (Gordon 2016), and the study of rarities in general (Wohlgemuth/Cysouw 2010).

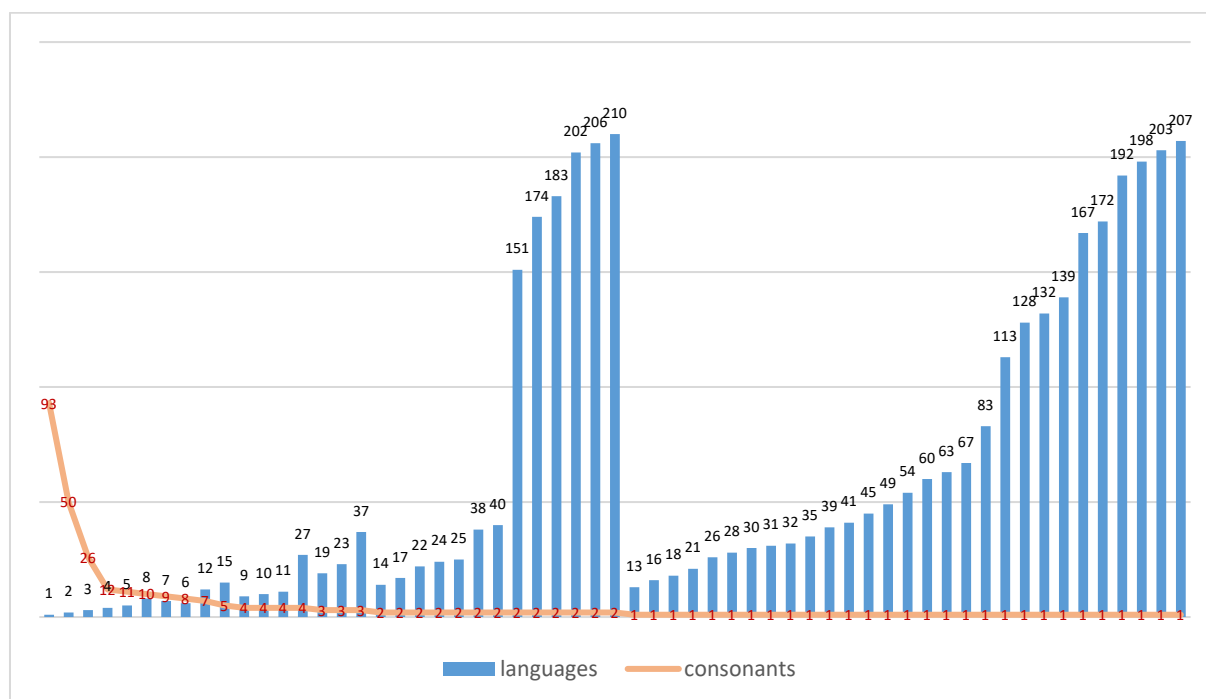


Diagram 1: Distribution of consonants over members of the European sample

References

- Gordon, Matthew K. (2016), *Phonological Typology*, Oxford: Oxford University Press.
- Guzmán Naranjo, Matías & Miri Mertner (2023), Estimating areal effects in typology: a case study of African phoneme inventories, *Linguistic Typology* 27 (2), 455–480.
- Haspelmath, Martin (1993), *A Grammar of Lezgian*, Berlin & New York: Mouton de Gruyter.
- Haspelmath, Martin (2001), The European Linguistic Area: Standard Average European, in Martin Haspelmath et al. (eds.), *Language Typology and Language Universals*, Berlin & New York: De Gruyter, 1492–1510.
- Hyman, Larry M. (2018), What is phonological typology?, in Larry M. Hyman & Frans Plank (eds.), *Phonological Typology*, Berlin & Boston: De Gruyter Mouton, 1–20.
- König, Ekkehard & Martin Haspelmath (1999), Der europäische Sprachbund, in Norbert Reiter (ed.), *Eurolinguistik. Ein Schritt in die Zukunft*, Wiesbaden: Harrassowitz, 111–128.
- Maddieson, Ian (1984), *Patterns of Sounds*, Cambridge: Cambridge University Press.
- Maddieson, Ian, (2005a), Absence of common consonants, in Martin Haspelmath et al. (eds.), *The World Atlas of Language Structures*, Oxford: Oxford University Press, 78–81.
- Maddieson, Ian (2005b), Presence of uncommon consonants, in Martin Haspelmath et al. (eds.), *The World Atlas of Language Structures*, Oxford: Oxford University Press, 82–85.
- Maddieson, Ian (2023), Investigating the ‘what’, ‘where’ and ‘why’ of global phonological typology, *Linguistic Typology* 27 (2), 245–266.
- Moran, Steven & Daniel McCloy (eds.) (2019), *PHOIBLE 2.0*, Jena: Max Planck Institute for the Science of Human History.
- Ternes, Elmar (2010), Phonetische Eigenschaften der Sprachen Europas, in Uwe Hinrichs (ed.), *Handbuch der Eurolinguistik*, Wiesbaden: Harrassowitz, 577–596.
- Wohlgemuth, Jan & Michael Cysouw (eds.) (2010), *Rara & Rarissima. Documenting the Fringes of Linguistic Diversity*, Berlin & Boston: De Gruyter Mouton.

General Session : Pragmatics

When Contrast Overrides Animacy: Overt Personal Pronouns as Subjects in Anaphoric Contexts in Adult Romanian

Adina Camelia Bleotu & Ekaterina Levina
(University of Bucharest; University of Vienna; & University of Vienna)

Keywords: Romanian, overt subjects, pronouns, null pronouns, experimental linguistics

The current study shows that contrast and animacy lead to an increase in anaphoric personal pronoun subjects over null ones in inter-sentential contexts in adult Romanian. We test experimentally the proposal by Dobrovie-Sorin & Giurgea (2013) that pronominal subjects in Romanian are used to mark contrast. Previous work targeted focus particles (Istrate et al. 2024) but the current study looks at contrastive entities/sets introduced with *dar* ‘but’ (Table 2). In line with Chomsky’s (1981) **Avoid Pronoun** principle and Ariel’s (1990) **Accessibility Theory**, in non-contrastive contexts (Table 1), speakers should prefer null subjects over overt pronouns when the referent is recoverable. We explore whether pronominal subjects can refer to inanimate entities in contrastive contexts, given the animacy preference (or restriction) observed in other languages (e.g., Perlmutter & Oresnik 1973 for Slovenian, Jaeggli 1982, Schroten 1992 for Spanish, Cardinaletti & Starke 1999 for Italian, German, Slovak, Hungarian, Hebrew). In Romanian, Giurgea & Ivan (2023) argue for a similar animacy preference, potentially indicating a tendency rather than a strict restriction. We ask whether the need to mark contrast can override animacy.

To address these questions, we conducted two tasks: a **No Contrast Task** (N = 31) and a **Contrast Task** (N = 26). In both tasks, participants were introduced to a character, Maria, who was learning Romanian. In the No Contrast Task, Maria provided a single sentence with a first-person null subject, a past tense verb, and a third-person nominal object. In the Contrast Task, Maria further added a second incomplete sentence introducing a contrast between a general and a present state. Participants then had to choose between continuations with an overt personal pronoun subject or a null subject, both referring back to a third-person object antecedent (see Tables 1 and 2). Materials included 2 training items, 32 test items, and 16 fillers. Apart from the *between subjects* contrast manipulation, the test items involved the *within subjects condition* Animacy (animate vs. inanimate for the object antecedent). Participants favored overt pronoun subjects over null pronouns in contrastive contexts and for animate antecedents (Figure 1). Notably, many participants also selected overt personal pronouns for inanimate antecedents in contrastive contexts, describing the overt pronouns as natural during debriefing. We fit the data into a generalized linear-mixed (logit) model with Answer as a dependent variable, the fixed effects Task, Animacy, and their interaction, and random effects for Participant and Item. The model shows significant effects per Task and Animacy ($p < .05$).

These findings highlight the robust role of contrastive contexts in overt pronoun use, suggesting a strong requirement for a filled contrastive Focus position (Rizzi 1997). Interestingly,

this requirement appears to outweigh animacy preferences for many speakers. Thus, the association between overt pronouns and animacy in Romanian is not absolute (not lexically or semantically encoded). Instead, it reflects a probabilistic tendency: while overt personal pronoun subjects are more likely to refer to animates, given their higher discursive and thematic role prominence (as potential Agents), the requirement to express contrast can take precedence.

(496 words)

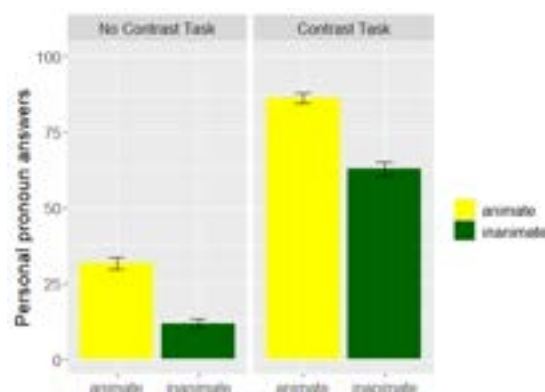
Table 1: Test items crossing Animacy of the object antecedent in the No Contrast Task

Animacy of the antecedent	Lead	Pro continuation	Personal pronoun continuation
Animate DP	Am auzit o bufniță. have.1SG heard an.F.SG owl.F.SG	Era zgomotoasă. was noisy.F.SG	Ea era zgomotoasă. she was noisy.F.SG
Inanimate DP	Am plantat un copac. have.1SG planted a.M.SG tree.M.SG	Era înalt. was tall.M.SG	El era înalt. He was tall.M.SG

Table 2: Test items crossing Animacy of the object antecedent in the Contrast Task

Animacy of the antecedent	Lead	Pro continuation	Personal pronoun continuation
Animate DP	Am auzit o bufniță. have.1SG heard an.F.SG owl.F.SG Ador bufnițe-le tăcut-e dar adore.1SG owls-the.F.P silent-F.PL but	era zgomotoasă. was noisy.F.SG	ea era zgomotoasă. she was noisy.F.SG
Inanimate DP	Am plantat un copac. have.1SG planted a.M.SG tree.M.SG Detest copaci-i scunz-i dar hate.1SG trees-the.M.PL short-M.PL but	era înalt. was tall.M.SG	el era înalt. he was tall.M.SG

Figure 1. Rate of personal pronoun answers in the No Contrast Task and in the Contrast Task



Acknowledgements

This research has been funded by the Austrian Science Fund (FWF) 10.55776/F1003. Both authors are currently supported by a postdoctoral position at the University of Vienna in the SFB project *Language between Redundancy and Deficiency*, specifically in subproject TP6 *The (Non-) Deficiency and (Non-)Redundancy of Clitic Pronouns* (led by PI Dalina Kallulli).

References

- Ariel, Mira (1990), *Accessing Noun-Phrase Antecedents*, London: Routledge.
- Cardinaletti, Anna and Michal Starke (1999), “The typology of structural deficiency: A case study of the three classes of pronouns,” in H. van Riemsdijk (ed), *Clitics in the Languages of Europe*, Berlin: Mouton de Gruyter, 145–233.
- Chomsky, Noam (1981), *Lectures on Government and Binding: The Pisa Lectures*, Dordrecht: Foris.
- Dobrovie-Sorin, Carmen and Ion Giurgea (2013), *A Reference Grammar of Romanian. Vol. 1: The Noun Phrase*, Amsterdam and Philadelphia: J. Benjamins.
- Giurgea, Ion and Rodica-Rudmila Ivan (2023), “On the internal structure of pronouns: From the perspective of noun ellipsis,” *Revue Roumaine de Linguistique* LXVIII (4), 315–351.
- Istrate, Fabian, Anne Abeillé and Barbara Hemforth (2024), “Subject alternation and antecedent preference in Romanian,” *Discours* [Online], No. 34.
- Jaeggli, Osvaldo (1982), *Topics in Romance Syntax*, Dordrecht: Foris.
- Perlmutter, David and Janez Orešnik (1973), “Language particular rules and explanation in syntax,” in *Festschrift for Morris Halle*, New York: Holt, Reinhart, and Winston, 419–459.
- Rizzi, Luigi (1997), “The Fine Structure of the Left Periphery,” in L. Haegeman (ed), *Elements of Grammar*, Dordrecht: Springer (Kluwer International Handbooks of Linguistics), 281–337.
- Schroten, Jan (1992), “On Spanish definite determiners: Personal pronouns and definite articles,” *Recherches de Linguistique Romane et Française d’Utrecht*, 9–24.

Italian *figuriamoci* as a discourse marker: A preliminary analysis of its functional domain

Cinzia Russi

(The University of Texas at Austin)

Keywords: discourse markers, Italian, pragmatic functions, grammaticalization, verb-based discourse markers.

Discourse Markers (DMs) are among “the most frequently used linguistic expressions in many languages” (Heine et al. 2021: 1). In the past thirty years, research on DMs has flourished considerably, focusing on different facets of DMs, including (a) their precise characterization regarding terminology, distinctive features, and functions (Degand et al. 2013, Fedriani and Sansò 2017); (b) their diachronic development, particularly related to grammaticalization (Ghezzi and Molinelli 2014a, Brinton, 2017, and Heine et al. 2021); (c) addressing a wider language sample, including Italian (Bazzanella 2001, 2010, and Fiorentini 2017). Yet, DMs remain a fertile field of inquiry and much about them remains to be uncovered or clarified. This study contributes to the ongoing research on DMs by addressing an Italian DM that (to the best of my knowledge) is still uninvestigated: *figuriamoci*, the 1PL imperative/exhortative form of the pronominal verb *figurarsi* ‘to imagine’, illustrated in (1) and (2) (from *CORIS/CODIS*).

- (1) “Avvocato Guerrieri, buongiorno!” [...] “Buongiorno Damiano [...] Ho bisogno di chiederti una cosa, ma devo parlarne a voce. Potresti farmi la cortesia di venire in studio?”. “Certo che posso, **figuriamoci**.”
“Dr. Guerrieri, good morning!” [...] “Good morning, Damiano [...] I need to ask you something, but I have to talk to you in person. Would you come to my office?” “Of course, **figuriamoci**.””
- (2) *E da quell'incontro non sono uscita affatto terrorizzata, figuriamoci.*
‘And that encounter did not terrorize me at all, **figuriamoci**.’

In (1) and (2) *figuriamoci* functions as a DM rather than a verb form because it neither affects the truth conditions of the utterances nor it adds anything to their propositional content: if omitted, the meaning of the utterances does not change. Yet, its omission significantly affects the speaker’s intended interpretation of the utterances.

This study offers a preliminary outline of *figuriamoci*’s discourse pragmatic functions. Sketching the synchronic functional domain of *figuriamoci* serves as starting point for a more comprehensive research project that analyzes other forms of *figurarsi* that also function as DMs, 2SG and 2PL imperative *figuratevi* and *figuratevi*, to (a) determine if they cover a similar array of pragmatic functions, and (b) situate them within the class of Italian DMs derived from verbs (Waltereit 2002, 2006, and Ghezzi and Molinelli 2014b) in terms of both functional domain and trajectory of diachronic development.

The analysis of an initial data set comprising about 100 tokens of *figuriamoci* reveals two distinct pragmatic functions, which I tentatively label ‘accommodating’ (1) and ‘contradicting’ (2). In (1) *figuriamoci* conveys the speaker’s willingness to comply with the interlocutor’s request (underscoring that it is not inconvenient at all); *figuriamoci*, then, functions essentially as a politeness marker. In (2), instead,

figuriamoci expresses the speaker's (contemptuous) rebuttal of their interlocutor's statement (i.e., that the encounter was terrorizing).

Research shows that a given DM typically fulfills multiple pragmatic functions. My present corpus, in fact, includes many tokens of *figuriamoci* to which the two functions proposed here cannot be satisfactorily applied, suggesting that further analysis of a larger data set will uncover others and/or lead to a refinement of the functions proposed here.

References

- Bazzanella, Carla (2001), I segnali discorsivi, in L. Renzi, G. Salvi, and A. Cardinaletti (eds), (2001), *Grande grammatica italiana di consultazione*, vol.3, Bologna: Il Mulino, 225–257.
- Bazzanella, Carla (2010), I segnali discorsivi, in G. Salvi, and L. Renzi (eds), (2010), *Grammatica dell'italiano antico*, Bologna: Il Mulino, 1339–1357.
- Brinton, Laurel J. (2017), *The Evolution of Pragmatic Markers in English: Pathways of Change*, Cambridge: Cambridge University Press.
- Corpus di Italiano Scritto (CORIS/CODIS)*: https://corpora.ficlit.unibo.it/coris_eng.html
- Degand, Liesbeth, Bert Cornillie, and Paola Pietrandrea (eds) (2013), *Discourse Markers and Modal Particles: Categorization and Description*, Amsterdam / Philadelphia: John Benjamins.
- Fedriani, Chiara, and Andrea Sansò (eds) (2017), *Pragmatic Markers, Discourse Markers and Modal Particles*, Amsterdam/Philadelphia: John Benjamins.
- Fiorentini, Ilaria (2017), Italian discourse markers and modal particles in contact, in Fedriani and Sansò (eds), 417–438.
- Ghezzi, Chiara, and Piera Molinelli (eds) (2014a), *Discourse and Pragmatic Markers from Latin to the Romance Languages*, Oxford: Oxford University Press.
- Ghezzi, Chiara, and Piera Molinelli (2014b), Italian *guarda, prego, dai*. Pragmatic markers and the left and right periphery, in Kate Beeching, and Ulrich Detges (eds), (2014), *Discourse Functions at the Left and Right periphery: Crosslinguistic Investigations of Language Use and Language Change*, Leiden/Boston: Brill, 117–150.
- Heine, Bernd, Gunther Kaltenböck, Tania Kuteva, and Haiping Long (2021), *The Rise of Discourse Markers*, Cambridge: Cambridge University Press.
- Waltereit, Richard (2002), Imperatives, interruption in conversation, and the rise of discourse markers: a study of Italian *guarda*. *Linguistics* 40(5), 987–1010.
- Waltereit, Richard (2006), The rise of discourse markers in Italian: a specific type of language change, in K. Fischer (ed), (2006), *Approaches to Discourse Particles*, Leiden/Boston: Brill, 61–76.

There's a new discourse marker use, so: On the recent emergence of final so

Gunther Kaltenböck (University of Graz)

Key words: final particle, grammaticalization, cooptation, prosody, Construction Grammar

While the discourse marker *so* has received considerable attention in the past decades (e.g. Raymond 2004, Bolden 2009, Denison 2020), its use as a clause/utterance-final particle has been noted, if at all, only in passing (e.g. cursory remarks in Schiffrin 1987, Cheshire & Williams 2002). This paper focuses on precisely this use, as illustrated by the example in (1), which is shown to have substantially increased in recent spoken (American) English.

- (1) AL-ROKER: And how are you celebrating your anniversary?
CRAIG-MELVIN: We're going to go down to D.C. for the Nationals game.
SHEINELLE-JONES: Oh, that's fun.
HODA-KOTB: Cool.
CRAIG-MELVIN: Going to hopefully go watch some history being made in Washington, see the Nats go to the World Series. She has covered the Nats for a number of years, so.
SHEINELLE-JONES: This is a special place for you guys.
AL-ROKER: Yeah (COCA:2019:Spoken)

Drawing on a number of corpora of mainly American English (*Corpus of Contemporary American English* (COCA), *Corpus of Historical American English* (COHA), *Fisher Corpus*; as well as the *British National Corpus*), the paper addresses the following research questions: (i) What are the discourse functions of this recent innovation?, (ii) what is its typical prosodic realization?, and (iii) how can we explain its recent emergence and development? The approach is thus corpus-based (both quantitative and qualitative) and the overall framework usage-based, discourse-analytic, and interactional.

The discourse function is identified as being both interpersonal and textual. On the interpersonal level, clause-final *so* signals that the speaker relinquishes their turn and tries to elicit a response from the interlocutor. The overwhelming majority of clause-final *so* (97% of the total of 979 instances in COCA Spoken) are in fact turn-final, involving full speaker change (rather than just backchannelling). On the textual level, it ties the host clause proposition to the preceding text by providing a relevance link, which is frequently that of a warrant or an explanation (e.g. in (1), where the underlined host clause explains why they are 'going down to D.C. for the Nationals game for their anniversary').

The prosody of *so*, which will be illustrated by PRAAT pictures, typically exhibits a falling (or, less frequently, level) contour and may be prosodically integrated with the host clause. In terms of its diachronic development clause-final *so* shows a significant increase in frequency particularly since the early 2000s, from 0.8 occurrences per million words in the period 1990-94 to 14.1 in 2015-19 in COCA Spoken (this corresponds to a significant rise also if measured against the baseline of all types of final *so*). It is argued that the emergence of clause-final *so* can be explained in terms of cooptation (Heine et al. 2015) of result subordinator uses of *so* and subsequent grammaticalization to different degrees.

References

- Bolden, Galina B. (2009), Implementing incipient actions: The discourse marker 'so' in English conversation. *Journal of Pragmatics* 41, 974–98.
- Cheshire, Jenny and Ann Williams (2002), Information structure in male and female adolescent talk. *Journal of English Linguistics* 30(2), 217-238.
- Denison, David (2020), Explaining explanatory *so*, in E. Jonsson and T. Larsson (eds.), (2009), *Voices Past and Present: Studies of Involved, Speech-related and Spoken Texts. In Honor of Merja Kytö*, Amsterdam: John Benjamins, 207–225.
- Heine, Bernd, Tania Kuteva and Gunther Kaltenböck (2015), On the evolution of final particles, in S. Hancil (ed.), (2015), *Final Particles*. Amsterdam: Benjamins, 111-140.
- Raymond, Geoffrey (2004), Prompting action: The stand-alone 'so' in ordinary conversation. *Research on Language and Social Interaction* 37(2), 185-218,
- Schiffrin, Deborah (1987), *Discourse markers*. Cambridge: Cambridge University Press.

Cracking the idiom code: an investigation of the dynamic interplay between idiom features and contexts

Irene Pagliai
(University of Göttingen)

Keywords: idioms, Italian, acceptability ratings, creativity, experimental linguistics

Idioms are a diverse class of multiword expressions (MWEs) characterized by a figurative meaning that deviates from their literal-compositional meaning (Cacciari and Tabossi 2014; Libben and D. A. Titone 2008). Because of this semantic peculiarity, idioms typically occur in contexts that do not reference their literal-compositional meaning.

In certain contexts, however, the literal-compositional meaning of an idiom may play a pivotal role. This occurs either when the intended interpretation of the MWE is its literal-compositional meaning (“literal-compositional context”) or when speakers creatively manipulate the literal meaning to engage in idiomatic wordplay (“figurative-creative context”; Langlotz 2006). Furthermore, idioms with certain inherent characteristics are arguably better suited to contexts that exploit the MWE literal meaning. Such idioms typically exhibit high levels of literal plausibility, decomposability and transparency (Moreno 2005; Vulchanova et al. 2019; Wagner 2021).

The present study experimentally investigates whether the ease of integrating idiom literal completions across fully literal-compositional and figurative-creative contexts is influenced by idiom-specific characteristics. These characteristics were operationalized using the *Potential Idiomatic Ambiguity* (PIA) index from the “Normed lexicon of English and Italian idioms” (Pagliai 2023), which assigns each idiom a PIA score based on the sum of mean values for literal plausibility, decomposability and transparency. Sixteen Italian idioms were selected from the lexicon based on their PIA index and categorized as either high- or low-PIA idioms.

Eighty-four native Italian speakers were recruited via Prolific to complete a fully online acceptability rating experiment using PennController for Internet-Based Experiments (PCIBEX, Zehr and Schwarz 2018). Idioms were embedded in dialogues between two fictitious speakers, A and B. Initially, only A’s line appeared, designed to bias participants toward either a literal or figurative interpretation. By clicking, participants revealed B’s reply, which consistently included the MWE followed by its literal completion, i.e., a context explicitly referring to the idiom’s literal constituents. Participants rated the acceptability of B’s response on a Likert scale from 1 to 7, based on A’s context. The staggered presentation of A’s and B’s lines also enabled the recording of reading times. Table 1 provides two English-translated examples of the experimental stimuli, illustrating one high- and one low-PIA idiom.

Building on hybrid, multidetermined models of idiom representation and processing (Libben and D. A. Titone 2008; Senaldi and D. Titone 2024), we expected higher ratings and shorter reading times for high-PIA idioms in the literal-compositional context. As for the figurative-creative context, this is, to our knowledge, the first attempt to experimentally investigate the systematicity of idiomatic wordplay. Thus, an exploratory approach was adopted to examine whether such wordplay exhibits systematicity rooted in distinct idiom features.

Results indicate that the ease of integrating idiom literal completions varies with PIA: high-PIA idioms tend to have higher acceptability ratings (clmm: $b = 0.55$, 95% CI [-0.02, 1.12], $p = .058$) and shorter reading times across both contexts compared to low-PIA idioms (lmm: $b = 31.27$, 95% CI [5.47, 57.00], $p = .025$). Notably,

Idiom	PIA	Condition	A	B
prendere per mano (take by hand)	High	FIG	You know, at school I have to follow a child with learning disabilities. I have to guide her in learning basic math skills.	Well, if you have to take her by hand, make sure you have a firm grip.
		LIT	You know, next week we are going with the whole family to Tokyo. I'm terrified of those huge street crossings, especially for my 4-year-old daughter.	
fare acqua da tutte le parti (make water from all sides)	Low	FIG	You know, I thought I had a good idea for the doctoral project. But at the moment, I don't think it's working.	Well, if it has to make water from all sides, make sure you have large buckets.
		LIT	You know, yesterday a pipe burst in our bathroom and it's continuing to leak. The floor is a mess, it's completely wet!	

Table 1: Examples of experimental items for both high- and low-PIA idioms

the positive effect of PIA on the ratings is more pronounced in the literal-compositional context, likely reflecting the greater difficulty of figuratively reinterpreting the literal completion in the figurative-creative context.

Overall, these findings provide further support for hybrid models of idiom representation and processing, highlighting the role of idiom-specific features in contextual integration (Beck and Weber 2020; Libben and D. A. Titone 2008; Senaldi and D. Titone 2024). Additionally, the study offers preliminary evidence for systematicity in idiomatic wordplay, suggesting that even creative idiomatic language may follow experimentally investigable principles, challenging the view of idiomatic wordplay as non-systematic (Bargmann et al. 2021; Fellbaum 2019).

Acknowledgements

Funded by the *Deutsche Forschungsgemeinschaft* (DFG, German Research Foundation) - GRK2636.

References

- Bargmann, S., Gehrke, B., & Richter, F. (2021). Modification of literal meanings in semantically non-decomposable idioms. In B. Crysmann & M. Sailer (Eds.), *One-to-many relations in morphology, syntax, and semantics* (pp. 245–279). Language Science Press. <https://doi.org/10.5281/zenodo.4729808>
- Beck, S. D., & Weber, A. (2020). Context and literality in idiom processing: Evidence from self-paced reading. *Journal of Psycholinguistic Research*, 49(5), 837–863. <https://doi.org/10.1007/s10936-020-09719-2>
- Cacciari, C., & Tabossi, P. (2014). *Idioms: Processing, structure, and interpretation*. Psychology Press.
- Fellbaum, C. (2019). How flexible are idioms? a corpus-based study. *Linguistics*, 57(4), 735–767.
- Langlotz, A. (2006). *Idiomatic creativity*. 1–339.
- Libben, M. R., & Titone, D. A. (2008). The multidetermined nature of idiom processing. *Memory & cognition*, 36, 1103–1121.
- Moreno, R. E. V. (2005). Idioms, transparency and pragmatic inference (tech. rep.). UCL Working Papers in Linguistics, 17: 389–425, 2005. 25.
- Pagliai, I. (2023). Bridging the gap: Creation of a lexicon of 150 pairs of english and italian idioms including normed variables for the exploration of idiomatic ambiguity. *Journal of Open Humanities Data*. <https://doi.org/10.5334/johd.123>

- Senaldi, M. S. G., & Titone, D. (2024). Idiom meaning selection following a prior context: Eye movement evidence of L1 direct retrieval and L2 compositional assembly. *Discourse Processes*, 61(1–2), 21–43. <https://doi.org/10.1080/0163853x.2024.2311637>
- Vulchanova, M., Milburn, E., Vulchanov, V., & Baggio, G. (2019). Boon or burden? the role of compositional meaning in figurative language processing and acquisition. *Journal of Logic, Language and Information*, 28(2), 359–387. <https://doi.org/10.1007/s10849-019-09282-7>
- Wagner, W. (2021). Idioms and ambiguity in context. De Gruyter. <https://doi.org/doi:10.1515/9783110685459>
- Zehr, J., & Schwarz, F. (2018). Penncontroller for internet based experiments (ibex) [OSF Preprints]. <https://doi.org/10.17605/OSF.IO/MD832>

Insubordinate constructions in Sardinian: Gerund clauses between syntactic and contextual (in)dependency

Roberta Caddeo
(University of Bergamo)

Keywords: insubordination; gerund; Campidanese Sardinian; grammaticalization; mirativity

The talk aims to illustrate some syntactic constructions attested in Campidanese Sardinian in which the gerund occurs as an autonomous predicate, not dependent on a matrix clause and used as a finite verb. These constructions typically convey modal meanings (especially evaluative-expressive and mirative ones) and exhibit an indexical character since they are strongly anchored in the pragmatic context, as exemplified as follows:

- (1) ['zɛmpiri βa'p:ɛŋɖi o ma'ri:a]?
always eating VOC Maria
'(Are you) eating all the time, Maria?'
context: the speaker sees Maria eating again after several times
- (2) ['vɛ:mu in s:a 'dɔ:mu e 'mari:a e 'is:a ɣistjo'nɛŋɖi
I.was in the house of Maria and she talking
e 'dɛ:u moreŋɖi'mi e s:u 'zon:u]
and I dying=REFL.1SG of the sleep
'I was at Maria's house and she (kept) talking and I (was) dying of sleep'

As far as Sardinian is concerned, the phenomenon has not yet been extensively investigated in the literature (Jones 1993; Cuzzolin 2005). We therefore propose to partially fill this gap in the studies, inscribing the phenomenon within the broader theoretical framework of insubordination (Evans 2007; Beijering, Kaltenböck & Sansiñena 2019).

Moving from empirical data obtained from approximately 10 hours of spontaneous and colloquial speech, we interpret the concept of finiteness and (in)subordination in scalar terms (Givón 1990; Bisang 2007) and argue that the various functional 'types' of independent gerunds attested in Campidanese Sardinian correspond to more or less advanced stages of grammaticalization of insubordinate processes, which can be placed along a *continuum* proceeding from prototypical subordination to prototypical insubordination. We will illustrate the pragmatic-discursive mechanisms that trigger the phenomenon and show how Sardinian resort to the inherent imperfectivity of the gerund to project narrated events – deictically past – into a space perceived as contemporary to the moment of enunciation, reducing the metaphorical distance between speakers and subjects in discourse space or time. Therefore, we will propose the application of the notions of "phantasmatic deixis" (Bühler 1934) and "proximity device" (Caffi & Janney 1994), that is, indexical strategies that guide the speaker in reconstructing the coordinates of spatio-temporal *representation* of an imagined or remembered event to stimulate his/her emotional and psychological proximity.

References

- Beijering, Karin, Kaltenböck, Gunther & Sansiñena, María Sol (eds.), *Insubordination. Theoretical and Empirical Issues*, Berlin / Boston: De Gruyter, 1-28.
- Bisang, Walter (2007), Categories that make finiteness: discreteness from a functional perspective and some of its repercussions, in I. Nikolaeva, (ed.), *Finiteness. Theoretical and Empirical Foundations*, Oxford: Oxford University Press, 115-137.
- Bühler, Karl (2011) [1934], *Theory of Language*, Amsterdam / Philadelphia: John Benjamins Publishing.
- Caffi, Claudia & Janney, Richard W. (1994), Toward a pragmatics of emotive communication, *Journal of pragmatics* 22, 325-375.
- Cuzzolin, Pierluigi (2005), Some remarks on gerund in Sardinian, *Sprachtypologie und Universalienforschung* 58(2/3), 176-187.
- Evans, Nicholas (2007), Insubordination and its uses, in I. Nikolaeva, (ed.), *Finiteness. Theoretical and Empirical Foundations*, Oxford: Oxford University Press, 366-431.
- Givón, Talmy (1990), *Syntax: A Functional-Typological Introduction*, Vol. 2, Amsterdam: Benjamins.
- Jones, Michael Allan (1993), *Sardinian Syntax*, London, Routledge.

Subject-auxiliary inversion in embedded and appositive questions in TED talks: A semantic and pragmatic analysis across contextual categories

Sepideh Nazari & Agnes Celle

University of Paris

Abstract

This study investigates the syntactic, semantic, and pragmatic features of Subject-Auxiliary Inversion (SAI) in embedded and appositive inquiries as they show in the semi-formal speech of TED Talks. While SAI has been investigated in direct questions, little is known about how it occurs in more complex syntactic patterns, including embedded and appositive inquiries. E.g., "She wondered why the meeting had been postponed" shows an example of SAI in an embedded question, while in an appositive question it could seem as, "The question, why had the meeting been postponed, puzzled everyone." Leonarduzzi & Herment (2020) have looked at this hybrid mix's syntactic and prosodic effects. This work aims to offer a new understanding of SAI's functional adaptability and communicative value: Why is SAI used in appositive and embedded questions? In which contextual categories does SAI occur most frequently? In which situations are contextual interactions conducive to SAI use? Inspired by the constructional model of grammar (Goldberg, 1995), which helps to view SAI as a discourse-sensitive structure shaped by communicative function rather than syntactic position alone, the analysis is grounded in a discourse-pragmatic and usage-based framework.

The study looks at SAI both quantitatively and qualitatively using a mixed-methods approach. Ten contextual categories adapted from the TransQuest framework were applied in the quantitative analysis, expanded to include speaker- and addressee-sensitive dimensions: Tense, Aspect, Polarity, Modality, Answerhood, Interlocution, Relation to Preceding Context, Relation to Subsequent Context, Speaker Role, and Addressee Function. These parameters taken together provide a discourse-pragmatic explanation of SAI in embedded and appositive questions intended not to generate answers. In the qualitative study, the semantic and pragmatic consequences of SAI were examined, with particular focus on how inversion marks rhetorical stance, improves discourse coherence, and involves the audience via conceptual alignment rather than direct interaction. These approaches taken together provide a comprehensive understanding of SAI's role in complex syntactic systems. Transcripts of 15 TED Talks, including presenters from many linguistic and cultural backgrounds, were utilized for a broad representation of contextual variance. Preprocessing involved cleaning the transcripts, manually identifying SAI occurrences, and tagging them with contextual information to allow for more complete analysis.

The results reveal SAI is more likely in appositive inquiry, especially when rhetorical goals and discourse relations are stressed. In contrast, embedded questions have limited, context-dependent SAI uses. These statistical tendencies are supported by significant differences between the two types of questions. Qualitative studies reveal that appositive inquiries use SAI to emphasize questions, enhancing audience engagement and rhetorical effectiveness. Since they are more context-sensitive and specialized, embedded inquiries use SAI to handle difficult syntactic and semantic limitations. This study advances linguistic theory by exploring SAI in circumstances other than direct inquiry and highlighting how it interacts with contextual and communicative elements in spoken English.

Keywords: Subject-Auxiliary Inversion (SAI), Embedded and Appositive Questions, TED Talks Discourse, Syntactic and Pragmatic Analysis, Rhetorical Strategies

References

- Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. University of Chicago Press.
- Leonarduzzi, L., & Herment, S. (2020). Subordinate Interrogatives and Subordination in Oral Speech: Syntax and Prosody. *Linx. Revue des linguistes de l'université Paris X Nanterre*, (80).

Was um Gottes willen ist daran Besonderes?

Comparing the different [*um* NP_{Gen} *willen*] phrases in the WHX Construction in German

Steven Schoonjans

Alpen-Adria-Universität Klagenfurt & KU Leuven

Keywords: WHX Construction, German, *um* NP_{Gen} *willen*, wh questions, syntax

Among the expressions that can be used as intensifiers (the H slot) in the WHX Construction (cf. Hugou 2017) in German are not just phrases such as *zum Teufel* ‘the devil’, *zur Hölle* ‘the hell’, and *in aller Welt* ‘in all world’, but also a number of phrases that instantiate the [*um* NP_{Gen} *willen*] pattern, such as *um Gottes willen* ‘for God’s sake’ and *um Himmels willen* ‘for heaven’s sake’:

- (1) Was *um Gottes willen* ist daran Besonderes? (DeReKo E96)
‘What *for God’s sake* is so special about that?’
- (2) Was *um Himmels Willen* ist eine Bodhran oder eine Bombarde? (DeReKo RHZ01)
‘What *for heaven’s sake* is a bodhrán or a bombard?’

This pattern is mainly instantiated by the lexicalized variants in (1-2), but it is also marginally productive, with variants such as *um aller Welt willen* ‘for all world’s sake’, *um Satans willen* ‘for Satan’s sake’, or *um Merlins willen* ‘for Merlin’s sake’:

- (3) Was *um Merlins willen* tust du hier, Hermione? (HP ECH5)
‘What *for Merlin’s sake* are you doing here, Hermione?’

The aim of this contribution is to investigate to what extent the different phrases instantiating the [*um* NP_{Gen} *willen*] pattern behave similarly or differently from a syntactic and distributional point of view. To this end, all 3081 occurrences of this pattern in six corpora (DeReKo, DGD, GeRedE, CodE Alltag 2.0 and a self-compiled Harry Potter and SpongeBob corpus) are analyzed with regard to their position in the sentence (prefield/middlefield), their text type distribution, the question words they combine with, their use in verbless questions, the extent to which they are combined with the particle *nochmal* ‘again’, etc.

It will be shown, among other things, that while there are no major differences at the level of text type distribution, the phrases do behave differently in other respects. For instance, *um Gottes willen* can more easily be used in the middlefield (i.e. following the finite verb), while *um Himmels willen* and even more so the occasional variants show a stronger tendency immediately follow the question word. Similarly, *um Gottes willen* can more easily be used in questions without a verb and can more easily be separated from the rest of the sentence by means of commas than *um Himmels willen* and the occasional variants. On the other hand, *um Himmels willen* stands out in that it is less commonly used in questions with *warum* ‘why’ than the other phases (note that *warum* is one of the most typical question words in the WHX Construction in German).

In the remainder of the talk, I will reflect on possible correlations between these factors and possible explanations for these observations, looking, among other things, at the different origins of the different phrases and at the influence from other patterns used in the WHX Construction (e.g. *in Gottes/Himmels Namen* ‘in God’s/heaven’s name’).

References

Hugou, Vincent (2017), The WHX construction (what the hell...?) and intensity: A corpus-based study, *Lexis* 10(1), 1-30. DOI: 10.4000/lexis.1103

Questioning the homogeneity of move analysis: Segmentation and annotation of legal opinions

Mary C. Lavissière & Warren Bonnard

(Nantes Université, CRINI, & Université de Lorraine, ATILF)

Move analysis (Swales, 1990, 2004) is a dominant framework in discourse analysis for modeling the rhetorical structure of specialized discourse, especially research articles (RAs). This study questions its applicability to highly variable genres, particularly legal decisions published by American courts. These lack a uniform structure, are often long, and pose challenges to discourse analysis and natural language processing (NLP), including difficulties in identifying consistent move boundaries and calculating interrater agreement across diverging segmentations.

Move analysis has been applied to various specialized genres, including judicial opinions (Lavissière & Bonnard, 2024; Bhatia, 1993; Goźdz-Roszkowski, 2020). It defines a two-tiered organization of discourse units: “moves,” units with one communicative function, and “steps,” the concrete ways moves are carried out through language (Biber et al., 2007). Despite the use of corpora, the framework remains influenced by the rhetorical organization of RAs, which typically follow a linear structure (IMRD) aimed at distinguishing the researcher’s contribution. This RA-centered model does not fully reflect the rhetorical variability in other specialized genres (Rau, 2021), particularly in legal opinions for which prior analyses often identify broad discourse units—e.g., “Arguing the case” (Mazzi, 2007)—that overlook finer distinctions important for rhetorical or computational modeling.

This study presents a new framework for analyzing legal discourse and other argumentative prose, particularly for NLP applications. It includes (1) a segmentation framework prioritizing syntactic cues over functional ones and (2) an annotation scheme informed by alternative discourse models (Charolles, 2020; Charolles & Pietrandrea, 2012) and dialogic linguistics (Bres et al., 2016). It integrates the narrative and argumentative dimensions of judicial opinions, as well as the interplay of multiple voices (Bres, 2017). This framework builds on corpus-based methodology (see Moreno & Swales, 2018), but applies it to finer-grained syntactic units annotated for communicative function. These functions are further interpreted dialogically—for instance, indicating whether a party or a judge is speaking (see examples (1) and (2))—refining large analytic sections in previous move models for legal opinions. The framework also accounts for how the Court mobilizes authoritative sources, such as prior decisions, through rhetorical functions like *Recalling a SCOTUS decision* (see example (3)). Findings suggest that, unlike scientific texts, legal reasoning integrates itself into existing precedent rather than presenting novel contributions.

- (1) <Recalling an argument from the respondent> = “Before this Court, as they have throughout this litigation, respondents assert only one justification for their use of race in the admissions process.”
- (2) <Recalling an argument from a justice authoring a separate opinion> = “THE CHIEF JUSTICE believes that the Law School’s policy conceals an attempt to achieve racial balancing, and cites admissions data to contend that the Law School discriminates among different groups within the critical mass.”
- (3) <Recalling a SCOTUS decision> = “As Justice Powell made clear in *Bakke*, truly individualized consideration demands that race be used in a flexible, nonmechanical way.”

The framework was used to annotate a representative sample ($n = 180$) of U.S. Supreme Court majority opinions, selected by author-judge and theme (Lavissière & Bonnard, 2024) to represent cases from 1945 to 2020. An acceptable interrater reliability (Cohen's Kappa = .67) shows the framework minimizes fuzziness in identifying moves and steps ($n = 26,328$ segments).

These findings challenge the dominance of move analysis as a universal tool for analyzing rhetorical structures, especially in non-standardized genres. This study expands methodological tools in genre analysis by enabling finer-grained, dialogically informed annotation. A larger corpus will be analyzed using machine learning to uncover rhetorical patterns in judicial reasoning.

Acknowledgements

This study has been funded by the French National Research Agency (ANR), project ANR-22-CE38-0004.

References

- Bhatia, V. K. (1993). *Analysing genre: Language use in professional settings*. Routledge.
- Biber, D., Connor, U., & Upton, T. A. (2007). *Discourse on the move: Using corpus analysis to describe discourse structure*. Benjamins.
- Bres, J. (2017). Dialogisme, éléments pour l'analyse. *Recherches en didactique des langues et des cultures. Les cahiers de l'Acedle*, 14(2), Article 2. <https://doi.org/10.4000/rdlc.1842>
- Bres, J., Nowakowska, A., & Sarale, J.-M. (2016). Anticipative interlocutive dialogism: Sequential patterns and linguistic markers in French. *Journal of Pragmatics*, 96, 80–95. <https://doi.org/10.1016/j.pragma.2016.02.007>
- Charolles, M. (2020). Discourse topics and digressive markers. *Journal of Pragmatics*, 161, 57–77. <https://doi.org/10.1016/j.pragma.2020.01.005>
- Charolles, M., & Pietrandrea, P. (2012). En réalité: De la modalisation à l'organisation du discours. *Travaux de linguistique*, 64(1), 111–142. <https://doi.org/10.3917/tl.064.0111>
- Goźdz-Roszkowski, S. (2020). Move Analysis of Legal Justifications in Constitutional Tribunal Judgments in Poland: What They Share and What They Do Not. *International Journal for the Semiotics of Law - Revue Internationale de Sémiotique Juridique*, 33(3), 581–600. <https://doi.org/10.1007/s11196-020-09700-1>
- Lavissière, M. C., & Bonnard, W. (2024). Who's really got the right moves? Analyzing recommendations for writing American judicial opinions. *Languages*, 9(4), 119.
- Mazzi, D. (2007). The construction of argumentation in judicial texts: Combining a genre and a corpus perspective. *Argumentation*, 21(1), 21–38.
- Moreno, A. I., & Swales, J. M. (2018). Strengthening move analysis methodology towards bridging the function-form gap. *English for Specific Purposes*, 50, 40–63.
- Rau, G. (2021). Development of component analysis to support a research-based curriculum for writing engineering research articles. *English for Specific Purposes*, 62, 46–57. <https://doi.org/10.1016/j.esp.2020.12.001>

- Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.
- Swales, J. (2004). *Research Genres: Explorations and Applications*. Cambridge University Press.
<https://doi.org/10.1017/CBO9781139524827>

Yael Maschler & Anna Inbar
(University of Haifa & University of Haifa)

Acknowledgements: This research is generously supported by the Israel Science Foundation, grant no. 572/24 to Yael Maschler.

Our data are drawn from the *Haifa Multimodal Corpus of Spoken Hebrew* (Maschler et al. 2025), which consists of naturally occurring casual conversations among friends and relatives. Using Interactional Linguistic methodology (Couper-Kuhlen & Selting 2018) and Multimodal Conversation Analysis (Goodwin 2018), we focus on insubordinate causal clauses comprising an ironic utterance prefaced by the Hebrew causal conjunction *ki* ‘because,’ demonstrating that in naturally occurring interactions, such insubordinate clauses are deployed to ridicule previous utterances, as illustrated in the following excerpt.

Ex. (1):

- 221

06 ...**ki** **xaser** **lahem kesef.**
 because lack.PRS.3SGM to-3PLM money
 ...because they lack money.

...

09 Yair: 'az lama hem lo kanu plata,
 so why they NEG buy.PST.3PL platter
 so why didn't they buy a platter,

10 'im lo xaser lahem kesef?
 if NEG lack to-3PLM money
 if they don't lack money?

Hanan's utterance is ironic—the acquaintances are sufficiently well off—as attested by Yair's immediate response 'if they don't lack money' (line 10). With this ironic construction, Hanan ridicules Yair's and Omri's assessments as inappropriate since the acquaintances' not having a Sabbath platter is not a matter of lack of money, as they should know, being familiar with their economic situation. This causal clause cannot be interpreted as dependent, even when considering *structural latency* (Auer 2015). Yet, similar to other *ki*-causal clauses, it can be viewed as a justification (Inbar & Maschler 2023), in this case, of the ridiculing action. In this talk, we will explore the emergence of this construction in interaction, examining its formal and functional characteristics, as well as its degree of autonomy based on syntactic, prosodic, semantic, and pragmatic features.

References

- Auer, Peter (2015), The temporality in interaction: Projection and latency, in A. Deppermann, and S. Günthner (eds), (2015), *Temporality in Interaction*, Amsterdam/Philadelphia: John Benjamins, 27–56.
- Beijering, Karin, Kaltenböck, Gunther, & Sansiñena, Maria S. (2019) Insubordination: Central issues and open questions, in K. Beijering, G. Kaltenböck, and M. S. Sansiñena (eds), (2019), *Insubordination: Theoretical and Empirical Issues*, Berlin: De Gruyter Mouton, 1–28.
- Couper-Kuhlen, Elizabeth, & Selting, Margret (2018), *Interactional Linguistics: Studying Language in Social Interaction*, Cambridge: Cambridge University Press.
- Dwyer, Arienne (2016), Ordinary insubordination as transient discourse, in N. Evans, and H. Watanabe (eds), (2016), *Insubordination [Typological Studies in Language 115]*, Amsterdam/Philadelphia: John Benjamins, 183–208.
- Evans, Nicholas (2007), Insubordination and its uses, in I. Nicolaeva (ed.), (2007), *Finiteness: Theoretical and Empirical Foundations*, Oxford: Oxford University Press, 366–431.
- Goodwin, Charles (2018), *Co-operative Action*, Cambridge: Cambridge University Press.
- Inbar, Anna, & Maschler, Yael (2023), Shared knowledge as an account for disaffiliative moves: Hebrew *ki* 'because'-clauses accompanied by the Palm Up Open Hand gesture, *Research on Language and Social Interaction* 56(2), 141–164.
- Kaltenböck, Gunther, & Keizer, Evelien (2022), Insubordinate *if*-clauses in FDG: Degrees of independence, *Open Linguistics* 8(1).
- Maschler, Yael, Polak-Yitzhaki, Hilla, Aghion, Gallit, Fofliger, Ophir, Wildner, Nikolaus, Ben-Moshe, Yotam M., Lagil, Rotem, Saar, Shira, Inbar, Anna, & Geva, Yuval (2025), The Haifa Multimodal Corpus of Spoken Hebrew, Haifa: University of Haifa.

Communication when the local context changes

Yoshie Yamamori
(Doshisha University)

Keywords: <communication, common ground, presupposition, metalinguistic negation, local context change>

The Stalnakerian view of presupposition and common ground (CG) (Stalnaker, 1978, 2014) is a well-known context-directed theory of communication. The CG is the set of propositions, the truth of which is taken for granted, that constitute part of the background of conversation. CG can play this role only insofar as all interlocutors share the CG as public information. Thus, updating of context is the essential aim of a communicative act that seeks to retain the continuity/coherence of contexts. This idea informed Roberts' theory (2004, 2012, 2018), in which communicative acts are interpreted as relevant only if their contents bear certain logical relations to the questions under discussion (QUDs). Within this framework, when interlocutors disagree about the QUDs, the context enters a state that leads to miscommunication. However, it is possible to communicate effectively even in such a situation. This study aimed to discover mechanisms in the background of conversation that continue even when the public nature of the QUDs is not guaranteed. I show that the performance of a communicative act requires amendment of the addressee's local contexts when an incongruity develops between interlocutors. I argue that Roberts' theory is an idealized model, based on my observations of Japanese data that follow.

Below, (1) and (2) include the same interrogative phrases. Although ordinary yes-no questions can be associated with true and false answers, (2b) is an exception to this rule.

(1) (A staff member of city hall A is filling out a family survey on behalf of B.)

a. A: Ootoo-san-wa moo shakaizin desu-ka?
Younger brother-Mr-TOP already working adult Copula-PRESENT-Q
'Is your younger brother already employed?'

b. B: *Ie, shakaizin zya-nai-desu*
No working adult Copula-Neg-Copula-PRESENT
'No, he's not employed.'

(2) a. A: *Hisashiku at-te-nai-desu ga,*
For a long time meet-te-Neg-Copula-PRESENT but
otooto-san-wa moo shakaizin-desu-ka?
'I haven't seen him in a while, but is your younger brother already employed?'

b. B: # *Ie, shakaizin zya-nai-desu*
(✓ *Ie, Mada, gakusei-desu*)
No still student- Copula-PRESENT
'No, he is still a student.'

The (2b) example indicates that a negative sentence cannot serve as an answer to (2a). The question of whether negative sentences are interpreted appropriately is determined by the properties of the interrogative phrases. In (1a), the information *p* - "your younger brother is employed" - is recognized as a survey question by both A and B. By contrast, in (2a), *p* presupposes only the mental state of A.

The important feature of the restricted negative answer in (2b) can be understood as a strategy by which B implies disagreement about the context presupposed by A, and repairs that context. The evidence includes the fact that the metalinguistic negation of (3), which represents an objection to a previous utterance (Horn, 1985), can serve as an answer to (2a) but not (1a).

(3) Ie, shakaizin na-no-dewa-nai-desu.
No, working adult Copula-COMP-Copula-Neg-Copula-PRESENT
'No, it is not the case that he is employed.'

In (3), *no* creates a CP-domain wherein the phrase *be employed* cannot be used when engaging in negation.

Acknowledgments: This work is supported by JSPS KAKENHI Grant Number JP25K04053.

References

- Horn, Lawrence R. (1985), Metalinguistic negation and pragmatic ambiguity, *Language* 61, 121-174.
- Roberts, Craige (2004), Context in dynamic interpretation, in L. Horn & G. Ward (eds), *The handbook of Semantics and Pragmatics* 5, Oxford: Blackwell, 197-220.
- Roberts, Craige (2012), Information structure in discourse: Toward an integrated formal theory of pragmatics, *Semantics and Pragmatics* 5, 1-69.
- Roberts, Craige (2018), Speech acts in discourse context, in D. Fogal, D. Harris & M. Moss (eds), *New work on speech acts*. Oxford: Oxford University Press, 317-359.
- Stalnaker, Robert C. (1978), Assertion, in P. Cole (ed), *Syntax and semantics* 9, New York: Academic Press, 315-332.
- Stalnaker, Robert C. (2014), Context. *Context and content*, Oxford: Oxford University Press.

General Session : Semantics

What kind of verb is TAKE?

The English and Swedish verbs *take/ta* and their correspondences in other European languages

Åke Viberg
(Uppsala University)

Keywords: lexical semantics, Possession verbs, Motion verbs, crosslinguistic study, European languages

The aim of this study is to describe the polysemy of English *take* and its major correspondences in a selection of other European languages with a focus on the most concrete meanings. In English, *take* is often regarded as the opposite of *give* but also as the opposite of *bring* or of *put* (Kopecka & Narashimhan 2012, Stathli 2023, Viberg 2010). In several languages, different verbs enter these oppositions. Data consist of all occurrences of English *take* and Swedish *ta* in the English-Swedish Parallel corpus (Altenberg & Aijmer 2000) and all 1600 occurrences of the Swedish verb *ta* 'take' and its translations into English, German, French, Spanish and Finnish extracted from the Multilingual Parallel Corpus (MPC), which consists of extracts from 22 Swedish novels, around 600,000 words in the Swedish original texts. In addition, data from other languages will be taken from Intercorp (Subtitles, Europarl) and ImagAct. TAKE can be described as a goal-directed action sequence (Viberg 2016): extending the arm and grasping an object as the first part of a chain of actions carried out for some specific purpose as in: *She took a pen and started writing*. This use is referred to as Handling (see Figure 1). Motion is involved but only within the region of interaction of the Agent. Earlier steps than Handling are occasionally focused, for example grasping: *She took my hand*. Physical contact between humans has several emotional and symbolic extended meanings. TAKE has usually developed from a verb meaning 'take hold of, seize' in Indo-European languages (Darling Buck 1949).

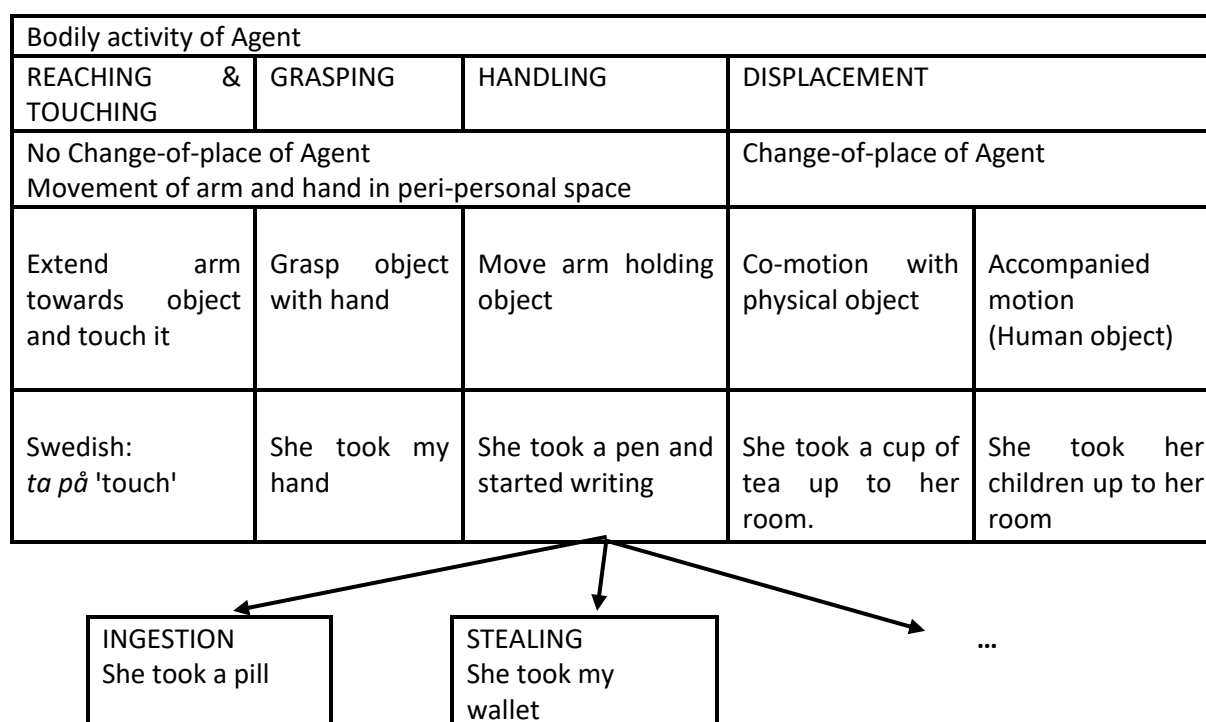


Figure 1. The meaning potential of TAKE (sketch)

Handling specific objects is associated with specific goals: *She took a pill* implies ingesting the pill. Extensions of this type can be interpreted as a form of metonymy: Sub-event for Event. Stealing can be regarded in a similar way. Displacement in Figure 1 will be discussed from a perspective where TAKE is regarded as the opposite of BRING (cf. Margetts et al. 2022). The result is shown schematically below:

<i>Languages</i>	Possession	Co-motion			
	TAKE	TAKE ^{CM}		BRING	
		Hum obj	Inan obj	Hum obj	Inan obj
English	take			bring	
Swedish	ta	ta med		ta med, komma med	
German	nehmen	mitnehmen		(mit)bringen	
French	prendre	emmener	emporter	amener	apporter
Spanish	tomar	llevar		traer	
Greek	pérno	páo 'go'		féro	
Finnish	ottaa	viedä, ottaa mukaan		tuoda	

In Swedish, the reflexive form *ta sig* is often used to refer to SelfMotion, a language-specific use. How abstract meanings can be accounted for will be briefly indicated, for example, Duration (*The journey took two hours*) and various uses as a support verb (*take a decision*).

REFERENCES

- Altenberg, Bengt & Aijmer, Karin (2000), The English-Swedish Parallel Corpus: A resource for contrastive research and translation studies, in C. Mair & M. Hundt, M. (eds), *Corpus Linguistics and Linguistic Theory*, Amsterdam and Atlanta: Rodopi, 15-33.
- Darling Buck, Carl (1949), *A Dictionary of Selected Synonyms in the Principal Indo-European Languages*. Chicago: The University of Chicago Press.
- Kopecka, Anetta and Bhuvana Narasimhan (eds. 2012), *Events of Putting and Taking. A Crosslinguistic Perspective*. [Typological Studies in Language 100] Amsterdam: Benjamins.
- Margetts, Anna, Sonja Riesberg & Birgit Hellwig (eds. 2022), *Caused Accompanied Motion: Bringing and Taking Events in a Cross-linguistic Perspective*. [Typological Studies in Language, 134] Amsterdam: Benjamins.
- Stathli, Katerina (2023), *Granularity in the Verbalization of Events and Objects. A Cross-linguistic Study*. Amsterdam: Benjamins.
- Viberg, Åke (2010), Basic Verbs of Possession, in M. Lemmens (ed.) Unison in multiplicity: Cognitive and typological perspectives on grammar and lexis. *CogniTextes* 4. <http://cognitextes.revues.org/308>
- Viberg, Åke (2016), Polysemy in action: The Swedish verb *slå* 'hit, strike, beat' in a crosslinguistic perspective, in P. Juvonen & M. Koptjevskaja-Tamm (eds), *The Lexical Typology of Semantic Shifts*. [Cognitive Linguistics Research 58] Berlin/Boston: de Gruyter Mouton, 177-222.

Electronic sources

<https://intercorp.korpus.cz/?lang=en>
[ImagAct. http://imagact.it/imagact/query/dictionary.seam](http://imagact.it/imagact/query/dictionary.seam)

Additive questionnaire

Alina Russkikh, Petr Rosseyaykin
(HSE University, Lomonosov Moscow State University)

Keywords: additivity, scalarity, additive particles, typology, questionnaire

The basic function of additives, such as English *also* and German *auch*, is to trigger an inference (often regarded as a presupposition) that there is at least one true alternative proposition. Beyond their additive meaning, such markers are often used in other functions (König 1991, 2017; Gast & van der Auwera 2011; Forker 2016). A systematic description of uses of a given additive marker is an important typological task, requiring a questionnaire that factors in relevant semantic and syntactic parameters of variation. We present a typological questionnaire for describing additive markers based on uses discussed in typological literature and identified by us through our fieldwork studies of additive markers in several languages of Central Eurasia (mainly from Turkic and Nakh-Dagestani families).

The questionnaire (available at rb.gy/9qncah) includes over 80 stimuli covering 12 functions in which additives are attested: additivity, scalar additivity, coordination, comitative, contrast, pragmatic, concessive, compound numerals, exhaustivity/distributivity with numerals, quantifiers, distributivity with nominals/pronouns, indefinite pronouns.

The primary goal of this grouping is to structure the questionnaire rather than to make a theoretical claim about what counts as a distinct “function”. The purpose of the questionnaire is to capture morphosyntactic and semantic properties of additives, and the stimuli are designed to serve this purpose. For example, the acceptability of VP-coordination with additives can depend on whether conjoined phrases are complex (1). As for semantics, the availability of collective/distributive reading can depend on mono- vs polysyndetic marking of conjunction or even external factors, such as the number marking (2). These and many other parameters of variation are investigated in the part of the questionnaire dealing with coordination.

(1) Ossetic (Belyaev & Khomchenkova 2022: 41)

- a. *alan bad-ə d3r 3m3 xuəss-ə d3r.
Alan seat-PRS.3SG ADD and lie-PRS.3SG ADD
Intended: ‘Alan seats and lies’
- b. alan bad-g3 d3r k3nə 3m3 xuəss-g3 d3r k3nə.
Alan seat-CVB ADD does and lie-CVB ADD does
‘Alan seats and lies.’

(2) Terek Kumyk (own data)

- a. xabib-de kerim-ne seksen kila bol-a.
Habib-ADD Kerim-ADD 80 kilogram be-IPFV
‘Habib and Kerim weigh 80 kg (together).’
- b. xabib-de kerim-ne seksen kila-lar bol-a.
Habib-ADD Kerim-ADD 80 kilogram-PL be-IPFV
‘Habib and Kerim (each) weigh 80 kg.’

There are several methodological aspects of collecting data on additives using the elicitation method. First, since additive interpretation depends on the context, a left context is provided for each stimulus. Compare the stimuli examples for additivity (3) and scalar additivity (4).

(3) [John left.] Bill left too.

(4) [Everyone knows this.] Even children know this.

Second, in several functions, additives are *optional* and can be omitted without an obvious change in meaning, as in (5). Given this possibility, we suggest checking each example with and without the additive and controlling the interpretation in both cases.

- (5) Poshkart Chuvash (own data)
 pëdëm atɕa(=da)
 all child=ADD
 'All the children.'

Finally, a distinct challenge is determining if an additive is used independently or *with other obligatory elements*. For example, in Poshkart Chuvash (6), numeral constructions with exhaustive meaning require only the additive particle, while in Zilo Andi (7) both an additive and an emphatic particle =*gu* are obligatory. Consequently, we suggest controlling the use of additional markers and their semantic contributions.

- (6) Poshkart Chuvash (own data)
 igë atɕa=da.
 two child=ADD
 'Both children.'
- (7) Zilo Andi (own data)
 č'e*(-gu)=lo moč'i.
 two-EMPH=ADD child
 'Both children.'

The talk will explore in greater depth the variation parameters relevant to the functions of additives under consideration, as well as the challenges encountered in studying additive particles in general.

List of abbreviations

3SG — 3rd person, singular, ADD — additive, CVB — converb, FUT — future tense, EMPH — emphatic particle, IPFV — imperfective, PL — plural, PRS — present tense.

Acknowledgments

Petr Rossyaykin is supported by the RSF grant #25-18-68012.

References

- Belyaev, Oleg and Irina Khomchenkova, (2022), Noun phrase conjunction in Ossetic, *Rhema*, 4, Pp. 32–54.
- Forker, Diana (2016), Toward a typology for additive markers, *Lingua* 180, 69–100.
- Gast, Volker and Johan van der Auwera (2011), Scalar additive operators in the languages of Europe, *Language* 87(1), 2–54.
- König, Ekkehard (1991), *The meaning of focus particles: a comparative perspective*, London, New York: Routledge.
- König, Ekkehard (2017), Syntax and semantics of additive focus markers from a cross-linguistic perspective: A tentative assessment of the state of the art, in A.-M. De Cesare, and C. Andorno (eds), (2017), *Focus on Additivity: Adverbial modifiers in Romance, Germanic and Slavic languages*, Amsterdam: John Benjamins, 23–44.

Expanding FrameNet Frames for a Valency Lexicon of Croatian

Ana Ostroški Anić & Ivana Brač
(Institute for the Croatian Language)

Keywords: verbs of cognition, valency patterns, verb lexicon, FrameNet, Croatian

The role of verbs in a sentence has traditionally been viewed as central, which is why there is a long list of resources covering various aspects of verbs, starting from those focusing on their semantics, e.g., WordNet (Fellbaum 1998), FrameNet (Ruppenhofer et al. 2016); semantics and syntax, e.g., VerbNet (Kipper et al. 2006), PropBank (Palmer, Gildea & Kingsbury 2005); or semantics, syntax, and morphology, e.g., VALLEX (Kettnerová & Lopatková 2012) and Walenty (Przepiórkowski et al. 2014). Although each resource is developed within a different theoretical framework, linking them or providing mappings between different description poses an extra step in building lexical resources for efficient applications, both human-oriented and machine-directed ones. For instance, VALLEX enriched its description by adding information from FrameNet (Kettnerová & Lopatková 2012), while the Unified Verb Index integrates links from diverse NLP projects such as VerbNet, PropBank, FrameNet, OntoNotes (Hovy et al. 2006), and the SynSemClass Lexicon (Straková et al. 2023). Given that there is no similar resource for Croatian yet (except for WordNet, to a certain extent), a valency lexicon is being developed to integrate different approaches and data into a semantic and syntactic description of verbs.

This paper describes how the frame-semantic description of verbs is incorporated into the valency lexicon. First, an overview of the lexicon is given. Verbs are defined on several levels in the lexicon: 1. verb sense, associated with semantic classes from Levin (1993), VerbNet, and WordNet; 2. valency frames, related to verb senses, which include examples from corpora and an analysis of participants at the syntactic level (phrase types), morphological level (cases, prepositions, and conjunctions), and semantic level (semantic roles mainly taken from VerbNet).

Following the principles of Frame Semantics (Fillmore 1985; Fillmore, Johnson & Petruck 2003) and applying the FrameNet's methodology (Ruppenhofer et al. 2016), a semantic frame is identified for each verb sense, while for each participant, the appropriate frame element (FE) is determined. Using verbs of cognition as an example, this paper describes issues in applying the frames from the Berkeley FrameNet in the frame-semantic description of Croatian verbs in the valency lexicon. The argument structure of Croatian verbs such as *misлити* 'think', *pomisлити* 'think, have a thought', *razmišljati*_{IMPF}/*razmisлити*_{PERF} 'think, think over, ponder', *smisлити* 'think of, come up with', *zamisliti* 'imagine, envision', *promisliti* 'think through, reflect on', *izmisлити* 'make up, invent, fabricate', etc., is explained, along with their semantic description within appropriate semantic frames of FrameNet, e.g., Opinion, Regard or Coming_up_with. Corpus examples of sentences are annotated for FEs in order to compare the verbs' valence patterns and address these key research questions: 1. Are FrameNet frames applicable to the description of Croatian verbs? 2. Which frames should be expanded, either to include more FEs or additional lexical units? Examples of verb senses for which there is no adequate frame, e.g. for the sense of *think* 'to have concern' (*I must think first of my family.*) are additionally discussed.

Acknowledgments: This work has been supported by the Croatian Science Foundation under the project *Semantic-Syntactic Classification of Croatian Verbs* (SEM-TACTIC) (IP-2022-10-8074) and the European Union – NextGeneration EU (SEF and Verbion).

References

- Fellbaum, Christiane (ed.) (1998), *WordNet: An Electronic Lexical Database*, MIT Press.
- Fillmore, Charles. J. (1985), Frames and the semantics of understanding, *Quaderni di semantica: Rivista internazionale di semantica teorica e applicata* 6, 222–254.
- Fillmore, Charles. J.; Johnson, Christopher R. & Petruck, Miriam R. L. (2003), Background to Framenet, *International journal of lexicography* 16(3), 235–250.
- Hovy, Eduard; Marcus, Mitchell; Palmer, Martha; Ramshaw, Lance & Weischedel, Ralph (2006), OntoNotes: The 90% Solution, in *Proceedings of the Human Language Technology Conference of the NAACL, Companion Volume: Short Papers*, New York City: Association for Computational Linguistics, 57–60.
- Kettnerová, Václava; Lopatková, Markéta & Bejček, Eduard (2012), Mapping Semantic Information from FrameNet onto VALLEX, *The Prague Bulletin of Mathematical Linguistics* 97, 23–41.
- Kipper, Karin, Korhonen, Anna, Ryant, Neville & Palmer, Martha (2006), Extending VerbNet with Novel Verb Classes, in N. Calzolari et al. (eds), *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06)*, ELRA, 1027–1032.
- Levin, Beth (1993), *English Verb Classes and Alternations*, The University of Chicago Press.
- Palmer, Martha; Gildea, Daniel & Kingsbury, Paul (2005), The Proposition Bank: A Corpus Annotated with Semantic Roles, *Computational Linguistics Journal* 31(1), 71–106.
- Przepiórkowski, Adam; Hajnicz, Elżbieta; Patejuk, Agnieszka; Woliński, Marcin; Skwarski, Filip & Świdziński, Marek (2014), Walenty: Towards a comprehensive valence dictionary of Polish, in *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, ELRA, 2785–2792.
- Ruppenhofer, Josef et al. (2016), *FrameNet II: Extended Theory and Practice*. <https://framenet2.icsi.berkeley.edu/docs/r1.7/book.pdf>.
- Straková, Jana; Fučíková, Eva; Hajič, Jan & Urešová, Zdeňka (2023), Extending an Event-type Ontology: Adding Verbs and Classes using Fine-tuned LLMs Suggestions, in *Proceedings of the 17th Linguistic Annotation Workshop*, Association for Computational Linguistics, 85–95.

Who put semantics in my syntax? Finding lexical-semantic effects on morphosyntactic variation with elastic net regression

Anthe Sevenants, Freek Van de Velde & Dirk Speelman
(KU Leuven)

Keywords: regression analysis, semantics, lexical semantics, morphosyntax, variation

Whenever there are “alternate ways of saying the same thing” (Labov 1972) in a language, linguists take great interest. As such, several approaches have been developed to try and explain the choice between these “alternate ways” as accurately as possible. A first approach focusses on high-level, general predictors with a limited number of levels, sometimes pooled from more narrow categories. Theoretically, this approach aligns most with “probabilistic grammar”, the idea that variation reflects competing constraints on linguistic choices (Grafmiller et al. 2018). To discover these constraints, (mixed-effects) regression analysis (e.g. Gries 2015) is typically used.

Another approach focusses on frequency only, hinging on the associations of lexical items towards a specific morphosyntactic variant. This practice fits within the larger theoretical framework of Construction Grammar (CxG) with its lexicon-syntax continuum and multi-word units (Hoffmann and Trousdale 2013). To discover associations between constructions, collostructional analysis or derivatives (Stefanowitsch 2013) are often used.

The issue with the aforementioned two methodological lines is that they are biased for their own predictors types, with both types never being present in the same models. Theoretically, this means linguistic conclusions are centred around only one type of model, and never paint a full picture. To bring these two methodological lines together, we showcase elastic net regression (Friedman, Tibshirani, and Hastie 2010). Elastic net allows lexical items to appear on the same level as traditional, high-level predictors, enabling fuller models of variation. To demonstrate the methodology, we apply elastic net regression to 1,296,574 Dutch verbal cluster tokens from the SoNaR corpus (Oostdijk et al. 2008), analysing the choice between the so-called “red” and “green” word orders (shown in 1 and 2), a morphosyntactic alternance in Dutch subordinate clauses (Bloem 2021). Our elastic net model reproduces the higher-level effects known from previous research, but also shows morphosyntactic preferences (red, green) for different verbs. This proves that lexical effects can be at play in cases of morphosyntactic variation.

(1) **Red word order (AUX + PTCP)**

omdat	hij	heeft _{AUX}	geroepen _{PTCP}
because	he	has	shouted

“because he has shouted”

(2) **Green word order (PTCP + AUX)**

omdat	hij	geroepen _{PTCP}	heeft _{AUX}
because	he	shouted	has

“because he has shouted”

To gain insight into the motivations behind the lexical effects we see, we attempted to cluster the semantic vectors of the verbs occurring in the verbal clusters with PAM (Kaufman and Rousseeuw 1990). We found that certain clusters have statistically significant preferences for either word order, but any further interpretation remained difficult. A subsequent analysis with a lexical-semantic database ('Cornetto', Vossen et al. 2013) showed very weak influences of valency and semantic type. An interpretation of the lexical patterns thus remains difficult. Nonetheless, our study proves the feasibility and usefulness of a monolithic elastic net model, and shows that it is possible to investigate the role of the lexicon in morphosyntactic variation on a large scale. In our case, the AUC (Area Under Curve) improved from 0.775 without lexical predictors to 0.803 with lexical predictors. We advocate for inclusion of lexical items as full predictors in future models of morphosyntactic variation, as some of the variation left unexplained by high-level predictors could potentially be explained in lexical terms.

Acknowledgements

This work was supported by Fonds Wetenschappelijk Onderzoek, grant G059922N.

References

- Bloem, Jelke. 2021. *Processing Verb Clusters*. LOT International Series 586. Amsterdam: LOT. <https://doi.org/10.48273/LOT0586>.
- Friedman, Jerome, Robert Tibshirani, and Trevor Hastie. 2010. 'Regularization Paths for Generalized Linear Models via Coordinate Descent'. *Journal of Statistical Software* 33 (1): 1–22. <https://doi.org/10.18637/jss.v033.i01>.
- Grafmiller, Jason, Benedikt Szendrői, Melanie Röthlisberger, and Benedikt Heller. 2018. 'General Introduction: A Comparative Perspective on Probabilistic Variation in Grammar'. *Glossa: A Journal of General Linguistics* 3 (1). <https://doi.org/10.5334/gjgl.690>.
- Gries, Stefan Thomas. 2015. 'The Most Under-Used Statistical Method in Corpus Linguistics: Multi-Level (and Mixed-Effects) Models'. *Corpora* 10 (1): 95–125. <https://doi.org/10.3366/cor.2015.0068>.
- Hoffmann, Thomas, and Graeme Trousdale. 2013. 'Construction Grammar: Introduction'. In *The Oxford Handbook of Construction Grammar*, edited by Thomas Hoffmann and Graeme Trousdale, 1–9. Oxford: Oxford University Press.
- Kaufman, Leonard, and Peter Rousseeuw. 1990. 'Partitioning Around Medoids (Program PAM)'. In *Finding Groups in Data*, 68–125. John Wiley & Sons, Ltd. <https://doi.org/10.1002/9780470316801.ch2>.
- Labov, William. 1972. *Sociolinguistic Patterns*. Conduct and Communication. Philadelphia (Pa.): University of Pennsylvania Press.
- Oostdijk, Nelleke, Martin Reynaert, Paola Monachesi, Gertjan van Noord, Roeland Ordelman, Ineke Schuurman, and Vincent Vandeghinste. 2008. 'From D-COI to SoNaR: A Reference Corpus for Dutch', 9.
- Stefanowitsch, Anatol. 2013. 'Collostructional Analysis'. In *The Oxford Handbook of Construction Grammar*, edited by Thomas Hoffmann and Graeme Trousdale, 0. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780195396683.013.0016>.
- Vossen, Piek, Isa Maks, Roxane Segers, Hennie van der Vliet, Marie-Francine Moens, Katja Hofmann, Erik Tjong Kim Sang, and Maarten de Rijke. 2013. 'Cornetto: A Combinatorial Lexical Semantic Database for Dutch'. In *Essential Speech and Language Technology for Dutch*, edited by Peter

Spyns and Jan Odijk, 165–84. *Theory and Applications of Natural Language Processing*. Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-30910-6_10.

Signs for present day linguistics: beyond Peirce

Arie Verhagen

Leiden University Centre for Linguistics

Keywords: Sign theory, Peirce, Symbols/icons/indexes/signals/symptoms, Evolution

The ultimate goal of scientific inquiry is explanation: to answer the question *why* things are the way they are (Tinbergen 1963; Chomsky 2022). Explaining something involves reducing it to more elementary phenomena and showing how it emerges from interactions between these elementary phenomena (Verhagen 2021). For language as a bio-cultural phenomenon, this explanatory goal is hampered by the fact that scholars differ greatly in the ways they conceive of crucial characteristics of the object of study: of what it is that has evolved –culturally or biologically– and has to be explained (cf. Jackendoff 2010). This is true for grammar, but also for the basic concept of “linguistic sign”. Despite the consensus that linguistic signs are characteristically “symbols”, the answer to the question what makes symbolic signs special, and how they may have evolved out of forms of non-symbolic communication, varies from one theorist to the next.

Around the turn of the 19th and 20th century, C.S. Peirce (1991, 2020) developed a theory of signs that has since been repeatedly invoked in discussions on the nature of signs, in particular his way of distinguishing symbols from indexes and icons. Deacon (1997) reconstructs Peirce’s classification in terms of cognitive complexity, and analyses the emergence of symbols in terms of the evolution of learning and cognitive capacities. But Hurford (1998) argued that Deacon was not using the term *symbol* in the same way as linguists and philosophers of language. Still, Planer & Sterelny (2021) recently criticized the Peircean framework as irrelevant to evolutionary explanations based on their reading of Deacon, and proposed to replace it by a “sender-receiver framework” in terms of ‘signaling games’ (Skyrms 2010). However, Planer & Sterelny do not provide a positive characterization of symbols as distinct from, or subsumed under, signals, and signals do not include all of Peirce’s indexes, while these are arguably relevant in explaining language evolution, especially pointing (cf. Tomasello 2008). Linguistically motivated approaches like Clark (1996) and Keller (1998) also differ in their application of Peirce’s typology; for example, Keller’s classification does not include pointing, while Clark’s ignores the role of some other indexes.

Using an extension of Keller’s (1998) framework, I argue that for a comprehensive account, we must replace the classification of signs into a hierarchically ordered set of super- and subordinate categories, with one in terms of distinctive features (some of which exhibit mutual constraints), along the following lines:

action	causal knowledge	simulation	conventional	communicative situation	shared personal history	(sign ‘type’)
–	+	–	–	–	–	symptom
+	–	–	–	+	–	indicator
+	–	–	+	+	–	deictic
+	–	+	–	–	–	icon
+	–	–	+	–	+	proper name
+	–	–	+	–	–	noun, verb, ...
+	–	+	+	–	–	(metaphor)
+	+	–	+	–	–	(metonymy)

(For the ‘formal’, observable dimension of signs, cf. Holler & Levinson (2019), Verhagen (2025)). What needs to be explained, is the evolution of certain bundles of features of signs, rather than a monolithic sign type “symbol”.

References

- Chomsky, Noam (2022), Genuine explanation and the strong minimalist thesis, *Cognitive Semantics* 8, 347-365.
- Clark, Herbert (1996), *Using Language*, Cambridge: Cambridge University Press.
- Deacon, Terrence W. (1997), *The Symbolic Species. The Co-evolution of Language and the Brain*, New York: W.W. Norton & Company.
- Holler, Judith & Stephen C. Levinson (2019), Multimodal language processing in human communication, *Trends in Cognitive Sciences* 23, 639-652.
- Hurford, James R. (1998), Review of Deacon (1997), *Times Literary Supplement*, October 23, 34.
- Jackendoff, Ray (2010), Your theory of language evolution depends on your theory of language, in Richard K. Larson, Viviane M. Déprez, Hiroko Yamakido (eds), (2010), *The Evolution of Human Language: Bilingual Perspectives*, Cambridge: Cambridge University Press, 63-72.
- Keller, Rudi (1998), *A Theory of Linguistic Signs*, Oxford, etc.: Oxford University Press.
- Peirce, Charles S. (1991), *Peirce on Signs. Writings on Semiotic by Charles Sanders Peirce*, edited by James Hoopes, Chapel Hill/London: University of North Carolina Press.
- Peirce, Charles S. (2020), *Selected Writings on Semiotics 1894-1912*, edited by Francesco Bellucci, Berlin/New York: De Gruyter Mouton.
- Planer, Ronald J. & Kim Sterelny (2021), *From Signal to Symbol. The Evolution of Language*, Cambridge, MA/London: The MIT Press.
- Skyrms, Brian (2010), *Signals. Evolution, Learning, & Information*, Oxford: Oxford University Press.
- Tinbergen, Niko (1963), On aims and methods of ethology, *Zeitschrift für Tierpsychologie* 20, 410-433.
- Tomasello, Michael (2008), *Origins of Human Communication*, Cambridge, MA: MIT Press.
- Verhagen, Arie (2021), *Ten Lectures on Cognitive Evolutionary Linguistics*, Leiden/Boston: Brill.
- Verhagen, Arie (2025), Construction grammar, multimodal communication, and design features of language: Preliminaries to a consistent research program, in Kiki Nikiforidou and Mirjam Fried (eds), (2025), *Multimodal Communication from a Construction Grammar Perspective*, Amsterdam/Philadelphia: John Benjamins Publishing Company, 26-37.

Alternations between underspecified functions as bridging contexts for semantic and syntactic reanalysis

Björn Wiemer
(Mainz University)

Keywords: reanalysis, Slavic syntax, clause combining, directive-optative markers, oscillation

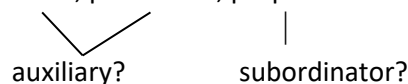
All Slavic languages employ uninflected markers indicating optative or directive illocutionary force (henceforth: 'DIR-units'): Czech *ať*, Russian *pust'*/*puskaj*, Polish *niech(aj)* + cognates of the latter in the remaining languages (ex. 1-2, 4). While the directive-optative function represents the initial stage still belonging to the main usage types in all Slavic languages, the functions of DIR-units have considerably expanded sometimes comprising even concessive/conditional or hearsay uses (Wiemer 2023a). With some of these functional changes the original combinatorial restriction to the indicative present of the predicate disappears (see 3, concessive). Concomitantly, DIR-units may be variably described as particles, auxiliaries or subordinators, particularly as complementizers (ex. 2); e.g., Kramer (1986), Xrakovskij/Volodin (1986), Topolińska (2008), Uhlik (2018), Dobrushina (2019), Sonnenhauser (2021). However, these different syntactic functions are usually assigned without sufficient methodological and/or empirical grounding (Wiemer 2023b, 2023c).

- (1) Polish *Ja tam wierzę swoim metodom, a komputerami **niech** się zajmują^{IPFV.PRS} geniusze.*
'I believe in my methods; as for computers, **may** geniuses deal with them.'
→ optative, non-curative?; auxiliary?

- (2) Serbian *Majka mi reče, **neka** mu skočim^{PFV.PRS} na leđa.*
'Mother told me to jump on his back.' (lit. '...**may** I jump...')
→ directive, optative?; complementizer?

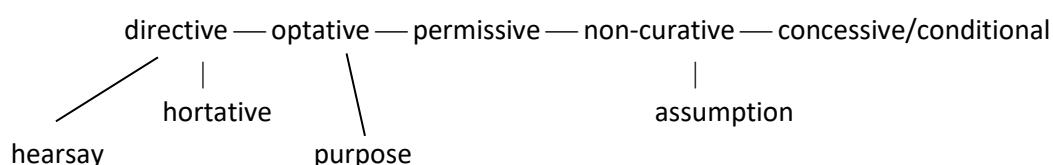
- (3) Russian ***Pust'** nedostavalo^{IPFV.PST} komforta na bortu, no mačta byla, vesla byli, parus xot' kakoj-to byl.*
'**Although** there was a lack of comfort on board, there was a mast, there were oars, there was at least some kind of sail.'

- (4) Czech *Nechal jsem ji, **ať** vybere^{PFV.PRS} a objedná^{PFV.PRS} nám oběma.*
'I let her (so that) **may** she choose and order for both of us.'
→ directive, optative, permissive, purpose?



More importantly, particular functions of DIR-units discussed in the literature (see References) often cannot be disentangled in discourse (ex. 1, 2, 4); the same applies to their variable syntactic status. That is, often DIR-units oscillate between semantic and/or syntactic values that are otherwise treated as distinct; see a tentative connectivity map based mainly on Holvoet/Konickaja (2011), Dobrushina (2008, 2019).

Function range of Slavic DIR-units (tentative)



Oscillation signals underspecified functions that can co-exist within the same utterance without requiring “resolution” (Mendoza/Sonnenhauser 2023). On the basis of samples from synchronic and diachronic corpora (see References), the talk will show that, on discourse level, oscillation differs from ambiguity (‘either—or’). It will specify the bridging contexts which favor either oscillation or ambiguity, and address the following issues:

1. Can we draw a diachronic connectivity map (Georgakopoulos/Polis 2018) of DIR-units that integrates “areas” of oscillation? This implies revising arguments for different functions postulated in the literature (see References).
2. Which types of correspondences between semantic and syntactic oscillation, or of ambiguity, are recurrent?
3. How time-stable are different types of oscillation, or of ambiguity?

Preliminary findings suggest that

- (i) complementizer-like use occurs since the earliest stages, although through all stages it is rare.
- (ii) generally, persistence does not depend on frequency.
- (iii) The path taken by Slovene *naj* differs from all other DIR-units.

References

- Dobrushina [Dobrušina], Nina R. (2008), *Naklonenie i diskursivnyj režim teksta: na primere upotreblenij časticy pust’* [Mood and discourse register: the example of the uses of the particle *pust’*], in V. A. Plungjan, V. Ju. Gusev, and A. Ju. Urmančieva (eds), (2008), *Issledovanija po teorii grammatiki 4: Grammatičeskie kategorii v diskurse* [Investigations in the theory of grammar 4: Grammatical categories in discourse], Moskva: Gnozis, 167–193.
- Dobrushina [Dobrušina], Nina (2019), Status konstrukcij s časticami *pust’* i *puskaj* v ruskom jazyke [The status of constructions with the particles *pust’* and *puskaj*], *Russian Linguistics* 43, 1–17.
- Georgakopoulos, Thanasis and Stéphane Polis (2018), The semantic map model: State of the art and future avenues for linguistic research. *Lang Linguist Compass* 12.
<https://doi.org/10.1111/lnc3.12270>
- Holvoet, Axel & Jelena Konickaja, 2011, Interpretive deontics: A definition and a semantic map based mainly on Slavonic and Baltic data, *Acta Linguistica Hafniensia* 43(1), 1–20.
- Kramer, Christina Elizabeth (1986), *Analytic Modality in Macedonian*, München: Sagner.
- Mendoza, Imke and Barbara Sonnenhauser (2023), Oscillation and oscillating structures in syntax. *Zeitschrift für Slawistik* 68(2), 261–281.
- Sonnenhauser, Barbara (2021), Slovene *naj*: an (emerging) clausal complementizer?, in B. Wiemer, and B. Sonnenhauser (eds), (2021), *Clausal Complementation in South Slavic*, Berlin / Boston: De Gruyter Mouton, 443–475.

- Topolińska, Zuzanna (2008), 'Neka'-konstrukcije i nivniot status vo slovenskite glagolski sistemi ['Neka'-constructions and their status in the verb systems of Slavic languages], in Z. Topolińska, Z. Polski do Macedonii. *Studia językoznawcze, tom 1: Problemy predykcji* [From Poland to Macedonia. Linguistic studies, vol. 1: Problems of predication], Kraków: Lexis, 217–223.
- Uhlik, Mladen (2018), O *naj* in *pust'* v slovensko-ruski sopostavitvi [About *naj* and *pust'* in Slovene-Russian comparison], *Slavistična revija* 66(4), 403–419.
- Wiemer, Björn (2023a), Directive-optative markers in Slavic: observations on their persistence and change, *Linguistica Brunensia* 71(1), 5–45.
- Wiemer, Björn (2023b), Between analytical mood and clause-initial particles – on the diagnostics of subordination for (emergent) complementizers, *Zeitschrift für Slawistik* 68(2), 187–260.
- Wiemer, Björn (2023c), Clause-initial connectives, bound and unbound: Indicators of mood, of subordination, or of something more fundamental? *Slavia Meridionalis* 23 (Special issue: *Comparative and typological approaches to Slavic languages*. Ed. by Jakub Banasiak, Julia Mazurkiewicz-Sułkowska, Bożena Rozwadowska, Dorota Klimek-Jankowska). DOI: 10.11649/sm.3194
- Xrakovskij, Viktor S., and Aleksandr P. Volodin (1992), *Semantika i tipologija imperativa (Russkij imperativ)* [The semantics and typology of the imperative (The Russian imperative)], Leningrad: Nauka.

Corpora

Bulgarian

EU DGT-UD: <https://www.sketchengine.eu/nosketch-engine/>

Czech

Czech National Corpus: <https://www.korpus.cz/> (syn 2020; diachronní)

Croatian

ENGRI; Croatian news portals <https://www.sketchengine.eu/nosketch-engine/>

Macedonian

CLASSLAWiki-mk <https://www.sketchengine.eu/nosketch-engine/>

Polish

Electronic Respository of Greater Poland Oaths: <https://rotha.ehum.psnc.pl/>

Korpus staropolski: <https://ijp.pan.pl/publikacje-i-materialy/zasoby/korpus-tekstow-staropolskich/>

The Polish Diachronic Research Corpus **PoIdi**:

<http://westslang.sprachen.hu-berlin.de:8080/annis/poldi>

The Electronic Corpus of 17th- and 18th-century Polish Texts (up to 1772):

https://korba.edu.pl/query_corpus/

Polish National Corpus (NKJP): <https://nkjp.pl/>

Russian

Russian National Corpus: <https://ruscorpora.ru/>

Slovak

hist-6.0: Historical corpus of Slovak: <https://bonito.korpus.sk/>

HSSJa: *Historický slovník slovenského jazyka* z r. 1991–2008: <https://www.juls.savba.sk/hssj.html>

Slovene

IMP language resources for historical Slovene: <https://nl.ijs.si/imp/>

Gigafida: <https://www.clarin.si/repository/xmlui/handle/11356/1035>

Ukrainian

GRAC v.16: <https://uacorporus.org/Kyiv/en>

Upper Sorbian

hotko (via CzNC: <https://www.korpus.cz/>)

Competing strategies for Median path encoding in two satellite-framed languages

Christine Lamarre & Anetta Kopecka
(Inalco and CRLAO ; Université Lumière Lyon 2 and DDL)

We focus here on the encoding of the Median path observed in Polish, a Slavic language, and in Standard Chinese, a Sinitic language, two unrelated languages making wide use of Satellite-framed constructions. Both languages have several distinct sets of spatial grams encoding Path: verb satellites (see Imbert et al. 2011), realized as verb prefixes in Polish and as postverbal directionals in Chinese, which belong to the adverbial type of encoding (Wälchli 2001), and prepositions (Chinese and Polish), case markers (Polish), and relational nouns (Chinese), that belong to the adnominal type and require a locative Ground NP. They all contribute to the encoding of Median Path, defined by Luraghi (2003: 22) as a state of affairs in which the figure “moves along a trajectory through, or across a landmark”. In English, one of the best documented satellite-framed languages, the Median path is expressed through a range of prepositions such as ‘past’, ‘across’, ‘through’, ‘via’, and ‘over’, each conveying different aspects of Median meaning. (Bennett 1972: 78-82).

Our study investigates constructions expressing translational spontaneous motion that comprise a selected sample of manner-of-motion verbs (e.g., ‘walk’, ‘run’, ‘fly’, ‘squeeze’, ‘swim’, etc.), in contemporary literary texts. Our data show that unlike English, Chinese and Polish typically express these different Median meanings through the combination of the median satellite (*prze-* in Polish and *guò* in Chinese) on one hand, and of prepositions (Polish and Chinese), morphological case (Polish), and postnominal relational nouns (Chinese), on the other hand, as illustrated in (1-4).

POL: Prefix *prze-* ‘across, through’ with preposition *przez* ‘across, through’ (NP_{ACC})

- (1) *Dziecko prze-biegło przez hall, (...)*
child.NOM.SG.N SAT_{MED}-run.3SG.N.PST through hall.ACC.SG.M
‘The child passed (running) through the hall’

POL: Prefix *prze-* ‘across, through’ with a locative preposition (here *nad* ‘above’) + NP_{INSTR}

- (2) *Bardzo nisko, tuż nad dachami prze-latywał dwupłatowy samolot*
very low just above roof.INSTR.PL.M SAT_{MED}-fly.3SG.M.PST biplane.NOM.SG.M plane.NOM.SG.M
‘A biplane flew very low, just above the rooftops.’

ZH: Directional *-guo* ‘across, through’ with preposition *cóng* ‘from’ and relational noun *=shang* ‘upside’

- (3) *Jípǔchē cóng jiē=shang shǐ-guò, ...*
Jeep from street=on drive-SAT_{MED}
‘The jeep drove down (/through) the street, ...’

ZH: Directional *-guo* ‘across, through’ with preposition *cóng* ‘from’ and relation noun *biān* ‘side’

- (4) *Tā-mén cóng tā shēn-biān zǒu-guò, méi rén kàn tā yī yǎn.*
3-PL from 3SG body-side walk-SAT_{MED} not.have person look 3SG one eye
‘They walked by [her side] without looking at her.’

Our observations support Luraghi’s remark (2003:22) on the comparatively greater complexity of Median Path (labelled “Path” in her study), “mirrored in the fact that a separate morphological case for Path, so-called perlativ, is less frequently attested in the world’s languages” than separate cases for Location,

Source or Goal (see also Creissels 2008 on spatial cases, and Yamaguchi 2004: 131-138 for a similar claim for prepositions). In Chinese and Polish, this complexity is reflected in the interplay between distinct linguistic devices (adverbial and adnominal) and the types of Median Grounds referred to. Svorou (1993:28) thus remarks that “motions such as ALONG, AROUND, ACROSS and THROUGH are [...] prototypically understood by reference to elongated objects [...], spherical objects, concave and transversal objects, or, finally, objects with prominent parallel boundaries, which we may cross.” By investigating the interplay between different sets of linguistic devices used to express *Median Path* and its sub-types in Polish and Standard Chinese, this study contributes to documenting intratypological variation within satellite-framed languages (cf. Filipović & Ibarretxe-Antuñano 2015; Filipović 2013). It also illustrates the point made in Sinha and Kuteva (1995) that languages vary on “the extent to which the distributed semantics of spatial relations receive explicit information”: like Japanese and Bulgarian, both Polish and Chinese overtly express spatial semantics through several form classes.

Keywords: median path, ; intratypological variation; verb satellite; preposition; relational noun

Acknowledgments: This research has been supported by the Salta project (Spatial Asymmetries across Languages: A Typological Approach (ANR: ANR-20-CE27-0015).

References

- Creissels, Denis (2008), Spatial cases, in A. Malchukov and A. Spencer (eds.), *The Oxford Handbook of Case*, Oxford: Oxford University Press, 609-625.
- Bennett, David (1972), Some observations concerning the Locative-Directional distinction, *Semiotica* 5(1), 58-88.
- Filipović, Luna (2013), Typology as a continuum: Intratypological evidence from English and Serbo-Croatian, in J. Goschler and A. Stefanowitsch (eds), *Variation and Change in the Encoding of Motion Events*, Amsterdam/Philadelphia: John Benjamins, 17–38.
- Filipović, Luna and Ibarretxe-Antuñano, Iraide (2015), Motion, in E. Dąbrowska and D. Divjak (eds), *Handbook of Cognitive Linguistics*, Berlin/Boston: Walter de Gruyter, 527-546.
- Imbert, Caroline, Grinevald, Colette and Sörös, Anna (2011), Pour une catégorie de "satellite" de Trajectoire dans une approche fonctionnelle-typologique, *Faits de Langues* 38(2), 99-116.
- Luraghi, Silvia (2003), *On the meaning of prepositions and cases: The expression of semantic roles in Ancient Greek*, Amsterdam/Philadelphia: John Benjamins.
- Przybylska, Renata (2006), *Schematy wyobrażeń i semantyka polskich prefiksów czasownikowych do-, od-, prze-, roz-, u-* [Image schemas and the semantics of Polish verb prefixes *do-, od-, prze-, roz-, u-*], Kraków: Universitas.
- Sinha, Chris, and Kuteva, Tania (1995), Distributed spatial semantics, *Nordic Journal of Linguistics* 18, 167-199.
- Svorou, Soteria (1993), *The Grammar of Space*, Amsterdam/Philadelphia: John Benjamins.
- Wälchli, Bernhard (2001), A typology of displacement (with special reference to Latvian), *STUF - Language Typology and Universals* 54(3), 298-323.
- Wei, Pei-Chuan (2013), “V-*guo-lai/qu*” de lìshǐ fāzhǎn [The historical development of “V-*guo-lai/qu*”], *Bulletin of Chinese Linguistics* 7(2), 1–34.
- Yamaguchi, Kazuyuki (2004), *Typological, historical, and functional Study of adpositions in the languages of the World*, University of New Mexico PhD dissertation.

Proving that mirativity exists: A reassessment of putative miratives in Tibetan and Ladakhi

Eric M  lac
Universit   Paul Val  ry - Montpellier 3

This paper examines the evidence that led to establish mirativity as a crosslinguistic notion (Aikhenvald 2012, Mexas 2016 *inter alia*). All languages can refer to ‘unexpected information’, so mirativity as a semantic domain is not denied. However, the existence of grammatical markers dedicated to this domain is disputed (Hill 2012, DeLancey 2012, Peterson 2017).

I will reinvestigate the semantics of two markers identified as ‘mirative’ in previous studies: ‘*dug*’ in Lhasa Tibetan and *tshug* in Leh Ladakhi (DeLancey 1997, Zeisler 2018). I argue that ‘*dug*’ is not inherently mirative while *tshug* is a definite example of a fully grammatical mirative marker. I adopt a two-step methodology (i) an in-depth inductive assessment from large corpora of authentic conversations (TSC, 2010-2011; Leh Corpus, 2022-2023), and (ii) thorough semantic testing by elaborating a felicity assessment questionnaire of sentences containing the targeted morphemes in mirative and non-mirative contexts.

The annotation of all the instances of ‘*dug*’ in the TSC indicates that the marker only appears in clear contexts of unexpected information approximately 3% of the time (n=20/676). The postulated mirative interpretation of ‘*dug*’ seems to result from its contrast with morphemes in the same paradigm that are only felicitous with assimilated information, namely egophorics and factuais. However, ‘*dug*’ can be used for unsurprising, familiar information, as in (1):

- (1) *nga tshor* *lha sa-la yar* *stabs bde kha lag* *za-sa’i* *za khang*
1:PL:DAT Lhasa-LOC fast_food eat-LOC:GEN restaurant
mang po *cig* ‘*dug-ga*
many DM POSS.DPERC-PHAT
‘We in Lhasa have many fast food restaurants.’ (TSC)

Consequently, the very marker which established mirativity as a research field was indeed misanalysed from the start (see also Hill 2012, Zeisler 2018).

Ladakhi *tshug* is presented as a marker of surprise in Koshal (1979) and Zeisler (2018). Contrary to Tibetan ‘*dug*’, this marker is not evidential, as it is compatible with all types of access to information. Examples from the Leh Corpus and felicity tests also show that it does not encode modality, emphasis, exclamation, or the time of information acquisition. All evidence points to a mirative function of *tshug*, provided that mirativity is clearly defined. The marker *tshug* does not encode ‘surprise’ or a ‘sudden revelation’ but ‘mis-expected information’, as in (2):

- (2) *a chaa* *sgog pa* *yin tshug* *ngas* *ra bar* *yin grag* *bsam*
INJ garlic COP.MIR 1SG:ERG rubber COP:NONVIS think:PFV
‘Oh, it’s garlic! I thought it was a rubber...’ (Leh Corpus)

Because the mirative meaning of *tshug* passes the cancellation test, it is inherent and not implied, as shown in (3):

- (3) ‘*i bo* *so sman* *yin tshug*
DEM toothpaste COP:MIR
‘This is toothpaste!’ [the speaker expected it to be lip balm] (Leh Corpus)
?? *snga’a* *rgyus* *yod-pin*
before:DAT knowledge EXIST.EGO-PAST
‘...I knew it.’

My study concludes that morphemes dedicated to mirativity may not be as common as previously thought, but Ladakhi does possess a grammatical marker which can only be described as mirative. Only

with extensive corpus data and precise semantic tests may mirative markers be authenticated, and the controversy over the notion be settled.

Abbreviations

1 first-person, COP copula, DAT dative, DEM demonstrative, DM discourse marker, DPERC direct perception evidential, EGO egophoric, ERG ergative, EXIST existential verb, GEN genitive, INJ interjection, LOC locative, MIR mirative, NONVIS non-visual evidential, PAST past tense, PFV perfective aspect, PHAT phatic, PL plural, POSS possessive verb, SG singular, TSC Tibetan Student corpus.

References

- Aikhenvald, Alexandra Y. 2012. The essence of mirativity. *Linguistic typology* 16(3). 435–485.
- DeLancey, Scott. 1997. Mirativity: The grammatical marking of unexpected information. *Linguistic typology* 1. 33–52.
- DeLancey, Scott. 2012. Still mirative after all these years. *Linguistic Typology* 16(3). 529–564.
- Hill, Nathan W. 2012. “Mirativity” does not exist: ḥdug in “Lhasa” Tibetan and other suspects. *Linguistic Typology* 16(3). 389–433.
- Koshal, Sanyukta. 1979. *Ladakhi grammar*. Delhi: Motilal Banarsidass.
- Mexas, Haris. 2016. *Mirativity as realization marking: A cross-linguistic study*. University of Leiden MA thesis.
- Olbertz, Hella. 2009. Mirativity and exclamatives in functional discourse grammar: Evidence from Spanish. *The London papers I, Special Issue of Web Papers in Functional Grammar* 82. 66–82.
- Peterson, Tyler. 2017. Problematizing mirativity. *Review of Cognitive Linguistics* 15(2). 312–342.
- Zeisler, Bettina. 2018. Don’t believe in a paradigm that you haven’t manipulated yourself! Evidentiality, speaker attitude, and admirativity in Ladakhi. *Himalayan Linguistics* 17(1). 67–130.

What can be in the scope of subjectivity marking?

Jan Nuyts

University of Antwerp

Keywords: subjectivity, modality, evidentiality, TAM, scalarity

This paper deals with the question of what can be in the scope of subjectivity marking in the sense of Nuyts (2001, 2012, 2025). In this concept subjectivity involves indicating that the speaker is expressing a strictly personal view – as opposed to intersubjectivity, indicating that the view is shared by others. Research on the ‘attitudinal’ categories, including deontic and epistemic modality and inferentiality, has shown that the dimension is often expressed in combination with these categories, and is even partly coded structurally on their markers (e.g. subjective *it seems probable (to me) that XYZ* vs intersubjective *it is probable that XYZ*).

Intuitively, subjectivity marking is not confined to these attitudinal categories, however, especially when it is coded by independent lexical forms – cf. e.g. *that sounds like an excellent proposal to me*, with subjectivity marker *to me* concerning a kind of quality assessment. But its range is not unlimited, as is demonstrated by the infelicity of combinations such as ??*it has 30 degrees centigrade to me* or ??*in my view he is in Paris*.

What semantic categories can actually be subject to the dimension has not been investigated yet, though. This paper makes a first attempt to explore this matter, by means of a corpus-based case study of one subjectivity marker in Dutch, viz. the mental-state-like verb *vinden* ‘find’, as in *ik vind dat wel leuk* (literally) ‘I find that rather pleasant’.

The analysis is based on a sample of 250 subjectivity marking instances of the verb, randomly selected from the *Corpus Gesproken Nederlands* (2006) (subjectivity marking is most common in spoken language). The sample is confined to 1SG present tense *ik vind* ‘I find’, as the verb form most likely to expresses subjectivity (occurrences of the form with other meanings were excluded from the sample).

The data show that the verb, as expected, significantly often concerns expressions involving the attitudinal categories. Yet it also commonly combines with a range of other semantic categories – among others markings of quality (*strong, chaotic, boring*, etc.), of efficiency or feasibility (such as *useful, optimal, complicated*), and of quantity in a wide sense (e.g. *sufficient, much/little*, but also *big/small, early/late*). Very many of these in their context imply an attitudinal stance of (dis)liking or (dis)approving though (e.g. the speaker states that something is chaotic and thereby contextually indicates/suggests that (s)he disapproves of it). In any case, the subjectivity marker appears to exclusively affect scalar categories, i.e. ones involving values plotted on a scale, which moreover allow or require speaker deliberation (whence the infelicity of the centigrades example above: temperature is a scale, too, yet not one involving deliberation).

These observations offer food for thought on the relation of subjectivity to the semantic system of categories ‘qualifying’ states of affairs (Van Valin & LaPolla 1998, Hengeveld & Mackenzie 2008, Nuyts

2025), on the position of the categories affected by subjectivity marking in this system, including, prominently, of the attitudinal ones relative to the others, and on the role of scalarity in all of this.

References

Corpus gesproken Nederlands (2006), Version 2.0, Leiden: TST-Centrale INL.

Hengeveld, Kees and Lachlan Mackenzie (2008), *Functional discourse grammar*, Oxford: Oxford University Press.

Nuyts, Jan (2001), Subjectivity as an evidential dimension in epistemic modal expressions, *Journal of Pragmatics* 33, 383-400.

Nuyts, Jan (2012), Notions of (inter)subjectivity, *English Text Construction* 5, 53-76.

Nuyts, Jan (2025), *Modality in mind*, Cambridge: Cambridge University Press.

Valin, Robert D. Van and Randy J. LaPolla (1997), *Syntax*, Cambridge: Cambridge University Press.

The semantic-pragmatic differences between discourse-secondary *I don't think* and *I think...not*

Jiqiang Lu & Kristin Davidse
KU Leuven (University of Leuven)

jiqiang.lu@kuleuven.be kristin.davidse@kuleuven.be

Keywords: discourse-secondary complement-taking-predicate clause; Neg-raising; *I think ... not*; *I don't think*; semantic-pragmatic effects

This paper seeks to pinpoint the semantic-pragmatic differences between grammaticalised *I don't think* and *I think...not*. We address three problems previous research has tended to suffer from. Firstly, lexical uses like (1), which entail that the thought in the complement clause was not formed, have been either ignored (Thompson and Mulac 1991a, 1991b) or cases like (2), which does not mean 'I've not formed the idea it's debatable', but 'in my view, it is NOT debatable' have been wrongly analysed as lexical.

(1) I've never thought Becks wouldn't sign a new deal (WB)

(2) B. That's surely debatable

A: *I d\on't think* it's deb/atable, n\o. (Conversational Corpus, quoted Simon-Vandenberghe 1998: 319)

With Boye and Harder (2007, 2012) we operationalise grammaticalised clauses as discourse-secondary clauses, which, unlike lexical uses, allow NEG-raising, i.e., raising negation of the proposition into the matrix, e.g., (2).

Secondly, we reject the equation of discourse-secondary with epistemic status, restricting the latter to the *degree* of likelihood of (non-)occurrence (Davidse et al. 2022), e.g., (3), which conveys considerable likelihood that she is not married but allows the possibility that she is.

(3) *I don't think* she's married yet. (WB)

We also recognise opinative uses, e.g., (2), which mark the proposition as "strictly personal opinion" (Janssens and Nuyts 2021: 231), and uses with interactional functions like refusing an offer, e.g., (4).

(4) A: is there anything else that...you'd like to clarify?

B: No, *I don't think* so. (WB)

Thirdly, we reject the assumption that the raised variant (typically) has weaker rhetorical force (Horn 1978; Bublitz 1992). Rather, emphatic markers, e.g., *d\ont ...n\o* in (2), and rhetorical upscaling of discourse-secondary clauses, e.g., *I think* abandoned for *I don't think...in* (5), suggest that these raised variants *strongly disalign* with preceding (implied) positive propositions.

(5) A: Have you had any contact with the outreach workers here?

B: No. I don't want...I think...*I don't think* there's anything they can do for me (WB)

Conversely, *mitigating* clues, e.g., hesitation markers and (*not*) *particularly* in B's complex sentence in (6), suggest the down-pedalled force of *I don't think*.

(6) A: Mm. Erm <pause> only snag is that erm it stops your antibiotics from working

<B: Yeah. C: Oh. B: It's not a good idea>

A: but *I don't think* it has any particularly harmful effects (WB)

We ground our own account in a qualitative-quantitative corpus study of the two variants in the British spoken subcorpus of WordbanksOnline (WB), which contains 8,012 and 870 potentially relevant hits of *I don't think* and *I think not*, from which we removed lexical and other irrelevant contexts. We analysed random samples of 152 and 120 contexts respectively in terms of the following parameters: (i) lexical -- discourse-secondary; (ii) epistemic, opinative, interactional; (iii) initial, medial, final; (iv) disalignment, alignment, no (dis)alignment with preceding propositions; (v) strong -- weak rhetorical force, (vi) contextual clues. Based on obtained results, we draw up quantitatively-based profiles, which reveal, amongst others, that *I don't think* has larger proportions of opinative and rhetorically stronger uses than *I think not*.

References

- Boye, Kasper & Peter Harder, (2007), Complement-taking predicates: Usage and linguistic structure, *Studies in Language, International Journal sponsored by the Foundation "Foundations of Language"* 31(3), 569–606.
- Boye, Kasper & Peter Harder, (2012), A usage-based theory of grammatical status and grammaticalization, *Language*, 1–44.
- Bublitz, Wolfram, (1992), Transferred negation and modality, *Journal of Pragmatics* 18, 551–577.
- Davidse, Kristin, An Van linden & Lieselotte Brems, (2022), A semiotic approach to grammaticalization: Modelling representational and interpersonal modality expressed by verbonominal patterns, *Language Sciences* 91: 101473
- Horn, Laurence R., (1978), Remarks on neg-raising, in Peter Cole (ed.), *Syntax and semantics, Volume 9: Pragmatics*, New York: Academic Press, 129–220.
- Janssens, Karolien & Jan Nuyts, (2021), On the origins of the epistemic, evidential, and subjectivity meanings in the mental state predicates: The case of Dutch, *Jezikoslovlje* 22(2), 227–250.
- Simon-Vandenberghe, Anne-Marie, (1998), The modal metaphor I don't think: System and text, in Johan van der Auwera, Frank Durieux, and Ludo Lejeune (eds.), *English as a human language: To honour Louis Goossens* (LINCOM Studies in Germanic Linguistics 4), Miinchen: Lincom Europa, 312–324.

Thompson, Sandra A. & Anthony Mulac, (1991a), A quantitative perspective on the grammaticization of epistemic parentheticals in English, in Elizabeth Traugott and Bernd Heine (eds.), *Grammaticalization II*, Amsterdam: John Benjamins, 313–339.

Thompson, Sandra A. & Anthony Mulac, (1991b), The discourse conditions for the use of complementizer that in conversational English, *Journal of Pragmatics* 15, 237–251.

Perception-based inferences: A compositional account to German evidential *an*PPs

Judith Lauterbach
(University of Tübingen)

Keywords: <perception, inferential reading, evidence, preposition, trope>

The reading of perception verbs (PVs) depends on ontological features of the complement. Direct perception is bound to concrete entities that are localized in space. Inferential readings, in contrast, are restricted to propositional complements that cannot be sensorily perceived. In this case, the PV presupposes a direct perception stimulus argument as information source for the experiencer's inference, that is, the PV presupposes perceptive evidence. In German, implicit stimulus arguments can be overtly realized by PPs headed by the locative preposition *an* 'at, on' (Müller 2020). For example, in (1), Mary's inference that it is autumn is based on her perception of the leaves, resulting in an evidential reading of the *an*PP.

- (1) Maria sieht am Laub, dass es Herbst ist.
Mary sees at.on-the leaves that it fall is
 'Mary can tell that it is autumn by (seeing) the leaves'

Evidential *an*PPs are restricted to perception-based inferential contexts. Reportative *hören* 'hear' as well as perception-independent inferential verbs like *schlussfolgern* 'infer' both block evidential *an*PPs for the reference to information sources, see (2)-(3).

- (2) Reportative reading of PV:
- a. Maria hört im Rundfunk, dass das Fußballspiel beginnt.
Mary hears in-the broadcast that the soccer-match begins
 'Mary learns from the broadcast that the soccer match is starting'
- b. *Maria hört am Rundfunk, dass das Fußballspiel beginnt.
Mary hears at.on-the broadcast that the soccer-match begins
 intended: 'Mary can tell from (hearing) the broadcast that the soccer match is starting'
- (3) *Maria schlussfolgert am Laub, dass es Herbst ist.
Mary infers at.on-the leaves that it fall is
 intended: 'Mary can tell that it is autumn by inferring it from the leaves'

Although largely mentioned in the literature on German PVs, and attested for few other verbs like *erkennen* 'recognize' or *erahnen* 'guess, surmise', evidential *an*PPs have not been studied thoroughly. Crucially, it is unclear how their reading is derived compositionally and how the stimulus relates to the propositional content.

I argue that evidential *an*PPs basically establish a conceptual relation: As shown in (4), trope-referring terms like *Peters Müdigkeit* 'Pete's tiredness' are compatible with evidential *an*PPs. According to Bücking (2012), Pete's tiredness is only perceivable via localized tiredness-manifestations, e.g. having

rings under his eyes. Perceiving such a concrete manifestation constitutes evidence for the claim that Pete is tired. This is introduced by an *anPP*, as in (1).

- (4) Maria sieht Peters Müdigkeit an seinen Augenringen.
 Mary sees Pete's tiredness at.on his circles-under-the-eyes
 'Mary can tell that Pete is tired by (seeing) the circles under his eyes'

In (1) and (4), the reference objects (RO, internal argument) of *an* are underspecified. In these examples, it is a particularized property that manifests in the leaves or the circles under the eyes, respectively, that justifies Mary's inference, i.e. their specific color and/or size. Indeed, only coloured leaves may be conceptually considered as a manifestation of autumn.

Along these lines, I argue that the evidential use of *an* is due to the spatial concept coded in the lexical semantics of the preposition and the physiology of perception. *An* encodes a non-holistic conceptualization of the RO by focusing on its physical boundaries (e.g. Carstensen 2015), which is decisive in perceptual contexts: We perceive and discern concrete things by their outer appearances, which constitute spatial boundaries in the case of physical bodies.

Treating evidential *anPPs* as event-internal modifiers in the sense of Maienborn (2003) yields the representation of (1) in (5). The implicit stimulus argument *c* provided by the PV is identified as external argument of *an* via pragmatic reasoning and related to specific aspects of its RO: *c* is underspecified but located at the boundary of the leaves. As a visually perceivable entity, *c* might contextually be enriched as their colorfulness.

- (5) $\exists s \exists s' \exists c$ [see'(s) & experiencer(s,Mary) & stimulus(s,c) & belief(s') & experiencer(s',Mary) & content(s', 'it is fall') & loc(c,boundary(def-leaves))]

The evidential use of *an* is thereby compositionally fully derivable from its primary spatial meaning. Generally, perception-based inferences are sourced in the exterior appearances of concrete entities.

References

- Bücking, Sebastian (2012), Müdigkeit und Müde-Sein: Zur Semantik adjektivbasierter Zustandsnominalisierungen im Deutschen, *Linguistische Berichte* 232, 361-397.
 Carstensen, Kai-Uwe (2015), A cognitivist attentional semantics of locative prepositions, in Marchetti/Giorgio/Benedetti, Giulio/Alharbi, Ahlam (eds), *Attention and Meaning. The Attentional Basis of Meaning*. Hauppauge, NY: Nova Science Publishers, 93-132.
 Maienborn, Claudia (2003), *Event-internal modifiers: Semantic underspecification and conceptual interpretation*, Universitätsbibliothek Johann Christian Senckenberg, 475-509.
 Müller, Kalle (2020), Perception verbs and finite complement clauses. *Empirical Issues in Syntax and Semantics* 13, 55-79.

The effects of semantic relations in the source and recipient language on the development of borrowings: A case study of SEXUALITY borrowings in Dutch

Marijn Boomars
(University of Manchester)

Keywords: lexical borrowing, lexical relations, connotations, anglicism, Dutch

Acknowledgments: This paper is part of my PhD thesis which is funded by the NWCDTP

Semantic relations, such as synonymy and antonymy, can influence the semantic development of the members of these relations (Lehrer 1970, 1978). For example, although Lehrer (1970, p. 353) considers the emotional meaning of *hot* ‘passionate’ to be metaphorically motivated, she argues that the meaning ‘placid’ of *cold* is not and that this meaning must have developed as a result of the antonymous relation of *cold* with *hot*. More recently, linguists (e.g. Paradis et al. 2009) have pointed out the need to differentiate conventional (e.g. *slow/fast*) and non-conventional (e.g. *slow/sudden*) pairings, with only the former capable of motivating semantic change (Murphy 2003). This study investigates how these ideas interact with borrowing. Unlike native formations, borrowings do not form semantic relations, both conventional and non-conventional pairings, wholly anew in a recipient language: the borrowed words are already part of semantic relations in the source language and enter a network of relations in the recipient language. It has been noted that the very process of lexical borrowing can influence the semantic development of a borrowing in the recipient language (Alexieva 2008, and Winter-Froemel 2013). I show that the semantic relations within both the recipient and source language also have the potential to impact the development of borrowings.

Borrowings feature prominently in the lexical field of SEXUALITY in many languages (Vecchio 2021) and Dutch is no exception to this. I use a corpus of *De Gay Krant* ‘The Gay Newspaper’ and *SoNaR* to trace the development of sexuality borrowings in the twentieth and twenty-first century and supplement this data with native speaker judgements. I demonstrate that semantic relations in the source language influence the pairings that borrowings can form with native and non-native items in the recipient language and impact the subsequent semantic change that the borrowings undergo. For instance, it was only after *gay* ‘homosexual’ was introduced from English into Dutch, that *straight*, borrowed earlier in the meaning ‘honest’, began to be used in the sense ‘heterosexual’. Since *gay* and *straight* are frequently combined by Dutch speakers, I hypothesise that this change is triggered by the fact that speakers were aware that the pairing of *gay/straight* is more conventional than *gay/heterosexual* in English. In the recipient language, relations also trigger change. It was observed already by Bréal (1897) that if two synonyms are retained, they often undergo semantic differentiation. Using questionnaire and corpus data, I show that sexuality terms with similar meanings carry different connotations and/or belong to different registers. This differentiation is again well illustrated by *gay*: this term is more strongly associated with the gay liberation movement than its synonyms *homoseksueel*, *homofiel* and *homo* ‘homosexual’.

This study shows the influence of native and non-native semantic relations on the development of borrowings, which highlights the complexity of the process of lexical borrowing.

References

- Alexieva, Nevena (2008), How and why are anglicisms often lexically different from their English etymons?, in R. Fischer, and H. Pulaczewska (eds), (2008), *Anglicisms in Europe: Linguistic diversity in a global context*, Newcastle upon Tyne: Cambridge Scholars Publishing, 42-51.
- Bréal, Michel (1897). *Essai de Semantique: Science des Significations*, Paris: Hachette.
- Lehrer, Adrienne (1970), Static and dynamic elements in semantics: *Hot warm cool cold*, *Paper in Linguistics* 3(2), 349-373.
- Lehrer, Adrienne (1978), Structures of the lexicon and transfer of meaning, *Lingua* 45(2), 95-123.
- Murphy, M. Lynne (2003), *Semantic relations and the lexicon: Antonymy, synonymy and other paradigms*, Cambridge: Cambridge University Press.
- Paradis, Carita, Willners, Caroline, and Jones, Steven (2009), Good and bad opposites: Using textual and experimental techniques to measure antonym canonicity, *The Mental Lexicon* 4(3), 380-429.
- Van der Sijs, Nicoline (2010), *Etymologiebank*, retrieved December 15 2024, from <https://etymologiebank.nl>.
- Vecchio, Nicholas Lo (2021), Borrowing and the historical LGBTQ lexicon, *Pragmatics & Cognition* 28(1), 167-192.
- Winter-Froemel, Esme (2013), Formal variance and semantic changes in borrowing: Integrating semasiology and onomasiology, in E. Zenner, and G. Kristiansen (eds), (2013), *New perspectives on lexical borrowing: Onomasiological, methodological and phraseological innovations*, Berlin: De Gruyter Mouton, 65-100.

The mirative use of the possessive suffix 2SG in Solon

NARGIL

(Dynamique du langage & Université Lumière Lyon 2, France)

Keywords: Northern Tungusic, Mongolic, personal possessive, complex clause, Hulunbuir

This presentation explores the mirative function of the second-person singular possessive suffix (2SG.POSS) *-fi* in Solon. Solon is a Northern Tungusic language spoken in Hulunbuir, China, by approximately 20,000 speakers (Nargil 2025). The analysis is based on 26,234 words of narrative data collected through fieldwork.

While personal possessive suffixes typically mark possession and add referent tracking, non-possessive functions, such as associative relationships and particularization are also found in Uralic (Nikolaeva 2003) and Tungusic (Nikolaeva & Tolskaya 2001; Aralova & Pakendorf 2023) languages.

In Mongolic languages, the 2.POSS is known to mark contrastive topics (Janhunen 2012). Through contact with Mongolian and Dagur, the contrastive topic function appears to have transferred into Solon as shown by the fact that according to the Solon consultant *-fi* corresponds to the Mongolian 2.POSS. Furthermore, as will be shown in this talk 2SG.POSS has developed a mirative function and its frequency reached 51.2% (83/163).

To investigate this mirative function, nouns are categorized as either possessable or non-possessable:

Possessable items: inanimate objects, non-human animates, body parts, kinship terms

Non-possessable items: personal pronouns, demonstratives, proper names, natural phenomena, temporal nouns

When the 2SG.POSS attaches to a possessable noun, it expresses possession (01). When attached to a non-possessable noun, the marked noun become the main point of the surprising and focal point of the narrative expresses mirativity (term from König 2013), such as surprise, sudden discovery, counterexpectation, or new information relevant to the speaker, addressee, or character (based on the definition of mirativity from Aikhenvald 2012) (02).

(01) *uŋkəŋ-fi ant^{hi} pal-lti-tfi-ra-ŋ?*
cow-2SG.POSS how give.birth-RECP-PROG-NPST-3
'How are **your cows** calving?'

(02) ... *ajja ajja=fi pəjə ant^{hə} fikk^hul!*
INTRJ PROX=2SG.POSS person NEG ghost
'Oh, **she is not a human**, she is a ghost!' (counterexpectation)

The mirative use of 2SG.POSS also occurs in complex clause structures where an initial event (X) is followed by an unexpected event (Y) that triggers a mirative interpretation. In such cases, the predicate of X typically involves a past participle followed by 2SG.POSS (03). Similar constructions are found in Mongolian, further supporting a contact-induced origin (04).

(03) Solon

taxin pəŋkən pa:p kai əmə-sə=ʃi mi:sa-tə-ni tʃəmu a:ʃin
Then big bear come-**PST.PTCP=2SG.POSS** gun-DAT-3 bullet NEG
'Then **when a big bear came**, the (hunter's) gun ran out of bullets. (surprise)

(04) Mongolian (Leipzig Corpora Collection)

mašin_terge-tai bolo-xson=čini qiyiu toson-o ün-e üs-čü jam
vehicle-PROP become-**PST.PTCP-2.POSS** petrol oil-GEN price increase-CVB road
büglere-jü aɣar boqirdu-ba
block-CVB air pollute-PST

'**When (he) got a vehicle**, (he suddenly found) the petrol price is high, there are traffic jams, the air is polluted.' (counterexpectation)

The origin of the mirative use of the 2SG.POSS is unclear. One possibility is that some Tungusic languages (e.g., Bystraja Even, Brigitte Pakendorf, p.c) have a =*ʃi* particle to mark contrastive topic. Potentially, Solon once had =*ʃi* particle and due to the contact influence from Mongolic languages, the =*ʃi* particle and 2SG.POSS -*ʃi* are merged. The development of the mirative use of 2SG.POSS may have originated from contrastive meaning, a cross-linguistic pathway (Aikhenvald 2012). This change likely reflects multiple causations: internal grammaticalization and external influence from contact.

Acknowledgments: I would like to express my sincere appreciation to the LabEx ASLAN for funding my fieldwork projects; the invaluable contribution provided by the Solon consultants and speakers; the valuable suggestions from my supervisor Brigitte Pakendorf.

References

- Aikhenvald, Alexandra Y. 2012. The essence of mirativity. *Linguistic Typology*. De Gruyter Mouton 16(3). 435–485.
- Aralova, Natalia & Brigitte Pakendorf. 2023. Non-canonical possessive constructions in Negidal and other Tungusic languages: a new analysis of the so-called “alienable possession” suffix. *Linguistics* 61(6). 1563–1592.
- Janhunen, Juha A. 2012. *Mongolian*. Amsterdam: John Benjamins Publishing Company.
- König, Christa. 2013. Source of Information and Unexpected Information in !Xun—Evidential, Mirative and Counterexpectation Markers. In Alexandra Aikhenvald & Anne Storch (eds.), *Perception and Cognition in Language and Culture*, 69–94. Leiden: Brill.
- Leipzig Corpora Collection 2011. *Mongolian newspaper corpus based on material from 2011*. Leipzig Corpora Collection. Dataset.
- Nargil. 2025. The Language Ecology and Endangerment of Solon, a Tungusic Language Spoken in China. *Sibirica* 24(1). 1–37.
- Nikolaeva, Irina. 2003. Possessive affixes in the pragmatic structuring of the utterance: Evidence from Uralic. In Comrie Bernard & Pirkko Suihkonen (eds.), *International symposium on deictic systems and quantification in languages spoken in Europe and North and Central Asia*. Collection of Papers, 130–145. Izhevsk & Leipzig: Udmurt State University, Max Planck Institute for Evolutionary Anthropology.
- Nikolaeva, Irina & Maria Tolskaya. 2001. *A Grammar of Udihe*. Berlin: Walter de Gruyter.

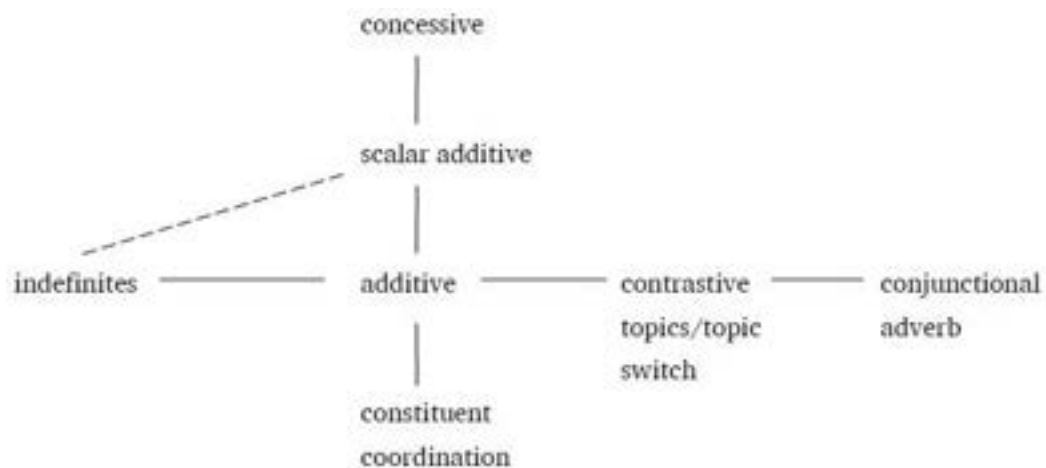
Extended semantic map of additives

Petr Rosseyaykin & Alina Russkikh
(Lomonosov Moscow State University & HSE University)

Keywords: additivity, scalarity, quantification, contrast, semantic maps

Additive (like English *too*) and scalar additive (like English *even*) markers have a number of typologically stable functions, for which Forker (2016) proposed a semantic map in (1).

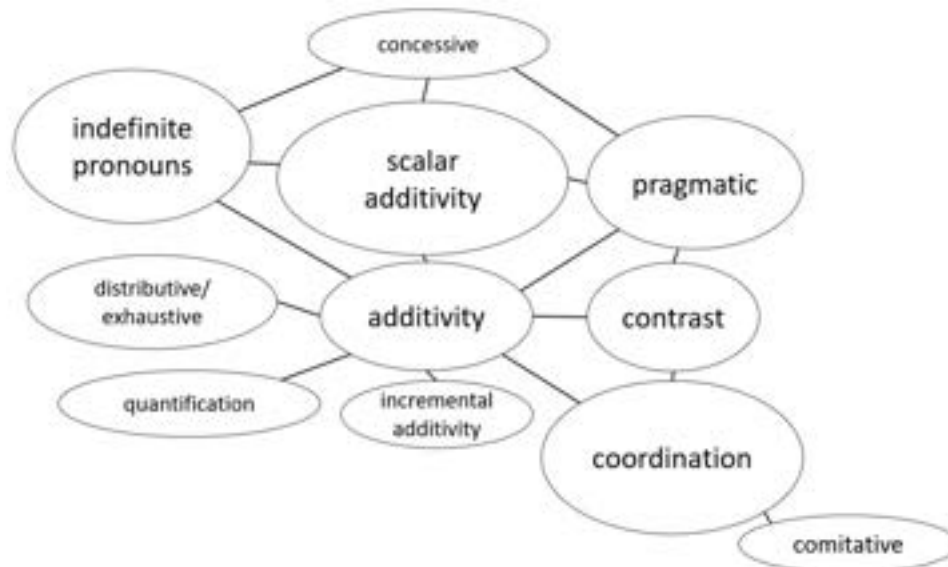
(1)



A semantic map of additives (Forker 2016)

Based on typological and language-specific literature and our own fieldwork studies of several languages of Central Eurasia (mainly from Turkic and Nakh-Dagestani families), we propose an extended version of (1) shown in (2). In our version new functions (in terms of Haspelmath 1997: 60–62, 2003) are added and some functions from (1) are reanalyzed as clusters of functions. They are represented by ellipses with larger ellipses corresponding to clusters with a greater number of (sub)functions.

(2)



An extended semantic map of additives

To illustrate what a “cluster of function” means let us zoom in to the scalar additive cluster. (3a) shows that in Karachay-Balkar (Malkar variety) there are two particles (*da*, *oquna*) which can associate with the (contextual) high endpoint of a scale (*kamčatka* ‘Kamchatka’), but only *oquna* can associate with low endpoints in imperatives (3b) yielding the weak scalar (concessive) reading. In Telugu a single scalar additive particle can even have either weak or strong scalar reading in non-modal environments (4). Thus, the “scalar function” of additives actually covers a semantic map of its own, on which we will not elaborate here, see Gast & van der Auwera 2011.

- (3) a. xabib kamčatka-da oquna/da bol-ɣan-di.
Khabib Kamchatka-LOC even/ADD be-PFCT-3SG
‘Khabib have even been to Kamchatka.’
b. bir zat oquna/%da aša!
one thing even/ADD eat.IMP
‘Eat at least one thing!’ (Malkar Balkar; own data)
- (4) id-ai-naa tin-Taatu.
this-be-IF.EVEN eat-will
‘I will eat even / at least this.’ (Telugu; Balusu 2020: 7)

In this contribution we focus on peripheral and understudied (clusters of) functions not present in (1), in particular bottom left (distributive/exhaustive, quantification) and upper right (pragmatic, contrast) parts of the map (2).

As for the quantificational use, it has been observed that some additives can be attached to universal quantifiers, often without any obvious semantic contribution (5a). A subset of those additives is also compatible with completeness (‘whole’) quantifiers (5a). Combinations with weaker quantifiers, e.g. ‘most’, are marginally attested (5b). Interestingly, combinations with even weaker quantifiers, e.g. ‘half’, ‘few’, are (to our knowledge) unattested at all.

- (5) a. ragu {ella.t-Ø-ai-um / muRu.tt-Ø-ai-um} paattaa.
Raghu every.OBL-Ø-ACC-ADD entire.OBL-Ø-ACC-ADD see.PST
‘Raghu saw { everything / the entire thing }.’ (Tamil; Iyer 2017: 11)
b. mikka kuTTi-kaL-um maaNGaa thinnu.
most child-PL-ADD mango ate
‘Most children ate a mango.’ (Malayalam; Aravind 2013 via Iyer 2017: 10)

A typologically rarer and practically undescribed (although see Forker 2016 for some examples; see also Schwenter & Waltireit 2010; Kim & Jahnke 2011) set of functions is associated with mirativity, contrast and emphasis. We combine these functions in the pragmatic cluster in (2). E.g. in Poshkart Chuvash the additive particle =*ta*/*da* can mark sudden realization (6a) and verum (6b). It also has the enimitive function (6c), which Panov (2020) discusses without connecting it to additives.

- (6) a. o, mënle jorl-at=ta!
oh which sing-NPST[3SG]=ADD
‘Oh, the way she sings!’
b. [The child said he wasn't feeling well. The woman takes his temperature and says:]
atɕa sumar=da
child sick=ADD
‘The child is actually sick.’
c. [context: Why can't you fall asleep?]
jep eş-se=de
I drink-CV_SIM=ADD
‘Because I drank!’ (Poshkart Chuvash; own data)

In our talk we will examine the peripheral functions of additives and general composition of the map (2) in more detail.

List of Abbreviations

3SG — 3rd person singular; ACC — accusative; ADD — additive; CV_SIM — simultaneous converb; IMP — imperative; LOC — locative; NPST — non-past tense; OBL — oblique; PFCT — perfect; PL — plural; PST — past tense

Acknowledgments

Petr Rossyaykin is supported by the RSF grant #25-18-68012.

References

- Aravind, Athulya (2013), *Quantifier morpheme um and overt domain restriction in Malayalam*, ms., Massachusetts Institute of Technology.
- Balusu, Rahul (2020), *Unifying FCIs, NPIs, and Unconditionals in Dravidian*, ms. EFL-University, Hyderabad.
- Gast, Volker and Johan van der Auwera (2011), Scalar additive operators in the languages of Europe, *Language* 87(1), 2–54.
- Haspelmath, Martin (1997), *Indefinite pronouns*, Oxford: Oxford University Press.
- Haspelmath, Martin (2003), The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison, in M. Tomasello (ed), (2003), *The new psychology of language*. Mahwah, New Jersey & London: Erlbaum, 211–242.
- Kim, Min-Joo and Nathan Jahnke (2011), The meaning of utterance-final *even*, *Journal of English Linguistics*, 39(1), 36–64.
- Iyer, Jyoti (2017), *Towards a universal analysis of Tamil -UM NPIs: Evidence from unconditionals*, ms., UMass.
- Panov, Vladimir (2020), The marking of uncontroversial information in Europe: presenting the enimitive, *Acta Linguistica Hafniensia*, 52(1), 1–44.
- Schwenter, Scott A. and Richard WALTEReit (2010), Presupposition accommodation and language change, in K. Davidse, L. Vandelanotte, H. Cuyckens (eds.), *Subjectification, Intersubjectification and Grammaticalization*. De Gruyter Brill, 75–102.

***Aeglane nagu tigu* ‘slow as a snail’: Similes for speed in Estonian**

Piia Taremaa, Kärolin Veerpalu & Ann Veismann
(University of Tartu)

Keywords: speed, similes, semantics, asymmetries, corpus study

The everyday human experience involves speed, which combines space and time. While space and time are well-studied in linguistics, speed is less explored. As a result, there is limited knowledge about how speed is expressed in language. It is well known that speed adjectives (e.g. fast and slow in English) are polar antonyms, with fast serving as the representative of the whole scale of speed (Cruse 1986). It has also been shown that if several resources are available in a language within a word class, fastness tends to be represented with more numerous and more frequent items than slowness. For instance, in Russian, there are several fast adjectives, but only one main slow adjective (Plungian & Rakhilina 2013). Thus, in expressing speed-related information, there seem to be asymmetries.

The current exploratory study on Estonian focuses on simile constructions of X as Y which contain a target (X), speed adjective/adverb, the linker ‘like, as if’, and source (Y) as the comparison reference; e.g. *ta on aeglane nagu tigu* ‘(s)he is as slow as a snail’. The target and source in similes can occur as NPs, but also as other, e.g. verb structures (Moon 2008). The study aims to reveal whether there are differences between similes of slow and fast words regarding frequencies and usage patterns.

We examine similes with the four main Estonian speed words: *aeglane* ‘slow’, *aeglaselt* ‘slowly’, *kiire* ‘fast’, *kiiresti* ‘fast, quickly’. With each of them, search queries were performed in the Estonian National Corpus 2023 (Koppel et al. 2023), with the condition that the search word was followed by one of the two linking words, *kui* or *nagu* ‘like, as (if)’. Then, up to 500 sentences were randomly selected for further analysis and were annotated for: 1) the compared items, X and Y, and their semantics, 2) the form of the comparison item (NP vs. VP), and 3) verb semantic type (e.g. cognitive verb, motion verb).

Corpus searches resulted in an unbalanced sample of similes: with *aeglaselt* ‘slowly’ 741 hits, *aeglane* ‘slow’ 381 hits, *kiiresti* ‘fast, quickly’ 9609 hits, and *kiire* ‘fast’ 1974 hits. These numbers reflect the overall frequencies of these speed words (fast words are more frequent than slow words and speed adverbs are more frequent than speed adjectives in Estonian).

As for sentence patterns, preliminary results indicate that slow words (*aeglane*, *aeglaselt*) are most likely to combine with motion verbs, whereas fast words (*kiire*, *kiiresti*) with state or change-of-state verbs. Slow words also tend to occur with animate subjects, and comparisons often involve a concrete entity (e.g. ‘he moves as slowly as a snail’). With fast words, comparison is frequently made with a state, process, earlier situation, etc. (e.g. ‘he was as fast as in Australia’). These results suggest that slow words in simile constructions have a more restricted usage (often expressing physical motion), whereas fast words can be used in more varied semantic and morphosyntactic contexts. This can be interpreted as an indication of a *fast-over-slow* asymmetry (see also Taremaa & Kopecka 2023), whereby fastness in language, compared to slowness, is characterised by greater diversity and usage flexibility.

References

- Cruse, D. Alan. 1986. *Lexical semantics* (Cambridge Textbooks in Linguistics). Cambridge/New York/Melbourne: Cambridge University Press.
- Koppel, Kristina, Jelena Kallas, Madis Jürviste & Helen Kaljumäe. 2023. *Eesti keele ühendkorpus 2023*. Lexical Computing Ltd. / Eesti Keele Instituut.
- Moon, Rosamund. 2008. Conventionalized *as*-similes in English: A problem case. *International Journal of Corpus Linguistics* 13(1). 3–37. <https://doi.org/10.1075/ijcl.13.1.03moo>.
- Plungian, Vladimir & Ekaterina Rakhilina. 2013. Time and speed: Where do speed adjectives come from? *Russian Linguistics*. Springer 37(3). 347–359. <https://doi.org/10.1007/s11185-013-9117-7>.
- Taremaa, Piia & Anetta Kopecka. 2023. Manner of motion in Estonian: A descriptive account of speed. *Studies in Language* 47(1). 32–78. <https://doi.org/10.1075/sl.21038.tar>.

Semantic extensions of evidentials in languages with or without egophoricity

Satu Keinänen
(University of Helsinki)

Keywords: Evidentiality, egophoricity, first-person effects, mirativity, typology

While first-person effects of evidentials are well-studied, there are fewer studies that explicitly discuss how first-person effects or other extensions of evidentials occur in different evidential systems. In this paper, I compare evidential systems in languages with or without egophoricity: do similar evidentials acquire similar extensions, or is the nature of the attested extensions determined by the presence or lack of egophoricity? In several languages with egophoricity, like Guambiano and Spiti, a firsthand/visual evidential indicates non-volitionality in first-person contexts (Norcliffe 2018, Hein 2001), whereas otherwise that function typically occurs on non-firsthand, nonvisual, or inferential evidentials, and a visual can emphasise the speaker's volitional action, like in Tucano (Ramirez 1997).

Languages with egophoricity have differential marking for the first person's and other persons' personal knowledge (egophoric vs. allophoric) (see San Roque et al. 2018, Widmer 2020). The marking can be specialised copulas or inflections, for instance. Besides the differential treatment, the dimension of epistemic authority or privileged knowledge has consequences for how evidentials are used with the first person. Thus, the occurrence of the first-person effects, i.e. additional meanings of evidentials when used with the first person (see e.g. Curnow 2002, Aikhenvald 2004, Sun 2018), can differ from those found in languages without egophoricity. Languages with egophoricity can further have other usage patterns of evidentials but they require more comparison to each other and to languages without egophoricity.

To study evidentials and egophoric marking, I created a pilot sample of 68 languages, of which 18 mark both egophoricity and evidentiality. The descriptions of evidentials and egophoric markers and their uses were compared within typological-functional framework. The analysis was focused on first-person effects, especially on the expression of non-volitionality. The use of evidentials with unexpected information (mirativity) and general knowledge was also analysed, to find out if there are other distinct usage patterns in either group of languages. Mirativity and general knowledge were chosen because they both convey the speaker's attitude towards the knowledge, which can coincide with the notion of epistemic authority.

The results show that the languages with egophoricity have tendencies that differ from the languages without egophoricity. In languages with egophoricity, firsthand/visual or allophoric evidentials typically have non-volitional first-person effects, whereas in languages without egophoricity the expression of non-volitionality with firsthand/visual is not common. When it comes to the marking of mirativity or general knowledge, the languages with egophoricity use diverse evidentials (or do not use evidentials at all), but there are still differences to uses found in languages without egophoricity. Often mirativity is expressed with non-firsthand evidentials and general knowledge is marked with firsthand/visual evidentials in languages without egophoricity, but in languages with egophoricity these uses of non-firsthand and firsthand/visual are quite rare.

This study shows that egophoricity is relevant to our understanding of evidentiality. Evidentiality could be understood in a broader sense as expression of epistemic authority, encompassing egophoricity (see Bergqvist & Grzech 2023), but even then, dedicated egophoric marking changes how other evidentials are used when first-person perspective or the speaker's attitude is concerned.

Acknowledgments

This research was partially funded by Finnish Cultural Foundation.

References

- Aikhenvald, Alexandra Y. (2004), *Evidentiality*, Oxford: Oxford University Press.
- Bergqvist, Henrik & Karolina Grzech (2023), The role of pragmatics in the definition of evidentiality, *STUF - Language Typology and Universals*, 76(1), 1–30.
- Curnow, Timothy J. (2002), Types of interaction between evidentials and first-person subjects, *Anthropological Linguistics* 44(2), 178–196.
- Hein, Veronika (2001) The Role of the Speaker in the Verbal System of the Tibetan Dialect of Tabo/Spiti, *Linguistics of Tibeto-Burman Area* 24(1), 35–48.
- Norcliffe, Elisabeth (2018), Egophoricity and evidentiality in Guambiano (Nam Trik), in Simeon Floyd, Elisabeth Norcliffe & Lila San Roque (eds), (2018), *Egophoricity*, Amsterdam/Philadelphia: John Benjamins, 305–345.
- Ramirez, Henri (1997), *A fala Tukano dos Ye'pâ-Masa, tomo 1: Gramática, Manaus: Inspetoria Salesiana Missionária da Amazônia and Centro "Iauareté" de Documentação Etnográfica e Missionária.*
- San Roque, Lila, Simeon Floyd & Elisabeth Norcliffe (2018), Introduction, in Simeon Floyd, Elisabeth Norcliffe & Lila San Roque (eds), (2018), *Egophoricity* (Typological Studies in Language 118), Amsterdam: John Benjamins, 1–77.
- Sun, Jackson T.-S. (2018), Evidentials and person, in Alexandra Aikhenvald (ed), (2018), *The Oxford handbook of evidentiality*, Oxford: Oxford University Press, 47–64.
- Widmer, Manuel (2020), Same same but different: On the relationship between egophoricity and evidentiality, in Henrik Bergqvist & Seppo Kittilä (eds), (2020), *Evidentiality, egophoricity, and engagement* (Studies in Diversity Linguistics 30), Berlin: Language Science Press, 263–287.

General Session : Sociolinguistics

Subject inversion in French relative clauses: Experimental results and region-based influences

Gabriel Thiberge, Dinél Badeau & Céline Pozniak

(CLLE, CNRS, U. Toulouse Jean Jaurès; SFL, CNRS, U. Paris 8; & SFL, CNRS, U. Paris 8)

Keywords: French; relative clause; verb-subject inversion; experimental; sociolinguistics

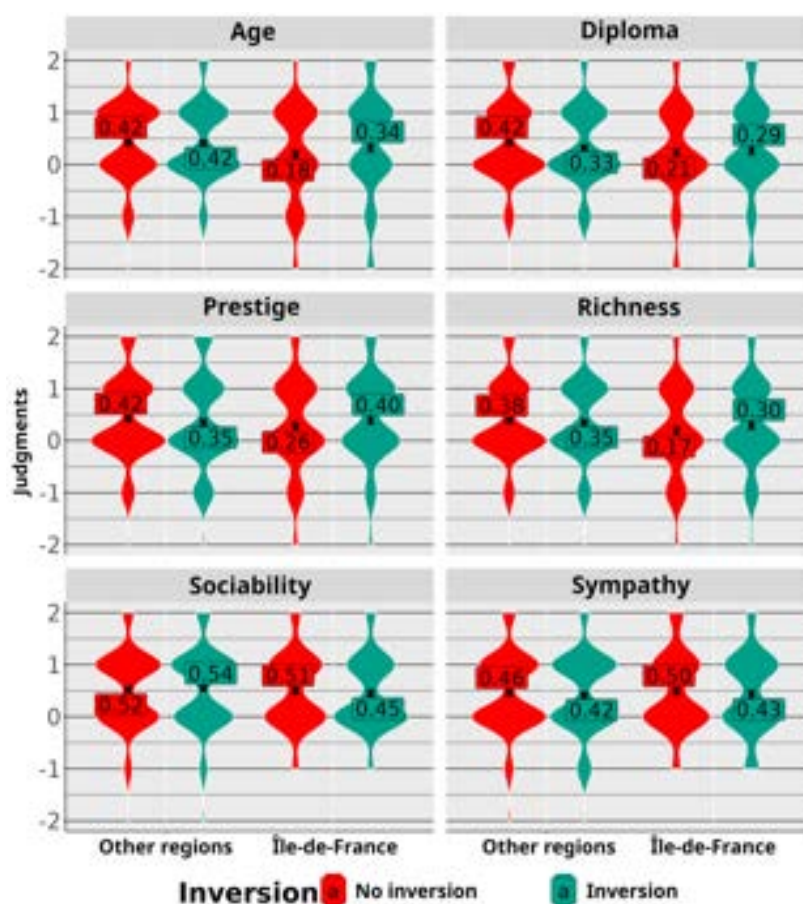
Introduction. In French, object relative clauses can feature a preverbal subject (e.g., *The manager that Manon hates goes to the restaurant*) or a postverbal subject (e.g., *The manager that_{obj} hates Manon goes to the restaurant*). Subject inversion is often considered optional and is referred to as stylistic inversion, a form favored by standard norms (Kayne and Pollock 1978). Corpus studies by Pozniak et al. (2021) found that postverbal subjects are as frequent as preverbal subjects in object relatives. However, postverbal subjects tend to be longer and have non-agentive and non-intentional properties. Experiments by the same authors showed that postposition is preferred when the subject is long and non-agentive, indicating that linguistic factors influence the choice between these variants. In experimental sociolinguistics, other studies have shown that interrogative sentences with subject inversion (e.g., *When do you come?*), conforming to standard French norms, evoke different social representations compared to variants without inversion (e.g., *You come when?*) (Thiberge 2020). These inverted forms are associated with a particular social meaning (Beltrama 2020), such as higher age, wealth, and education levels. Our study extends this analysis of inverted subject to object relatives using a matched-guise experiment (Kircher 2015).

Method. 42 French speakers participated in an online experiment via Prolific. They were presented with short dialogues where person A introduced a context (formal or informal), and person B responded with a sentence containing an object relative, either with or without subject inversion. For example, in a formal context, person A might say *"The phone hasn't stopped ringing since this morning"*, and person B would respond with either *"The salesperson that hates Laurent keeps calling"* (inversion) or *"The salesperson that Laurent hates keeps calling"* (non-inversion). After each dialogue, participants evaluated person B on six social dimensions (prestige, age, likeability, education, wealth, sociability) using a five-level Likert scale. 16 distractors and 16 target items were presented in random order (4,032 total observations).

Results. We ran Bayesian inferential models on the data. Responses on different scales were correlated with wealth, education, and prestige grouped together, and likeability and sociability grouped separately. Generally, there were few differences in social dimension evaluations based on subject inversion, except on the likeability scale, where inversion led to marginally lower judgments. The formal/informal context contrast was negligible, contrary to findings in Thiberge (2020). However, a post-hoc analysis

revealed significant differences based on participants' regions. Those from Île-de-France (around Paris) associated inversion with higher prestige, age, education, and wealth, especially in formal contexts, while participants from other regions sometimes associated inversion with lower prestige and education.

Discussion. Unlike previous work on subject inversion, the sociolinguistic perception of subject inversion in relatives seems to be conditioned by geographical and sociolectal factors. Further research is needed to better understand these factors.



Results by scales and inversion status, by participant region (Île-de-France vs Other regions)

References

- Beltrama, Andrea. (2020). Social meaning in semantics and pragmatics. *Language and Linguistics Compass*, 14(9):e12398.
- Kayne, Richard and Pollock, Jean-Yves. (1978). Stylistic inversion, successive cyclicity, and move np in french. *Linguistic inquiry*, 9(4), 595–621.
- Kircher, Ruth. 2015. The matched-guise technique. In Hua Zhu (ed.), *Research methods in intercultural communication: A practical guide*. Hoboken: Wiley-Blackwell. 196–211.
- Pozniak, Céline, Anne Abeillé and Barbara Hemforth (2021). Subject inversion in French object relatives: What's your preference? In B. Crysmann and M. Sailer (Eds). *One-to-many relations in morphology, syntax, and semantics*. Berlin: Language Science Press. 151-173.
- Thiberge, Gabriel (2020). *Acquisition et maîtrise des interrogatives partielles en français: La variation sociolinguistique comme outil interactionnel*. Thèse de doctorat, Université Paris Cité.

What's writing got to do with it? Measures of language vitality in review

Jenna Sorjonen
(University of Helsinki)

Keywords: language vitality, language policy and planning, literacy development

Literacy development, the elaboration of writing systems and literacy practices, has long been considered a vehicle for enhancing the status and use of marginalised languages, and orthography development continues to be something of a mandatory chapter in handbooks on language endangerment and revitalisation (Grenoble & Whaley 2006; Cahill 2018). This is in line with the logic of strengthening the vitality of languages by introducing their use in education and other formal domains where literacy practices play an important part of interaction. As such top-down approaches to language vitality and maintenance are criticised (e.g. Lüpke 2015), the role of vernacular writing becomes open to question, too: If language vitality overall is not always best protected with top-down institutional measures, then what does writing have to do with vitality?

In this talk, I present findings from a study reviewing the role of literacy for language vitality with a qualitative multimethod approach. A first analysis examines the design principles behind vitality scales used in major cross-linguistic databases: the Expanded Graded Intergenerational Disruption Scale used in the Ethnologue (Lewis & Simons 2010), the UNESCO vitality scale used in the UNESCO Atlas of World Languages (Brenzinger et al. 2003), the Language Endangerment Index used in the Catalogue of Endangered Languages (Lee & Van Way 2018), and the Agglomerated Endangerment Status used in the Glottolog (Hammarström 2018). Second, this review of the vitality scales' design principles is linked with a case study of 40 languages from around the world contrasting their vitality ratings in said databases with a bottom-up analysis of qualitative sociolinguistic descriptions from academic publications on these languages.

The analysis of vitality scale design principles shows that despite differences in applied vitality indicators, all scales are founded on a formal approach to language policy and planning, taking domain expansion as a goal and assuming literacy as part of this development, ignoring informal literacy practices. Comparing the vitality ratings of my case languages confirms claims frequently made on an anecdotal level: the ratings in different databases point to different interpretations of individual language situations and are frequently not in line with qualitative descriptions of language vitality. Three kinds of outliers in the ratings appear, explained by the different design principles of the scales: in a number of cases, vitality is systematically exaggerated due to 1) literacy status, 2) official recognition, and 3) speaker numbers. The bottom-up analysis of the case languages' vitality factors and literacy development as presented in the qualitative descriptions also suggests that languages across societies differ in what the realistic functions for vernacular literacy are and how these relate to the goals of language maintenance. A major division appears between contexts in the global north and south, differing particularly in language transmission, presence of top-down language policies for marginalised languages, existing literacy practices and relevant literacy development goals.

References

Brenzinger, Matthias, Yamamoto, Akira, Aikawa, Noriko, Koundioubu, Dmitri, Minasyan, Anahit, Dwyer, Arianne, Grinevald, Colette, Krauss, Michael, Miyaoka, Osahito & Sakiyama, Osamu. 2003. Language vitality

and endangerment. UNESCO document submitted to the International Expert Meeting on UNESCO Programme Safeguarding of Endangered Languages, Paris, 10–12 March 2003.

Cahill, Michael. 2018. Orthography Design and Implementation for Endangered Languages. In Rehg, Kenneth L. & Campbell, Lyle (eds.), *The Oxford Handbook of Endangered Languages*, 326–346. Oxford: Oxford University Press.

Grenoble, Lenore A. & Whaley, Lindsay J. 2006. *Saving languages: an introduction to language revitalization*. Cambridge: Cambridge University Press.

Hammarström, Harald. 2018. Simultaneous Visualization of Language Endangerment and Language Description. *Language Documentation & Conservation* 12. 359–392.

Lee, Nala H. & Van Way, John R. 2018. Assessing Degrees of Language Endangerment. In Rehg, Kenneth L. & Campbell, Lyle (eds.), *The Oxford Handbook of Endangered Languages*, 47–65. Oxford: Oxford University Press.

Lewis, M. Paul & Simons, Gary F. 2010. Assessing endangerment: expanding Fishman's GIDS. *Revue roumaine de linguistique* 55(2). 103–120.

Lüpke, Friederike. 2015. Ideologies and typologies of language endangerment in Africa. In Essegbey, James & Henderson, Brent & Mc Laughlin, Fiona (eds.), *Language Documentation and Endangerment in Africa*, 59–106. Amsterdam: John Benjamins Publishing Company.

Variation and change in lexical diversity: Does synchronic change mirror diachronic leveling?

Karlien Franco
(KU Leuven Brussels Campus)

Keywords: dialectology, sociolinguistics, lexical diversity, Dutch, variation and change

Why do we have several synonyms for ‘winking’ (*to wink*, *to blink*, *to bat an eye* etc.) but only one for ‘nodding’ (*to nod*), according to the OED? This paper focuses on this type of variation in lexical diversity, i.e. variability in the *amount* of lexical variation a concept shows. As a case study, we investigate the lexicon of the Limburgish dialects, spoken in Belgium and the Netherlands.

In earlier work, Franco et al. (2019) showed that in historical data (roughly 1960-1980) from the Limburgish dialect area, variation in lexical diversity is influenced by characteristics of the concepts being expressed, such as familiarity and semantic field. More specifically, they showed that concepts that are less familiar for dialect users show a higher amount of lexical diversity than more familiar concepts. For example, a concept with low familiarity such as ‘to wink’ has 36 variants across the Limburgish dialects, whereas for a highly familiar concept like ‘knee’, only 1 lexical item occurs. Moreover, Pickl (2013) also found differences between semantic fields, in the sense that concepts from semantic fields related to “local” life, such as household items, show more variation (dialectal variants) than concepts that are universal and socially relevant on a supralocal level, such as salutations.

The aim of this paper is to shift from a synchronic to a diachronic perspective. More specifically, we analyze whether these concept characteristics also affect the amount of diachronic leveling, or the speed of dialect loss, in the Limburgish dialects. The hypothesis is that concept characteristics that were shown to decrease synchronic lexical diversity also cause a larger degree of diachronic lexical leveling. In particular, we hypothesize that more lexical leveling will occur for more familiar concepts and for concepts from universal semantic fields.

We investigate this hypothesis by relying on data from a large-scale digital survey that ran in August 2024. Participants to the survey were native speakers of a Limburgish dialect. The survey was designed to be completely parallel to the historical data analyzed in Franco et al. (2019). We specifically collected lexical variants in use for twenty-four concepts from four semantic fields. We investigate two culturally variable semantic fields (clothing and household items) and two universal semantic fields (animals and body parts). We analyze the data using linear regression.

Preliminary results show a consistent effect for lack of familiarity, with more levelling for more familiar concepts, as expected. However, the results for semantic field are not significant. This indicates that the same degree of lexical levelling is found across semantic domains. Moreover, the results seem to be complicated by the fact that many respondents are less proficient in the dialect nowadays than in the historical data. As a result, they often resort to standard language variants. Further analyses may indicate whether age differences between participants can shed more light on these results.

References

- Franco, Karlien, Dirk Geeraerts, Dirk Speelman & Roeland Van Hout. 2019. Concept characteristics and variation in lexical diversity in two Dutch dialect areas. *Cognitive Linguistics* 30(1). 205–242.
- Pickl, S. 2013. Lexical meaning and spatial distribution. Evidence from geostatistical dialectometry. *Literary and Linguistic Computing* 28(1). 63–81.

Exploring the effects of second language (L2) proficiency, gender, and first language on complexity, accuracy, and fluency in L2 writing

Khaled Barkaoui & Ibtissem Knouzi
York University, Canada

Keywords: writing, accuracy, complexity, fluency, gender, first language, proficiency

This study examines how second language (L2) proficiency, first language (L1), and gender affect writing quality and measures of linguistic accuracy, syntactic and lexical complexity, and fluency (CALF) in L2 writing. Understanding how these learner characteristics shape writing performance can inform L2 writing theory and more effective instructional practices. While much research has focused on the role of L2 proficiency in writing performance, less attention has been given to how L1 and gender influence CALF measures. Previous studies show that more proficient L2 writers tend to produce longer texts with greater syntactic complexity and lexical diversity (e.g., Barrot & Agdeppa, 2021; Jiang et al., 2019), though some research suggests that the relationship between L2 proficiency and CALF measures can be negative, with its strength potentially influenced by other learner attributes and task complexity (e.g., Liao, 2020; Verspoor et al., 2012). Studies also highlight the influence of L1 background, with learners from different L1 backgrounds exhibiting varying patterns of syntactic and lexical complexity (e.g., De Clercq & Housen, 2017; Lu & Ai, 2015). Gender differences are also evident, with females often outperforming males in syntactic complexity and overall writing quality (e.g., Martínez, 2020; Shirzad et al., 2013). Despite these findings, limited research has investigated the combined effects of L2 proficiency, L1, and gender on CALF measures and writing quality, which this study addresses.

The written responses of 1,569 L2 learners from four L1 backgrounds (Mandarin, Spanish, English, and Arabic) to an independent writing task were computer-analyzed for nine CALF measures (e.g., error ratio, global syntactic complexity, lexical sophistication). The CALF measures were then statistically analyzed to examine the direct and interaction effects of L2 proficiency, gender, and L1 on writing quality (measured by writing scores) and CALF measures. Results revealed that L2 proficiency, gender, and L1 all significantly impacted CALF measures, which in turn influenced writing quality. Specifically, L2 proficiency positively impacted fluency, accuracy, and syntactic complexity, but negatively affected lexical complexity, indicating a trade-off between CALF components (cf. Skehan, 2016). L1 significantly influenced fluency and complexity, suggesting that learners from different L1s display distinct CALF profiles, even at similar L2 proficiency levels. Gender significantly affected lexical complexity, with female students producing less lexically complex responses than males. L2 proficiency and L1 influenced writing quality indirectly through fluency and accuracy. For example, higher proficiency led to better writing scores via increased fluency and accuracy. Lexical complexity mediated the effects of both L2 proficiency and gender, with female students and those with higher proficiency achieving better grades through simpler lexical choices. Finally, L2 proficiency and L1 moderated the relationships between CALF dimensions, but gender did not. For example, correlations between fluency and accuracy were stronger at lower proficiency levels and among Mandarin speakers. These results support the tenets of Complex Dynamic Systems Theory, which

highlights the complex, interdependent nature of learner factors in shaping CALF measures and writing quality (e.g., Larsen-Freeman, 2009; Verspoor et al., 2012). The findings and their implications for theory, instruction, assessment, and future research will be discussed.

References

- Barrot, J. S., & Agdeppa, J. Y. (2021). Complexity, accuracy, and fluency as indices of college-level L2 writers' proficiency. *Assessing Writing*, 47, 100510.
- De Clercq, B., & Housen, A. (2017). A cross-linguistic perspective on syntactic complexity in L2 development: Syntactic elaboration and diversity. *The Modern Language Journal*, 101(2), 315-334.
- Jiang, J., Bi, P., & Liu, H. (2019). Syntactic complexity development in the writings of EFL learners: Insights from a dependency syntactically annotated corpus. *Journal of Second Language Writing*, 46, 100666.
- Larsen-Freeman, D. (2009). Adjusting expectations: The study of complexity, accuracy, and fluency in second language acquisition. *Applied linguistics*, 30(4), 579-589.
- Liao, J. (2020). Do L2 lexical and syntactic accuracy develop in parallel? Accuracy development in L2 Chinese writing. *System*, 94, 102325.
- Lu, X., & Ai, H. (2015). Syntactic complexity in college-level English writing: Differences among writers with diverse L1 backgrounds. *Journal of second language writing*, 29, 16-27.
- Martínez, A. C. L. (2020). Analysis of accuracy in the writing of EFL students enrolled on CLIL and non-CLIL programmes: The impact of grade and gender. *The Language Learning Journal*, 48(2), 121-132.
- Shirzad, F., Musavi, K., Atmani, S., Khanchobni Ahranjani, A., & Iraj, S. (2013). Gender differences in EFL academic writing. *International Journal of Academic Research*, 5(4).
- Skehan, P. (2016). Tasks versus conditions: Two perspectives on task research and their implications for pedagogy. *Annual Review of Applied Linguistics*, 36, 34-49.
- Verspoor M., Schmid M., & Xu, X. (2012) A dynamic usage-based perspective on L2 writing. *Journal of Second Language Writing* 21: 239-63.

Property vs Person Centered Language Across Dialects of French

Marie Flesch & Heather Burnett

Laboratoire de Linguistique Formelle, CNRS-Université Paris Cité

This paper introduces a syntactic alternation which to our knowledge, has never before been studied as such: *property vs person-centered language*. In both English and French, noun phrases referring to humans through descriptive property can be formed in a variety of ways. One way, which we will call the *property-centered* variant, involves using a nominalized property, such as *the handicapped* or *les handicapés*. Another way, which we will call the *person-centered* variant, involves using a general human noun such as *personne/person* with the property encoded in a modifier, such as *handicapped people* or *les personnes handicapées*. The use of person-centered language over property-centered language has been the topic of much discussion in social justice language policy in the English-speaking world, where it is often claimed that person-centered variants are less essentializing and less dehumanizing than property-centered ones (Granello & Gibbs 2016). However, to our knowledge, there have not yet been any detailed quantitative linguistic studies of how speakers use these different variants. Furthermore, despite being prominent in anglophone discourses on language and social justice, person-centered language has been mostly absent from the corresponding normative discourses in French, particularly in France. In order to contribute to filling the empirical gaps, we present a quantitative study of person vs property centered language online in two countries: France and Québec, Canada. It is based on a 700-million word corpus of comments posted on the Reddit forums *r/france* and *r/Quebec* between 2008 and 2022. We extracted all occurrences of the nouns *personne(s)*, *gens*, and *individu(s)*, and manually coded them to only retain person-centered language variants. Then, we created a subset with all variants which had a relative frequency of at least 0.2 per million words, and extracted and coded the corresponding property-centered variants. The final dataset contains 1,196,097 observations, including 46,082 examples of person-centered language variants and 1,150,015 property-centered variants corresponding to 116 properties (*pauvres*, *montréalais*, *musulmans*, *blancs*, *retraités*, etc.).

In our results, we find that property vs person-centered language is conditioned by both social and grammatical factors: To investigate geographical variation in the use of person-centered variants, we created a mixed-effects binomial logistic regression model with the R package *lme4* (Bates et al. 2015). It shows that that person-centered language is more common in Québec ($OR=1.39$; $p<.001$), which we analyze as the result of francophone Canada's proximity to anglophone North American social justice movements. We also find that person-centered language varies according to the kind of property, with properties which are the focus of social justice movements (like race and disability) being more likely to favour the person-centered variant (see Figure 1). Restricting our attention to singular nouns, we find that definiteness also conditions the alternation: indefinite NPs are much more likely to appear in the person-centered variant than definite ones ($OR: 3.55$, $p<.001$). We therefore conclude that property vs person-centered language offers new opportunities to study the interaction between grammatically conditioned and politically motivated syntactic variation.

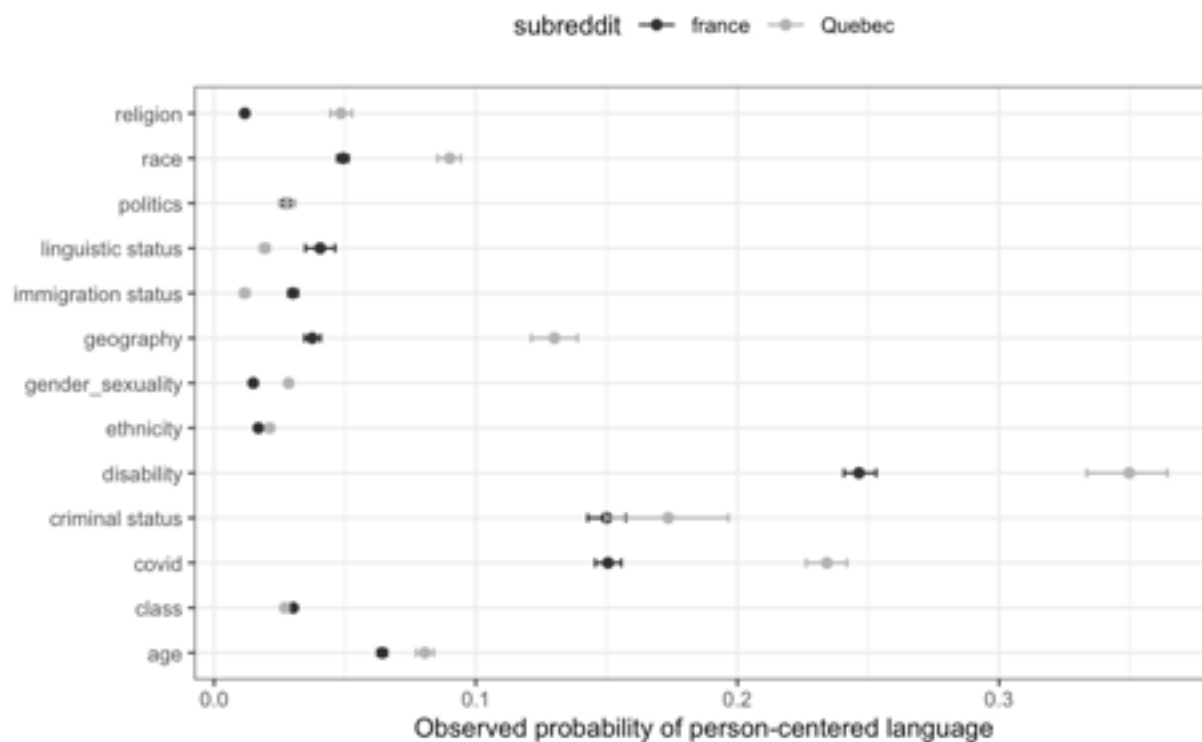


Figure 1.

References

- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1). <https://doi.org/10.18637/jss.v067.i01>
- Granello, D. H., & Gibbs, T. A. (2016). The power of language and labels: “the mentally ill” versus “people with mental illnesses”. *Journal of Counseling & Development*, 94(1), 31-40.

Wave, Wind, Tide, or Storm ? Metaphorical representations of the #MeToo movement in three Taiwanese newspapers

Pei-Ci Li & Sheng-Fu Wang
University of Lorraine & Academia Sinica, Taiwan

This study examines the metaphorical representations related to natural phenomenon used in the context of #MeToo movement in Taiwan through an analysis of three local newspapers. The movement #MeToo has emerged late in Taiwan in 2023, compared to other countries. One of the catalysts for this movement was the Taiwanese TV series *Wave Makers*, which addresses workplace sexual harassment. The series sparked considerable resonance and provided a shared social platform, empowering victims to express themselves. Furthermore, scandals involving several candidates prior to the presidential and parliamentary elections created a political crisis, prompting the ruling party to amend relevant laws and further encourage victims to speak out (Chen 2021, Chen & Huang 2023).

This article collects news containing the keyword *MeToo* from three major local newspapers in Taiwan: Liberty Times, United Daily, and China Times. A total of 3544 articles have been collected. We conduct a quantitative and qualitative analysis by combining corpus linguistics and critical metaphor analysis (CMA: Charteris-Black 2004, 2018). CMA is a blend of Conceptual Metaphor Theory (Lakoff & Johnson 1980) and Critical Discourse Analysis (Dijk 2015) that aims at identifying how metaphors are used as an argumentative tool for communicating attitudes, evoking emotions, and persuading audiences (Charteris-Black 2004). CMA also focuses on the predicative function, which highlights how socially sensitive issues, such as sexual harassment, are communicated (Cibulskienė 2020), reflecting speakers' positive or negative attitudes (Charteris-Black 2018, Musolff 2006).

Firstly, we analyze the frequency of the term *MeToo* in the three newspapers: from the global context in 2017 to the emergence of the movement in Taiwan in 2023, how does the frequency relate to social events during this period? What positions do these newspapers express?

Secondly, we analyze five metaphorical expressions that frequently co-occur with the keyword *MeToo*, along with their common collocations:

- | | | |
|-----|---|--------------------------|
| (1) | 浪 <i>lang</i>
wave
“wave” | 潮 <i>chao</i>
tide |
| (2) | 熱 <i>re</i>
hot
“wave” | 潮 <i>chao</i>
tide |
| (3) | 風 <i>feng</i>
wind
“wave” | 潮 <i>chao</i>
tide |
| (4) | 風 <i>feng</i>
wind
“wave/controversy” | 波 <i>bo</i>
wave |
| (5) | 風 <i>feng</i>
wind
“storm” | 暴 <i>bao</i>
violence |

Their main elements of these metaphors include *wave*, *wind*, and *tide*. Although all the expressions could be translated as “wave” or “movement” depending on the context, their significations and implications differ subtly. By analyzing both the global structure (theme) and the local structure (words and phrases), we identify frequent linguistic patterns in large amounts of data, while also enabling in-depth qualitative analysis.

The results show that these movements can be distinguished based on several aspects. First, the topic: the movement is categorized as social, political, legal, or entertainment news. Second, the scope: it is perceived either as a structural social movement or as individual gossip. Third, the intensity and duration: the strength and longevity of the movement are represented with varying emphasis. Fourth, the implied meaning: the movement is portrayed either as an active agent with a profound societal impact or as an ephemeral event. Lastly, the purpose: the movement is framed as aiming for social progress, political attack, or sensational reporting.

These differences reveal the perspective and attitude of the journalist or the newspaper toward the movement and how different newspapers with different political stances frame the movement and the implications for public perception and social change in Taiwan.

Keywords: #MeToo movement, Taiwan, critical metaphor analysis, corpus linguistics, media representation

References

- Charteris-Black, Jonathan (2004), *Corpus Approaches to Critical Metaphor Analysis*, New York, NY: Palgrave Macmillan
- Charteris-Black, Jonathan (2018), *Analysing political speeches: Rhetoric, discourse and metaphor*, Bloomsbury Publishing.
- Chen, Chao-ju (2021), The Limits and Power of Law: What the Absence of# MeToo in Taiwan Can Tell Us about Legal Mobilization. *Politics & Gender*, 17(3), 514-519.
- Chen, Chao-ju & Chang-ling Huang (2023), Can justice come after the victim’s belated credibility?, *Journal of Women’s Research*, (119), 34-47. [陳昭如, & 黃長玲. (2023). 在遲來的受害者公信力之後, 正義能隨之而到嗎?. 婦研縱橫, (119), 34-47.]
- Cibulskienė, Jurga (2020), Cross-linguistic metaphorical representation of the# MeToo movement: Communicating attitudes. *Respectus philologicus*, 38(43), 54-66.
- Lakoff, George & Mark Johnson (1980), *Metaphors We Live By*, University of Chicago press.
- Musolff, Andreas (2006), Metaphor scenarios in public discourse, *Metaphor and symbol*, 21(1), 23-38.
- Van Dijk, Teun A (2015), Critical discourse analysis, *The handbook of discourse analysis*, 466-485.

On the interaction of cognitive styles and social circles in syntactic change: the case of progressive [BE Ving]

Peter Petre

This talk discusses the weight of social and cognitive variables in syntactic change, and argues that inter-individual differences help understand its timing (the actuation problem). The case study is [BE Ving] ('the progressive'), and its growth in the 17th/18th centuries. Throughout this growth, its two sources, a verbal and a nominal one, continued to exert influence. Both go back to Old English, and are represented in (1) and (2).

- | | | |
|----------------------|-----------------------------------|--------------------------------------|
| (1) [BE Vende] | <i>Ic eam huntende (deor)</i> | 'I am hunting (deer)' |
| (2) [BE on Ving/ung] | <i>Ic eam on huntunge (deora)</i> | 'I am on hunting (of deer)' |

While (1) is verbal, coding objects as direct objects and assigning (active) voice, (2) inherits from the nominal gerund, coding the object as a postmodifier ('of deer'), and remaining underdetermined for voice (e.g. in *John is in a training* John is likely the trainee but may also be the trainer).

From late Middle English onwards, the two sources merge, the preposition *on* being reduced and reanalyzed as the prefix *a-*. While this blend boosted the unified [BE Ving] construction, it also led to increased complexity. One unexpected result was the emergence of the hybrid PASSIVAL, which was active in form but passive in meaning, as in (3):

- (3) *Whilst **Troy was Sacking** by the Greeks...* 'Whilst Troy was being sacked by the Greeks' (Dryden, 1693)
- (4) *... at the Debates about [those Bills]_i in the Privy-Council of England. Every Clause in them was so strenuously oppos'd and vindicated, that **they_i were debating** ('that they were being debated') in Council near Two Years.* (Oldmixon, 1716)

Some passivals were probably hard to parse even for contemporaries, such as (4), where *they* might (mistakenly) be taken to refer to the members of the privy-council.

To see how language users deal with this complex input I examine 25 individuals from the EMMA corpus (Petré et al. 2019). Analysis reveals two major groups: (a) 'Grammar-informed thinkers', who no longer use postmodifiers (the outgoing 'of deer' variant), and constrain the passival, aiming for a context-independent distribution of voice and subject type: e.g., they are fine with (3), where the *by*-agent formally marks the passive voice, but avoid the context-dependency of (4). Close scrutiny reveals that they not only share a cognitive style, but are also all educated and London-based. (b) 'Associative thinkers', who still display (relic) uses of nominal objects, and do not constrain the passival, relying on context for interpretation.

Generally, the results suggest that 'grammar-informed thinkers' systematically neo-analysed their input. This resulted in an overly constrained passival and the growing need for an alternative, which the progressive passive (*he was being bullied*) filled in the late 18th century. While therefore conservative towards incoming uses, they are quicker to abandon outgoing uses. Associative thinkers, by contrast, promote incoming uses more (though here their influence was not sufficient). The evidence complements work in the vein of Dąbrowska (2012), which argues that some language users have a higher impact in shaping community grammar, but without differentiating between outgoing and incoming variants.

References

- Petré, Peter, Lynn Anthonissen, Sara Budts, Enrique Manjavacas, Emma-Louise Silva, William Standing & Odile A. O. Strik. 2019. Early Modern Multiloquent Authors (EMMA): Designing a large-scale corpus of individuals' languages. *ICAME Journal* 43. 83–122.
- Dąbrowska, Ewa. 2012. Different speakers, different grammars: Individual differences in native language attainment. *Linguistic Approaches to Bilingualism* 2(3). 219–253.

Sociolinguistic variation through a typological lens: relative clauses in spoken Italian

Silvia Ballarè (University of Bologna)

Italian displays a complex set of relativization strategies: different syntactic functions can be realized through different strategies and more strategies can be employed for the same syntactic function (Cerruti 2017). From a structural point of view, these strategies include: (i) relative (1) and (ii) resumptive pronouns (2), (iii) gap constructions (3) (Comrie & Kuteva 2013a, 2013b) and (iv) double encoding (4) (Murelli 2011). Not surprisingly, they display a different sociolinguistic characterization and, in this paper, I aim to thoroughly analyze the factors underlying these patterns of variation empirically.

(1) KIParla, TOD1008

<i>Fa</i>	<i>parte</i>	<i>di</i>	<i>una</i>	<i>collana</i>	<i>della</i>	<i>quale</i>	<i>fa</i>
do:PRS.3SG	part	of	INDEF:SG.F	series:F.SG	of.DEF:SG.F	REL	do:PRS.3SG
<i>parte</i>	<i>anche</i>	<i>questo</i>	<i>libro</i>	<i>qui</i>			
part	also	DEM.M.SG	book	here			

‘It is part of a series which also includes this book here’

(2) KIParla, PTD012

<i>Una</i>	<i>strada [...]</i>	<i>dove</i>	<i>ci</i>	<i>passa</i>	<i>molta</i>	<i>più</i>	<i>gente</i>
INDEF:F.SG	street:SG.F	REL	LOC	pass.by:PRS.3SG	much:SG.F	more	people:SG.F

‘A street where much more people pass by’

(3) KIParla, PBA017

<i>Il</i>	<i>giorno</i>	<i>che</i>	<i>è</i>	<i>libero</i>	<i>lui</i>	<i>lavoro</i>	<i>io</i>
DEF:SG.M	day:SG.M	REL	be:PRS.3SG	free	OBJ.3SG.M	work:PRS.1SG	SUBJ.1SG

‘The day he is free, I work’

(4) itTenTen20

<i>Sembravo</i>	<i>un</i>	<i>bambino</i>	<i>a</i>	<i>cui</i>	<i>gli</i>	<i>era</i>	<i>stato</i>
seem:PST.1SG	INDEF:SG.M	child:SG.M	to	REL	DAT.3SG	AUX.3SG	PST.PTCP
<i>fatto</i>	<i>il</i>	<i>regalo</i>	<i>che</i>	<i>da</i>	<i>sempre</i>	<i>aveva</i>	
do:PST.PTCP	DEF:SG.M	gift:SG.M	REL	from	always	AUX.3SG	

desiderato
desire:PST.PTCP

‘I looked like a child who had been given a gift that he had always desired’

This study examines relativization strategies in spoken Italian through a typological lens, given that “the patterns of variation and change found in [...] a particular language are in many cases simply instances of patterns of variation and change found across languages” (Croft 2022: 27).

The analysis draws from the KIParla corpus (Mauri et al. 2019), by contrasting subcorpora representing speakers with high (A) and low (B) educational achievements (335.916 tokens). All the relativizing elements are extracted (n= 2.898) and manually annotated according to several linguistic parameters (structural strategy; syntactic function; restrictive/non-restrictive value; argument/adjunct nature of the relativized element; definiteness of the antecedent). Data are analyzed by adopting a conditional inference tree and a random forest (Tagliamonte & Baayen 2012), to shed light on the influence of extra- and intra-linguistic factors over the selected structural strategy.

Key findings (Ballarè 2024) reveal that A speakers use a wider variety of strategies, favoring more explicit case-marked structures; while B speakers rely predominantly on the gap strategy. More precisely, while nominative and accusative relativizations exhibit uniformity across social groups, oblique relativizations display significant variation in substandard realizations: while B speakers produce them by simplifying the structure and over-extending the gap strategy, A speakers realize them complexifying the structure and employing the double-encoding strategy.

The analysis highlights the interplay between syntactic function and relativization choices, showing that, with respect to other obliques, genitives and datives, due to their syntactic and semantic connections to the

main clause, are more frequently expressed through case-marked strategies, contradicting the explicitness scale proposed by Comrie (1989).

These patterns suggest that sociolinguistic factors, alongside linguistic typology, play a critical role in shaping the variability of relative clause constructions in Italian.

References

- Ballarè, Silvia. 2024. Relativization strategies and sociolinguistic variation in spoken Italian: a typological account. *Linguistic Typology at the Crossroads* 2024(2).125-157.
- Cerruti, Massimo. 2017. Changes from below, changes from above. Relative constructions in contemporary Italian. In Massimo Cerruti, Claudia Crocco e Stefania Marzo (eds.), *Towards a new standard. Theoretical and empirical studies on the restandardization of Italian*, 32-61. Berlin-New York: Mouton de Gruyter.
- Comrie, Bernard. 1989. *Language universals and linguistic typology. Second edition*. Chicago: University of Chicago Press.
- Comrie, Bernard & Tania Kuteva. 2013a. Relativization on Subjects. In Matthew Dryer & Martin Haspelmath (eds.), *World atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at <http://wals.info/chapter/122>, (Accessed 2025-01-09).
- Comrie, Bernard & Tania Kuteva. 2013b. Relativization on Obliques. In Matthew Dryer & Martin Haspelmath (eds.), *World atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at <http://wals.info/chapter/123> (Accessed 2024-01-09.)
- Croft, William. 2022. *Morphosyntax. Constructions of the World's Languages*. Cambridge: Cambridge University Press.
- Mauri, Caterina, Silvia Ballarè, Eugenio Gorla, Massimo Cerruti & Francesco Suriano. 2019. KIParla corpus: a new resource for spoken Italian. In Raffaella Bernardi, Roberto Navigli & Giovanni Semeraro (eds.), *Proceedings of the 6th Italian conference on Computational Linguistics CLiC-it*, Torino, Accademia University Press. <http://ceur-ws.org/Vol-2481/paper45.pdf>
- Murelli, Adriano. 2011. *Relative constructions in European non-standard varieties*. Berlin/New York: Mouton de Gruyter.
- Tagliamonte, Sali A. & R. Harald Baayen. 2012. Models, forests and trees of York English: Was/were variation as case study for statistical practice. *Language variation and change* 24(2). 135-178.

Gender marked job titles: A sociolinguistic survey of the attitude of Ukrainian and Modern Greek speakers to gender marked language units

Svitlana Pereplotchykova

Taras Shevchenko National University of Kyiv

Key words: gender marked job titles, Modern Greek, Ukrainian, sociolinguistic survey

Recently the Ukrainian society has witnessed a rather rapid increase in the amount of gender marked job titles used. One of the obvious, objective reasons is that due to the war Ukrainian women have had to undertake male roles in various professions as well as in the social spheres. Unlike their grandmothers who experienced similar consequences of the WWII, nowadays many Ukrainian women choose to be visible. This resulted in the necessity to introduce specific nominations, the appearance and wide usage of gender marked job titles, referred to in Ukrainian media and academic sphere as feminitives. However, not all the speakers eagerly, immediately or completely have accepted gender marked job titles with lots of ongoing discussions and debates on the issue.

The alike reaction of the society is still witnessed in Greece in which the appearance of new gender marked language units aiming at making women visible and equal with male members of the society accelerated because of the international economic crisis of 2009, and the Greece's crisis of 2015, after which the Greek family and the role of a Greek woman in the society changed.

This talk presents the results of the sociolinguistic surveys conducted among Ukrainian and Modern Greek speakers in 2024-2025. The Greek survey has one additional question, which is connected with some typological differences of the languages. According to the surveys about 44% of the respondents, speakers of Ukrainian, approve the usage of gender marked job titles, unlike 15% of Greek speakers who participated in the survey.

The arguments mentioned by the respondents in favour of or against the usage of gender marked job titles show their attitude to the place of a woman. In particular Ukrainian respondents refer to the crucial role of gender marked job titles in abandoning the norms of the patriarchal society with a leading role of a man. At the same time Greek respondents, who also referred to the patriarchal society still dominating in Greece, believe that the usage of gender marked job titles does the opposite. Introduction of gender marked units highlights the absence of gender equality, while the usage of masculine nouns for females makes women equal to men. Also, some amount of the respondents, speakers of both Ukrainian and Modern Greek, tend to stick to the types *φίλολογος* and *η φιλόλογος* as they believe them to be strongly associated with professionalism. In Modern Greek it is connected with the Katharevousa origin of the type (Αλβανούδη 2024), in Ukrainian it stresses the equality with male philologists (Дацишин 2020).

References

- Αλβανούδη, Αγγελική (2024, March 20), Τα προβλήματα μιας μεταδιγλωσσικής γλώσσας, in *Πολίτης: Παράθυρο*. <https://parathyro.politis.com.cy/stiles/762523/ta-provlimata-mias-metadiglossikis-glossas?fbclid=IwAR31CGuxyiVAXDQIY7T0Sceg-78taCd2jOwDHEacoA5kpdR4XJZ5CoAyUKo>
- Дацишин, Христина (2020), Засоби масової інформації і динаміка мовної норми: новітні фемінітиви як відображення суспільних викликів сьогодення. *Вісник Національного університету «Львівська політехніка»: журналістика», 4, 183-191.* <https://doi.org/10.23939/sjs2020.01.183>

Intensification of Adjectives: The Case of Estonian Young People's Language

Triin Aasa

(University of Tartu)

Intensifiers are used to scale a word's degree of quality upward or downward (Bolinger 1972: 17, Quirk et al. 1985: 590–599). For example, saying *I was extremely frustrated* instead of *I was frustrated* amplifies the quality of *frustrated* and saying *I was a little frustrated* diminishes it. Intensifiers are constantly changing, as they lose their intensity in use and need to be replaced by new stronger words (Stoffel 1901: 2). The versatility and constant renewal make intensifiers a popular topic in language variation and change research (Tagliamonte, 2012: 320).

Research in various languages has revealed that young people use intensifiers the most (e.g. Ito & Tagliamonte 2003; Stratton 2020; Stratton & Sundquist 2022). Although there are some studies on intensifiers in Estonian (e.g. Kotilainen 2006; Metslang 1997; Petron 2023), research on usage patterns in young people's language is insufficient. Moreover, despite change being a key feature of intensifiers, no studies have explored this aspect in spoken Estonian. The aim is to fill gaps in local research, continue the tradition of intensifier studies and contribute to the wider field of intensifier research, which has predominantly focused on Indo-European languages.

This study investigates intensifier usage in Estonian young people's language from 1996–2004 to 2020–2021, examining general patterns and change over time. Using the chi-squared test, standardised Pearson's residuals, Cramér's V, and logistic regression mixed models, I analyse and compare spoken conversations from 14–27-year-olds. More specifically, I focus on vocabulary, intensification rate, and the functions of intensifiers. Additionally, to identify the circumstances in which intensifiers are most often used, I analyse the effects of the gender of the speaker, syntactic position, semantic category, and the type of modified adjective. The data is drawn from the Corpus of Spoken Estonian (CoSE) and the Estonian Teenage Language Corpus (TeKE). The conversations from CoSE consist of 31 612 words (111 speakers) and the conversations from TeKE consist of 203 512 words (39 speakers) in total. This study tests three hypotheses: (1) intensifiers in Estonian function similarly to those in Indo-European languages (e.g. Stratton 2020; Stratton & Sundquist 2022); (2) intensifier vocabulary has changed, other aspects remain relatively stable; (3) intensifiers borrowed from English appear only in the newer data (TeKE).

The results show that the changes are lexical. The most frequently used intensifiers in CoSE are *nii* 'so', *väga* 'very' and *hästi* 'well', while in TeKE, they are *nii* 'so', *väga* 'very' and *suht* 'relatively'. Additionally, six intensifiers borrowed from English occurred in the TeKE conversations (*kind of*, *fucking*, *so*, *full*, *hella*, *next level*), while none appeared in CoSE. The intensification rate has remained consistent (31.9% in CoSE, 34.4% in TeKE).

Analysis of independent variables' effect shows that intensifiers most often modify scalar adjectives, predicative adjectives, and adjectives in independent phrases, words from semantic categories *human propensity* (e.g. *smart*, *happy*) and *evaluation* (e.g. *beautiful*, *difficult*). Amplifying intensifiers

appeared most frequently. These patterns have remained consistent over time and support the hypothesis of similarity between Estonian and Indo-European intensifiers in the aspects analysed.

References

Bolinger, D. (1972). *Degree Words*. Berlin/Boston, Germany: De Gruyter, Inc.

CoSE = Corpus of Spoken Estonian. 10.15155/1-00-0000-0000-0000-00077L.

Ito, R., & Tagliamonte, S. (2003). Well Weird, Right Dodgy, Very Strange, Really Cool: Layering and Recycling in English Intensifiers. *Language in Society*, 32(2), 257–279.

Kotilainen, J. (2006). *Adjektiivide intensiivisusprefiksoidide moodustamine ja kasutus eesti keeles* (Bachelor's thesis). University of Tartu.

Metslang, H. (1997). Maksimaalsuse ja minimaalsuse väljendamisest eesti püsiühendites. *Pühendusteos Huno Rätsepale, Tartu Ülikooli eesti keele õppetooli toimetised* 7, 139–154.

Petron, E. (2023). *Omadussõnad ja intensiivistajad eesti eakate keeles* (Master's thesis). University of Tartu.

Quirk, R, Greenbaum, S, Leech, G, Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. London, New York: Longman.

Stoffel, C. (1901). *Intensives and down-toners; a study in English adverbs*. Heidelberg, C. Winter's universitätsbuchhandlung.

Stratton, J. M. (2020). Adjective Intensifiers in German. *Journal of Germanic Linguistics*, 32(2), 183–215. <https://doi.org/10.1017/S1470542719000163>

Stratton, J. M., & Sundquist, J. D. (2022). A Variationist Sociolinguistic Analysis of Intensifiers in Oslo Norwegian. *Journal of Germanic Linguistics*, 34(4), 385–419. <https://doi.org/10.1017/S1470542722000022>

Tagliamonte, S. A. (2012). *Variationist Sociolinguistics*. Wiley-Blackwell.

TeKE = Vihman, Virve-Anneli, Maarja-Liisa Pilvik, Aive Mandel, Annika Kängsepp, Mari Aigro, Kadri Koreinik, Kristiina Praakli, Liina Lindström. (2023). Estonian Teen Language Corpus v.1.0. Institute of Estonian and General Linguistics, University of Tartu. <https://doi.org/10.23673/re-455>.

General Session : Syntax

On intensional genitive in Polish

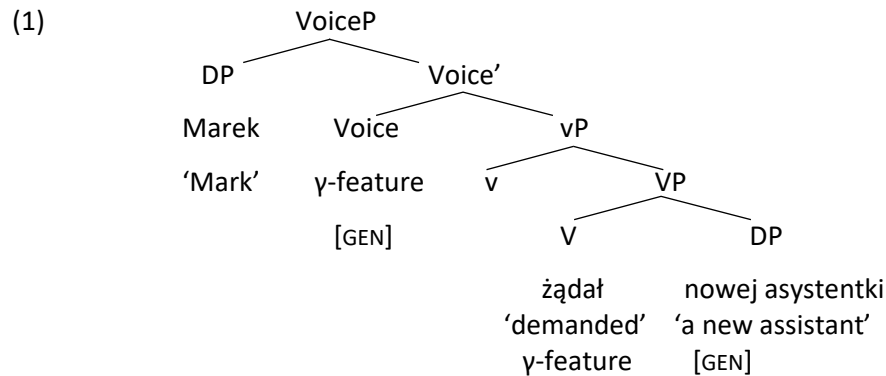
Anna Bondaruk & Anna Prażmowska
The John Paul II Catholic University of Lublin, Poland

Keywords: intentional verbs, intensional genitive, structural case, the Minimalist Program, the Polish language

Cross-linguistically, intensional verbs display intensionality effects in their complement, including: (i) the availability of non-specific readings, (ii) the failure of truth preservation under extensional substitution, and (iii) the lack of existential import (Moltmann 1997, Larson 2002, Hinzen et al. 2014, Schwarz 2021). In Balto-Slavic languages, such as Lithuanian (Šereikaitė 2020, Sigurðsson and Šereikaitė 2024), Russian, Ukrainian and Polish (Kagan 2013), complements of intensional verbs may be marked for genitive case, which is then called intensional genitive.

In this paper, we examine the distribution of intensional genitive in Polish and provide a syntactic analysis of this case, couched in the Minimalist Program of Chomsky (2000, 2008). In the literature, there is scanty information about the distribution of intensional genitive in Polish (Heinz 1955, Rudzka-Ostyn 2000), and no syntactic analysis of this case is available. The verbs that select genitive complements in Polish were first extracted from *Walenty*, an online valency dictionary of Polish (<https://walenty.ipipan.waw.pl>). All these verbs were then manually screened to guarantee that only verbs with genitive complements have been selected. Afterwards, these verbs were subjected to the three intensionality diagnostics, mentioned above. This has yielded 49 intensional verbs, which can take complements marked for intensional genitive and which belong to the following four semantic classes (Schwarz 2021): (i) verbs of absence: *brakować* ‘lack’, *potrzebować* ‘need’, *żądać* ‘demand’, etc., (ii) verbs of desire and volition: *chcieć* ‘want’, *pragnąć* ‘desire’, *życzyć* ‘wish’, etc., (iii) verbs of search: *szukać* ‘seek’, *upatrywać* ‘search for’, etc., and (iv) verbs of expectation: *spodziewać się* ‘expect’, *oczekiwać* ‘expect’, *czekać* ‘wait’, etc.

We argue that intensional genitive in Polish represents structural case. Although intensional genitive rarely alternates with accusative (this is only possible with *chcieć* ‘want’ and *potrzebować* ‘need’), it shows properties typical of structural case, viz. (i) the genitive complement may be replaced with *dużo* ‘a lot’-type phrases, found only in structural case positions (Przepiórkowski 1999: 112), (ii) the genitive complement may be replaced with distributive *po*-phrases, restricted to structural case positions (Franks 1995, Przepiórkowski 1999), and (iii) the genitive complement may turn into nominative under passivisation. Intensional genitive thus contrasts with inherent genitive, which shows opposite behaviour as regards the three properties mentioned above. We propose that with transitive intensional verbs, intensional genitive is assigned in a way similar to accusative, viz. by an active Voice with a thematic external argument. However, the Voice head that assigns intensional genitive is equipped with a special feature, which forces it to select a special class of predicates (for an idea that there is selection between Voice heads and verbal roots, cf. Alexiadou et al. 2008 and Šereikaitė 2020). Intensional case assignment is schematised in (1):



In (1), the Voice is equipped with an uninterpretable γ -feature, which is checked by the corresponding feature on the verb via Agree (Chomsky 2000). The Voice with the γ -feature combines with genitive case, which is then assigned to the complement DP *nowej asystentki* 'a new assistant'. If there is no Agree between the verb and Voice, the Voice can only value accusative.

References

- Alexiadou, Artemis, Elena Anagnostopoulou and Florian Schäfer (2008), The properties of anticausatives crosslinguistically, in M. Frascarelli (ed.), (2008), *Phases of interpretation*, Berlin: Mouton de Gruyter, 187–211.
- Chomsky, Noam (2000), Minimalist inquiries: The framework, in R. Martin, D. Michaels, and J. Uriagereka (eds.), (2000), *Step by step*, Cambridge, Mass.: MIT Press, 89–155.
- Chomsky Noam (2008), On phases, in R. Freidin, C.P. Otero, and M. L. Zubizarreta (eds.), *Foundational issues in linguistics theory. Essays in honour of Jean-Roger Vergnaud*, Cambridge, Mass.: MIT Press, 134–166.
- Franks, Steven (1995), *Parameters of Slavic morphosyntax*, Oxford and New York: Oxford University Press.
- Heinz, Adam (1955), *Genetivus w indoeuropejskim systemie przypadkowym*. [Genitive in the Indoeuropean case system], Warszawa: Państwowe Wydawnictwo Naukowe.
- Hinzen, Wolfram, Michelle Sheehan and Ulrich Reichard (2014), Intensionality, grammar and the sententialist hypothesis, in P. Kosta, S. L. Franks, T. Radeva-Bork, and L. Schrücks (eds.), *Minimalism and beyond: Radicalizing the interfaces*, Amsterdam: John Benjamins, 315–349.
- Kagan, Olga (2013), *Semantics of genitive objects in Russian. A study of genitive of negation and intensional genitive case₂*, Dordrecht: Springer.
- Larson, Richard K. (2002), The grammar of intensionality, in G. Preyer, and G. Peter (eds.), *Logical Form and language*, Oxford: Oxford University Press, 228–262.
- Moltmann, Frederike (1997), Intensional verbs and quantifiers, *Natural Language Semantics* 5(1), 1–52.
- Przepiórkowski, Adam (1999), *Case assignment and the complement-adjunct dichotomy: A non-configurational constraint-based approach*, Tübingen: Universität Tübingen dissertation.
- Rudzka-Ostyn, Brygida (2000), *Z rozważań nad kategorią przypadku* [Reflections on the category of case] (translated by Elżbieta Tabakowska), Kraków: Universitas.
- Schwarz, Florian (2021), Intensional transitive verbs “I owe you a horse”, in D. Gutzmann, L. Matthewson, C. Meier, H. Rullmann, and T. E. Zimmermann (eds.), *The Wiley Blackwell companion to semantics*, Oxford: Wiley-Blackwell.

- Sigurðsson, Einar Freyr and Milena Šereikaitė (2024), The dual face of structural object case: On Lithuanian genitive of negation, *Journal of Linguistics* 60(1), 161–212.
- Šereikaitė, M. (2020), Voice and case phenomena in Lithuanian morphosyntax. Ph.D. dissertation, University of Pennsylvania.

Clausal Complements of ‘Perception’ Verbs in English – From Perception to Cognition

Christelle Lacassain
(Sorbonne University, CELISO, Paris)

Keywords: perception verbs; clausal complementation; physiological perception; perceptual inference; cognition

In English, transitive perception verbs accept all types of complements: noun phrases (1), preposition phrases (2), finite subordinate clauses (content clauses, interrogatives and fused relatives, as in [3-7]) and non-finite subordinate clauses (8-12). Copular perception verbs license the same types of constituents as predicative complements, to which can be added adjective and adverb phrases.

This study, which is situated at the syntax-semantics interface, focuses on the clausal complementation of transitive perception verbs, more particularly non-agentive verbs, in contemporary English. Indeed, not all perception verbs take all possible types of complements (Viberg 1984; Huddleston & Pullum 2002; Lacassain-Lagoin 2020), except for *see* (which, however, does not have a copular use, unlike *feel*, which accepts almost all types of complements). For instance, agentive perception verbs hardly license a finite complement clause, particularly a *that*-content clause (Lacassain-Lagoin 2014). To account for the syntactic and semantic differences in clausal complementation for transitive perception verbs, a qualitative, corpus-based analysis is conducted, using the *British National Corpus* and the *Corpus of Contemporary American English*, as well as occurrences from a personal database of extracts from different novels.

It is observed that certain types of clausal complements are favoured in sentences expressing perception (non-finite subordinate clauses, for example, as in [8-10]), while others are favoured in sentences denoting perceptual inference, or even cognition ([5-7]). Dealing with non-finite clausal complements, many linguists (e.g. Jespersen 1909-1949; Declerck 1981; Quirk et al. 1985; Felser 1999) consider that the difference between the three types of non-finite subordinate clauses in (8-10) is merely aspectual. However, there is more to it than that; other inferences may be present, as shown by Kirsner and Thompson (1976).

It is then argued that on the one hand, the form of the clausal complement is motivated by the type of – perceptive or cognitive – object present in the experience or the mental representation and that, on the other hand, the varied meanings of both verbs and sentences result from the association between a given verb and such or such type of complement. The semantic range of verb meanings, from perception to concepts such as inference, knowledge or judgement, seems to align with the use of finite and non-finite subordinates as complements to verbs that express an original meaning of non-agentive perception.

My analysis shows that sentences with a physiological perception (rather than ‘direct perception’) reading can feature three types of non-finite clausal complements, and thus display various shades of meaning, whereas sentences with a perceptual inference or cognition (rather than ‘indirect perception’) reading mainly feature finite subordinates, but also *to*-infinitivals. It can be concluded that syntax is motivated by meaning, representation and conceptualisation of the situation (following Cotte 2006, 2012), and that the label ‘perception’ verb is therefore not always appropriate; nor are those of ‘direct perception’ and ‘indirect perception’, along the lines of Austin (1962).

Examples

- (1) The woman nodded and **touched** something under the table. (COCA)
- (2) Jay **peered** at Kevin’s computer screen, checking the details and looking through other possibilities, hoping against hope that Kevin was mistaken. (COCA)

- (3) As the weeds were growing, and among them, the lilacs in full bloom, I discovered all these goings on by accident, when I **noticed** where Gogu was keeping the things he had stolen. From there he was distributing them whenever he found buyers. (COCA)
- (4) Please would you look into the matter and **see** what can be done to improve the situation. (BNC)
- (5) "I'd like you to look at some photos, Miranda, to **see** if you can identify the girl who called herself Susan." (COCA)
- (6) She looked to **see** whether her words had offended Sharpe, and was evidently reassured. (BNC)
- (7) I was sorry to **hear** that my uncle, my only surviving relation, was dead. However, the inheritance would give me independence for life, and I was glad of that. (BNC)
- (8) Ty **watched** her go for a moment, before straightening up and turning to his girlfriend. (COCA)
- (9) Vincent **could smell** the gravy cooking in the house and he was meanly hungry, but he didn't want to be rude to Grandpa. (COCA)
- (10) Tony nodded; his senses felt drained. He **felt** his body pulled out of the simulator by gentle hands, and he **saw** the real world replace the virtual world just before he fell into unconsciousness. (COCA)
- (11) Gudrun detected the tang of mockery in him, and she looked up and smiled into his face. He **felt** his nerves caressed. (WIL)
- (12) When they came to the woman's camp, Inman **saw** it to be a construction that had evidently begun life nomadic but had taken root. It was a little rust-colored caravan standing in a clearing among the canted trees. (COCA)

References

- Austin, John Langshaw (1962), *Sense and Sensibilia*, Comp. G. J. Warnock, Oxford: Oxford University Press.
- Cotte, Pierre (2006), De la cognition à la syntaxe: le complément dans la genèse du sens, in C. Delmas (ed.), (2006), *Complétude, cognition, construction linguistique*, Paris: Presses de la Sorbonne Nouvelle, 29-34.
- Cotte, Pierre (2012), Hiérarchies, *E-rea* 9.2, URL: < <http://erea.revues.org/2652> >.
- Davies, Mark (2004), *British National Corpus* (from Oxford University Press), URL: < <https://www.english-corpora.org/bnc/> >. (BNC)
- Davies, Mark (2008-), *The Corpus of Contemporary American English*, 1990-present, URL: < <https://corpus.byu.edu/coca/> >. (COCA)
- Declerck, Renaat (1981), On the Role of Progressive Aspect in Non-Finite Perception Verb Complements, *Glossa* 15.1, 83-114.
- Felser, Claudia (1999), *Verbal Complement Clauses. A Minimalist Study of Direct Perception Constructions*, Amsterdam: John Benjamins Publishing Company.
- Huddleston, Rodney D. and Geoffrey K. Pullum (eds) (2002), *The Cambridge Grammar of the English Language*, Cambridge: Cambridge University Press.
- Jespersen, Otto (1909-1949), *A Modern English Grammar on Historical Principles*, 7 vols, London: G. Allen & Unwin Ltd / Copenhagen: Ejnar Munksgaard.
- Kirsner, Robert S. and Sandra A. Thompson (1976), The Role of Pragmatic Inference in Semantics: A Study of Sensory Verb Complements in English, *Glossa* 10.2, 200-240.
- Lacassain-Lagoin, Christelle (2014), Perception, cognition et perspicacité: énoncés à lecture inférentielle avec les verbes de perception non agentive, in F. Buisson, C. Lacassain-Lagoin and F. Marie (eds.), (2014), *Perception, Perspective, Perspicacité / Perception, Perspective and Perspicacity*, Paris: L'Harmattan, 55-75.
- Lacassain-Lagoin, Christelle (2020), La complémentation propositionnelle des verbes de "perception" en anglais contemporain: de la perception à la cognition, in I. Thomières-Shakovskaya (ed.), (2020), *La*

perception: langue, discours, cognition, Actes du colloque international 6-7 décembre 2019, Paris-Sorbonne Université, 56-68.

Quirk, Randolph, Sydney Greenbaum, Geoffrey Leech and Jan. Svartvik (1985), *A Comprehensive Grammar of the English Language*, London: Longman.

Viberg, Åke (1984), The verbs of perception: a typological study, in B. Butterworth, B. Comrie and Ö. Dahl (eds.), (1984), *Explanations for Language Universals*, Berlin / Boston: De Gruyter Mouton, 123-162.

Valency (mis)matches in Baltic languages: Evidence from parallel subtitles

Daria Alfimova
(University of Potsdam)

Keywords: valency patterns, parallel corpus, Baltic, language contact, transitivity

Latvian and Lithuanian are two Eastern Baltic languages within the larger Balto-Slavic clade that are not mutually intelligible, although their speakers have inhabited neighboring territories for centuries. The use of verbs and valency patterns appears to play a significant role in the lack of mutual intelligibility. This study examines what features trigger mismatches in Baltic valency patterns based on data extracted from the parallel corpus of film subtitles ParTy-1.0 (Levshina 2016), and reveals how they compare to those in the major contact languages — Russian, Polish, and German.

The original English script contains 2105 lines, 755 of which are unique. The Latvian, Lithuanian, German, Polish and Russian lines are parallelly aligned with the English ones. Manual annotation identified 393 clauses as comparable between Latvian and Lithuanian, based on one crucial assumption: the overt coding of corresponding referents in both languages. Valency is therefore understood in its broad sense, i.e. including adjuncts (see Forker 2014, Lander 2015, Grossman & Witzlack-Makarevich 2019), and is formally defined as the set of all referable participants of a single lexical verb. Among other features, I annotate Latvian and Lithuanian valency patterns; etymological relation (if Latvian and Lithuanian verb roots are of the same origin); prefixation (if both Latvian and Lithuanian verbs have any prefixes). In case of a Baltic valency mismatch, I annotate the corresponding valency pattern from the German, Polish, and Russian lines (if available).

I tested three hypotheses. H1: *“Parallel cognates exhibit fewer valency mismatches than etymologically unrelated verbs”* was not confirmed ($\chi^2 = 1.6902$, $df = 1$, $p > 0.1$), see Figure 1. The difference between cognates and non-cognates is insignificant for explaining Baltic valency mismatches. This means that the genealogical effect, when considered as an explanatory factor for valency correspondences, does not necessarily imply that the languages retain a large number of verbs of shared origin and matching valencies. Rather, it indicates that they tend to follow similar principles in assigning marking devices to specific semantic roles.

H2: *“Parallel prefixation increases the chances of a valency match”* was confirmed ($\chi^2 = 5.0142$, $df = 1$, $p < 0.03$), see Figure 2. The prefix type does not appear to be a significant factor, since all prefixes—semantic or aspectual—promote valency matching. The Baltic languages maintain the Slavic-style aspectual system (Arkadiev 2012) with highly productive aspectual prefixes (which are often word-formational at the same time). So, the most frequent prefix

correspondences *at-/at-*, *iz-/su-*, *no-/pa-*, *pa-/pa-*, *no-/nu-* are strongly associated with resultative meaning and used with the accusative case. This emphasizes the crucial role of verb semantics for the valency analysis and the cross-linguistic comparison, confirming the close relationship between resultativeness and transitivity (Levin 2015, Goddard 2015).

H3: “If a valency mismatch occurs between the two Baltic languages, there is a greater-than-chance probability that one of them will match with at least one of its contact languages” was confirmed for Lithuanian, which matches with at least one of the contact languages—while not matching with Latvian—in 51% of all mismatches (42 clauses out of 82). Latvian does so in 13% (11 clauses). Two isoglosses can be drawn between Latvian and the four other languages based on (1) Goal/Target role encoding and (2) transitive verb sets. At the same time, Lithuanian valency patterns involving the instrumental case account for more than half of the overall mismatches (N = 19) among all five languages, highlighting the distinctive organization of valency classes within particular languages.

These results open up a discussion on the intricate interplay of intra- and cross-linguistic factors shaping valency distribution in authentic conversational contexts. A usage-based perspective is essential for understanding the distribution of valency classes and uncovering important language-specific characteristics.

Figure 1. Parallel use of cognates in Baltic

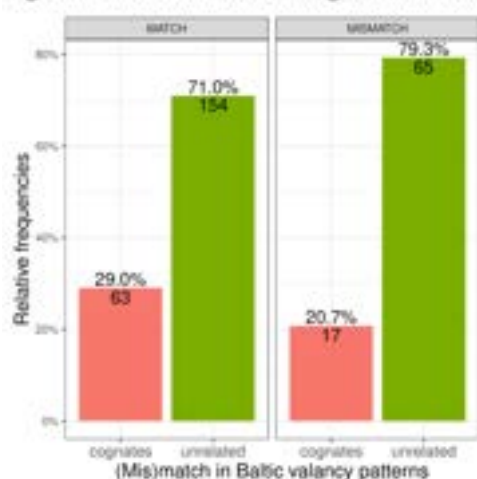
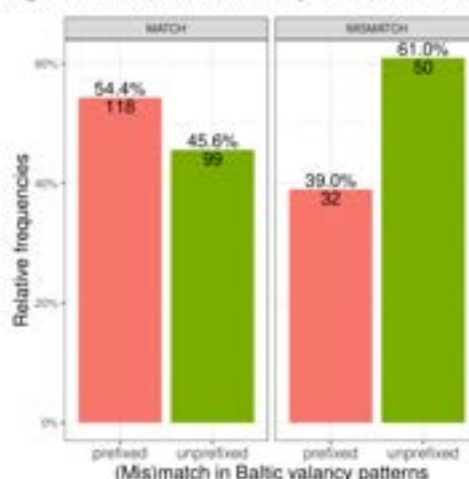


Figure 2. Parallel use of prefixes in Baltic



References

Arkadiev, Peter M. (2012), Аспектуальная система литовского языка (с привлечением ареальных данных) [Aspectual system of Lithuanian (on the areal background)]. In Vladimir Plungian (ed.), *Исследования по теории грамматики. Выпуск 6: Типология аспектуальных систем и категорий* [Studies in the Theory of Grammar. Vol. 6: Typology of Aspectual Systems and Categories] = *Acta Linguistica Petropolitana* 8–2, 2012, 45–121.

Forker, Diana (2014), A canonical approach to the argument/adjunct distinction. *Linguistic Discovery*, 12 (2), 27–40.

Goddard, Cliff (2015), 40. Verb classes and valency alternations (NSM approach), with special reference to English physical activity verbs". In Andrej Malchukov and Bernard Comrie (eds.), *Case Studies from Austronesia, the Pacific, the Americas, and Theoretical Outlook*, Volume 2, Berlin, München, Boston: De Gruyter Mouton, 1671–1702.

Grossman, Eitan and Witzlack-Makarevich, Alena. 2019. Valency and Transitivity in Contact: An Overview. *Journal of Language Contact*, 12(1), 1–26.

Lander, Yury (2015), Arguments and adjuncts in morphology and syntax of West Circassian (Adyghe) [Актанты и сирконстанты в морфологии и в синтаксисе адыгейского языка]. *Вестник РГГУ. Серия: История. Филология. Культурология. Востоковедение* (1): 7–31.

Levin, Beth (2015), 39. Verb classes within and across languages. In Andrej Malchukov and Bernard Comrie (eds.), *Case Studies from Austronesia, the Pacific, the Americas, and Theoretical Outlook*, Volume 2, Berlin, München, Boston: De Gruyter Mouton, 1627–1670.

Levshina, Natalia (2016), Verbs of letting in Germanic and Romance languages: A quantitative investigation based on a parallel corpus of film subtitles. *Languages in Contrast* 16(1): 84–117.

ParTy-1.0. Parallel Corpus for Typology. Available online at <https://github.com/levshina/ParTy-1.0> (accessed on 15 April 2024).

Retentions and refunctionalization of Spanish copular verbs in Zamboanga Chabacano

Eduardo Tobar
(Escuela Oficial de Idiomas de Vigo)

Keywords: copulas, grammaticalization, creoles, Chabacano, refunctionalization

Zamboanga Chabacano (ZC) is a Spanish-lexified creole spoken in the southern Philippines. This study draws on data from social media, other online sources, and a few available publications to examine the evolution of copulas of Spanish origin in this language. ZC is predominantly characterized as a zero-copula language (Forman 1972:160), a feature likely reflecting the influence of Philippine languages while also aligning with broader trends in creole languages. Example (1) illustrates a typical equational construction in ZC, marked by the absence of a copula and a predicate-subject word order, mirroring the structure of its Philippine substrate/adstrate languages:

- (1) *Buen amígo éle di-mío*
 good friend 3SG of-1SG.POSS
 ‘He/She is a good friend of mine.’

The overall picture is, however, more nuanced. On the one hand, Zamboanga Chabacano (ZC) has incorporated the particles *amó* from Bisaya and *ay* from Tagalog, which can optionally function as copulas, marking word order in subject-predicate constructions (Aoto 2002). On the other hand, some Spanish copular verb forms have been retained in specific contexts, assuming additional meanings and functions in some cases. This study provides an overview of the distinct pathways of functional, pragmatic, formal, and semantic change of these remnants of third-person singular forms. Specifically, we examine the following five elements:

- *Es* (< Sp. *es* ‘BE.COP.PRS.3SG’): optional copula.
- *Está* (< Sp. *está* ‘BE.LOC.PRS.3SG’): occasional locative copula.
- *Éra* (< Sp. *era* ‘BE.COP.IMPFV.3SG’): marker of the desiderative mood + occasional past copula (Fernández 2002).
- *Estába* (< Sp. *estaba* ‘BE.LOC.IMPFV.3SG’): past locative copula + preposition + adverb + adjective, consistently indicating past tense or origin. (cf. Frake 1980:308)
- *Taquí, tallí, tallá* (< Sp. *está* + *aquí* ‘here’, *allí* ‘there (near)’, *allá* ‘there’ (far)): locational existential pseudoverbs (Forman 1972:161).

The first four items occur, occasionally rather than systematically, as formal and functional retentions from Spanish, but only *éra* and *estába* have developed new meanings and functions. The retentions of *es* and *está* appear to be idiolectal and/or socio-indexical, although their distribution in relation to zero-copula constructions, copulas of Philippine origin, and pseudoverbs warrants further investigation. Notably, *éra* has adopted the functions of Tagalog *sána* (‘if only, I wish’), though the reasons for this shift remain unclear. The refunctionalization of *estába* is particularly striking, as it can modify nouns and verbs or even function as a

preposition. Consider, for example, (2), where *estába* introduces a prepositional phrase accompanied by the oblique case marker *na*. This suggests that *estába*, as a preposition, may have originated from a relative clause.

- (2) *Al llegár el táta estába na pescáda...*
 on arrive DEF father FROM OBL fishing
 ‘When the father arrived from fishing.’

While some of these processes may be purely language-internal, we believe others exemplify cases of contact-induced grammaticalization as defined by Matras (2011:281). In this study, we aim to shed light on the processes of semantic change, accelerated functionalization, and coalescence involved in grammaticalization (Michaelis & Haspelmath 2020), particularly within the context of creole languages.

References

- Aoto, Seiichi (2002), La cópula *amo* del chabacano de Zamboanga, *PAPIA-Revista Brasileira de Estudos Crioulos e Similares* 12(1), 84-106.
- Fernández, Mauro (2002), Materiales para el estudio del modo irreal en el chabacano de Zamboanga: la partícula *era*, in X. A. Fernández Roca and M. J. Martínez López (eds), (2009), *Vir bonus docendi peritus: Homenaxe a José Pérez Riesco*, A Coruña: Servicio de Publicacións da Universidad de A Coruña, 105-121.
- Forman, Michael (1972), *Zamboangueño Text with Grammatical Analysis: A Study of Philippine Creole Spanish*, Ithaca, NY: Cornell University PhD dissertation.
- Frake, Charles (1980), Zamboangueño verb expressions, in A. S. Dill (ed), (1980), *Language and Cultural Description, Essays by Charles O. Frake*, Stanford: Stanford University Press, 277-310.
- Matras, Yaron (2011), Grammaticalization and language contact. In B. Heine and H. Narrog (eds), (2011), *The Oxford Handbook of Grammaticalization*, Oxford: Oxford University Press, 279–290.
- Michaelis, Suzanne Maria & Martin Haspelmath (2020), Grammaticalization in creole languages: Accelerated functionalization and semantic imitation. *Grammaticalization scenarios: Cross-linguistic variation and universal tendencies*, 2, Berlin/Boston: De Gruyter Mouton, 1109-1128.

A syntactic analysis of Cantonese pseudocleft-like construction

Esther Lam
University of Edinburgh

Keywords: Pseudocleft, Cantonese, Predication, Specificational copular clause, Relative clause

Cantonese has a construction which looks similar to pseudoclefts in English and European languages:

- (1) Mary gam-ziuzing ge hai ni go daangou
Mary today-morning make GE COP this CL cake
'What Mary made this morning is this cake.'

The pre-copula constituent in (1) consists of a "headless relative clause" terminating in the functional item *ge*; it does not have any nominal head or *wh*-pronoun. In this talk, I will provide an analysis of this pseudocleft-like construction in Cantonese. I argue that the "headless relative clause" is a predicate, so that the construction shown in (1) is an inverted predication copular clause, with the predicate at the pre-copula position, and the subject at the post-copula position.

In pseudoclefts of English and European languages, the *wh*- clause that appears at the initial position is generally analysed as either a free relative or an embedded question (den Dikken 2017). Yet, the headless relative clause in Cantonese is neither. Since there is no question marker in a Cantonese headless relative, it is not plausibly an embedded question. Moreover, whereas free relatives can function as syntactic arguments of verbal and adjectival predicates, Cantonese headless relatives cannot. As (2) and (3) show, a head noun must be present when a relative clause occurs at the subject or the object position, and when it functions as the subject of an adjectival predicate.

- (2) a. Mary gam-ziu zing ge *(je) faat-zo mou
Mary today-morning make GE thing grow-PFT mould
'What Mary made have turned mouldy.'
b. John sik-zo Mary gam-ziu zing ge *(je)
John eat-PFT Mary today-morning make GE thing
'John ate what Mary made this morning.'
- (3) Mary gam-ziu zing ge *(je) hou leng
Mary today-morning make GE thing very beautiful
'What Mary made this morning is very beautiful.'

While (2) and (3) show that Cantonese headless relatives cannot appear in argument positions, (4) demonstrates that they can appear in the typical position for predicates. In (4), the headless relative denotes a property ascribed to the subject 'this cake'.

- (4) ni go daangou hai Mary gam-ziu zing ge
this CL cake COP Mary today-morning make GE
'This cake is made by Mary this morning.'

Given that headless relative clauses are predicates, (4) is an inverted predication copular clause. Yet, as (5) shows, predicative bare nouns cannot occur at the pre-copula position in Cantonese.

- (5) *naamzai hai John
 boy COP John
 Intended: 'John is a boy.'

To account for why a headless relative can occur at the pre-copula position, I propose that a pre-copula headless relative is a definite predicate. Given that Cantonese generally disallows sentence-initial indefinite phrases (Cheng & Sybesma 1999), predicative bare nouns, which are indefinite, cannot occur at this position; yet, a definite headless relative can. The claim that pre-copula headless relatives are definite is supported by their exhaustivity requirement. For instance, (1) cannot be followed by the Cantonese counterpart of (6), which suggests that (1) requires exhaustivity.

- (6) Mary also made biscuits this morning!

The claim that predicates can be indefinite or definite is in line with Cheng, Heycock and Zamparelli (2017) and Coppock and Beaver (2015), who argue that the weak definiteness meaning (i.e. uniqueness/maximality) is carried by predicates, and definiteness is not limited to phrases of argumental semantic types.

References

- Cheng, L.L., Heycock, C., & Zamparelli, R. (2017). Two levels for definiteness. In *Proceedings of GLOW in Asia XI: MIT working papers in linguistics*, MIT, 307-346.
- Chen, L.L., & Sybesma, R. (1999). Bare and not-so-bare nouns and the structure of NP, *Linguistic inquiry*, 30(4), 509-542.
- Coppock, E., & Beaver, D. (2015). Definiteness and determinacy. *Linguistics and philosophy*, 38(5), 377-435.
- Den Dikken, M. (2017). Pseudoclefts and other specificational copular sentences. In M. Everaert & H. van Riemsdijk (Eds.), *The companion to syntax* (2nd ed.). Malden, MA: Blackwell Publishing, 292-409

An empirical approach to coordinated *wh*-questions in Romanian

Gabriela Bîlbîie

(University of Bucharest & LLF, Paris)

Keywords: multiple *wh*-questions, coordinated *wh*-questions, multiple *wh*-fronting, Romanian, corpus

Romanian, like other languages (Bulgarian, Hungarian, Serbo-Croatian, Russian, etc., see Citko & Gračanin-Yuksek 2013), allows the alternation between coordinated *wh*-questions (CWQs) (1a) and ‘paratactic’ multiple *wh*-questions (MWQs) (1b), where two (or more) *wh*-phrases are fronted with (1a) or without (1b) a conjunction, regardless of their syntactic function (in particular, in both constructions, the fronted *wh*-phrases can be arguments of a verbal head, like in (1)).

- (1) a. **Cine și ce** a mâncat? (CWQ) b. **Cine ce** a mâncat? (MWQ)
 who and what has eaten who what has eaten
 ‘Who ate what?’

Previous literature (Comorovski 1996, Rațiu 2011, Citko & Gračanin-Yüksek 2013) assumes that the two patterns (CWQs and MWQs) are distinct constructions, with different semantic and syntactic properties, as schematized in Table 1.

	Coordinated <i>wh</i> -questions (CWQs)	Multiple <i>wh</i> -questions (MWQs)
Semantic interpretation	single-pair reading	pair-list reading
Ordering constraints (superiority effects)	no ordering constraints	strict ordering constraints

Table 1. Properties of CWQs vs. MWQs

However, previous works are based solely on introspection data, many examples being artificial and lacking an appropriate context. Our main goal is to confront previous research with corpus data in order to account for the behaviour of CWQs in actual usage. Here we present the preliminary results of a corpus investigation based on the authentic uses from the *CoRoLa (The Reference Corpus of the Contemporary Romanian Language*, Barbu Mititelu et al. 2018) that comprises both a written and an oral part.

One of the most striking properties of these two constructions in corpus is their preference for embedded contexts. In the previous research, the prototypical pattern is in a main clause configuration. Our corpus data show, however, a main/subordinate clause asymmetry, as noted by Gazdik (2011) for Hungarian.

Moreover, there is no clear-cut correlation between the construction type (CWQs vs. MWQs) and semantic interpretation: both single-pair and pair-list readings are available with both constructions. In particular, CWQs are compatible not only with single-pair reading, but also with pair-list reading (*contra* Citko & Gračanin-Yüksek 2013).

Concerning so-called superiority effects, although CWQs do not impose strict ordering constraints as MWQs, there are still preferences for maintaining the same ordering as in MWQs. We

observe a preference for ‘animate-first’ (and in particular ‘human-first’) if one of the *wh*-phrases is an animate subject, in line with the animacy hierarchy (Silverstein 1976).

Crucially, our corpus data show that CWQs in particular do not form an homogeneous class in Romanian, but they give rise to three potential syntactic analyses: (i) mono-clausal CWQs with a subclausal coordination; (ii) bi-clausal CWQs with ellipsis in the first clause (i.e. the first *wh*-phrase is a fragmentary ‘short question’) and (iii) bi-clausal CWQs with ellipsis in the second clause (i.e. the second *wh*-phrase is a fragmentary ‘short question’ and has a parenthetical status, while the first *wh*-phrase belongs to the full clause). The bi-clausal analyses with ‘sharing’ proposed by Citko & Gračanin-Yüksek (2013) are challenged by those attested data that involve mismatch, where the verbal head is not compatible with the first *wh*-clause. This kind of data is problematic for any syntactic approach that appeals to a syntactic reconstruction mechanism, but not for approaches assuming a semantic reconstruction and a fragmentary syntax (Ginzburg & Sag 2000).

References

- Barbu Mititelu, Verginica, Tufiş, Dan & Irimia, Elena (2018), The reference corpus of the contemporary Romanian language (CoRoLa), in N. Calzolari, K. Choukri, C. Cieri, T. Declerck et al. (eds.), *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, European Language Resources Association (ELRA), 1178-1185.
- Citko, Barbara & Gračanin-Yüksek, Martina (2013), Towards a typology of coordinated *wh*-questions, *Journal of Linguistics* 49(1), 1-32.
- Comorovski, Ileana (1996), *Interrogative phrases and the syntax-semantics interface*, Dordrecht: Kluwer.
- Gazdik, Anna (2011), *Multiple questions in French and in Hungarian: A Lexical-Functional analysis with special emphasis on the syntax-discourse interface*, Université Paris Diderot – Paris 7, PhD dissertation.
- Ginzburg, Jonathan & Sag, Ivan A. (2000), *Interrogative investigations: The form, meaning and use of English interrogatives*, Stanford: CSLI Publications.
- Raţiu, Dafina (2011), A multidominance account for conjoined questions in Romanian, in J. Herschensohn (ed.), *Romance linguistics 2010*, John Benjamins, 257-270.
- Silverstein, Michael (1976), Hierarchy of features and ergativity, in R.M.W. Dixon (ed.), *Grammatical categories in Australian languages*, Canberra: Australian Institute of Aboriginal Studies, 112-171.

Syntax of backward anaphora: a view from Polish

Jacek Witkoś (wjacek@amu.edu.pl)
Adam Mickiewicz University, Poznań, Poland

Keywords: syntax, cataphora, background, topic, focus

This study examines backward binding (BB, where the pronoun linearly precedes its antecedent) in Polish. Building on Biskup (2011), focusing on subject pronouns in Czech. Our research concerns object and prepositional object pronouns. Additionally, it adopts Bruening's (2014) phase-command framework over Chomsky's (1981, 1986) c-command to regulate pronoun-antecedent placement.

Reinhart (1983) identifies two key conditions for BB: (a) the pronoun may not c-command the antecedent and (b) the antecedent must be backgrounded/topical. The condition in (b) aligns with the topic-antecedent hypothesis in Bianchi (2009), which posits that pronouns can reference antecedents only if these serve as sentence topics. This is further confirmed through Biskup's (2011) Background Adjunct Coreference Principle (BACP): the antecedent in BB must be backgrounded, with tangible syntactic consequences in Czech (and Polish). In these languages a backgrounded NP (here the subject) must occupy a preverbal position, while in English both the subject and the object of active sentences can be backgrounded:

- (1) a. When he entered the room, ^{ok}Max greeted ^{ok}Bill.
b. When he entered the room, ^{ok}Max was greeted by *Bill.
- (2) a. [zanim *pro*₁ wyjechał] ^{ok}Piotr₁ pocałował Marię₂.
before *pro*₁ left Piotr_{1.NOM} kissed Maria_{2.ACC}.
b. *[zanim *pro*₁ wyjechał] Marię₂ pocałował *Piotr₁
before *pro*₁ left Maria_{2.ACC} kissed Piotr_{1.NOM}.

In Polish, object pronouns in (3) allow BB with right adjunct clauses but not object clauses, as the former are beyond in the object's c-command domain. Interestingly, prepositional object pronouns in (4) exhibit identical behavior:

- (3) a. Piotr₂ ją₁ pocałował, [zanim ^{ok}Maria₁ wyjechała].
Peter₂ her₁ kissed before Maria₁ left
b. *Piotr₂ ją₁ pocałował [zanim wyjechała *Maria₁].
c. *Piotr₂ ją₁ przekonał [że Maria₁ będzie gwiazdą]
Peter₂ her₁ convinced that Maria₁ will-be star
- (4) a. Piotr₂ spojrział na nią₁, [zanim ^{ok}Maria₁ wyjechała].
Peter₂ looked at her₁ before Maria₁ left
b. *Piotr₂ spojrział na nią₁ [zanim wyjechała *Maria₁].
c. *Piotr₂ [vp mówił [pp o niej₁] [cp że Maria₁ nie umie prowadzić]].
Peter₂ spoke about her₁ that Maria₁ cannot drive

But if c-command determines BB, ex. (4c) should be fine, as PP is a genuine constituent that embraces the pronoun. This problem is solved when BB is not determined by the classic c-(onstituent) command but by Bruening's (2014: 343) 'precede and phase-command':

- (5) A phase-commands B iff A precedes B and the first phase node dominating A dominates B (phase nodes include: vP, CP and DP/NP but not PP).
- (6) [CP [TP [TP *pro*₁ [vp .. obj₂ [vp V obj₃ [pp obj₃] [CPobj ..R*_{1/2/3}..]]]] [CPadj..R*_{1/2/3}..]]]

CP and **vp** delimit phase-command domains; CP_{adj}, an adjunct clause, is adjoined to TP on the right; CP_{obj}, the object clause, is embraced by **vp**. Only obj₂ and obj₃ tolerate BB by R*_{1/2/3} because the right adjunct is beyond their phase-command domain; however, the subject *pro*₁ does not, as it both precedes and phase-commands it. No BB by R*_{1/2/3} is possible, since the object clause is inside the **vp** phase and it is preceded and phase-commanded by *pro*₁, obj₂ and obj₃. Polish data confirm that BB is sensitive to both 'precede and phase-command' (Bruening 2014) and the BACP (Biskup 2011).

References

- Bianchi, Valentina. 2009. A Note on Backward Anaphora. *Rivista di Grammatica Generativa* 34. 3-34.
- Biskup, Peter. 2011. *Adverbials and the Phase Model*. Berlin: John Benjamins.
- Bruening, Benjamin. 2014. Precede-and-Command Revisited. *Language* 90(2). 342-388.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Kluwer.
- Chomsky, Noam. 1986. *Barriers*. Cambridge, MA.: The MIT Press.
- Diesing, Molly. 1992. *Indefinites*. Cambridge, MA.: The MIT Press.
- Reinhart, Tanya. 1983. *Anaphora and semantic interpretation*. London: Croom Helm.
- Reinhart, Tanya. 2011. Processing or pragmatics? Explaining the coreference delay. In Gibson, E. & N. Perlmutterl (eds.), *The Processing and Acquisition of Reference*. Cambridge, MA: The MIT Press, 157-194.

“As said, then you can change your way of speaking”: mono-lexemic adverbials and the Extra-left position in Danish

Jan Heegård Petersen
(University of Copenhagen)

Keywords: left-dislocation, resumptive elements, prefield, summarizing adverbials, communicative adverbials

This paper explores mono-lexemic adverbials in Extra-left position in modern spoken Danish. Mono-lexemic adverbials consist of one word or of lexicalized two-, three- or four-words combination. Extra-left position is the syntactic construction where an adverbial is followed by a resumptive element, *så* 'so, then' (in the 'prefield') and a finite verb. Examples (1)-(5) illustrate this common construction in Danish (and other mainland Scandinavian languages), all examples are also grammatical without the resumptive *så*.

- (1) (*man går ikke ned i løn*) **i stedet for** *så* får man et personligt tillæg
 instead **then** get one a personal supplement
‘(you don’t go down in salary,) **instead** you get an individual supplement’
- (2) **som sagt** *så* kan man ændre sin talemåde
 as said **then** can one change one’s way of speaking
‘**as said**, you can change the way you speak’
- (3) **heldigvis** *så* starter sæsonen næste måned
 luckily **then** begins season-DEF next month
‘**luckily**, the season begins next month’

A study of 6,000 mono-lexemic adverbials reveals an almost equal distribution between sentence-initial monolexemic adverbials with (49%) and without (51%) *så* (Heegård Petersen, accepted). The study also shows that the semantics of the adverbials play an important role for the distribution: contrastive and summarizing adverbials, (1)-(2), are most likely to occur with *så* (49-50%), epistemic or attitude adverbials, (3) are least likely (17%), disjunctive and additive adverbials like *alternativt* 'alternatively' and *desuden* 'besides', etc. take an in-between position (39-40%).

The paper presents the data behind this observation which to some extent can be explained by Nordström's (2010) principle of 'a new point of departure' and of Ekerot's (1992) principle of verbal dependency. According to Nordström, (Swedish) *så* marks the preceding adverbial as a new point of departure for the following proposition. Some adverbials are more likely to occur as new points of departure, for example, (Swedish) *för min del* 'for my part' and *förövrigt* 'besides'. Following Ekerot's principle of verbal dependency, adverbials that are closely linked to the main predicate, adverbial arguments or time and locative adverbials, cannot or are least likely to occur with *så* (in Swedish). This

principle may explain why epistemic and attitude adverbials are unlikely to occur with *så*. In contrast to e.g. adverbials of contrast, which are strongly related to a preceding proposition, epistemic and attitude adverbials express the speaker's attitude etc. to the proposition denoted by the main predicate, and not to a preceding proposition. Hence, there is no obvious reason for marking 'a new departure'.

However, neither principles can explain why epistemic adverbials occur unproblematically with *så* and why contrastive adverbials often occur without *så*. To explain this variation, I discuss the applicability of Ekerot's (1988) principle of 'communicative adverbial function': With extra-posed adverbials and a resumptive element, the speaker expresses a comment to the proposition. Thus, (3) should be interpreted as 'luckily, I'm in a position to state that the season begins'; without *så* the interpretation would be propositional, 'it is lucky that the season begins ...'.

References

- Heegård Petersen, Jan (accepted), Monoleksemske adverbialer i ekstraposition til venstre [Mono-lexemic adverbials in extra-left position], *Ny forskning i grammatik* [New research in Grammar].
- Ekerot, Lars-Johan (1988), *Så-konstruktionen i svenskan: konstruktionstypen "Om vädret tillåter, så genomföres övningen" i funktionellt grammatiskt perspektiv* [The so-construction in Swedish: the construction type "If the weather permits, then the exercise will be carried out" in a functional-grammatical perspective]. (Lundastudier i nordisk språkvetenskap 42). Lund University.
- Ekerot, Lars-Johan (1992), The Swedish *så*-construction in a functional and contrastive perspective, In *The Nordic Languages and Modern Linguistics* 7, vol. 1. Tórshavn, 191–203.
- Nordström, Jackie (2010), The Swedish *så*-construction, a new point of departure, *Working Papers in Scandinavian Syntax* 85, 37–63.

The Dativus Ethicus in German and the *gei wo*-Construction in Chinese: A semantic and pragmatic comparison

Jian Wang & Karin Pittner
(Tongji University & Ruhr University Bochum)

Keywords: Dativus ethicus, *gei wo*-construction, directive speech, modal particle, speaker affect

This paper examines two parallel pronominal constructions in German and Chinese: the Dativus ethicus and the *gei wo*-construction, which both occur in directive speech acts and incorporate first-person pronominal elements that function to intensify illocutionary force while encoding speaker affect and stance (cf. Wegener 1985, Mollica 2014, Yin 2014, Sun 2019).

(1) Dativus ethicus:

Vergiss	mir	den	Honig	nicht!
forget.IMP	1SG.DAT	DEF.ACC	honey	NEG

‘Don't forget the honey!’

(2) *Gei wo*-construction:

ni	gei	wo	bi-zui!
2SG	give	1SG	close-mouth

‘Shut up!’

While prior analyses have characterized the Dativus ethicus as primarily encoding speaker concern and the *gei wo*-construction as marking threat or encouragement, the syntactic status of these elements—whether as pronouns or modal particles—remains disputed (Hole 2014, Panther 2021). Through corpus-based investigation, this study examines the semantic-pragmatic properties of both constructions to clarify their functions within the broader system of directive constructions in their respective languages, and explores the interaction between the syntax and the pragmatics.

Our study is based on 131 sentences containing a dativus ethicus from DWDS corpora and 1689 sentences containing *gei wo*-constructions from the BCC corpus. The corpus data indicate that both constructions predominantly express scolding or power-marked commands, particularly in conflictive situations. Their co-occurrence with modal adverbs, particles, and negation words serves to heighten directive urgency while mimicking features of spoken discourse in literary texts. The *gei wo*-construction exhibits a stronger tendency toward face-threatening acts, though such usage appears primarily oriented toward pragmatic effect rather than genuine aggression.

This cross-linguistic comparison has several key findings: First, both constructions demonstrate significant grammaticalization, with their literal meanings giving way to primarily pragmatic functions through blending of first-person pronouns and imperative constructions. Second, they share core features in expressing speaker stance and affect, particularly in contexts of urgency or emotional investment. Both constructions display variation in interlocutor relations and emotional expression,

demonstrating their pragmatic flexibility. Third, their distribution patterns suggest similar constraints on register and speech act types, being primarily associated with informal, directive contexts.

Furthermore, this study provides insights into the relationship between grammatical structure and pragmatic function in directive speech acts. The findings suggest that distinct grammatical mechanisms can serve parallel pragmatic functions across languages, highlighting potential universals in the linguistic encoding of speaker affect and stance. By exploring the interplay between grammar and pragmatics, this research contributes to our understanding of how urgency, power, and affect are linguistically constructed and conveyed in directive speech.

References

- Hole, Daniel (2014), *Dativ, Bindung und Diathese*. Berlin: De Gruyter.
- Gutzmann, Daniel (2007), Eine Implikatur konventioneller Art: Der Dativus Ethicus, *Linguistische Berichte* 211, 277–308.
- Mollica, Fabio (2014), Der Dativus ethicus im Deutschen aus konstruktionsgrammatischer Sicht, *Zeitschrift für Germanistische Linguistik* 42(3), 349–378.
- Sun, Tianqi (2019), *Patterns and Licensing Mechanisms of Non-core Argument Realization in Modern Chinese*, Shanghai: Zhongxi Press.
- Panther, Klaus-Uwe (2021), Der freie Dativ *mir* in direktiven Sprechakten. In Anja Binanzer, Jana Gamper & Verena Wecker (eds.), *Prototypen – Schemata – Konstruktionen: Untersuchungen zur deutschen Morphologie und Syntax*, 77–92. Berlin, Boston: De Gruyter.
- Wegener, Heide (1985), *Der Dativ im heutigen Deutsch*. Tübingen: Narr.
- Yin, Hailiang (2014), A Study on the “gei wo + VP” Construction as a Strong Directive Marker, *Chinese Language Learning* 1, 51–60.

Corpora:

BCC (BLCU Corpus Center): *Multidisciplinary Corpus (Duolingyu-Corpus)* <http://bcc.blcu.edu.cn/>

DWDS (Digitales Wörterbuch der deutschen Sprache): *Kernkorpus 20th and 21st century, newspaper corpora (Berliner Zeitung, Tageszeitung)*. <https://www.dwds.de/>

How to approach the extent of directional verb agreement in sign languages

Kevin Behrens (University of Bremen)

Keywords: sign language, agreement, verb classes, typology

In the sign space, sign languages have a unique feature that sets them apart from oral languages. This area in front of the signer's body is grammaticalised in most sign languages to serve specific functions. One of them is the representation of first, second, and third persons, with each occupying a distinct area, i.e., different loci, within the sign space. A special class of verbs, often referred to as directional or agreement verbs, can incorporate those loci into the root to mark the argument structure of the sentence. In such a case, a verb like CALL would modify the starting location (L1) of the sign to be within the locus of the subject argument and the ending location (L2) in the object's locus, glossed as 2-CALL-1 for 'You call me' (Schembri et al. 2018: 1–2).

This property is said to be universal among community sign languages like American Sign Language or German Sign Language (Meir et al. 2010: 13). Previous studies have often put the focus on only a handful of sign languages as the majority of them is hardly described or documented at all. A cross-linguistic description of directionality as illustrated above for verbal agreement is yet to be done. The same goes for a comprehensive account of the extent of the agreement verb class. Estimates of the number of such verbs range from 10 to 20 in Taiwan Sign Language to more than 80 in American or Russian Sign Language (SIGN-HUB, Quer & Cecchetti 2016–2020). While some theories like Meir's (2002: 423) assume that agreement verbs always have a semantic component that refers to a "transfer" (e.g., GIVE or TAKE), others hypothesise that directional agreement in sign languages is purely syntactic as some verbs do not show any semantic component of "transfer", like SEE, DEFEAT, or KILL (Pfau et al. 2018: 15–16). When assuming the first hypothesis, there should be a great overlap of verbs that allow directional agreement among the sign languages. A comparative study between a number of sign languages can give insights into this.

In this talk, I present the results of a cross-linguistic comparison between 20 sign languages worldwide. Emanating from grammatical descriptions of the extent of the verb class inventories, I create a model of over 1,000 agreement verbs and their cross-linguistic overlap. It shows that while a very small number of verbs seem to be almost universally directional among the sign languages like GIVE, the majority varies with most of the verb meanings being unique to only sign language each, like REPAIR, FORGIVE, IMPROVE, or SUGGEST. My data shows that the membership to the respective agreement verb class is conditioned by language internal processes. The inventory sizes vary largely as do their semantic make-ups, showing that a necessary semantic "transfer" component cannot be confirmed with the opposite being true.

Abbreviations:

- 1 index for first person on a verb
- 2 index for second person on a verb
- L1 starting position of a sign
- L2 ending position of a sign

References

- Meir, Irit (2002),. A cross-modality perspective on verb agreement, *Natural Language and Linguistic Theory* 20, 413–450.
- Meir, Irit, Wendy Sandler, Carol Padden, Mark Aronoff (2010), Emerging Sign Languages, in Irit Meir, Wendy Sandler, Carol Padden, Mark Aronoff (eds.), *Oxford Handbook of Deaf Studies, Language, and Education*, vol. 2, 268–280, Oxford: Oxford University Press.
- Pfau, Roland, Martin Salzmann & Markus Steinbach (2018) The syntax of sign language agreement: Common ingredients, but unusual recipe, *Glossa: a journal of general linguistics* 3(1): 107, 1–46.
- Quer, Josep & Carlo Cecchetto (2016–2020), SIGN-HUB, Online: <https://thesignhub.eu/project>
- Schembri, Adam, Kearsy Cormier and Jordan Fenlon (2018), Indicating verbs as typologically unique constructions: Reconsidering verb ‘agreement’ in sign languages, *Glossa: a journal of general linguistics* 3(1): 89, 1–40. DOI: <https://doi.org/10.5334/gjgl.468>

Are some Estonian adverbs pointlessly redundant? A comparative analysis of adverbs “kohe” and “koheselt” expressing immediacy

Lydia Risberg, Maarja-Liisa Pilvik & Liina Lindström

(Institute of the Estonian Language and University of Tartu; University of Tartu & University of Tartu)

Keywords: temporal adverbs, usage-based linguistics, quantitative methods, language planning, Estonian

Acknowledgments: R&D project “Morphosyntactic variation in Estonian” (EKKD-TA2; Project executor: University of Tartu; Funder: Ministry of Education and Research)

Some adverbs with the *-lt* suffix are not recommended by Estonian language planning for over half a century (e.g., see DSE 1960). In particular, the formation of *-lt* adverbs from *-ne* adjectives has been considered redundant. A common example given by language planners is the adverb *koheselt*, derived from the adjective *kohene*, which itself originates from the adverb *kohe*. Language planners have advised against using *koheselt*, claiming that the shorter adverb *kohe* conveys the exact same meaning, ‘immediately’, which is why *koheselt* is redundant and unnecessary (e.g., Mäearu 2000). While new words are typically not created if their meaning is already covered in the lexicon, this restriction does not always apply to productive derivational types, such as adverbs with the *-lt* suffix (Kasik 2015: 47).

Although language reference materials should indicate that certain expressions or grammatical forms are suitable for specific situations, many prescriptive sources have made such claims based only on small and limited text samples (Lüdeling et al. 2022: 4, 9). Similarly, handbooks and dictionaries provided by Estonian language planning that discuss *-lt* adverbs have been prescriptive (e.g., Erelt, Erelt & Ross 2020), which is why these adverbs are viewed as negatively redundant and characteristic of bureaucratic language. Although redundancy is a typical feature of language and serves various functions within it (Bazzanella 2011).

Do the temporal adverbs *koheselt* and *kohe* really express the same meaning, or is the longer word needed for something else – for instance, to add intensity or emphasize temporal succession? In this presentation, we provide an overview of a comparative study of *koheselt* and *kohe*, aimed at understanding their current usage patterns and drawing conclusions about the reasons behind choosing the derived form *koheselt* over the shorter *kohe*.

We formulated two hypotheses. First, that the usage contexts of *koheselt* emphasize urgency and temporal immediacy more than those of *kohe*. This hypothesis is motivated by the vagueness of *kohe* in expressing time (e.g., alongside ‘at that exact moment’ it can also mean ‘soon’) and by the observation that for stressing immediate temporal succession, more complex forms like *otsekohe*, *kohemaid*, and the repetition *kohe-kohe* are also used instead of just *kohe*. Second, we assume that the use of *koheselt* correlates with morphosyntactic and contextual features expressing directivity, and higher degree of textual complexity.

The variation between *kohe* and *koheselt* is analysed using random samples taken from the *Estonian National Corpus* (Koppel et al. 2023) via the corpus tool Sketch Engine (Gilgarrieff et al. 2014). These random samples were drawn from three genres: blogs, forums, and periodicals, ensuring that the sentences originate from both authentic, unedited texts and edited texts. We analyzed the correlation between word choice and factors such as genre, modality, sentence structure, and

contextual complexity using common methods of language variation analysis (chi-square tests and conditional decision trees). To understand the historical background of *koheselt* usage and prescriptive recommendations, we also examined how the usage of *koheselt* developed in the 20th century, based on Estonian articles in the DIGAR archive.

References

- Bazzanella, Carla (2011). Redundancy, repetition, and intensity in discourse. *Language Sciences* 33, 243–254.
- Díaz-Campos, Manuel, and Sonia Balasch, S. (2023). Introduction. In M. Díaz-Campos, and S. Balasch (eds.), (2023). *The Handbook of Usage-Based Linguistics*. John Wiley & Sons, Ltd, 1–6.
- DSE 1960 = *Õigekeelsuse sõnaraamat [Eng: Dictionary of Standard Estonian]*. E. Nurm, E. Raiet, and M. Kindlam (eds.), (1960). ENSV TA Keele ja Kirjanduse Instituut. Tallinn: Eesti Riiklik Kirjastus.
- Erelt, Mati, Erelt, Tiiu, and Ross, Kristiina (2020). *Handbook of Estonian Language*. Tallinn: Eesti Keele Sihtasutus.
- Estonian articles in the DIGAR archive = DIGARi Eesti artiklid. <https://dea.digar.ee/>
- Kasik, Reet (2015). *Sõnamoodustus*. Tartu: Tartu Ülikooli Kirjastus
- Kilgarriif, Adam, Baisa, Vít, Bušta, Jan, Jakubíček, Miloš, Kovář, Vojtěch, Michelfeit, Jan, Rychlý, Pavel, and Suchomel, Vít (2014). The Sketch Engine: Ten years on. *Lexicography* 1, 7–36.
- Koppel, Kristiina, Kallas, Jelena, Jürviste, Madis, and Kaljumäe, Helen (2023). *Estonian National Corpus 2023*. Lexical Computing Ltd., Eesti Keele Instituut.
- Lüdeling, Anke et al. (2022). Register: Language Users' Knowledge of Situational-Functional Variation. Register Aspects of Language in Situation. *REALIS: Register aspects of language in situation* 1(1), 1–58.
- Mäearu, Sirje (2000). *It-liitelised määrsõnad. Keelenõuanne soovitab 2*. Tallinn: Eesti Keele Sihtasutus.

Towards an account of non-finite complementation in Estonian

Michael Green

(The University of Manchester)

Keywords: syntax; Role and Reference Grammar; Estonian; infinitive; complementation

There are two verb forms commonly described as infinitives in Estonian grammars: the *ma*- and the *da*-form, named for their characteristic endings. All Estonian verbs have both forms; some examples are provided in (1).

(1)

<i>da</i> -form	<i>ma</i> -form	Meaning
<i>Süüa</i>	<i>Sööma</i>	EAT
<i>Juua</i>	<i>Jooma</i>	DRINK
<i>Olla</i>	<i>Olema</i>	BE
<i>Öelda</i>	<i>Ütlema</i>	SAY
<i>Teha</i>	<i>Tegema</i>	DO

Matrix verbs overwhelmingly take non-finite complements in only one of the two forms (Green, 2023). For example, with *pidama* ‘must’, only the *ma*-form is grammatical, while with *oskama* ‘know how to’, only the *da*-form is.

- (2) Tüdruk pea-b [söö-ma/*süüa] jäätis-t
 girl must-3SG.PRS [eat-MA/*eat-DA] ice.cream-PRT
 ‘The girl must eat the ice cream’

- (3) Tüdruk oska-b [*sööma/süüa] jäätis-t
 girl know.how-3SG.PRS [*eat-MA/eat-DA] ice.cream-PRT
 ‘The girl knows how to eat the ice cream’

Currently, there is no coherent analysis of the distribution of the two forms. Previous research has approached the issue in terms of semantic notions such as volitionality or relative futurity or relegated the *-ma* vs. *-da* alternation to a vestige of the different diachronic developments of the two infinitive forms. I argue that the former approach is inadequate on its own, while the latter does not explain the distribution synchronically. What is needed is an analysis which looks beyond individual matrix verbs and provides a coherent account whereby different linkages (in the sense of Van Valin, 2023) can be derived from the semantics of the predicates which license them.

The data presented in the talk looks at non-finite complement structures for 13 matrix verbs, encoding a variety of semantic relations. The research used randomly-selected data samples extracted from the etTenTen21 corpus (Jakubíček et al., 2013) using CQL.

Two findings are reported. First, matrix verbs which take complements with the *ma*-form participate in semantically cohesive linkages, such as phasal and implicative causative relations, while those which take the *da*-form encode comparatively looser relations, such as desiderative and jussive ones (Silverstein, 1976; Givón, 1980; Van Valin & LaPolla, 1997). Second, matrix verbs which share their second argument with the complement encode it differently depending on the form in their complement. Matrix verbs which take the *ma*-form mark this argument with a core case as with *veenma* ‘persuade’ in **Error! Reference source not found.**, while those which take the *da*-form encode it in the adessive case and omit it much more frequently, as with *soovitama* ‘recommend’ in **Error! Reference source not found.**

- | | | | | | |
|-----|---|-------------------|-----------|------------------|---------------|
| (4) | Ma | veen-an | tüdruk-u | [sööma/*süüa] | jäätis-t |
| | 1SG | persuade-1SG.PRS | girl-GEN | [eat-MA/*eat-DA] | ice.cream-PRT |
| | ‘I persuade the girl to eat the ice cream’ | | | | |
| | | | | | |
| (5) | Ma | soovita-n | tüdruk-ul | [*sööma/süüa] | jäätis-t |
| | 1SG | recommend-1SG.PRS | girl-ADE | [*eat-MA/eat-DA] | ice.cream-PRT |
| | ‘I recommend the girl to eat the ice cream’ | | | | |

I argue that these patterns are indicative of two different complementation types. *Da*-forms are indicative of a core subordination relationship (Van Valin, 2023:84-8): the non-finite complement functions as an argument of the matrix predicate, and the shared argument takes the adessive, third argument, case as a result. *Ma*-forms are indicative of the more cohesive core cosubordination linkage type (ibid.). Here, the shared argument is case-marked as an argument of the matrix predicate and the complement is dependent on the matrix predicate for the expression of grammatical categories (e.g., deontic modality).

References

- Givón, T. (1980). The Binding Hierarchy and the Typology of Complements. *Studies in Language*, 4(3), 333–377. <https://doi.org/10.1075/sl.4.3.03giv>
- Green, M. (2023). *A Corpus-Assisted Study of Non-Finite Complementation Patterns in Estonian* [Unpublished MA Dissertation].
- Jakubíček, M., Kilgariff, A., Kovář, V., Rychlý, P., & Suchomel, V. (2013). The TenTen corpus family. *7th International Corpus Linguistics Conference CL*, 125–127. http://www.sketchengine.co.uk/wp-content/uploads/The_TenTen_Corpus_2013.pdf
- Silverstein, M. (1976). Hierarchy of features and ergativity. In R. Dixon (Ed.), *Grammatical Categories in Australian Languages* (pp. 112–171). Australian Institute of Aboriginal Studies.
- Van Valin Jr, R. D. (2023). Principles of Role and Reference Grammar. In D. Bentley, R. M. Usón, W. Nakamura, & R. D. Van Valin Jr (Eds.), *The Cambridge handbook of role and reference grammar*. Cambridge University Press.

Van Valin, R. D., & LaPolla, R. J. (1997). *Syntax: Structure, meaning, and function*. Cambridge University Press.

Is null complement anaphora reducible to clausal ellipsis in Russian?

Experimental study

Mikhail Knyazev

Institute for Linguistic Studies RAS & HSE University

Keywords: null complement anaphora, clausal complementation, argument ellipsis, experimental syntax, Russian

Null complement anaphora (NCA) refers to zero realization of the clausal complement (CP) in contexts like (1).

- (1) I suggested the price was too high, and she **agreed** \emptyset .

NCA differs from ellipsis among others in that it is *lexically governed* (*She complained/*boasted* \emptyset) (Depiante 2019). Yet, in languages like Russian CP omission appears to be fairly unrestricted, e.g. it is licensed with predicates that do not license it in English, as shown in (2).

- (2) Letajut v ličnyx samoletax s pozoločennymi unitazami, eščë i **xvastajutsja** \emptyset .
fly.3PL in personal planes with gilded toilets still PTCL boast.3PL
'They are flying in planes with gilded toilets, and are also boasting about it.' (RNC)
cf. *...and are also boasting \emptyset

This raises the possibility that CP omission in NCA-type contexts like (2) is an instance of CP ellipsis (a type of argument ellipsis involving clausal arguments, see Landau 2023) rather than true NCA, given that CP ellipsis is independently attested in Russian based on extraction diagnostics, as in (3).

- (3) Vasju₁ ja ne uveren, čto ona priglasit __₁, a Petju **uveren** \emptyset .
Vasya.ACC I not sure that she will.invite and Petya.ACC sure
'Vasya, I will not sure that she will invite, but Petya I am sure she will.'

This paper argues that despite appearances CP omission in examples like (2) cannot be derived by CP ellipsis and that NCA must still exist in languages with CP ellipsis.

The argument comes from the class of verbs that disallow CP omission in NCA-type contexts, as in (4), but still allow CP complements, as in (5) ("CP/*NCA verbs"). Examples include *utverdit'sja* 'become firm', *ssylat'sja* 'refer' and others.

- (4) Èto lučšij bank? Da, ja segodnja okončatel'no **utverdilsja** *(v ètom).
this best bank yes I today finally became.firm in this.LOC
'Is this the best bank? Yes, today I became firmly convinced of this.'

- (5) Ja segodnja okončatel'no **utverdilsja** (v tom), čto èto lučšij bank.
I today finally became.firm in that.LOC that this best bank
'Today, I became firmly convinced that this is the best bank.'

Finding such verbs is tricky because higher-frequency verbs tend to allow CP omission. Further, it is not always clear whether and to what extent verbs (especially less frequent ones) generally allow CP complements, given that less frequent verbs tend to prefer nominalized (*to,čto*) as opposed to bare (*čto*) CPs, cf. (5). After inspecting two corpus-based lists of verbs that take *to,čto*-clauses inside *na* ‘on’ (N = 162) and *v* ‘in’ PPs (N = 282), I selected 12 candidate CP/*NCA verbs for the experiments.

Two acceptability rating studies were conducted contrasting six CP/*NCA verbs (from each list) with six CP/NCA verbs in three conditions (bare CP vs. NCA vs. overt anaphora, cf. (4)). The results confirmed that NCA was rated lower than CP with CP/*NCA verbs, with no such difference for CP/NCA verbs, suggesting that CP omission is not predictable from CP licensing. This argues against the CP ellipsis analysis and in favor of the NCA analysis, presumably indicating that NCA-type contexts are not sufficiently parallel to license CP ellipsis (e.g. due to the lack of an overt CP antecedent).

A follow-up study directly compared CP omission in NCA-type vs. contexts that involve syntactic parallelism like (6) using a subset of the verbs used in the experiments above (and relative to overt anaphora and no-ellipsis baselines).

- (6) Ja ešče raz utverdilsja, [_{CP} čto ličnost’ učitelja mnogo značit dlja
I more time became.firm that personality teacher.GEN a.lot means for
učenika]. I moi druž’ja tože utverdilis’ Ø.
student and my friends also became.firm
‘I was once again became firmly convinced that the teacher's personality means a lot to the
student. And my friends also did.’

As predicted, CP/*NCA verbs showed lower ratings for CP omission in NCA-type contexts compared to parallelism contexts, with no such difference for CP/NCA verbs, supporting the conclusion that CP ellipsis is not the right analysis for (2), since it fails to predict the observed contrast with CP/*NCA verbs. (Interestingly, CP omission with CP/*NCA verbs received below-average ratings. However, this resulted from CP/*NCA verbs showing a general dispreference for CP relative to PP complements, as well as generally lower ratings for conditions like (6), including with overt anaphora.)

The results have implications for the understanding of CP ellipsis and the syntactic identity/parallelism conditions associated with it (Landau 2023).

Acknowledgments

This research is supported by Russian Science Foundation, project 24-28-01873 realized at HSE University, <https://rscf.ru/en/project/24-28-01873/>.

References

- Depiante, Marcela A. 2019. Null complement anaphora. In Jeoren Van Craenenbroeck & Tanja Temmerman (eds.), *The Oxford handbook of ellipsis*, 657–680. Oxford: Oxford University Press.
Landau, Idan. 2023. Force mismatch in clausal ellipsis. *The Linguistic Review* 40 (3): 419–460.

Discursive distancing with the preposition *med* ‘with’ in Danish

Peter Juul Nielsen
(University of Southern Denmark)

Keywords: presupposition, information structure, reference, discourse, prepositions

As part of a larger project on the functions of the comitative preposition *med* ‘with’ in Danish, this paper presents the first analysis of how *med* is used for marking a proposition as contextually accessible without being presupposed, a hitherto undescribed information-structural option that allows speakers to distance themselves from a viewpoint that is present in the discourse. The analysis is based on an empirical study of this use of *med* in the 56-million-word corpus of Modern Danish *KorpusDK*, and the theoretical framework is structurally informed sign-based functionalism (Engberg-Pedersen et al. 1996, 2005, and Harder 1996).

In Danish, a nominal clause marked by *at* may be subordinate to the pronoun *det* ‘that’. This construction – the *det at* construction, DA – presents the *at* clause as a presupposition, making the DA compatible only with factive predicates such as *forståeligt* ‘understandable’, not with truth-evaluating predicates such as *løgn* ‘lie’ (1). When the *at* clause is prepositionally subordinate to *det* with *med* – creating the *det med at* construction, DMA – the propositional content of the clause is not encoded as presupposed by the speaker, and this makes the DMA compatible with truth-evaluating predicates (2). What is presupposed by the DMA is the status of the proposition as a discursive element: a piece of text, such as a claim or an expression of a viewpoint, that is in one way or another present in the discourse context, but the speaker is not (necessarily) committed to holding this view. Indeed, the DMA may be used by the speaker to express distance and non-commitment vis-à-vis the proposition, cf. (2).

- (1) *det at han græd, er forståeligt/*løgn*
that COMP he cried is understandable/lie
‘(the fact) that he cried is understandable/*a lie’
- (2) *det med at han græd, er løgn*
that with COMP he cried is lie
‘(the proposition/claim) that he cried is a lie’

The paper examines the range of textual and discursive uses of the DMA with a primary focus on its use for discursive distancing, the most significant function of DMA in contrast to DA. It is argued that the meaning contributed by the DMA construction is the combination of non-presupposition and recognitional reference: reference to specific knowledge shared by speaker and addressee (Himmelmann 1996, 1997, and Therkelsen 2003). Without committing themselves to the truth of the proposition, speakers can signal assumption of the addressee’s familiarity with the proposition, in some regards akin to the use of English *you know* as a metalinguistic monitor (Erman 2001).

The interpretation of the DMA as a marker of discursive distancing relies on the paradigmatic contrast to the DA: it is the non-selection of the presupposition-coding DA that leads to the inference of non-commitment and distancing. Thus, DA vs. DMA can be regarded as a constructional paradigm (Nørgaard-Sørensen et al. 2011), and the case of the DMA demonstrates the importance of the concept of the paradigm

and the paradigmatic contrast in the shaping of the meaning of syntactic constructions (cf. Diewald & Politt 2022).

References

- Diewald, Gabriele & Katja Politt (eds) (2022), *Paradigms regained: Theoretical and empirical arguments for the reassessment of the notion of paradigm*, Berlin: Language Science Press.
- Engberg-Pedersen, Elisabeth, Michael Fortescue, Peter Harder, Lars Heltoft and Lisbeth Falster Jakobsen (eds) (1996), *Content, Expression and Structure: Studies in Danish Functional Grammar*, Amsterdam: John Benjamins.
- Engberg-Pedersen, Elisabeth, Michael Fortescue, Peter Harder, Lars Heltoft, Michael Herslund and Lisbeth Falster Jakobsen (2005), *Dansk Funktionel Lingvistik*, University of Copenhagen, Copenhagen Business School & Roskilde University.
- Erman, Britt (2001), Pragmatic markers revisited with a focus on *you know* in adult and adolescent talk, *Journal of Pragmatics* 33, 1337-1359.
- Harder, Peter (1996), *Functional Semantics: A Theory of Meaning, Structure and Tense in English*, Berlin & New York: Mouton de Gruyter.
- Himmelmann, Nikolaus P. (1996), Demonstratives in narrative discourse: A taxonomy of universal uses, in B. Fox (ed.), (1996), *Studies in Anaphora*, Amsterdam & Philadelphia: John Benjamins, 205-254.
- Himmelmann, Nikolaus P. (1997), *Deiktikon, Artikel, Nominalphrase: Zur Emergenz syntaktischer Struktur*, Tübingen: Niemeyer.
- Nørgård-Sørensen, Jens, Lars Heltoft & Lene Schøsler (2011), *Connecting grammaticalisation. The role of paradigmatic structure*, Amsterdam & Philadelphia: John Benjamins.
- Therkelsen, Rita (2003), Talesprogets og skriftsprogets bestemthedsparadigmer, *Ny forskning i grammatik* 10, 189-206.

Dutch intransitive verbs reanalyzed as transitive verbs

Petra Sleeman
(University of Amsterdam)

Keywords: relative pronoun, adverbial function, reanalysis, Dutch

In recent work, I have shown that in a Grammaticality Judgment Task, native speakers of Dutch seemed to reinterpret a noun modified by an unaccusative or unergative participial attribute (Sleeman 2022) and the partitive pronoun *er* replacing the head of a quantificational adverbial NP (Sleeman 2023) as a direct object. In this paper I present the results of a test investigating the acceptance of two types of relative pronouns with intransitive verbs in Dutch and I show that the results point in the same direction.

In Dutch, the form *dat* is used as a complementizer. It may also function as a relative pronoun introducing a relative clause that modifies a neuter singular noun (1). With a non-neuter singular noun the form *die* is used. The form *die* is also used with a plural noun (2). In (1)-(2) the pronoun has the function of direct object of the embedded verb:

- (1) *het boek dat ik heb gelezen*
the.N.SG book.N.SG that.N.SG I have read
'the book that I have read'
- (2) *de boeken die ik heb gelezen*
the.PL book.N.PL that.PL I have read
'the books that I have read'

According to prescriptive rules (such as those of the official Dutch and Belgian institution the Dutch Language Union), with an adverbial function the form *dat* should be used, irrespective of the gender or number of the antecedent:

- (3) *de maanden dat / *die je afwezig was*
the months.M.PL that / that you absent were
'the months that you were absent'

Since this rule is not taught in schools, my research question was if native speakers accept the use of *dat* in noun phrases as in (3) or if they prefer *die*. I submitted a test to a group of 27 native speakers of Dutch on their judgments of noun phrases such as (3) with the form *dat* or *die*.

The results show that in 74% of the cases, the participants accepted the *die* form, as if it were a direct object, whereas according to Dutch prescriptive rules it should be *dat*, because of the adverbial function. The form *dat* was accepted in only 34% of the cases. The results also show that the acceptance of *die* or *dat* depends on the type of intransitive verb that was used and that with some verbs (unergatives like 'work' or 'swim', unaccusatives like 'remain') the preposition-less measure noun phrase indicating a period or a distance was more likely to be reanalyzed than with others (such as *niets doen* 'do nothing', which already contains an object).

Coupled with my recent work, the results suggest that in certain constructions, at least in a Grammaticality Judgment Task, a reanalysis of the verb takes place, in accordance with Van Gelderen's

(2004) *Head Preference Principle* in language change: “Be a head, rather than a phrase”. An analysis of the verb as a V, in a V-NP complement structure, is preferred above the analysis of the verb as a VP, in a VP-NP adjunct structure.

References

Gelderen, Elly van (2004), *Grammaticalization as Economy*, Amsterdam: John Benjamins.

Nederlandse Taalunie (Dutch Language Union). <https://taaladvies.net/de-maanden-die-of-dat-je-afwezig-was/>

Sleeman, Petra (2022), Partitive pronouns and quantified adverbial NPs: A labeling account, in G. Alboiu, D. Isac, A. Nicolae, M. Tănase-Dogaru, and A. Tigău (eds), (2022), *A Life in Linguistics: A Festschrift for Alexandra Cornilescu on her 75th Birthday*, Bucharest: Bucharest University Press, 513–525.

Sleeman, Petra (2023). Partitive pronouns in intransitive contexts in Italian and Dutch, in S. Luraghi, and P. Sleeman (eds), (2023), *Partitives cross-linguistically. Dimensions of variation*. Special issue of *Linguistic Variation* 23.1, 217–243.

Investigative adverbial clauses

Rieke Scheffner & Sune Gregersen
(Kiel University & Uppsala University)

Keywords: subordination, purpose clauses, complementation, adjuncts, embedded questions

This paper presents and discusses an overlooked type of subordinate clause in the world's languages, which we will term 'investigative adverbial clauses' (IACs). IACs express a matter under investigation which the situation in the matrix clause is supposed to answer, as in (1) with the Wangerooge Frisian subordinator *wut*:

(1) Wangerooge Frisian (Indo-European; Germany)

yaa wult weg un wult naa hirii kantoor too, wut yaa daa
they will.PL away and will.PL after her office to to.see.if they SDEF.PL

breíver nich fiin kant
letter.PL not find.INF can.PL

'They will leave and go to her office **to see if** they might be able to find the letters' (Versloot 1996: 449.112)

In (1), the matrix state-of-affairs 'go to her office' is carried out in order to determine the truth value of the proposition expressed by the subordinate clause. The subordinator *wut*, which may also introduce various types of complement and relative clauses, is best rendered with English 'to see if' or 'to find out if' in IAC contexts like (1).

In our paper, we first present a functional analysis of IACs in Wangerooge Frisian, after which we provide a cross-linguistic overview of IACs in a convenience sample of eight additional languages where they are attested (see Table 1).

Table 1. Languages in the sample

Language	ISO code	Family	Marker	References
Ancient Greek	grc	Indo-European	<i>ei</i>	Wakker (1994)
German	deu	Indo-European	<i>ob</i>	Imo (2017)
Japanese	jpn	Japonic	<i>-ka</i>	Kim & Tomioka (2014)
Korean	kor	Koreanic	<i>-ci</i>	Kim & Tomioka (2014)
Kosraean	kos	Austronesian	<i>lah</i>	Lee (1975)
Latin	lat	Indo-European	<i>si</i>	Cabrillana (2009), Pinkster (2021)
Turkish	tur	Turkic	<i>diye</i>	Özyıldız (2018)
Upper Tanana	tau	Na-Dene	<i>de' kah</i>	Lovick (2020)

IACs have been described under various headings in a number of individual languages, but appear to have been overlooked in cross-linguistic work on clause linking and adverbial subordination (e.g. Kortmann 1996, Hengeveld 1998, Cristofaro 2003, Dixon 2009, Hetterle 2015). We argue that while IACs

show similarities with complement, ‘potential circumstance’ (Hengeveld 1998), and purpose clauses, they are best regarded as a distinct type of clausal adjunct in the languages that have them. This is particularly clear in cases like the Upper Tanana subordinator *de’ kah* (glossed ‘to.check’ by Lovick 2020), as in (2), which is dedicated to the expression of IACs:

- (2) Upper Tanana (Na-Dené; Alaska, Yukon)
Ntthi’ elok de’_kah dqq’ djjdj’
 2SG:head hot to.check thus 3SG:do:IPV
 ‘Do like this **to check if** your forehead is hot.’ (Lovick 2020: 1, 266)

We end our paper with a brief discussion of the attested diachronic sources of IAC subordinators, such as complementizers, conditional subordinators, and prepositions.

Abbreviations

INF = infinitive; IPV = imperfective; PL = plural; SDEF = strong definite article; SG = singular.

References

- Cabrillana, Concepción (2009), Purpose and result clauses, in P. Baldi & P. Cuzzolin (eds.), *New perspectives on historical Latin syntax*, vol. 4, Berlin: De Gruyter, 19–92.
- Cristofaro, Sonia (2003), *Subordination*, Oxford: Oxford University Press.
- Dixon, R. M. W. (2009), The semantics of clause linking in typological perspective, in R. M. W. Dixon & A. Y. Aikhenvald (eds.), *The semantics of clause linking*, Oxford: Oxford University Press, 1–55.
- Hengeveld, Kees (1998), Adverbial clauses in the languages of Europe, in J. van der Auwera (ed.), *Adverbial constructions in the languages of Europe*, Berlin: De Gruyter, 335–419.
- Hetterle, Katja (2015), *Adverbial clauses in cross-linguistic perspective*, Berlin: De Gruyter.
- Imo, Wolfgang (2017), *Ob-Sätze in der mündlichen und schriftlichen Interaktion*, *Deutsche Sprache* 45, 1–30.
- Kim, Jooyoung & Satoshi Tomioka (2014), Two types of unselected embedded questions, *West Coast Conference on Formal Linguistics (WCCFL)* 31, 276–284.
- Kortmann, Bernd (1997), *Adverbial subordination: A typology and history of adverbial subordinators based on European languages*, Berlin: Mouton de Gruyter.
- Lee, Kee-dong (1975), *Kusaiean reference grammar*, Honolulu: The University Press of Hawaii.
- Lovick, Olga (2020), *A grammar of Upper Tanana*, vols. 1–2, Lincoln: University of Nebraska Press.
- Özyıldız, Deniz (2018), Unselected questions. Paper presented at the Workshop on Altaic Formal Linguistics (WAFL) 14, Massachusetts Institute of Technology, 19–21 Oct.
- Pinkster, Harm (2021), *The Oxford Latin syntax*, vol. 2, Oxford: Oxford University Press.
- Versloot, Arjen, ed. (1996), *“Mittheilungen aus der Sprache der Wangeroger”: De neilittenskip fan H.G. Ehrentraut oangeande it Eastfryske dialekt fan it eilân Wangereach út it argyf fan it Mariengymnasium yn Jever*, Leeuwarden: Fryske Akademy.
- Wakker, Gerry (1994), *Conditions and conditionals: An investigation of Ancient Greek*, Amsterdam: Gieben.

English *much less* construction: A corpus-based perspective

Seulkee Park & Jong-Bok Kim
(Kyung Hee University)

Keywords: *much less*, focus-sensitive coordination, scalarity, negative polarity item (NPI), *let alone*

The expression *much less* is generally used to indicate contrast or emphasize an even lower likelihood than what was previously mentioned in a negative context, functioning as a coordination between two situations (Hulsey 2008, Toosarvandani 2010, Harris 2016: among others). Consider the following two illustrative examples:

- (1) a. I can't drink [beer], much less [vodka].
b. I can't [drink beer], much less [make wine]. (Harris & Carlson 2019: (6)-(7))

A remnant complement of *much less* corresponds to its (bracketed) correlate through a contrastive focus relation, forming a focus-sensitive coordination structure, such as *let alone* or *not to mention*, within a negative context. Besides negative environments, focus-sensitive coordinators are also allowed in interrogative contexts or pragmatic adversity, similar to the distribution of negative polarity items (NPIs) (Fillmore et al. 1988).

Furthermore, the negative antecedent strongly implies or entails the negated ellipsis, suggesting a scalar relationship between a remnant and its correlate (Fillmore et al. 1988, Toosarvandani 2010: a.o.).

- (2) a. I don't have [any clients], much less [any work]. (2009 TV)
b. I don't have any clients. $\xrightarrow{\text{entail}}$ I don't have any work.

Example (2) indicates that the contextually weaker element must appear in the first conjunct; reversing this order results in semantic incoherence, rather than merely a simple constituent (Harris & Carlson 2019).

In the previous analyses, a focus-sensitive coordination construction is similar to stripping or sluicing, as it remains a remnant after ellipsis and requires coordination (Hulsey 2008; Harris 2016).

- (3) I can't drink beer, much less [_{FocP} vodka]₁ <I drink *t*₁>. (Harris 2016: 74)

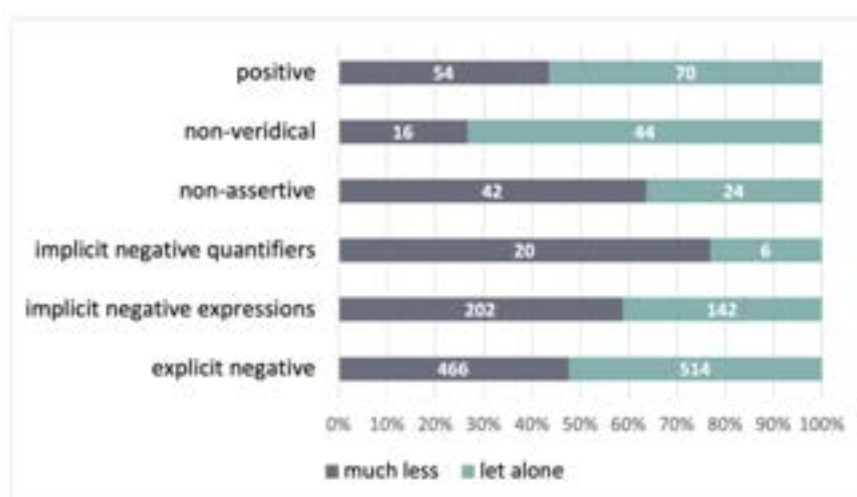
However, while the move-cum-delete analysis of the *let alone* construction seems plausible, when applied to *much less*, it raises doubts about whether clausal derivations can convey the implicit meaning in question.

A comparative analysis of *much less* construction and another form of focus-sensitive coordination, i.e., *let alone*, was undertaken by examining corpus data from the Corpus of Contemporary American English (COCA) with 800 tokens both from written and spoken sections to identify its constructional idiosyncrasies. Each instance was systematically annotated with syntactic and semantic variables. For instance, we classified the licensing environment of *much less* construction according to negative types, as in Table 1:

Table 1: Types of licensing environment and the examples

Licensing environment		Licensors	Raw freq. (%)
Explicit negatives	Negators	<i>not, never, no, ...</i>	466 (58.25%)
Implicit negatives	Negative predicates or adverbials	<i>incapable, barely, less, bad, difficult, refuse, ...</i>	202 (25.25%)
	Quantifiers	<i>few, most, ...</i>	20 (2.5%)
No licensors	Non-veridical context	modals, intensional context	16 (2%)
	Non-assertive context	interrogatives, imperatives, conditionals, comparatives, ...	42 (5.25%)
	No overt licensors	–	54 (6.25%)
Total			800

The findings indicated that the majority of polarity context licensors were explicit or implicit negative expressions. Interestingly, a statistically significant difference was observed between *much less* and *let alone* constructions, particularly within contexts of implicit negative expressions and non-veridical environments (p -value < 0.05), as shown in Figure 1.

Figure 1: Comparison of polarity licensing environment between *much less* and *let alone* constructions

Specifically, it indicates that *much less* is more contextually linked to negative meanings, whereas *let alone* may have broader range of uses, where the meaning is not strictly negative. Moreover, for all scalar relationships between a remnant and its correlate, *much less* appears to be more context-dependent, as evidenced by its higher frequency in context-dependent situations (p -value < 0.01). Consider the following distributional comparison in Table 2:

Table 2: Scalar relations between a remnant and its correlate and their context dependencies

Scalarity	Context-Independent	Context-Dependent	Examples
More general	196 (24.5%)	300 (37.5%)	<i>I can't stand to think about it, much less talk about it.</i> (1999 TV)
Stronger	64 (8%)	74 (9.25%)	<i>There is no way he could ever hurt someone, Much less kill them.</i> (2009 TV)
Reverse	26 (3.25%)	72 (9%)	<i>Mr. Wiseman isn't a pedagogue and doesn't wield a hammer, much less a wagging finger.</i> (2015 NEWS)
Relevant	8 (1%)	60 (7.5%)	<i>Family says he's never been to Reno, much less Sioux City.</i> (2019 TV)
TOTAL	294 (36.75%)	506 (63.25%)	

These findings significantly implies the understanding of how the expression *much less* may inherently convey an implicitly negative meaning to propositional content, rather than being limited to a negative expression in the first conjunct. The research provides insights into much less functioning as a strong Negative Polarity Item (NPI) operator, similar to Anti-Additive elements such as *few*, *less than*, *not every*, or an additive *either* within the scope of negation, as suggested by Zwarts (1998), in contrast to *let alone* as a weak NPI. This extensively supports the idea that the much less construction functions as a felicitous focus-sensitive response to an implicitly negative Question Under Discussion (QUD).

- (4) a. Family says he's never been to Reno, much less Sioux City. (2019 TV)
 b. $\lambda w[-\text{been-to}_w(\text{reno})(\text{he}) \wedge \sim \text{likely-to-visit}_w(\text{sioux-city})(\text{he})]$

In (4), both negation and a likelihood predicate (or possibly adopting a contextual operator) from contextual information may be required to express the comparative improbability by *much less*. It also indicates that the syntactic licensing of the constructional meaning does not originate from the initial conjunct, including the correlate. Instead, it may be derived from the inherent properties of the *much less* construction and the syntactic and contextual environment of the utterance, particularly as it interacts with the domain of relevant contextual information with a strong NPI operator.

References

- Fillmore, Charles J., Paul Kay, & Mary Catherine O'Connor (1988), Regularity and idiomatcity in grammatical constructions: The case of *let alone*, *Language* 64, 501–538.
- Harris, Jesse (2016), Processing *let alone* coordination in silent reading, *Lingua* 169, 70–94.
- Harris, Jesse & Katy Carlson (2019), Correlate not optional: PP sprouting and parallelism in *much less* ellipsis, *Glossa: a journal of general linguistics* 4(1), 83.
- Hulsey, Sarah (2008), *Focus Sensitive Coordination*. Cambridge, MA: Massachusetts Institute of Technology (Doctoral dissertation).
- Toosarvandani, Maziar (2010), *Association with Foci*. University of California, Berkeley (Doctoral dissertation).
- Zwarts, Frans (1998), Three types of polarity, in F. Hamm & E. W. Hinrichs (eds.), *Plurality, Quantification*, Kluwer, 177–238.

That or Zero? Diachronic insights into the syntactic alternation of the verb ‘think’

Vassiliki Geka

National & Kapodistrian University of Athens

Keywords: zero/that complementation, diachrony, Construction Grammar

Drawing on diachronic corpus evidence (COHA) spanning Late Modern English (ca. 1820) to Present-day English (ca. 2010), the paper discusses syntactic alternation through the lens of Construction Grammar (Traugott 2005, Traugott & Trousdale 2013, Barðdal et al., 2015). In particular, it examines the variation that the verb ‘think’ exhibits between *that* and *zero* complementation, as in 1-2.

(1) “I did not **think that** you had no faults, or were not liable to the infirmities of human nature;”

(COHA, FIC: Logan: A Family History, 1822)

(2) “I **think** she’s an intense young woman. Smart and aloof...”

(COHA, FIC: Analog Science Fiction & Fact, 2010)

To empirically investigate the two variants and their constructional differences, the study examines 7200 tokens randomly collected across the above-mentioned 190-year spectrum. Random sampling follows a stratified approach per decade and genre, the latter being restricted to fiction texts only, which consistently constitute the most multitudinous category in the corpus. With respect to annotation, the data are first manually tagged for several contextual, matrix and complement clause related parameters, including – inter alia – verbal morphology, subject correferentiality, pronominality, *pre*- and *post*-subject intervening material, harmony of polarity and coterminality. The annotated data are then statistically analysed using Pearson’s test, linear regression, and stepwise logistic regression with two main objectives: (a) to detect significant correlations holding between the data and the annotation parameters, and (b) to assess whether these parameters may predict (or possibly condition) the emergence of each variant, as proposed in prior studies (e.g., Thompson & Mulac 1991; Rissanen 1991; Finegan & Biber 1995; Diessel & Tomasello 2001; Shank et al., 2016a, 2016b; Shank & Plevoets 2018). In light of the empirical evidence collected, the study ultimately argues for applying a constructionist approach – definitionally sensitive to minimal constructional synonymy (Goldberg 1995) – to the analysis of syntactic alternation phenomena and specifically seeks to respond to the following: (a) what meaning differences can be detected between the constructional variants despite their partial overlap, (b) whether and to what extent diachronic trends favouring each alternative can be established, (c) what morpho-syntactic, contextual, and/or other factors may be argued to jointly or independently motivate each variant, and d) to which degree these variants could be considered predictive.

References

- Barðdal, Jóhanna, Elena Smirnova, Lotte Sommerer, and Spike Gildea (Eds.) (2015), *Diachronic Construction Grammar*. Amsterdam: John Benjamins. <https://doi.org/10.1075/cal.18>.
- Diessel, Holger, and Michael Tomasello (2001), The Acquisition of Finite Complement Clauses in English: A Corpus-Based Analysis, *Cognitive Linguistics* 12(2), 97–141. <https://doi.org/10.1515/cogl.12.2.97>

- Finegan, Edward, and Douglas Biber (1995), That and Zero Complementizers in Late Modern English: Exploring ARCHER from 1650–1990, in B. Aarts, and C. F. Meyer (eds), (1995), *The Verb in Contemporary English*, Cambridge: Cambridge University Press, 241–257.
- Goldberg, Adele (1995), *A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.
- Rissanen, Matti (1991), On the History of That/Zero in Object Clause Links in English, in K. Aijmer, and B. Altenberg (eds), *English Corpus Linguistics: Studies in Honour of Jan Svartvik*, London: Longman, 272–289. <https://doi.org/10.4324/9781315845890>
- Shank, Carl, Johan Van Bogaert, and Kristof Plevoets (2016a), The Diachronic Development of Zero Complementization: A Multifactorial Analysis of the That/Zero Alternation with *Think*, *Suppose*, and *Believe*, *Corpus Linguistics and Linguistic Theory* 12(1), 31–72. <https://doi.org/10.1515/cllt-2015-0074>.
- Shank, Carl, Kristof Plevoets, and Johan Van Bogaert (2016b), A Multifactorial Analysis of That/Zero Alternation: The Diachronic Development of the Zero Complementiser with *Think*, *Guess* and *Understand*, in J. Yoon, and S. Th. Gries (eds), *Corpus-Based Approaches to Construction Grammar*, Amsterdam: John Benjamins, 201–240. <https://doi.org/10.1075/cal.19.08sha>.
- Shank, Carl, and Kristof Plevoets (2018), Investigating the Impact of Structural Factors upon That/Zero Complementizer Alternation Patterns in Verbs of Cognition: A Diachronic Corpus-Based Multifactorial Analysis, *Research in Corpus Linguistics* 6, 83–112. <https://doi.org/10.32714/ricl.06.07>.
- Thompson, Sandra, and Anthony Mulac (1991), “The Discourse Condition for the Use of Complementizer *That* in Conversational English.” *Journal of Pragmatics* 15: 237–51. [https://doi.org/10.1016/0378-2166\(91\)90012-M](https://doi.org/10.1016/0378-2166(91)90012-M)
- Traugott, Elizabeth Closs (2005), Lexicalization and Grammaticalization, in D. A. Cruse, F. Hundsnurscher, M. Job, and P. R. Lutzeier (eds), (2005), *Lexikologie/Lexicology*, Vol. 2, Berlin: Walter de Gruyter, 1702–1712.
- Traugott, Elizabeth Closs, and Graeme Trousdale (2013), *Constructionalization and Constructional Changes*, Oxford: Oxford University Press.
- Van Bogaert, Johan (2011), *I Think* and Other Complement-Taking Mental Predicates: A Case of and for Constructional Grammaticalization, *Linguistics* 49(2), 295–332. <https://doi.org/10.1515/LING.2011.009>.

Verbal nouns and stativization in the encoding of event structure. The case of Irish passives and progressive events

Viviana Masia
(Roma Tre University)

Keywords: <Irish, verbal nouns, stativization, passivization>

In Irish, certain passive types and active progressive events (henceforth, APEs) display nominalized verb forms, known as *verbal nouns* (Stenson 2008). The Irish passive types formed with verbal nouns (VN) are the PROGRESSIVE (1), used to describe an event in its unfolding and thus not focusing its endpoint, and the PROSPECTIVE (2), typically denoting a future event (Greene 1966; Ó Siadhail 1989; Graver 2011; Nolan 2012).

- (1) *Tá an doras á chiceáil ag Máire*
Be.AUX DET door kick.VN at Mary
'The door is being kicked by Mary'
- (2) *Tá an doras le ciceáil ag Máire*
Be.AUX DET door (with) kick.VN at Mary
'The door is to be kicked by Mary'

The presence of verbal nouns is what these constructions have in common with APEs such as that in (3).

- (3) *Tá Máire ag ciceáil an dorais*
Be.AUX Mary kick.VN DET door.GEN
'Mary is kicking the door'

By and large, progressive and prospective passives, on the one hand, and constructions conveying APEs, on the other, essentially surface as clauses with nominalized predicates; and, despite denoting dynamic events, they lack the formal properties of fully transitive clauses, usually having full and non-nominalized verbs in sentence-initial position. In structures like (3), the genitive marking on the direct object - when preceded by a definite article - is indicative of the nominal nature of the predicate as well as of an incomplete affectedness of the undergoer participant, due to the more stative and weakly transitive nature of progressive events, cf. Hopper & Thompson 1980). On the whole, this aspect opposes the Irish case to that of other European languages in which direct objects of APEs do not receive distinct surface realizations (see the Italian pair *Mario ha colpito Luca* 'Mario hit Luca' vs. *Mario sta colpendo Luca* 'Mario is hitting Luca'). The use of verbal nouns thus entails a more stative representation of an event (cf. Bertinetto 1994 on the relation between progressive and stative events) by which sentences like (1) or (2) generally induce an underlying - almost locative - construal such as "the kicking of the door is/will be at Mary". Such an interpretation resembles the semantic structure of Irish sentences encoding emotional, psychological, or physical states, which are intrinsically stative and generally surface with non-canonical realizations of the obligatory argument(s). Notably, the main participant involved in these event types is not realized as subject but as a sentence-final prepositional phrase (Stenson 2008).

Drawing on Lazard's (1994, 2015) distinction between Major Biactant Construction (MaBC) – a typically transitive construction used to encode dynamic events - and Minor Biactant Construction (MiBC) – mostly encoding stative events and displaying less canonical argument marking – I will argue that while some languages have extended the use of MaBCs to both dynamic and stative events, in Irish these latter (including prospective, progressive passives and APEs) are linguistically expressed by MiBCs, through the use of nominalized verbs and less canonical argument surface realization, thus overall allowing for a *more motivated and less arbitrary form-function mapping* in event representation.

References

- Bertinetto, Pier Marco (1994), Statives, progressives and habituals: analogies and differences, *Linguistics* 32, 391-423.
- Graver, Jenny (2011), The Syntax and Development of the Old Irish Autonomous Verb, in Cairnie, Andrew (ed.), *Formal Approaches to Celtic Linguistics*. Cambridge, Cambridge Scholars Publishing, 41–64.
- Greene, David H. (1966), *The Irish Language*. An Ghaeilge, Dublin: Three Candles.
- Hopper, Paul J. & Thompson, Sandra A. (1980), Transitivity in Grammar and Discourse, *Language* 56(2), 251–299. doi: 10.2307/413757.
- Lazard, Gilbert (1994), *L'actance*. Paris: Presses Universitaires de France.
- Lazard, Gilbert (2015), Two possible universals: The Major Biactant Construction; The twofold notion of subject, *Linguistic Typology* 19(1), 111-130.
- Nolan, Brian (2012), *The Structure of Modern Irish. A Functional Account*, Bristol, Equinox.
- Ó Siadhail, Mícheál (1989), *Modern Irish*, Cambridge: Cambridge University Press.
- Stenson, Nancy (2008), *Intermediate Irish. A Grammar Book*, London: Routledge.

You Naughty Arabic Exclamatory Vocatives!

Youssef A. Haddad

Georgetown University Qatar

Keywords: Exclamatory vocatives; evaluative expressions; pragmatics; cross-linguistic variation; Lebanese Arabic

Exclamatory vocatives (EVs) were noted by Arab Grammarians over a millennium ago. They emerged in Western linguistic scholarship in early mentions by Welte (1980). EVs, exemplified in (1), have undergone varying terminologies: evaluative vocatives (Corver 2008), pseudo-vocatives (d'Avis & Meibauer 2013), and expressive vocatives (Gutzmann 2019). I expand on these vocatives, providing a cross-linguistic comparative framework, emphasizing their unique structural constraints and pragmatic functions.

1. a. You idiot! (English) b. Jij idioot! (Dutch) c. Du idiot! (German)

One characteristic of these vocatives, the cited authors argue, is the structural requirement to include a 2nd person pronoun in an initial position. Gutzmann (2019) further maintains that evaluative/expressive vocatives include not only stand-alone vocatives, (1), but also “integrated” and “parenthetical” vocatives. Both types are part of a larger utterance, as (2a-b, from Gutzmann 2019: 192), illustrate. The difference between the two types, Gutzmann maintains, is that an integrated expressive vocative serves as the subject of the utterance it occupies, while a parenthetical doesn't.

2. a. Integrated: You asshole stole my breakfast!
 b. Parenthetical: You idiot, the post office is closed tomorrow!

I present evidence from Lebanese Arabic (LA) to show that a 2nd-person pronoun is not a universal requirement of exclamatory vocatives. I further argue that the main requirement for exclamatory vocative is pragmatic rather than structural. I maintain that these vocatives are a unique linguistic phenomenon characterized by their use as complete, standalone expressions that convey evaluative attitudes. They, therefore, differ from regular vocatives which serve as calls or addresses and whose primary function is to establish or maintain contact with and ultimately seek a response from the addressee.

Consider (3) from the Lebanese movie *Where do we go now*. In this scene, Amal, a café owner in a divided Lebanese village, calls out to end a violent clash between local men. After multiple unsuccessful attempts, she resorts to EVs.

3. ja: ziʃra:n . ja: bala: ʔaxle:ʔ .
 VOC thugs VOC withoutmorals
 ‘You thugs! You degenerates!’

Three defining features of LA EVs are illustrated in (3). First, the vocative particle is essential. Second, EVs are syntactically independent and may stand alone as complete acts; they do not rely on surrounding discourse to clarify their communicative goal. Third, they may be preceded and followed by

the negative and positive interjections, *lah* and *ʔa:h* respectively, as in (4), which amplify the evaluative stance. Importantly, only EVs allow the flexibility of word order with interjections that we witness in (5).

4. *lah* , *ja:* ʔibn l-ħara:m , *lah* !
 N-INT VOC son of-sin N-INT
 ‘How dare you, you son of a bitch!’
5. a. *lah* *ja:* ʔibn l-ħara:m
 b. *lah* *ja:* ʔibn l-ħara:m *lah*
 c. *lah* *ja:* ʔibn l-ħara:m *lah* *ja:*
 d. *ja:* ʔibn l-ħara:m *lah* *ja:*
 e. ʔibn l-ħara:m *lah* *ja:*

This study explores the syntax-pragmatics interface of exclamatory vocatives, focusing on their structure, function, and cross-linguistic variations. This comparative approach to vocatives across languages sheds light on the broader syntactic and pragmatic strategies that underlie language-specific methods for conveying attitudinal stances.

References

- Corver, Norbert (2008), Uniformity and diversity in the syntax of evaluative vocatives, *Journal of Comparative Germanic Linguistics* 11, 43–93.
- d’Avis, Franz and Jörg Meibauer (2013), Pseudo-vocative constructions and insults in German (and Swedish), in Barbara Sonnenhauser, and Patrizia Noel Aziz Hanna (eds), (2013), *Vocative! Addressing between system and performance*, Berlin: Mouton de Gruyter, 189–218..
- Gutzmann, Daniel (2019), *The grammar of expressivity*, Oxford: Oxford University Press.
- Welte, Werner (1980), Zur Syntax, Semantik und Pragmatik exklamatorischer Vokative, *Indogermanische Forschungen* 85, 1–34.

General Session : Typology

Perception-finding polysemy: Chukchi and beyond

Aleksey Starchenko

(HSE University, Moscow)

Keywords: perception, lexical typology, lexical semantics, verbs of finding, Chukchi

The verb *#ʔuk* ‘see’ in Chukchi (< Chukotka-Kamchatkan) demonstrates wide polysemy. In addition to the perceptual interpretation (1), it can be interpreted as ‘find’ (2).

- (1) *#uut=e pʔompʔonə-t tə-#ʔu-ne-t #yi-nə-mkə-qin*
suddenly=PTCL mushroom-ABS.PL 1SG.S/A-see-3SG.O-PL INTS-ST-multiple-ST.3SG
‘I suddenly saw very many mushrooms.’

- (2) *ɣəm-nin mane-josyə-n tejə#iŋə-tku-te ʔotsoj tə-#qərir-ɣʔe-n*
I-GEN money-CONT-ABS.SG group-ITER-INS for.long 1SG.S/A-search-TH-3SG.O
tejusyə-k ənkʔam tə-#ʔu-ɣʔe-n
bag-LOC and 1SG.S/A-see-TH-3SG.O
‘I groped for my wallet in the bag for long and found it.’

Noticeably, the interpretation of *#ʔuk* in (2) cannot be reduced to seeing, as groping for an object is incompatible with visual perception.

In the present study I focus on the polysemy of the semantic fields of perception and search/finding. This pattern of regular polysemy has been discussed in literature (San Roque et al., 2018), but the ways of how the areas within these two fields relate to each other is not extensively studied.

The aim of the study is twofold. Firstly, I characterize in detail the properties of the Chukchi verb *#ʔuk* in both interpretations. Secondly, taking Chukchi data as a starting point, I proceed to the general discussion of the relations between perception predicates and finding or searching.

The semantic field of finding in Chukchi can be classified as exhibiting a dominant system, that is, the verb *#ʔuk* covers most of the possible distinctions (discussed in Ryzhova et al. [2018]) within this field. The only case in which *#ʔuk* competes with other verbs is finding of animate objects (that is, meeting).

Among the visual perception verbs *#ʔuk* can be classified as the default verb for non-agentive visual perception. It cannot be used in non-visual perception contexts. One argument for non-agentivity of *#ʔuk* is that it cannot be in imperative form in perceptual meaning (3).

- (3) *#qə-#ʔu-ɣʔe-n*
2.S/A.SUBJ-see-TH-3SG.O
‘Find (him)!’ / #‘Look (at him)!’

Turning to the wider typological point of view, non-agentivity of *#ʔuk* is crucially important with regard to its ‘see’/‘find’ polysemy. According to the database CLICS³ (Rzymiski et al., 2019), non-agentive concept SEE is regularly co-lexified with FIND, and agentive LOOK — with LOOK FOR, whereas the opposite is very rare. If excluded the cases in which ‘see’ and ‘look’ are co-lexified themselves, one can find almost one-to-one correspondence of two pairs of concepts (Table).

Table. Colexifications of visual perception and searching / finding

	SEE	LOOK
FIND	27	1
LOOK FOR	0	22

One can thus conclude that the Chukchi case is an instantiation of a more general principle of regular polysemy between the relation of verbs of non-agentive perception with finding and verbs of agentive perception with search.

Lastly, I hypothesize asymmetrical connection between perception, searching and finding verbs in diachronic perspective. For seeing and finding, it turns out to be difficult to find the direction of the semantic change. Turning back to Chukchi data, both meanings of *ʔuk* are present up to the level of Pra-Chukotko-Kamchatkan family (Mudrak 2000, Fortesque 2011). In contrast, one can regularly find that ‘look’ changes into ‘look for’ and not the other way around.

Acknowledgements

The research was supported by RSF (project No. 24-78-10199). The fieldwork to the Amguema village in 2018–2022 was supported by the Faculty of Humanities and the Fund for Educational Innovation, HSE (within the Program of Students’ Field Research / Expeditions “Rediscovering Russia”).

References

- Fortescue, M. (2011). *Comparative Chukotko-kamchatkan dictionary* (Vol. 23). Walter de Gruyter.
- Mudrak, O. A. (2000). *Ėtimologičeskij slovar' čukotsko-kamčatskich jazykov* [Etymological dictionary of Chukotko-Kamchatkan languages]. (Studia philologica). M.: Jazyki russkoj kul'tury.
- Ryzhova, D. A., Dobrushina N. R., Bonch-Osmolovskaya A. A., Vyrenkova A. S., Kyuseva M. V., Orekhov B. V., Reznikova T. I. (eds.). (2018). *EVRika! Sbornik statei o poiskakh i nakhodkakh k yubileju E. V. Rakhilinoi* [EVRika! Collected papers on searching and finding in honor of E. V. Rakhilina]. Moscow: Labirint.
- Rzymiski, C., Tresoldi, T., Greenhill, S. J., Wu, M. S., Schweikhard, N. E., Koptjevskaja-Tamm, M., ... & List, J. M. (2020). The Database of Cross-Linguistic Colexifications, reproducible analysis of cross-linguistic polysemies. *Scientific data*, 7(1), 13.
- San Roque, L., Kendrick, K., Norcliffe, E. & Majid, A. (2018). Universal meaning extensions of perception verbs are grounded in interaction. *Cognitive Linguistics*, 29(3), 371-406.

Romanian Future Formation at the Romance/Balkan Crossroad

1. Objectives of the paper. In line with the general typological and descriptive nature of the workshop, the paper examines the rich inventory of grammaticalized future paradigms in modern Romanian. Its objectives include: presenting the full inventory of futures (distinguishing between fully grammaticalized and less cohesive structures); exploring sources of future formations; examining the relationship between the future and irrealis category, particularly the presumptive mood (a grammaticalized mood); and analyzing the role of genealogic (Romance family) and areal factors (Balkan Sprachbund) in shaping the Romanian system of futures.

2. The Romanian futures. In old Romanian (16th–18th c.), synthetic futures existed, but modern Romanian (19th c.–present) uses only analytic/periphrastic futures. There are four types: standard (1a) and regional (1b) futures with auxiliary *voi* or *oi* plus the infinitive (1); colloquial futures with the lexical verb in the subjunctive, and distinct functional elements: the verb *have* in the indicative present (with its full inflection) (2) or (4) the invariable particle *o*.

- | | | | | |
|-----|-------------------|--------------------|--------------|--------------|
| (1) | a. | <i>îl</i> | <i>voi</i> | <i>vedea</i> |
| | | CL.ACC.3SG | AUX.FUT | see.INF |
| | b. | <i>l-oi</i> | <i>vedea</i> | |
| | | CL.ACC.3SG-AUX.FUT | see.INF | |
| (2) | <i>am</i> | <i>să-l</i> | <i>văd</i> | |
| | have.IND.PRES.1SG | SUBJ= CL.ACC.3SG | see.SUBJ.1SG | |
| (3) | <i>o</i> | <i>să-l</i> | <i>văd</i> | |
| | PARTICLE.FUT | SUBJ= CL.ACC.3SG | see.SUBJ.1SG | |
| | 'I will see him' | | | |

Infinitive-based futures (1) are more cohesive than subjunctive-based ones ((2)-(3)): clitic climbing is mandatory in the former but absent in the latter.

Future perfect forms (4) use the same auxiliaries as infinitive-based forms ((1)-(2)) plus *be* and the past participle, while future-in-the-past forms (5) are parallel to the have-based subjunctive future in (2), and use *have* (or, occasionally, the verb *urma* 'follow') in the imperfect.

- | | | | | |
|-----|--|-----------------------|------------------|--------------|
| (4) | <i>voi / oi fi</i> | <i>văzut</i> | | |
| | AUX.FUT AUX.be | see.PPLE | | |
| | 'I will have seen' (/ 'I might have seen', see §4 below) | | | |
| (5) | <i>aveam (/ urma)</i> | <i>să-l</i> | <i>văd</i> | |
| | have.IND.IMPERF.1SG | follow.IND.IMPERF.1SG | SUBJ= CL.ACC.3SG | see.SUBJ.1SG |
| | 'I was going/(was about) to see him' | | | |

3. Sources for the Romanian futures. Romanian futures combine two sources, widely attested cross-linguistically (terminology from Kouteva et al. 2019): WANT (> FUTURE) for the infinitive-based futures ((1),(4)) and (most probably, for *o*+subjunctive future in (3), cf. (Zafiu 2013:40), and the H-POSSESSIVE (> FUTURE) for the *have*-based subjunctive futures ((2),(5)). Interestingly, verbs with the semantics of 'follow' (cf. Rom. *urma* in (5)) are not listed in Kouteva et al. (2019:481) as a possible source for the future.

4. Future > presumptive. The Romanian regional future (1b) and future perfect (4) became specialised for epistemic values and have come to be used mostly as present and past presumptives (cf. the alternative translation of (4)). Furthermore, in the structural pattern of the future perfect, the gerund is employed instead of the participle, giving rise to a periphrasis which is exclusively a present presumptive (5).

- (5) *voi / oi* *fi* *văzând*
 AUX.PRES AUX.be see.GER
 ‘I might be seeing’

5. The Romanian futures in a larger comparative perspective. The combination of multiple sources as well as the variation found in the internal structure of the Romanian periphrases has mixed sources: common Romance developments and contact-induced grammaticalization in the context of the Balkan Sprachbund.

Selected references

- Kouteva, Tania & Heine, Bernd & Hong, Bo & Long, Haiping & Narrog, Heiko & Rhee, Seongha. 2019. *World Lexicon of Grammaticalization* 2nd ed. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/9781316479704>
- Ledgeway, Adam & Maiden, Martin (eds.). 2016. *The Oxford Guide to the Romance Languages*. Oxford: Oxford University Press.
- Ledgeway, Adam & Maiden, Martin (eds.). 2022. *The Cambridge Handbook of Romance Linguistics*. Cambridge: Cambridge University Press.
- Mišeska Tomić, Olga. 2006. *Balkan Sprachbund Morpho-Syntactic Features*. Springer: Dordrecht.
- Zafiu, Rodica. 2013. Mood, tense, and aspect. In Pană Dindelegan, Gabriela & Maiden, Martin (eds.), *The Grammar of Romanian*, 24–65. Oxford, New York: Oxford University Press.
- Zafiu, Rodica. 2016. The syntax of moods and tenses. In Pană Dindelegan, Gabriela & Maiden, Martin (eds.), *The Syntax of Old Romanian*, 14–52. Oxford, New York: Oxford University Press.

Body part incorporation in Ainu: At the crossroads of syntax, morphology, and lexicon

Anna Bugaeva & Maria Koptjevskaja Tamm
(Tokyo University of Science, Stockholm University)

Studies of noun incorporation (NI) (Kroeber 1909, Sapir 1911, Mithun 1984) highlighted the frequent use of body-part terms (BPTs) in NI's, reflecting the cohesive nature of body-related activities and enabling affected individuals to assume primary roles, such as subject or object, rather than being limited to oblique possession. BPTs are sometimes the only nouns allowing NI in a language (Dahl 2004: 213–214); and there may be other properties distinguishing NI with BPTs from other types of NI. In this paper we focus on Ainu (an isolate polysynthetic language of Northern Japan), which extensively incorporates BPTs to generate new verbs, most of which are at least to some extent lexicalized. We aim to contribute to the discussion of the nature of NI as a morphological process without or with a syntactic dimension (Mithun 1984, Massam 2009) and emphasize the role of lexical semantics in it. BPTs in Ainu are 'obligatorily possessed', syntactically require a possessor and cannot occur independently without possessive marking, where the prefix/suffix encode the possessor resp. the possessed state, e.g., *ku/e/Ø-par-o* 'my/your/his mouth-POSS (Bugaeva et al. 2022). Following cross-linguistic tendencies (Mithun 1984: 859), incorporation of BPTs in Ainu typically results in the promotion of the possessor to subject status. However, the option of using NI with BPTs in Ainu shows a number of peculiarities.

On the one hand, BPTs have less restrictions on the valency class of the incorporating verb than other nouns. They are essentially almost the only nouns allowed in NI with intransitive verbs, where they retain their possessed suffixes (1b). They also participate (in their bare form) in the more general object-incorporation with transitive verbs (2b), cf. (2a), but only when the possessor is coreferential (in a broader sense) with the original subject as in (2b) and (3a). Non-coreferential possessors cannot be promoted to subject status (3b), but trigger quasi-incorporation, resulting in phrasal verbs (Satō 2001, 2022). These tight syntactic units are not single words, but do not allow insertion (3c). Although the marking of the possessor shifts, the non-coreferential possessor of BPTs is not promoted to the object of the verb either.

According to Mithun (1984: 856), "most incorporating languages do incorporate such terms as 'body' and 'mind', since they provide a device for qualifying V's which pertain to the physical or mental aspect of a person or animal". As many as 23 BPTs are attested in Hokkaido Ainu NI, including 35 MIND-based NI lexemes, 28 EYE-based NI lexemes etc. While possible combinations are dependent on the lexical semantics of the components, the syntactic ban on NI of possessors that are non-coreferential with the transitive verb subject applies automatically and without exception to any BPT. This makes NI in Ainu a special process that lies at the crossroads of syntax, morphology, and lexicon. We hope to see more research on this interaction and on the role of lexical typology in NI (Koptjevskaja-Tamm and Veselinova 2020).

Acknowledgements

This research is supported by the Japanese Government Scientific Research Category-A Grant "A Comprehensive Study of Minority Languages in Siberia: Focusing on Typological Diversity in the Early Stages" (#21H04346; PI: F. Ebata, Niigata U; Co-investigator: A. Bugaeva, Tokyo University of Science).

Examples (4th person: here, the protagonist's 'I' in folktales)

- (1) a. *an-hon-i* *arka* base clause
 4.A/POSS-belly-POSS hurt
 'My belly hurts (i.e. 'I am in labor.')
- b. *cip* *kamuy,* *hon-i-arka-an* NI
 boat god belly-POSS-hurt-4.S
 'The ship god! My belly hurts (i.e. 'I am in labor.')
- (2) a. *a-kema-ha* *a-huraye* base clause
 4.A/POSS-foot-POSS 4.A-wash
 'I washed my feet.'
- b. *kema-huraye-an* NI
 foot-wash-4.S
 'I wash my feet.'
- (3) a. *e-yay-ram-u-an* *e-yay-car-o-suke* NI
 2SG.S-REFL-mind-POSS-exist-4.S 2SG.S-REFL-mouth-at.APPL-cook
 '(Until) you understand things and cook for yourself, (I will do my best to raise you).'
 (lit. 'you cook at **your own mouth**')
- b. **e-en-par-o-suke*
 2SG.A-1SG.O-mouth-at.APPL-cook
 Intended meaning: 'You cook for me.' (lit. 'you cook at **my mouth**')
- c. *en-par* (**ka*) *e-o-suke* quasi-NI
 1SG.O/POSS-mouth (even/also) 2SG.A-at-APPL-cook
 'You cook for me.' (lit. 'you cook at **my mouth**')

References

- Bugaeva, Anna, Johanna Nichols & Balthasar Bickel (2022), Appositive possession in Ainu and around the Pacific, *Linguistic Typology* 26(1), 43–88. Available at <https://www.degruyterbrill.com/document/doi/10.1515/lingty-2021-2079/html>
- Dahl, Östen (2004), *The Growth and Maintenance of Linguistic Complexity* [Studies in Language Companion Series, 71]. Amsterdam / Philadelphia: John Benjamins.
- Kokuritsu Ainu Minzoku Hakubutsukan (eds.) (2017), *Ainugo ākaibu* [Ainu language archive]. Shiraoui: National Ainu Museum. Available at <https://ainugo.nam.go.jp/> (accessed 10 January 2025).
- Koptjevskaja-Tamm, Maria & Ljuba Veselinova (2020), *Lexical typology in morphology*. Oxford Research Encyclopedia of Linguistics. Oxford: Oxford University Press.
- Kroeber, Alfred L (1909), *Noun incorporation in American languages*. XVI Internationaler Amerikanisten Kongress, ed. by Frantz Heger, 569–76. Vienna & Leipzig: Hartleben.
- Massam, Diane (2009), Noun Incorporation: Essentials and Extensions, *Language and Linguistics Compass* 3(4), 1076–96. Available at <https://compass.onlinelibrary.wiley.com/doi/10.1111/j.1749-818X.2009.00140.x>
- Mithun, Marianne (1984), The evolution of noun incorporation. *Language* 60, 847–894.
- Okuda, Osami (1992), *Shizunai hōgen no denshō. Orita Suteno no kōshōbungei* 2 [Folktales in the Shizunai dialect. Oral tradition by Suteno Orita 2]. Shizunai: Shizunai Board of Education.
- Sapir, Edward (1911), The problem of noun incorporation in American languages. *American Anthropologist* n.s. 13, 250–82.

- Satō, Tomomi (2001), Ainugo Chitose hōgen no “daisanrui no dōshi” no kōzō to kinō [The structure and functions of the so-called “verbs of the third class” in the Chitose dialect of Ainu]. *Hokkaido-ritsu Ainu Minzoku Bunka Kenkyū Sentā Kenkyū Kiyō* 7, 51–71.
- Satō, Tomomi (2022), Noun incorporation in Ainu. In Anna Bugaeva (ed.), *Handbook of the Ainu language* (Handbooks of Japanese Language and Linguistics 12). Berlin/Boston: De Gruyter Mouton, 549–571.
- Tamura, Suzuko. 1996. *Ainugo Saru hōgen jiten* [Dictionary of the Saru dialect of Ainu]. Tokyo: Sōfūkan.

The Latvian perfect on Facebook: an analysis based on data from comments

Anna Daugavet & Danguolė Kotryna Kapkan
(Vilnius University)

Keywords: perfect, Latvian, Baltic, Facebook, comments

The Latvian perfect, formed with a BE auxiliary and a past active participle (Kalnača & Lokmane 2021, 232) (1), exhibits a range of functions that partially overlap with its weakly grammaticalized Lithuanian counterpart, but additionally demonstrates highly grammaticalized, cross-linguistically typical functions of the Perfect, such as current relevance or ‘hot news’ (Daugavet & Arkadiev 2021).

- (1) [Naīvie cilvēciņi, vai jūs tiešām ticat,
ka šim šakāļu kantorim rūp savu darbinieku labklājība?]
Viņi vienkārši ir izrēķināj-uš-i,
3PL simply BE.PRS.3 calculate-PST.ACT.PTCP-NOM.PL.M
[ka nav izdevīgi strādāt!]
‘[Naive people, do you actually believe that this business of jackals is concerned about the wellbeing of their employees?] They have simply figured out [that it is not profitable to work!]

Earlier typologically and theoretically informed studies on the Latvian perfect (Nau 2005; Daugavet & Arkadiev 2021) relied on data from questionnaires and from the *LiLa* corpus (Utka et al. 2013). Mostly comprised of fiction, *LiLa* represents a carefully edited written language variety, where some of the most grammaticalized functions of the Perfect may be absent or underrepresented. The only corpus-based study of the Latvian perfect to date, Daugavet & Arkadiev (2021) focused exclusively on feminine singular tokens in affirmative contexts, additionally excluding instances where the optional present tense auxiliary was omitted. All this highlights the necessity for a more comprehensive corpus-based investigation, along with a broader diversification of data sources.

Our study draws on a doculect (Wälchli & Cysouw 2012) of Latvian Facebook comments from the main Latvian news outlet pages, which have proven particularly suitable for research on perfects (Kapkan 2024). Using Facepager software (Jünger & Keyling 2020), we compiled a non-annotated corpus of around 300 000 words and implemented a filter to identify comments containing words with all possible Latvian past active participle suffixes. Instances of the present perfect, either with or without the auxiliary, are selected manually to produce a database of n≈2000 tokens, annotated for the relevant features.

The goal of our study is to give a comprehensive description of the range of functions of the Latvian perfect, as it is used in Facebook comments, thus allowing for a robust comparison of the Baltic perfects in a naturalistic and informal language variety, given the equivalent study on Lithuanian (Kapkan 2024). Adopting a form-to-function approach, we classify the tokens of the perfect by function considering a range of criteria, such as transitivity, telicity, and semantic class of the lexical input, accompanying adverbials, event-result metonymy (Rosemeyer 2022), situational anchoring of the past event (Holvoet 2020; 2022), and other semantic properties. We align our functions with those

described for Lithuanian in Kapkan (2024), while providing clearer, better structured definitions. From a typological perspective, a detailed qualitative analysis combined with quantitative data on the function frequency of a fairly grammaticalized Latvian BE perfect will shed further light onto cross-linguistic tendencies in the development of perfects in relation to their source construction.

References

- Daugavet, Anna & Peter Arkadiev. 2021. The perfects in Latvian and Lithuanian: A comparative study based on questionnaire and corpus data. *Baltic Linguistics* 12(12). 73–165. <https://doi.org/10.32798/bl.922>.
- Holvoet, Axel. 2020. *The Middle Voice in Baltic. vargreb.5*. John Benjamins Publishing Company. <https://benjamins.com/catalog/vargreb.5>. (29 July, 2022).
- Holvoet, Axel. 2022. Irrealis, aspect, and unanchoring in Slavonic and beyond. *Zeitschrift für Slawistik* 67(1). 60–76. <https://doi.org/10.1515/slav-2022-0003>.
- Jünger, Jakob & Till Keyling. 2020. Facepager. An application for automated data retrieval on the web. <https://github.com/strohne/Facepager/>.
- Kalnača, Andra & Ilze Lokmane. 2021. *Latvian Grammar*. Rīga: University of Latvia Press. https://www.apgads.lv/fileadmin/user_upload/lu_portal/apgads/izdevumi/2021/Latvian_Grammar-2021.pdf
- Kapkan, Danguolė Kotryna. 2024. *The Grammaticalization of BE Perfects and beyond: Case Studies in Lithuanian, Bulgarian and Barese*. Vilnius: Vilnius University PhD dissertation. <https://doi.org/10.15388/vu.thesis.659>.
- Nau, Nicole. 2005. Perfekts un saliktā tagadne latviešu valodā [Perfect and Compound Present in Latvian]. *Baltu filoloģija* (14.2). 137–154.
- Rosemeyer, Malte. 2022. Anteriors and resultatives in Old Spanish. In Mar Garachana Camarero, Sandra Montserrat Buendia & Claus Dieter Pusch (eds.), *IVITRA Research in Linguistics and Literature*, vol. 31, 149–170. Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/ivitra.31.08ros>.
- Utkā, Andrius, Kristīne Levāne-Petrova, Daira Vēvere, Guna Rābante-Buša, Jolanta Kovalevskaitė & Erika Rimkutė. 2013. Lithuanian-Latvian-Lithuanian Parallel Corpus (LILA). Vytautas Magnus University. <https://repository.clarin.lv/repository/xmlui/handle/20.500.12574/6>. (10 January, 2025).
- Wälchli, Bernhard & Michael Cysouw. 2012. Lexical typology through similarity semantics: Toward a semantic map of motion verbs. *Linguistics* 50(3). <https://doi.org/10.1515/ling-2012-0021>.

Tracing the grammaticalization of the Hey!-gesture in Papua: From beckoning to an attention-drawing device and beyond

Anna Inbar & David Gil

(University of Haifa & Max Planck Institute for Evolutionary Anthropology)

Keywords: attention drawing devices, gesture, grammaticalization, Papua

The present study explores the emergence, usage, and grammaticalization path of an attention-drawing device used in Papua. The device under consideration takes the form of a (repetitive) bending of the hand from the wrist or from the base of the fingers downward; in some cases, the gesture involves bending the forearm and/or extending the arm forward. This gesture is referred to here by its English gloss “of convenience” (Kendon 1980) as the “Hey!-gesture,” following Haviland (2015, 2022).

The data are based on video recordings of interactions conducted in both natural and experimental settings, encompassing exchanges among hearing participants as well as their communication with two deaf children using a gestural communication system (cf. Goldin-Meadow 2003), which will be further explored. Utilizing the methodologies of Interactional Linguistics (Couper-Kuhlen & Selting 2018) and Multimodal Conversation Analysis (e.g., Goodwin 2018; Mondada 2016), we show that the gesture is typically used to indicate that the speaker/signer needs positive evidence of the interlocutor’s attention (cf. summons, Pillet-Shore 2018; Schegloff 1968, 2002). This occurs most often when the interlocutor displays cues of disengagement from the interaction, as the gesture serves to draw their attention or elicit a response—a function particularly salient in interactions with the deaf children.

The study suggests that the attention-drawing function of the gesture may have its roots in a beckoning gesture (a gestured “come here!”) used by locals—a form that is prevalent across cultures and languages (see, for example, Haviland 2015, 2022; de Vos 2012; Bressem et al. 2017). As a beckoning gesture usually signals for someone to approach or come closer to the speaker, it constitutes an imperative, establishes a spatial relationship between the speaker and the addressee, and indicates a specific direction. We argue that the aforementioned semantic components enabled this form to evolve into an attention-drawing device. The phenomenon whereby linguistic means reflecting the speaker’s desire to physically draw the interlocutor closer evolve into attention-drawing devices is attested cross-linguistically in both the visual and verbal modalities (e.g., Haviland 2015, 2022, de Vos 2012, Rosenthal 2005, Trommer et al. 2022). For example, the imperative form of the Hebrew movement verb *bo* ‘come’ combined with the spatial adverb *hena* ‘(to) here’ forms the pragmatic marker *bona/boena*, which is used to attract attention in interaction, typically regarding a noteworthy issue (e.g., Rosenthal 2005).

Moreover, devices to elicit an interlocutor’s attention can undergo further development. For example, in a new sign language emerging within a single extended family of indigenous peasants in Mexico (“Z”, Haviland 2015, 2022), a similar gesture used to solicit attention—potentially originating in a beckoning gesture—has evolved into a pragmatic marker of turn-beginning, indicating merely the intention to initiate a conversational turn. A similar function is evident also in Papua. In the present talk, we will attempt to account for the reasons for and constraints upon these developmental paths, while also addressing further developments in the function of the gesture.

References

- Bressemer, Jana, Stein, Nicole, & Wegener, Claudia (2017), Multimodal language use in Savosavo: Refusing, excluding and negating with speech and gesture, *Pragmatics* 27(2), 173–206.
- Couper-Kuhlen, Elizabeth, & Selting, Margret (2018), *Interactional Linguistics: Studying Language in Social Interaction*, Cambridge: Cambridge University Press.
- De Vos, Connie (2012), *Sign-spatiality in Kata Kolok*, Ph.D. dissertation, Nijmegen: Radboud University.
- Goldin-Meadow, Susan (2003), *The Resilience of Language: What Gesture Creation in Deaf Children Can Tell Us About Language-learning in General*. New York NY: Psychology Press.
- Goodwin, Charles (2018), *Co-operative Action*, Cambridge: Cambridge University Press.
- Haviland, John B. (2015), Hey! *Topics in Cognitive Science* 7, 124–149.
- Haviland, John B. (2022), How and when to sign “Hey!” Socialization into grammar in Z, a 1st generation family sign language from Mexico, *Languages* 7, 80.
- Kendon, Adam (1980), A description of a deaf-mute sign language from the Enga Province of Papua New Guinea with some comparative discussion, Part II: The semiotic functioning of Enga signs, *Semiotica* 32, 81–117.
- Mondada, Lorenza (2016), Challenges of multimodality: Language and body in social interaction, *Journal of Sociolinguistics* 20(3), 336–366.
- Pillet-Shore, Danielle (2018), How to begin, *Research on Language and Social Interaction* 51(3), 213–231.
- Rosenthal, Ruvik (2005), *Dictionary of Israeli Slang*, Jerusalem: Keter Books.
- Schegloff, Emanuel A. (1968), Sequencing in conversational openings, *American Anthropology* 70(6), 1075–1095.
- Schegloff, Emanuel A. (2002), Opening sequencing, in J. E. Katz, and M. A. Aakhus (eds), (2002), *Perpetual contact: Mobile communication, private talk, public performance*, Cambridge: Cambridge University Press, 325–385.
- Trommer, Pnina, Gvura, Avi, & Manor, Rama (2022), *Ani holekh likro'a 'et ha'ir (I'm gonna tear up the town): Motion verbs in contemporary Hebrew*, Tel Aviv: The Mofet Institute.

Diversity of relativization strategies in Sinitic Languages: A Wave theory account

Buqian Li

INALCO-CNRS-EHESS, CRLAO

Keywords: relativization strategies; Sinitic languages; Wave theory; historical evolution; areal distribution

Relativization strategies in Sinitic languages, or Chinese dialectal varieties, exhibit a remarkable diversity. Using data from 52 documented Sinitic languages and comparing them with different historical stages of Chinese, this study follows the framework and methods of Givón (2001, 2012) and Creissels (2006), and proposes a new typology of syntactic variations in relative clause formation, and offers a possible explanation for their geographical distribution through the lens of Wave theory (*Wellentheorie*, see François 2015).

Previous studies on relativization in Chinese focus mostly on the relativizer/nominalizer *de* in Standard Chinese, also with several studies on Cantonese, such as Yip & Matthews (2007). Liu (2005) was the first to highlight additional relativization strategies in Sinitic languages, such as demonstratives, demonstrative-classifier complexes and bare classifiers. Arcodia (2017) presented a typology of relativization based on head position and strategies in 44 Sinitic languages, but did not clearly map their geographical distribution.

We propose that Sinitic languages exhibit three basic types of relativization markers, determined by the following parameters: whether nominalization and relativization markers are distinguished (I) or not (II); whether the relativizer comes from the demonstrative pronoun *dǐ* (II-a) or the general classifier *gè* (II-b); and if neither of these conditions is met, (III) whether it also functions as a demonstrative (III-a), a demonstrative-classifier complex (III-b), or a specific classifier (III-c).

Type I (wave 1), sporadically found in Min, Wu, and Hakka, retains strategies of Archaic or Pre-Medieval Chinese (with nominalizers *zhě* or *suǒ*, and *zhī* for linking two nominals).

(1) Liancheng dialect (Hakka) (Xiang 1997: 336)

a. 洗菜□面盆。 (DN: domain noun)

[[\square *sai*⁵¹ *ts^hiu*³] *a*³⁵ REL] [*mie*³ *p^han*⁵⁵ DN]_i
wash vegetables REL basin

‘(This is) a basin for washing the vegetables.’

b. 洗菜个。

[[\square *sai*⁵¹ *ts^hiu*³] *kuə*³ REL]_i
wash vegetables NMLZ

‘(This is) the one for washing the vegetables.’

Type II (wave 2) aligns with Medieval Chinese configurations and is attested in Mandarin dialects and Jin (II-a, using *dǐ* or its derived forms), as well as in Cantonese, Gan, Xiang, and Hui (II-b, using *gè* or its derived forms, as in example 2a). This type, where the relativization marker serves both nominalization and relativization functions, is the most frequent strategy in Sinitic languages (including *de* in Standard Chinese).

Type III, with a demonstrative (III-a, wave 4) or a demonstrative-classifier complex (III-b, wave 3, as in example 2b), found in Mandarin and Jin dialects, aligns with (Pre-)Modern Chinese. The strategy using

specific classifiers (III-c, wave 5), which only spreads in the southeast coastal area, is likely an innovation. All three subtypes share a connection to nominal reference.

(2) Hui'an dialect (Southern Min)

(Chen 2020: 419, 423)

a. 𠵿是個租其厝

tse^{24} si^{22} [[en^{33} $tsɔ^{33}$ ___] e^0 REL] [$ts^h u^{31}$ DN]_i
 this be 3PL rent REL/NMLZ house
 'This is the house they rented.'

b. 汝買迄本冊

[lu^{53} bue^{53} ___]_i REL **hit⁵⁴⁻³-pun⁵³⁻²⁴** [$ts^h e\gamma^{54}$ DN]_i
 2SG buy that-clf book
 'that book you bought'

We argue that the wave theory adequately accounts for this distribution (see waves 1-5 above). Dialectal varieties located farther from the innovation center tend to exhibit more coexisting strategies due to successive waves. Thus, the presence of multiple strategies within a single variety, such as in Min, results from overlapping waves of diffusion.

References

- Arcodia, Giorgio Francesco (2017), Towards a typology of relative clauses in Sinitic: Headedness and relativization strategies, *Cahiers de Linguistique Asie Orientale* 46(1), 32-72.
- Chen, Weirong (2020), *A Grammar of Southern Min: The Hui'an Dialect*. Berlin / Boston: De Gruyter Mouton.
- Creissels, Denis (2006), *Syntaxe Générale, une introduction typologique : La phrase, Tome 2*, Paris: Hermès Sciences Publications.
- François, Alexandre (2015), Trees, waves and linkages : Models of language diversification. In Claire Bowern & Bethwyn Evans (eds), *The Routledge Handbook of Historical Linguistics*, 161-189, New York: Routledge.
- Givón, Talmy (2001), *Syntax: An Introduction*, Amsterdam / Philadelphia: John Benjamins.
- Givón, Talmy (2012), Toward a diachronic typology of relative clause, In Bernard Comrie and Zarina Estrada-Fernández (eds), *Relative Clauses in Languages of the Americas : A typological overview*, 3-25, Amsterdam / Philadelphia: John Benjamins.
- Liu, Danqing 刘丹青 (2005), Hànyǔ guānxì cóngjù biāoji lèixíng chūtàn 汉语关系从句标记类型初探 [A first look at the typology of relative clause markers in Chinese], *Zhongguo Yuwen* 中国语文 [Studies of Chinese language] (1), 3-15.
- Xiang, Mengbing 项梦冰 (1997), *Liánchéng Kèjiāhuà yǔfǎ yánjiū* 连城客家话语法研究 [A Grammatical Study of the Liancheng Hakka], Beijing: Language and Culture Press 语文出版社.
- Yip, Virginia and Stephen Matthews (2007), Relative clauses in Cantonese-English bilingual children : Typological challenges and processing motivations, *Studies in Second Language Acquisition* 29(2), 277-300.

The role of named analogues in the lexicalization of introduced concepts: A typological study

Christine Troussart Van Bulck, Jean-Christophe Verstraete
(KU Leuven)

Keywords: Other (non-European) languages: Australia and Oceania, Semantics, Typology

The lexicalization of introduced concepts is a classic topic in lexical typology, sometimes described as ‘lexical acculturation’ (see Brown 1999). We have a good understanding of the formal mechanisms involved, with a commonly used typology of coinage (e.g. *baada=baada* ‘trousers’, lit. ‘leg-leg’ in Gaagudju (Harvey 2002)), extension (e.g. from *bumbir* ‘(skin) rug’ to ‘clothes’ in Batjala (Bell 2003)), borrowing (e.g. *kuruutha* ‘clothes’ from English *clothes* in Kayardild (Evans 1995)), and various combinations (see Haugen 1950). However, the semantic side of the question is less well studied. This paper addresses a specific semantic problem, viz. the degree to which concepts or objects introduced in culture contact are really new, and how this may influence lexicalization. In many instances, there are actually partial analogues in the receiving lexicon. For instance, in Aboriginal Australia, metal axes are an introduced concept, since there is no indigenous tradition of metalwork. However, people did have axes made from other materials, primarily stone, which means that the shape and function of the object were not new, only the material used. This implies that there is a named analogue in the receiving lexicon, which raises the question how this will affect lexicalization of the new concept.

In this study, we investigate this relation in the indigenous languages of Australia, where we have a good understanding of ‘traditional’ culture (e.g. Keen 2004), as well as recent contact situations (from what is now Eastern Indonesia, e.g. Schapper 2022, and from Europe, e.g. Simpson 2024). Specifically, we use a sample of 200 Australian languages to investigate the effects of named analogues in the receiving lexicon. We study the lexicalization of 10 introduced concepts from the domains of tools, clothing/adornment, foodstuffs and social roles for which partial analogues can be found, across the whole sample.

The results show three broad patterns. One is to extend the term for the conceptual analogue, resulting in polysemy (e.g. Batjala *bumbir* ‘(skin) rug, clothes’ (Bell 2003)). A second pattern is to use the term for the conceptual analogue for the introduced object, and innovate a new term for the old one (e.g. *aymala* ‘metal axe’, versus *mutu* ‘stone axe’, extended from ‘blunt’ in Umpithamu (Verstraete 2020)). A final option is for the introduced item not to capitalize on the conceptual analogue at all (e.g. terms for ‘flour’ hardly ever being based on terms for native grain, even where there are traditions of grain processing). We chart these patterns across the sample, and we argue that they relate to a combination of semantic and cultural factors, e.g. clear functional differentiation within the same domain, as with axes (the metal type being superior to the traditional ones, compare the extension from ‘blunt’), a sufficient referential distance between domains to allow for polysemy, as with Western

clothes versus skin rugs, and cultural and material factors blocking extension, as with flour (most traditional grain processing uses wet-milling, leading to a paste rather than a powder-like substance, see Hoogmartens & Verstraete 2020).

References

- Bell, J. (2003) *A sketch grammar of the Badjala language of Gari (Fraser Island)*. MA Thesis, University of Melbourne.
- Brown, Cecil (1999) *Lexical acculturation in Native American languages*. Oxford: OUP.
- Evans, Nick (1995) *A grammar of Kayardild*. Berlin: Mouton de Gruyter.
- Harvey, Mark (2002) *A grammar of Gaagudju*. Berlin: Mouton De Gruyter.
- Haugen, Einar (1950) The analysis of linguistic borrowing. *Language* 26: 210-231.
- Keen, Ian (2004) *Aboriginal economy & society: Australia at the threshold of colonisation*. Oxford: OUP.
- Schapper, Antoinette (2022) Beyond 'Macassans': Speculations on layers of Austronesian contact in northern Australia. *Australian Journal of Linguistics* 41: 434-452.
- Simpson, Jane (2024). After 1788: Contact varieties in the first sixty years of Australia's colonisation. *Journal of Pidgin and Creole Languages* 39: 71–124.
- Verstraete, Jean-Christophe (2020) *A dictionary of Umpithamu, with notes on Middle Paman*. Canberra: Aboriginal Studies Press.
- Vicky Hoogmartens & Jean-Christophe Verstraete (2020) Rations: Flour, sugar, tea and tobacco in Australian languages. *Australian Journal of Linguistics* 40: 444-474.

Beyond evidentiality: Epistemic authority and other effects on the use of grammatical evidentials in Southern Finnic

Denys Teptiuk, Miina Norvik & Petar Kehayov
(University of Tartu, University of Tartu & University of Tartu)

Keywords: quotative mood, reported speech, reported evidentiality, epistemic stance, Estonian, Livonian

The Southern Finnic languages (Standard & South) Estonian and Livonian possess grammatical evidentials that cover two meanings: reportative with the unknown/unspecified source, and quotative with the specified source (Aikhenvald 2004; Kehayov & Skribnik 2018). The additional epistemic meaning of uncertainty in the truth-value of the report is often ascribed to Estonian (Aikhenvald 2004: 193), while this additional meaning is less typical for the Livonian grammatical evidential (Kehayov et al. 2012). Furthermore, mirative overtones arise in contexts where the reported state-of-affairs was not expected by the reporter (1). Such reports are also based on first-hand accounts of other speakers, inaccessible to the reporter (2).

(1) Estonian

Suur oli minu üllatus kui kaal
big be:PST.3SG 1SG.GEN surprise when weights
näitas, et mina kaaluvat kokku
show:PST.3SG COMP 1SG weigh:REP.EVID altogether
72,4 kilo.
NUM kilogram.PRT

‘I was surprised a lot when the weights showed that I weighed 72,4 kilos.’ (etTenTen21)

(2) Livonian

ma u'm kūlōn ku oksāka'ļdi Rīgōs
1SG be.1SG hear:APP.SG COMP three.spined.stickleback:PL.PRT Riga:INE
sāl Vēnas ve'jjjid ja Liepās ma
there Daugava:INE catch.fish:REP.EVID:PL and Liepaja:INE 1SG
u'm īž nā'nd ku kešļdōks āt ve'jjjōnd
be.1SG self see:APP.SG COMP landing.net:PL.INS be.3PL catch.fish:APP.PL
oksāka'ļdi
three.spined.stickleback:PL.PRT

‘I've heard that three-spined sticklebacks have been caught in Riga, there in Daugava, and in Liepaja I've seen that three-spined sticklebacks were caught with a landing net.’ (Suhonen 1975: 26)

This study extends previous descriptions of the grammatical evidentials in these languages and scrutinizes additional connotations arising from the grammatical expression of evidentiality based on corpus material. Furthermore, we investigate how the asymmetry in epistemic authority (Grzech

2020; Bergqvist & Grzech 2023, i.a.) between the current speech participants is reflected in the use of grammatical evidentials in these languages.

Our preliminary results for Estonian show an effect of this parameter on the distribution of evidentials relative to person. While the grammatical evidential is relatively frequent in reports about the current speaker, who is also the (first person) reporter (ca. 1500 exx. in etTenTen21), it is rarely used in reports about the (second person) interlocutor (ca. 60 exx.). Furthermore, in reports about the current speaker, the grammatical evidential acquires quotative interpretation more frequently than reportative. The quotative reading often triggers additional epistemic and mirative effects (cf. First Person Effect in Aikhenvald 2004: 225), for which contextual cues and alternations with the epistemically neutral indicative mood are robust indexes. By contrast, reports about the interlocutor more frequently do not contain specification of the source. This can be viewed as a distancing strategy: to avoid potentially face-threatening situations, the speaker keeps information sources vague. A similar situation is observed in reports about the current speaker boasting about their own qualities: *ma pidavat olema ilus* [1SG must:REP.EVID beautiful] ‘People say, I am beautiful’.

Considering the rarity of epistemic overtones in the use of the Livonian evidential (cf. Kehayov et al. 2012), we expect to find differences in the use of the grammatical evidentials with the same meaning between these closely related languages, potentially stemming from different conventionalisation paths of the evidentials therein and their uses in the limited number of genres available for moribund Livonian.

References

- Aikhenvald, Alexandra Y. (2004), *Evidentiality*. Oxford: Oxford University Press.
- Bergqvist, Henrik & Karolina Grzech (2023), The role of pragmatics in the definition of evidentiality, *STUF* 76(1), 1–30.
- etTenTen21 = Estonian Web 2021. Accessed via <https://app.sketchengine.eu>.
- Grzech, Karolina (2020), Epistemic primacy, Common Ground management and epistemic perspective, in H. Bergqvist, and S. Kittilä (eds), (2020), *Evidentiality, egophoricity, and engagement*, Berlin: Language Science Press, 23–60.
- Kehayov, Petar, Helle Metslang & Karl Pajusalu (2012), Evidentiality in Livonian, *Linguistica Uralica* 48(1), 41–54.
- Kehayov, Petar & Elena Skribnik (2018), Evidentials in Uralic languages, in A. Y. Aikhenvald (ed), *The Oxford Handbook of Evidentiality*, Oxford: Oxford University Press, 525–553.
- Suhonen, Seppo (1975), *Liivin kielen näytteitä* (Castrenianumin toimitteita 5.). Helsinki: Helsingin yliopisto.

Towards a typology of Future Anterior outside Europe

Future anterior forms (cf. English *I will have done*) are widespread, although they lack a comprehensive typological description. Nor are they discussed as a separate topic in cross-linguistic sketches on futures (such as De Brabanter et al. (ed.) 2014), in books on futures within language groups (such as Barranzini (ed.). 2017), and in the diachronic study on grammaticalization paths of this semantic field in Indo-European (Wiemer et al. 2024).

Future anterior is often construed as a counterpart of the present perfect shifted into the future temporal frame. However as shown in Penkova 2018, 2019, the forms of future anterior in different European languages have uses that cannot be boiled down to the sum of their structural components. Their uses include evidential and modal functions, conditionals, or verificatives ('it is found out that...'). My previous paper (Sitchinava 2022) with primary focus on Romance also uses data driven for typological questionnaires for European future anterior, discerning some additional context as the one labeled by me "ignorative" ('we do not know if P').

In the present talk I research the foci of variation of form and semantics of future anterior outside the (core) European linguistic area (but including the Volga and Caucasus areas), complementing previous studies on Europe. Another research question is in what regard these data inform our knowledge on its uses driven from the European material.

The sources of data are descriptive grammars and edited collections of texts and corpora. A database of contexts is extracted from these sources and annotated by semantic type.

Markers expressing anteriority in the future often differ from a straightforward combination of two markers "future + anterior", in contrast to their counterparts in most European languages. It can be a dedicated marker, and even in a bipartite form, either component may have a default interpretation that is not explicitly future or anterior. Anteriority can be marked by completive features (Réunion Creole) or perfective features (Nzime / Bantu A). Futureness is often represented by more general irrealis forms, as seen in Northern Mao (Omotic) or Neverver (Oceanic). Languages with dedicated markers for future anteriority are found in the Eastern Sudanic (Shatt), Chadic (Mbuko), and Austronesian (Māori, Bonggi) language families.

Semantically, the data feature counter-factual uses of future anterior (in Omotic languages and Tagalog), scarcely attested in the European area. Inferentive uses (such as 'He must have seen us') find parallels in Alekano / Kainantu-Goroka and Aghul / Northeast Caucasian. A development towards conditional uses is found in Bantu A (Basaa) and Chadic (Mbuko).

Interplay of marking and semantics can be illustrated by a class of 'prospective shift markers' (symmetrical to the so-called 'discontinuous past' markers, cf. Plungian, Auwera 2006) discerned in different languages, including Dongolawi / Nubian and Hill Mari / Uralic. In the latter case, the resulting form develops verificative uses similar to those found in Old East Slavic. Bantu languages exhibit modal markers that evolve into future perfects; this pattern also has a Slavic parallel, specifically in the Balkan forms employing the epistemic marker *da*.

The typological data from outside the European area further confirm the cross-linguistic robustness of the future anterior as a distinct TAM category, characterized by a typologically

stable cluster of uses. They provide valuable insights for constructing a semantic map of future anterior. In particular, they demonstrate that conditional meaning is not a prerequisite for the development of evidential uses, contrary to what European data might suggest.

References

- 1) Barranzini L. (ed.). 2017. *Le futur dans les langues romanes*. Berne: Lang.
- 2) De Brabanter, Philippe, Kissine, Mikhail, Sharifzadeh, Saghie (eds). *Future Times, Future Tenses*. Oxford: Oxford University Press.
- 3) Pen'kova, Yana A. 2018. From Retrospective to Prospective: Grammaticalization of Future Anterior in the languages of Europe // *Voprosy Jazykoznanija* 2: 53–70.
- 4) Pen'kova, Yana A. 2019. Modal and evidential strategies for future anterior in European languages: Questionnaire and corpus data // *Voprosy Jazykoznanija* 6: 7–31.
- 5) Plungian, Vladimir, van der Auwera, Johan 2006. Towards a typology of discontinuous past marking. *Sprachtypologie und Universalienforschung — Language Typology and Universals* 4: 317–349.
- 6) Sitchinava, Dmitri V. 2022. Polysemy of Future Anterior in Romance in a typological context // *Voprosy Jazykoznanija* 4: 48–65.
- 7) Wiemer, Björn, Hill, Eugen, Kölligan Daniel, Linnemeier, Jan-Niklas. *Between the birth and death of future tenses: Related languages as a natural lab for research into grammatical change*. München: Lincom, 2024. (LINCOM Studies in Indo-European Linguistics 58)

How non-finite can infinitives be: The case of Maa

Doris L. Payne
(University of Oregon)

Keywords: finiteness, infinitive, dependent clause, Nilotic

Maa (Eastern Nilotic) has no dependent-clause verb type that is devoid of all morphological categories typically considered “inflectional”. Tucker & Mpaayei (1955) called certain dependent verb forms “infinitives”. This paper addresses how non-finite these forms are. Based on their use in texts plus elicitation, we conclude that subject-number, mood, and aspect are required in Maa infinitives.

What is “finite/non-finite” depends on definition and language (Nicolaeva 2007, 2010; Chamoreau & Estrada-Fernandez 2016; *inter alia*). Traditional approaches focus on morphological categories required for verbs to function as predicates of main clauses, versus what is missing in verbs that cannot function this way. Tense, aspect, mood, number, and (especially subject) person-indexation are typically considered, and in some languages valency, evidentiality, direction, politeness, or speech-act categories. Very frequently, verb agreement with subjects is lost in non-finite forms. Other perspectives simply consider “non-finite” to be any verb that cannot be the predicate of a simple independent sentence (Nedjalkov 1998: 421), or a clause that cannot assign subject case (Chomsky 1981).

Maa lacks tense (König 1993). Main-clause verbs index **person**. In most conditions, **number** of subject, **aspect** (non-perfective—usually zero marked, perfective, or progressive), and **mood** (subjunctive or indicative) are also required. Number is an especially salient feature: argument prefixes distinguish 1SG from 1PL; in certain aspects, final-syllable reduplication versus its absence distinguishes 2PL from 2SG and a final falling tone versus its absence distinguishes PL from SG subject; and a few verbs have suppletive stems for subject-number. Lexical and free-pronoun subjects in matrix clauses have overt nominative (NOM) case if post-verbal, seen in (1)–(3) below.

Maa infinitives (italicized in the following examples) do not mark person, nor license nominative-case or possessive-form subjects. They are understood to have the same subject as a preceding matrix verb. They must reflect number of the matrix subject via *a-* INF.SG and *áa-* INF.PL.

- (1) *n-é-man-áa* *dúóó* *ɔl=mɔrraní* *a-iguɛn-á*
 CONJ-3-surround-ITIVE previous MSG=warrior.NOM INF.SG-advise-MID+NONPFV
 ‘and the previously-mentioned warrior went/goes around discussing/advising’ (Arinkoi.036)

Like finite verbs, infinitives carry morphology for indicative vs. subjunctive mood, cf. (2); and aspect, cf. (1), (3). Maa thus violates the statistical hierarchy: if no person agreement, then no TAM or special TAM forms (Cristofaro 2007).

- (2) *peê* *e-tum* *ɪl=tɔŋaná* *l-án* *áa-to-nin* *o=róréí*
 so CVB.3-get.opportunity MPL=people.NOM M=our.NOM INF.PL-SBJV-hear MSG=word
 ‘so/when our people are able/get to hear the word’ (Camus1.085)
- (3) *e-tī* *nínyɛ* *ol=pûl* *a-ɪnɔs-ɪtá* *in=kírí*
 3-be.at 3SG.NOM MSG=meat.camp INF.SG-eat-PROG FPL=meats
 ‘he was/is in the meat-camp eating meat’ (DC.009)

One could argue that the INF.SG and INF.PL morphemes incorporate a type of anaphor, fitting with theories stipulating that non-finite clauses cannot be specified for subject agreement but allow bound anaphors (Chomsky 1981). But whether a form constitutes “agreement” versus is an “anaphor” is theory-dependent and some languages do allow subject indexation/agreement and/or nominative subjects in infinitives (Nikolaeva 2007, *inter alia*). In sum, Maa infinitives disallow person and nominative case, but subject-number, mood, and aspect occur. Typologically, they thus fall toward the more “(morphologically) finite” end of “infinitive” dependent-clause structures.

References

- Chamoreau, Claudine & Zarina Estrada-Fernández (2016), Finiteness and nominalization, An overview, in Claudine Chamoreau, and Zarina Estrada-Fernández (eds), (2016), *Finiteness and Nominalization*, Amsterdam: John Benjamins, 1-10.
- Chomsky, Noam (1981), *Lectures on Government and Binding: The Pisa Lectures*, Dordrecht: Foris.
- Cristofaro, Sonia (2007) Deconstructing categories: Finiteness in a functional-typological perspective, in Irina Nikolaeva (ed.), *Finiteness: Theoretical and Empirical Foundations*, Oxford: OUP, 91-114.
- König, Christa (1993), *Aspekt im Maa*. Cologne: Institut für Afrikanistik.
- Nikolaeva, Irina (ed.) (2007), *Finiteness: Theoretical and empirical foundations*. Oxford: OUP.
- Nikolaeva, Irina (2010), Typology of finiteness, *Language and Linguistics Compass* 4(12), 1176–1189.
- Tucker, Archibald N. and John T. Ole Mpaayei (1955), *Maasai Grammar, with Vocabulary*, London: Longman, Greens & Co.

The Northern Samoyedic predestinative: Quantitative and areal evidence of grammaticalization

Elia Calligari – University of Pavia

This work analyses the predestinative suffix of Northern Samoyedic languages — Nenets, Enets, Nganasan (Uralic; Russia) — through a corpus-based approach combined with areal and diachronic evidence.

This suffix has attracted the attention of typological and Uralic linguistics alike due to its atypical behaviour and debated history, which has earned it the label of typological *rarissimum* (see Creissels & Daniel 2006 and Khanina & Shluinsky 2014). Most always combined with core grammatical cases, it has been variously defined as a benefactive (Prokof'ev 1937, Khanina & Shluinsky 2014) or nominal TAM marker (Nikolaeva 2015, Leisiö 2014), as endorsed by prototypical uses like (1) and (2). In the latter, the recipient is encoded inside the predestinative-marked NP expressing the theme, whereby the ditransitive transfer verb is formally turned monotransitive (Creissels & Daniel 2006, Malchukov et al. 2010).

- (1) Tundra Nenets (Nikolaeva 2015:103)
ŋəno-də-mt° *temtaə-d°m*
 boat-PRD-ACC.2SG buy-1SG
 'I bought a boat for you [/your future boat].'
- (2) Nganasan (Wagner-Nagy 2018:211)
mənə kniga-ðə-mtu *mi-s'iə-m*
 1SG book-PRD-ACC.3SG give-PST-1SG
 'I gave him/her a book.'

The corpus-based analysis draws from the INEL Northern Samoyedic corpora (Brykina et al. 2018 for Nganasan, Shluinsky et al. 2024 for Enets, and Budzisch & Wagner-Nagy 2024 for Nenets), extracting 200 occurrences of the suffix from each to investigate, through distributional evidence, the semantic features of predestinative-marked nouns and of co-occurring verbs. The crosslinguistic trend shows that the suffix generally combines with the accusative case to encode recipient-benefactive or proper recipient values. The occurrence of the predestinative suffix with nominative and genitive, instead, displays language-dependent features: Nenets has the lowest score of nominative predestinative forms and the highest of genitive predestinative ones while Nganasan displays the opposite distribution; Enets stays in between.

Considering areal data for the Enisej zone of Northern Siberia, it is observed that Evenki (Tungusic; Russia) features a case suffix, labelled *indefinite accusative*, which encodes recipient-beneficiaries inside the theme's noun phrase like the Northern Samoyedic predestinative as can be seen in (3).

- (3) Evenki (Nedjalkov 1997:147)
d'av-ja-v *o:-kal*
 boat-ACCIN-1SG make-IMP.2SG
 'Make a boat for me.'

Unlike Northern Samoyedic, however, the Evenki suffix fits in the language's case system and has close functional cognates across many Tungusic languages. Therefore, a case of pattern borrowing (Sakel 2007) from Tungusic to Northern Samoyedic can be postulated. The native Samoyedic item targeted by pattern borrowing is internally reconstructed as Proto-Samoyedic **-tə*, resulting from Proto-Uralic **-ksi* (Janhunen 1989). A grammaticalization path is thus reconstructed (Figure 1) based on Luraghi's (2016) paths for beneficiaries. Northern Samoyedic languages would have first featured the reconstructed Proto-Samoyedic suffix **-tə* < P.U. **-ksi* as a translative marker combinable with possessive suffixes, which synchronically corresponds to the genitive form of the predestinative declension pattern (Salminen 2014). Subsequently, pattern borrowing from Tungusic would have prompted a reanalysis of this suffix into a non-prototypical marker for recipient-beneficiaries encoded NP-internally. Such a reanalysis ultimately led to the back-formation of accusative and nominative predestinative forms.

This study represents the first attempt to a corpus-based study of the predestinative suffix for all Northern Samoyedic languages, combining new corpus data and evidence from historical linguistics to shed light on the grammaticalization of this suffix.



Figure 1. Proposed grammaticalization of the predestinative from P.U. **-ksi*.

References

- Brykina, Maria & Gusev, Valentin & Szeverényi, Sándor & Wagner-Nagy, Beáta. 2018. *Nganasan Spoken Language Corpus (NSLC)*. Version 0.2. Archived in Hamburger Zentrum für Sprachkorpora. Publication date 2018-06-12. (<http://hdl.handle.net/11022/0000-0007-C6F2-8>) (Accessed 2025-01-05.)
- Budzisch, Josefina & Wagner-Nagy, Beáta. 2024. *INEL Nenets Corpus*. (Unpublished working version.)
- Creissels, Denis & Daniel, Michael. 2006. Monotransitivity in ‘give’-constructions (exploring the periphery of ditransitives). (Paper presented at the Rara and Rarissima Conference, Leipzig, 29 March - 1 April 2006.)
- Janhunen, Juha. 1989. Samojedin predestinatiivisen deklinaation alkuperästä [On the origin of the Samoyedic predestinative declension]. *Suomalais-Ugrilaisen Seuran Aikakauskirja* 82. 298–301.
- Khanina, Olesya & Shluinsky Andrey. 2014. A rare type of benefactive construction: Evidence from Enets. *Linguistics* 52(6). 1391–1431.
- Leisiö, Larisa. 2014. Imennye kategorii vremeni v nganasanskom i drugix severnosamodijskix jazykax [Nominal tense categories in Nganasan and other Northern Samoyedic languages]. *Voprosy jazykoznanija* 1. 39–59.
- Luraghi, Silvia. 2016. The mapping of space onto the domain of benefaction and some unpredicted trends in semantic change. *Linguistics* 54(2). 339–389.
- Malchukov, Andrej & Haspelmath, Martin & Comrie, Bernard. 2010. Ditransitive constructions: a typological overview. In Malchukov, Andrej & Haspelmath, Martin & Comrie, Bernard (eds.), *Studies in ditransitive constructions*, 1–64. Berlin / New York: De Gruyter Mouton.

- Nedjalkov, Igor. 1997. *Evenki*. London: Routledge.
- Nikolaeva, Irina. 2015. On the expression of TAM on nouns: Evidence from Tundra Nenets. *Lingua* 166. 99–126.
- Prokof'ev, Georgij Nikolaevič (ed.). 1937. *Jazyki i pis'mennost' narodov severa – Čast' I*. (Jazyki i pis'mennost' samoedskix i finno-ugorskix narodov.) [Languages and literature of the peoples of the North – Part I (Languages and literature of Samoyedic and Finno-Ugric peoples)]. Moscow / Leningrad: Gosudarstvennoe učebno-pedagogičeskoe izdatel'stvo.
- Sakel, Janette. 2007. Types of borrowing: Matter and pattern. In Matras, Yaron & Sakel, Janette (eds.), *Grammatical Borrowing in Cross-Linguistic Perspective*, 15–30. Berlin / New York: De Gruyter Mouton.
- Salminen, Tapani. 2014. Suomalais-samojedilaisia muotovertailuja [Comparison of Finnish and Samoyedic forms]. In Inaba, Nobufumi & Luutonen, Jorma & Hamari, Arja & Ahola, Elina (eds.), *Juuret marin murteissa, latvus yltää Uraliin: Juhlakirja Sirkka Saarisen 60- vuotispäiväksi 21.12.2014*, 291–302. Helsinki: Suomalais-Ugrilainen Seura.
- Shluinsky, Andrey & Khanina, Olesya & Wagner-Nagy, Beáta. 2024. *INEL Enets Corpus*. Version 1.0. Archived at Universität Hamburg. In: The INEL corpora of indigenous Northern Eurasian languages. (<https://hdl.handle.net/11022/0000-0007-FE1D-C>) (Accessed 2025-02-10.)
- Wagner-Nagy, Beáta. 2018. *A Grammar of Nganasan*. Leiden: Brill.

Alignment, Subjecthood and Transitivity Prominence in Indo-European

Eystein Dahl, Tania Ahmad & Krzysztof Stroński

Adam Mickiewicz University, Poznań

Keywords: alignment, subjecthood, transitivity prominence, Indo-European, Principal Components Analysis.

This paper explores the relationship between alignment, subjecthood features and transitivity prominence in the different branches of the Indo-European family. In the present context, alignment refers to the morphosyntactic realization of core arguments. Subjecthood is a cover term for the morphosyntactic properties of the argument referred to as subject, which in this context is defined as the generalized syntactically privileged argument, a notion borrowed from RRG (cf. Van Valin and Lapolla 1997, Van Valin 2005). Transitivity prominence is understood as the degree to which predicates in a language select the pattern characteristic of core transitive verbs (cf. e.g., Haspelmath 2015, Creissels 2018). These three morphosyntactic dimensions are taken to constitute the argument realization system of a language. The empirical data are drawn from the oldest attested stages of representative languages of each of the branches of Indo-European by means of a three-tiered typological questionnaire, which is designed to capture fine-grained differences in each of the three dimensions. The subjecthood properties represent a subset of the ones identified by Falk (2006). Relative transitivity prominence is measured by means of the predicate list in Creissels (2018b), which comprises 30 verb meanings, each of which is assigned a score based on three variables (voice marking, case marking of first argument, case marking of second argument), the scores being summarized for each language, cf. Table 4 below. Taken together, this questionnaire enables a multivariate comparative analysis of the argument realization systems in the languages under scrutiny (e.g., Bickel 2015). It is unclear, however, whether the three morphosyntactic dimensions are diachronically interrelated, so that changes in one dimension may cause changes in another dimension, or not.

Assuming that each of the variables in each of the dimensions represent independent features, we aim to establish whether there are any observable correlations between features. To this aim, we submit the combined data to Principal Components Analysis (PCA), which allows for examining the correlations between variables, cf. Figures 1a and 1b (Le et al. 2008). These figures indicate that certain subjecthood properties tend to cluster, e.g., control and raising constructions, but are weakly associated with certain alignment features, e.g., ergative agreement. Figure 2 plots the resultant clustering of the languages in our sample, indicating that there is a fundamental distinction between languages where control and raising represent subject properties, notably Latin, Greek and Lithuanian (in Hittite, only raising) and languages where they don't (Le et al. 2008). Interestingly, neutral agreement, which is found in Greek, Lithuanian and Hittite, seems to play an almost equally important role in the distribution of the languages. Moreover, raising and control constructions are not subject properties in the languages showing tense/aspect-based split ergative

alignment, i.e., Vedic Sanskrit and Classical Armenian. These brief observations are indicative of considerable variation within the languages under consideration and may seem to suggest that language-specific developments within one of the three abovementioned dimensions may condition or restrict developments in the other dimensions.

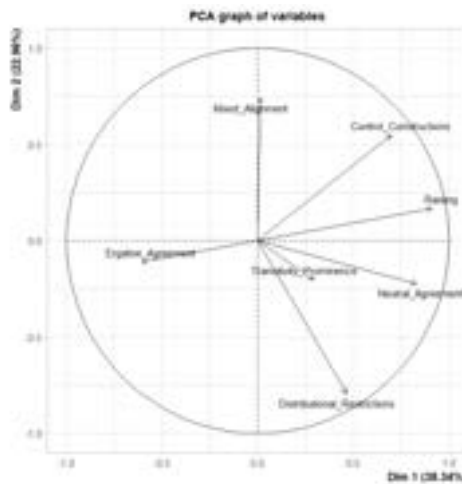


Figure 1a.

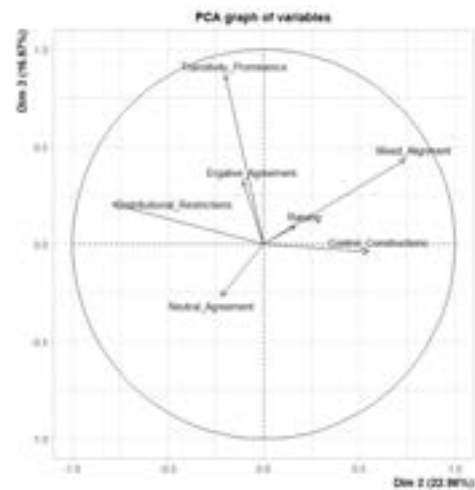


Figure 1b.

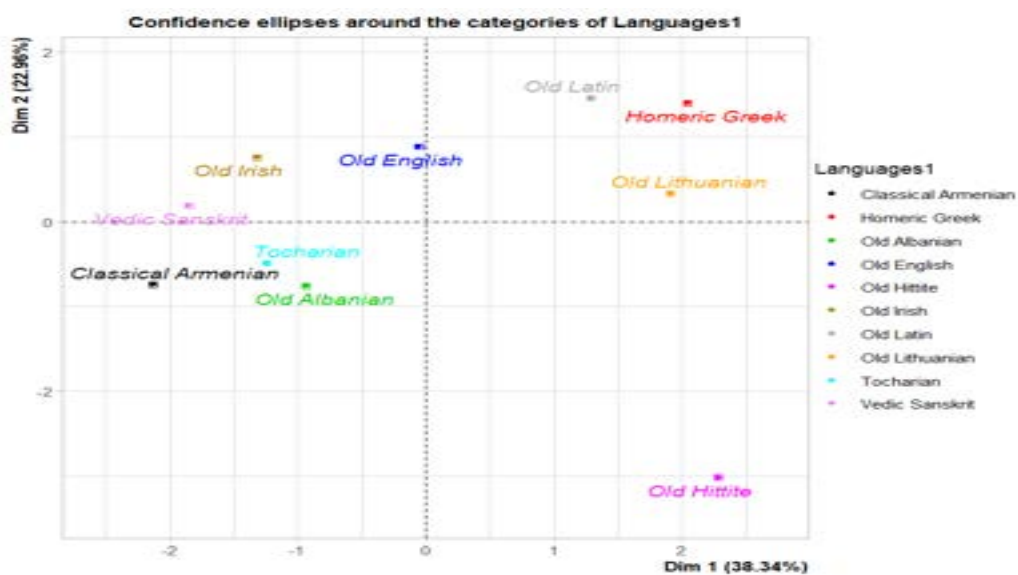


Figure 2.

This research is part of the project No. 2022/47/P/HS2/02564 co-funded by the National Science Centre and the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 945339.

References:

- Bickel, Balthasar. 2015. Distributional Typology: Statistical Inquiries into the Dynamics of Linguistic Diversity. In Bernd Heine and Heiko Narrog (eds.) *The Oxford Handbook of Linguistic Analysis (2nd Edition)*. Oxford: OUP, 901-924.
- Cotticelli, Paola and Eystein Dahl. 2022. 'Argument alternation and non-canonical argument marking in some archaic Indo-European languages.' In Eystein Dahl (ed.) *Alignment and Alignment Change in the Indo-European Family*. Oxford: OUP.
- Creissels, Denis. 2018a. The Obligatory Coding Principle in diachronic perspective. In Sonia Cristofaro and Fernando Zúñiga (Eds.). 2018. *Typological Hierarchies in Synchrony and Diachrony*. Amsterdam: Benjamins. 59–110.
- Creissels, Denis. 2018b. Transitivity prominence in typological perspective: the case of Basque. In Joseba A. Lakarra and Blanca Urgell (eds.) *Studia Philologica et Diachronica in honorem Joakin Gorrotxategi Vasconica et Aquitanica*. Special issue of *Anuario del Seminario de Filología Vasca «Julio de Urquijo» International Journal of Basque Linguistics and Philology LII: 1-2 (2018)*, 175-187.
- Dahl, Eystein. 2021. 'Pathways to split ergativity: The rise of ergative alignment in Anatolian and Indo-Aryan.' In *Diachronica*, 38:3 (2021), 413-456.
- Dahl, Eystein. 2022. 'Alignment in Proto-Indo-European'. In Eystein Dahl (ed.) *Alignment and Alignment Change in the Indo-European Family*. Oxford: OUP.
- Falk, Yehuda N. 2006. *Subjects and Universal Grammar. An Explanatory Theory*. Cambridge: CUP.
- Haspelmath, Martin. 2015. 'Transitivity Prominence.' In Andrej L. Malchukov & Bernard Comrie (eds.), *Valency classes in the world's languages: A comparative handbook*, vol. 1. Berlin: De Gruyter Mouton, 131–147.
- Le, Sebastien, Julie Josse and Francois Husson. 2008. FactoMineR: An R Package for Multivariate Analysis. In *Journal of Statistical Software*, 25(1), 1-18. 10.18637/jss.v025.i01
- Van Valin, Robert D. 2005. *Exploring the Syntax-Semantics Interface*. Cambridge: CUP.
- Van Valin, Robert D. and Randy Lapolla. 1997. *Syntax. Structure, meaning and function*. Cambridge: CUP.

Towards a typology of distributive markers: Evidence from North America

Filippo Maria Sergio & Simone Mattiola
(Alma Mater Studiorum – University of Bologna, University of Pavia)

Keywords: typology, distributives, plurality, pluractionality, morphosyntax

This presentation aims to provide a typological account of nominal and verbal distributive markers (DMs) by focusing on a genealogically balanced variety sample of 135 Indigenous languages of North America, including isolates, signed and mixed languages, pidgins and creoles for which a grammatical description is available (see Fig. 1).

Distributives convey “the separation of members of a group, whether entities, events, qualities or locations. Each is considered distinct in space, sort or time”, as Corbett (2000: 111) defines them (1-2).

- (1) Nominal DMs: Classical Nahuatl (Uto-Aztecan, Southern Uto-Aztecan; adapted from Andrews 2003:111)

<p><i>(cal)-li-</i> house-NC 'House.'</p>	>	<p><i>(cah~ca)-li-</i> DISTR~house-NC 'Houses located separately, various kinds of houses.'</p>
---	---	---

- (2) Verbal DMs: Northern Haida (Haida; Enrico 1988:56)

<p><i>'laa-.an</i> 3P-at 'They/People laughed at him.'</p>	>	<p><i>t'la k'ah-gan</i> INDEF laugh-PST 'They/People laughed at him (individually, not as a group).'</p>
--	---	---

Corbett's (2000: 111-117) description remains the most recent typological account of DMs to date. His purely synchronic contribution aims to place nominal DMs within his typology of number rather than exploring their cross-linguistic formal and functional properties. Crucially, it leaves unaccounted both the verbal domain, which previous studies highlighted as remarkably multifunctional (Mithun 1988, Mithun & Corbett 1995, Mithun 1999), and diachrony. Other studies, pursued through micro-typological, language-specific or formal approaches have dealt with DMs while addressing the broader phenomenon of pluractionality (Dressler 1968, Cusic 1981, Lasersohn 1995, Xrakovskij ed. 1999). Due to this wide-ranging scope of investigation, these studies ended up considering DMs' semantic-functional values as a mere semantic parameter, and hence overlooked their peculiar constructional properties. Therefore, no recent typological and sample-based study has yet addressed DMs' synchronic (forms, functions, cross-linguistic distribution) and diachronic (source and target constructions) properties.

This study addresses these gaps by providing a first typological survey of nominal and verbal DMs based on an ad-hoc built, genealogically balanced variety sample. The final dataset consists of 145 markers from 78 different languages. Each construction was tagged according to morphosyntactic (e.g., lexical category, marking strategy, specific marker) and semantic-functional parameters. Our work builds on recent typological research that sets distributives as a clear-cut class of pluractional markers (Wood 2007), adopts the findings and the standardized terminology of pluractionality based on a large-scale sample (Mattiola

2019) and acknowledges the importance of diachronic factors in shaping the cross-linguistic distribution and emergence of typological phenomena (Cristofaro 2014, 2019). For this reason, all the possible diachronic information found in the sources was also collected.

The study results in a first account of DMs' formal and functional properties, which also considers areal, genealogical, and diachronic information; ultimately, it constitutes the first building block towards the first typological survey of DMs in the languages of the world.



Fig. 1 Variety sample of 135 Native languages used for the study.

References

- Andrews, James Richard (2003), *Introduction to classical Nahuatl*. Norman: University of Oklahoma Press.
- Corbett, Greville G. (2000), *Number*. Cambridge: Cambridge University Press.
- Cristofaro, Sonia (2014), Competing motivation models and diachrony: What evidence for what motivations? In Brian MacWhinney, Andrej Malchukov & Edith Moravcsik (eds.), *Competing motivations in grammar and usage*, 282-298. Oxford: Oxford University Press.
- Cristofaro, Sonia (2019), Taking diachronic evidence seriously: Result-oriented vs. source-oriented explanations of typological universals. In Karsten Schmidtke-Bode, Natalia Levshina, Susanne Maria Michaelis & Ilja Seržant (eds.), *Explanation in typology: Diachronic sources, functional motivations and the nature of the evidence*, 25-46. Berlin: Language Science Press.
- Cusic, David (1981), *Verbal Plurality and Aspect*. PhD dissertation, University of Stanford.
- Dressler, Wolfgang (1968), *Studien sur verbalen Pluralität: Iterativum, Distributivum, Durativum, Intensivum in der allgemeinen Grammatik, in Lateinischen und Hethitischen*. Wien: Hermann Böhlau Nachf.
- Enrico, John (1988), *Aspect and Plurality in the Haida Verb*. Unpublished Ms. available at <https://www.uaf.edu/anla/record.php?identifier=HA975En1988b>.
- Laserson, Peter. 1995. *Plurality, Conjunction, and Events*. Dordrecht: Kluwer.

- Mattiola, Simone (2019), *Typology of pluractional constructions in the languages of the world*. Amsterdam: John Benjamins.
- Mithun, Marianne (1988), Lexical categories and the evolution of number marking. In Michael Hammond & Michael Noonan (eds.), *Theoretical morphology: Approaches in modern linguistics*, 211-234. San Diego, CA: Academic Press.
- Mithun, Marianne (1999), *The languages of native North America*. Cambridge: Cambridge University Press.
- Mithun, Marianne & Greville Corbett (1995), *Distributive morphology: its place in a typology of quantificational categories*. Presentation given at the Association for Linguistic Typology First Annual Meeting, Vittoria Gasteiz, Spain.
- Wood, Esther J. (2007), *The semantic typology of pluractionality*. PhD dissertation, University of California, Berkeley.
- Xrakovskij, Viktor S. (ed.) (1997), *Typology of Iterative Constructions*. Munich: Lincom.

The morphosyntax of (GOING) HOME in typological perspective

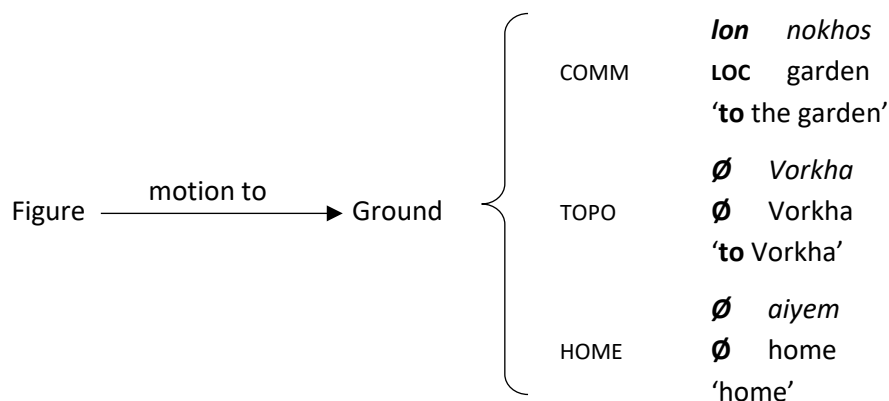
Julia Nintemann
(University of Bremen)

Keywords: spatial relations; toponyms; common nouns; home; differential place marking

“There is no place like home” – This sentiment is also reflected in many languages, as constructions involving the concept HOME often adhere to their own rules. The talk focuses on constructions where HOME is the goal of a movement in comparison to other common nouns (COMMS) and toponyms (TOPOS) in the same position. Linguistic research has already acknowledged that TOPOS often follow different rules than COMMS, especially in spatial constructions (cf., e.g., Stolz et al. 2014; Stolz & Nintemann 2024). Additionally, Haspelmath (2019: 322) observes that “languages sometimes give special treatment to a diverse set of nouns that denote concepts which are commonly used as spatial landmarks, such as ‘(one’s) house’ [...]”.

In Neverver [Austronesian], for example, COMMS must be marked with the general locative preposition *lon* in spatial constructions, whereas TOPOS are generally zero-marked. Furthermore, there is a restricted set of only three nouns that behave like TOPOS in that they are zero-marked, one of them being ‘(one’s) home’.

(1) Neverver [Austronesian] (Barbour 2012)



Matching Haspelmath’s (2019) term ‘topo-nouns’, *aiyem* ‘home’ receives special treatment by following the rules for TOPOS rather than those for COMMS. However, cross-linguistically, the special treatment of HOME does not necessarily imply the equal treatment of HOME and TOPOS. Sometimes, strategies deviating from both COMMS and TOPOS are employed, e.g., with adverbs as in Swedish *Jag går hem*, or verbs as in French *Je rentre*, both translating as ‘I go home’. The existence of constructions like French *chez moi* ‘at/to my place/home’ in many languages certainly supports the idea that the concept HOME receives special treatment. However, these constructions will not be included at this point.

As Haspelmath (2019: 319) postulates, “[i]f a language has asymmetric differential coding of place in common nouns and place names, the place-name marker will be shorter”. This is explained in terms of frequency-based predictability, as “place names (=toponyms) are more usually in a locative role than other nouns” (Haspelmath 2019: 315). It seems plausible that ‘(one’s) house/home’ is a spatial landmark that is frequented by most individuals of most speech communities (almost) on a daily basis, so that shorter marking due to a frequency-based predictability may also play a role here.

In this talk, I explore how HOME in Goal constructions is expressed cross-linguistically and if it behaves like COMMS, TOPOS, or neither. I demonstrate that the concept of (GOING) HOME indeed receives special treatment in languages all over the world, albeit by different means. Furthermore, I show that in accordance with Haspelmath's (2021: 2) form-frequency correspondence universal, Goal constructions involving HOME are on average shorter and less complex than constructions involving TOPOS or COMMS, respectively. For this purpose, I examine a sample of 100 areally and genealogically diverse languages from a functional-typological perspective. The data are primarily taken from descriptive grammars, complemented by primary sources and input from language experts.

Abbreviations

COMM = common noun, LOC = locative, TOPO = toponym

References

- Barbour, Julie (2012), *A grammar of Neverver* (Mouton Grammar Library 60), Berlin, Boston: De Gruyter Mouton.
- Haspelmath, Martin (2019), Differential place marking and differential object marking, *STUF - Language Typology and Universals* 72(3), 313–334.
- Haspelmath, Martin (2021), Explaining grammatical coding asymmetries: Form-frequency correspondences and predictability, *Journal of Linguistics* 57(3), 605–633.
- Stolz, Thomas & Julia Nintemann (2024), *Special Onymic Grammar in typological perspective: Cross-linguistic data, recurrent patterns, functional explanations* (Studia Typologica 34), Berlin, Boston: De Gruyter Mouton.
- Stolz, Thomas, Christel Stolz & Sander Lestrade (2014), *The Crosslinguistics of Zero-Marking of Spatial Relations* (Studia Typologica 15), Berlin, Boston: De Gruyter Mouton.

On transitivity and verbs without DOM in Estonian

Katrin Hiietam

(Univeristy College London)

Keywords: Transitivity, language typology, event type, case frames, DOM,

This paper looks at a small subset of Estonian transitive verbs that do not allow DOM and thus do not grammatically express a fully affected object in a telic event. We suggest their ungrammaticality together with objects marked with the case indicating a high degree of transitivity, may be due to the lexical semantics, rather than a true representation of a lower degree of transitivity. These verbs only express one specific event type – a continuous or spontaneous activity or event – and occur in only one specific verb frame (c.f. van Hout 2000).

In the literature on transitivity a highly transitive verb is expressing an event, where the subject acts volitionally and the object is totally affected by the activity expressed by the verb (e.g. Hopper and Thompson 1980). In Estonian, the distinction between higher and lower degree of transitivity is usually indicated by genitive vs the partitive case on objects in the singular. transitive constructions has been associated with reduced transitivity, in that it signals an object that is partially affected by the activity, and the genitive object indicates a fully affected object (e.g. Tamm 2012, Vaiss 2012, Metslang et al 2023), yet, the case frames in tables 2 and 3 suggest the grammatical marking of transitivity may be more nuanced.

The original basis for this study is the ValPaL list (75 basic verb meanings, Haspelmath 2015) which gave rise to the set of Estonian verbs with corresponding meanings. The majority of the studied verb meanings in Estonian were also encoded transitively 63.8% (53 verbs), i.e. with the unmarked subject and the marked object. About ¾ of them (75.5%) showed DOM on the object argument (PART v GEN/NOM) depending on the degree of transitivity and the number of the argument. Also, for a highest degree of transitivity, not only genitive marking is required, but also an adposition. Partitive on its own signals slightly reduced degree of transitivity. A sample of those verbs is given in Table 1.

Table 1. A sample of transitive verbs with DOM

Verb meaning	Case frame SG	Case frame PL
sööma (eat)	NOM – PART v GEN + ära	NOM – PART v NOM + ära/
kallistama (hug)	(PERF)/[välja (out), puhtaks	välja, puhtaks, valmis, vigaseks,
vaatama (look at)	(clean), valmis (ready), vigaseks	üles, kinni, lõpuni], tükkideks,
nägema (see)	(cripple), üles (up), kinni	lahti
lõhna tundma (smell)	(closed), lõpuni (until the end)],	
	tükkideks (into pieces), lahti	
	(open)	

However, our dataset also contained a group of verbs that do express either physical or cognitive events, yet do not accept the partitive v genitive marking alternation on the objects for signalling the

affectedness of the object or the telicity of the event for semantic or lexical reasons. These verbs, given in Tables 2 and 3, formed 13.2% of the set.

Table 2. Transitive verbs with only NOM-PART frame

Verb meaning	Case frame SG	Case frame PL
aitama (help)	NOM-PART	NOM-PART
järgima (follow)		
kohtama (meet)		
puudutama (touch)		
pilgutama (blink)		

Table 3. Transitive verbs with only NOM- GEN +adposition frame.

Verb meaning	Case frame SG	Case frame PL
karjuma (shout)	NOM- GEN +peale (on/at)	NOM- GEN +peale

All the verbs in Tables 2 and 3 express events with a possible end point, yet, they do not appear in the frames allowing for object case alternation. This could perhaps be explained by assuming their case frames are defined in the lexicon for continuous and/or spontaneous events. .

References:

- Haspelmath, Martin. 2015. 'Transitivity Prominence.' In Andrej L. Malchukov and Bernard Comrie (eds.), *Valency classes in the world's languages: A comparative handbook*, vol. 1. Berlin: De Gruyter Mouton, 131–147.
- Hopper, Paul J. & Thompson, Sandra A. (1980). Transitivity in grammar and discourse. *Language* 56, 251–299.
- Hopper, Paul J. and E. C. Traugott. 1993. *Grammaticalization*. Cambridge: Cambridge University Press.
- Van Hout, Angeliek. 2000. *Event Semantics in the Lexicon-Syntax Interface: Verb Frame Alternations in Dutch and their Acquisition*.
- Metslang, Helle et al. 2023. *Eesti Grammatika*. [The Grammar of Estonian] <https://dspace.ut.ee/items/7a3f1c4c-19fb-4ff3-9410-55810b7a8ef8?fbclid=IwAR2dyn9Ji9DvjaMN53NBQI3f4HKuP5BttB4s5imPflzCV4nCI5LsURcCTuY>
- Tamm, Anne. 2012. *Scalar Verb Classes. Scalarity, Thematic Roles, and Arguments in the Estonian Aspectual Lexicon*, 2012 (Biblioteca di Studi di Filologia Moderna; 14)
- Vaiss, Natalia. 2021. Verbide transitiivsuse kontiinumist eesti keeles. *Emakeele Seltsi Aastaraamat* 66 (344–386). Tallinn: Teaduste Akadeemia Kirjastus. <http://dx.doi.org/10.3176/esa66.15>
- Valency Patterns Leipzig (ValPaL) database - <https://valpal.info/>

Frequency effects in verbal argument marking: A spoken typology approach

The study of frequency effects in grammar has a long tradition in typology (e.g. Greenberg 1966, Haiman 1983, Bybee et al. 1994, Haspelmath 2021). They refer to the cross-linguistic tendency for a more frequent grammatical expression to be shorter than a comparable less frequent expression. Greenberg (1966: 31-32) already showed that singular verb forms are generally shorter than their plural and dual counterparts, and that third person forms are shorter than first and second person forms. He accounted for these differences in length by an inverse relation to the forms' frequencies in corpus data. In a large-scale quantitative typological study, Seržant & Moroz (2022) confirmed those length patterns with verbal subject markers, but did not measure frequency distributions. Additionally, two important empirical issues remain. First, typological studies usually use aggregated data, which does not capture the variation in exponence within languages. Second, the length of grammatical expressions is usually approximated by the number of (orthographic) segments, which excludes attested phonetic variation in speech. In this study, I test for frequency effects in verbal person marking addressing both issues by using data from spontaneous speech corpora. I use data from eight typologically distinct languages in DoReCo (Seifart et al. 2022): Beja, Baïnounk Gubëeher, Nafsan, Dalabon, Movima, Texistepec Popoluca, Kamas, Dolgan.

First, I model the overall effect of frequency on duration, predicting marker duration from marker frequency, while controlling for local speech rate, index position within the word and co-occurrence of silent pauses, semantic density, and individual speakers. I find robust overall effects of frequency on marker duration in the expected direction that hold up even when controlling for phonological shapes of the markers in the model, confirming frequency driven variation on the phonetic level (Fig. 1).

Second, I test the universal trends proposed by Greenberg (SG < PL, DU and 1 < 2, 3) by modeling marker duration and marker frequency across the nine types. According to Greenberg's proposal, we would expect to find a frequency distribution (Fig. 2 right) that is the mirror image of the durational patterns (Fig. 2 left), which is not the case. For instance, 3PL forms are predicted to be as frequent as 1SG markers, but their durations are predicted to be substantially longer. One possible explanation for such "deviations" from frequency-driven patterns is that phonological reduction is weaker for argument markers that are discontinuous. This is shown for Baïnounk Gubëeher in example (1b) for the first person plural (exclusive), which is expressed by a plural suffix *-min* in addition to the first person prefix *i-* used in singular contexts (1a).

To conclude, while I can confirm overall frequency effects in verbal argument marking with novel empirical methods, I discuss where frequency has less impact on shaping grammatical markers and what potential reasons for such deviations are.

- (1) Baïnounk Gubëeher (Cobbinah 2022)
- a. bare na i-ló-kénén tu hě-gěni i-wul
 but there 1-speak-3PL.POSS all AGR:HA-REL 1-see
 'but I'll tell them everything I saw'
- b. bala i-gěn-o-min
 before 1-show-O:2SG-1PL:EX
 'before we show you, ...'

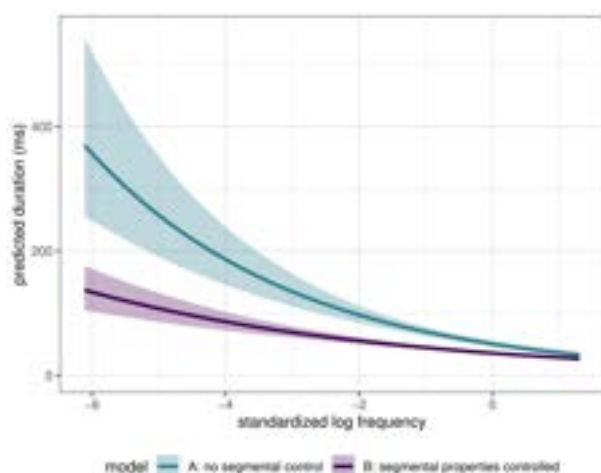


Figure 1: Conditional effects of frequency on marker duration.

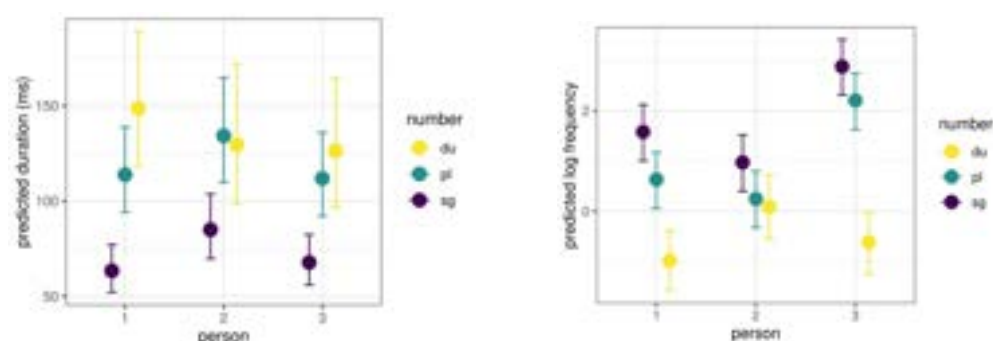


Figure 2: Conditional effects of marker type on duration (left) and frequency (right).

References

- Bybee, Joan & Perkins, Revere & Pagliuca, William. 1994. *The evolution of grammar. Tense, aspect, and modality in the languages of the world*. Chicago: The University of Chicago Press.
- Cobbinah, Alexander Yao. 2022. Bāinounk Gubēher DoReCo dataset. In Seifart, Frank & Paschen, Ludger & Stave, Matthew (eds.), *Language documentation reference corpus (DoReCo) 1.2*. Berlin & Lyon: Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2).
- Greenberg, Joseph. 1966. *Language universals: With special reference to feature hierarchies*. The Hague: Mouton.
- Haiman, John. 1983. Iconic and economic motivation. *Language* 59(4). 781–819.
- Haspelmath, M. 2021. “Explaining Grammatical Coding Asymmetries: Form–Frequency Correspondences and Predictability.” *Journal of Linguistics*, 1–29.
- Seifart, F., L. Paschen & M. Stave, eds. *Language Documentation Reference Corpus (DoReCo) 1.2*. Berlin & Lyon: ZAS & laboratoire Dynamique Du Langage, 2022.
- Seržant, Ilja & Moroz, George. 2022. Universal attractors in language evolution provide evidence for the kinds of efficiency pressures involved. *Humanities and Social Sciences Communications* 9(1). 1–9.

Violation of typological universals in Marwari

Liudmila Khokhlova

(Moscow University)

Key words: typological universal, accusative/ergative type of case marking, verbal concord, split ergativity

Indo-Aryan languages provide rich material for studying the ways of typological evolution: first from nominative-accusative towards ergative and then back towards nominative-accusative alignment. This paper examines violation of four typological universals in Marwari - an Indo-Aryan language spoken by about 13.2 million people mainly in the Indian state of Rajasthan. The research is based on prose texts in Marwari of the 14th -20th centuries as well as on the field work in Rajasthan. The violations described below concern the period when the language was evolving from ergative towards accusative strategy of syntactic coding. The situation when the old typological system is collapsing and the new one has not yet been established may have triggered the violation of some typological universals.

I. According to Anderson (1977) and Comrie (1978), any language cannot simultaneously have the ergative type of verbal concord and the accusative type of case marking - In Marwari the accusative type of case marking (A=S) is typical for many nominal paradigms. It goes well with the ergative type of verbal concord:

(1) Modern Marwari

gītā (S) *uṭh-ī*
Gita. NOM.F get up. PAST-F

‘Gita got up.’

(2) Modern Marwari

gītā (A) *rāwan* (O) *nai* *māriy-au* *h-ai*
Gita. NOM.F Rawan.M ACC beat-PP.M.SG be-PRES.3.SG

‘Gita has beaten Rawan.’

II. Dixon (1994:84) argues that ‘If pronouns and nouns have different systems of case inflection, then the pronoun system will be accusative, and the noun system ergative, never the other way around’. Compare similar marking of A and S in nominal paradigm [(3), (4) - accusative type] and different marking of A and S in pronominal paradigm [(5), (6) - ergative type]:

(3) 18-th century Marwari:

bādsāh (S) *uṇ* *darwes* *kanhai* *gay-ā*
king.NOM that.OBL dervish to go-Past.M.SG

‘The king went to that dervish’ (R.G.:67)

(4) 18-th century Marwari:

bādsāh (A) *kāl* *kiy-ā*
king.NOM death do-PAST.M.SG

‘The king died’ (R.G.:73)

(5) 18-th century Marwari

<i>hūⁿ</i> (S)	to	<i>tamām</i>	<i>haj</i>	<i>vec-ūⁿ</i>	<i>h-uⁿ</i>
I.NOM	part.	all	pilgrimage	sell-PRES.1.SG	be-PRES.1.SG

‘As for me, I am selling all pilgrimages’ (R.G.:67)

(6) 18-th century Marwari

<i>maiⁿ</i>	<i>pahlī bār</i>	<i>ek-hī</i>	<i>dāras</i>	<i>sūⁿ</i>	<i>piyāl-o</i>	<i>bhariy-o</i>	<i>th-o</i>
I.ERG	first time	one-EMPH	pomegranate	with cup-M.SG	fill-PP.M.SG	be-PAST.M.SG	

‘For the first time I filled the cup with (the juice of) one pomegranate’ (R.G.:76)

III. Dixon (1994:93) has shown that ergativity in split systems is usually most strongly marked in plural. Compare different marking of 1-st person singular pronoun (5), (6) and similar marking of 1-st person plural pronoun (7) in accusative and ergative domains:

(7) 18-th century Marwari

<i>mhe</i>	<i>ek</i>	<i>hāthī</i>	<i>liy-o</i>	<i>ch-ai,</i>	<i>tiṇ</i>
we.NOM	one	elephant	take-PP.M.SG	be-PRES.3.SG	so
<i>mhe</i>	<i>bhelī</i>	<i>asawārī</i>	<i>kar-āⁿ</i>		
we.NOM	together	riding	do-SUBJ.1.PL		

‘We have taken (only) one elephant, so we will ride this elephant together’. (R.G.:54)

IV. According to Trask (1979), Indo-Aryan languages belonging to ‘B’ ergative model do not allow verbal agreement in person. Verbal agreement with O in person was allowed in Old Marwari (8), while Modern Marwari sometimes allows agreement in person with A (9):

(8) 15-th century Marwari

<i>śrīpur-nagar-nāyak-i...(A)</i>	<i>tīhaⁿ</i>	<i>cor-ahaⁿ</i>	<i>māri-vā</i>	<i>nimittu</i>
Shripur-city-ruler-INSTR	this.OBL.PL	thief- OBL.PL	kill-INF.OBL	for
<i>amhe (O)</i>	<i>mokaḷ-iy-ā</i>	<i>ch-āⁿ</i>		
we.NOM	send-PP-M.PL	be-PRES.1.PL		

‘The ruler of Shripur city has sent us to kill these thieves’ (R.G.:11)

(9) Modern Marwari

<i>mhe</i>	<i>sapn-ai</i>	<i>meⁿ</i>	<i>ī</i>	<i>āp-rai</i>	<i>sāthe</i>	<i>dag-au</i>
we.NOM	dream-M.OBL.SG	in	EMPH	you-GEN	with	deception-M.SG
<i>karaṇ</i>	<i>r-au</i>	<i>vicār</i>	<i>kariy-au</i>	<i>vh-āⁿ</i>	<i>tau</i>	
do-INF.OBL	GEN-M.SG	thought	do-PP.M.SG	be-PRES.1.PL	then	
<i>mhāⁿ-nai</i>	<i>narak</i>	<i>meⁿ</i>	<i>ī</i>	<i>ṭhaur</i>	<i>nī</i>	<i>mil-ai</i>
we-DAT	hell	in	even	place	NEG	be obtained-SUBJ.3.SG

‘If I could think of deceiving you even in my dreams, there would be no place for me even in hell’ (Bahal 1989:107)

Further investigation of genetically and geographically different languages with different typological history will help to answer the question: to what extent the violation of typological universals depends on paths of morpho-syntactic evolution.

References

- Anderson, Stephen (1977), *On mechanisms by which languages become ergative*, in C. N. Li (ed), (1977), *Mechanisms of Syntactic Change*, Austin: University of Texas Press, 317 – 363.
- Bahal, Kali Charan (1989), *Ādhunik Rājasthānī kā sanrachnātmak vyākaraṇ* (Generative Grammar of Modern Rajasthani), Jodhpur: Rājasthānī sāhitya sansthān. (in Hindi).
- Comrie, Bernard (1978), *Ergativity*, in W. P. Lehmann (ed), (1978), *Syntactic Typology*, Austin: University of Texas Press, 329-95.
- Dixon, Robert M.W. (1994). *Ergativity*, Cambridge: Cambridge University Press.
- Trask, Robert L. (1979). *On the origins of ergativity*, in Frans Plank (ed.), (1979), *Ergativity: Towards a theory of grammatical Relations*, London: Academic Press, 385–404.

Referred texts

- R.G. – *Rājasthānī Gadya: vikās aur prakāś*. 1969. Narendra Bhanāvat (ed.), Agrā: Śrīrām Mehtā and Company.

When I imagine what I do not want you to do:
2nd person irrealis construction as a prohibitive function in Chepang

Languages employ irrealis forms to express imperative and prohibitive functions (Chafe 1995; Mauri & Sansò 2012; Van Der Auwera & Devos 2012). The use of an irrealis construction to convey a prohibitive function is typologically rare in absence of negative marker. This paper discusses an example of this rare construction in Chepang (Trans-Himalayan) where two prohibitives are found: a morpheme =*ljam* (1), and an irrealis construction combining the irrealis marker =*teja* and the 2nd person morpheme =*te* (2).

The irrealis marker =*teja* expresses a state of affairs that has not yet taken place in reality but could have been possible in the past (3) or could be possible in the future (4). This main function of irrealis corresponds to the definition found in Elliot (2000), Mithun (1999), van der Auwera and Devos (2012), and Haan (2012), that is, states of affairs that are imagined or not real, i.e., not actualized in reality.

The prohibitive function of the irrealis construction is attested when the function of 2nd person is S and A, but not O, as shown through (5) to (7). In an irrealis construction that functions as a prohibitive without an overt negation, the speaker₁ tells the interlocutor₂ what they₁ imagine they₂ could potentially do while wishing that it does not happen in reality. Thus, ex. (5) for instance, can be read as *si=te=teja* ‘You will die!’ but means ‘Don’t die!’

This type of construction sheds light on a possible diachronic development of irrealis and prohibitive markers under sociopragmatic pressures. The 2nd person irrealis prohibitive construction is less direct and more polite or soft than the use of the prohibitive marker =*ljam*. Chepang seeks renewal in second person form of address, as is the case for person indexation forms in Trans-Himalayan languages (DeLancey 2018).

- (1) *laʷ, maja, naŋ i=tʌ dahj=ljam,*
well love 2SG PROX=NMZ:ADV1 say=PROH
naŋ mu=sa=le pʌr=na.
2SG COP_stay_live=NMZ=DIS have.to=NPST
‘Well, love, don’t say this, you have to stay.’
(CH_MKW_RP_CHI_102519_1_Life_Archive)

- (2) - *naŋ kʌʷ=te=teja da ane !*
2SG flee=2=IRR PART PART
- *kʌʷ=ŋʌ=lʌ, kʌʷ=ŋʌ=lʌ !*
flee=1=NEG flee=1=NEG
‘- Don’t flee, come on!
- I won’t flee, I won’t flee!’
(CH_CTW_KMC_TAP_102520_2_The bat and the crab_Archive)

- (3) “*i mu=ma=l=o kʰe=ja ŋa=kaj ten*
PROX COP=NEG=COP=NMZ COP=COND 1SG=DAT today
sat=teja=u kʰe=to” dahj=ti=taŋ dahj=a kja !
kill=IRR=3O COP=REM:PST say=SEQ=ATT say=PST PART
‘S/he said “(The eagle) would have killed me today if he hadn’t been there.”’
(CH_MKW_CPR_BAN_102817_1_Mit_Co'_Archive)

- (4) *gʌ=tʌ haj=ti sjak=teja=ŋ ? dʒe=sa na=lʌ.*
INT=NMZ:ADV1 do=SEQ1 survive=IRR=1 eat=NMZ COP=NEG
‘How will I survive? There is no food.’
(CH_MKW_SBC_BGR_101719_1_Life_Archive)

- (5) *si=te=teja*, *si=sa* *k^he=l_Λ*, *pahj=ti* *way=l_Λ*.
 die=2=IRR die=NMZ COP=NEG leave=SEQ1 come=2SG.IMP.INTR
 ‘Don’t die, dying is not a solution, leave and come (here).’
 (CH_MKW_SC_SIL_010220_3_Life_Archive)
- (6) *ama* *pir=te* *ɖahj=teja=u*,
 mother worry=2 do_make=IRR=O
l_Λ, “*ŋa si=na=ŋ*” *m^hl_r=te=teja*.
 uh 1SG die=NPST=1SG think=2=IRR
 ‘Don’t worry mother, uh, don’t think “I will die.”’
 (CH_MKW_SC_SIL_010220_3_Life_Archive)
- (7) *mu=l_Λ*, *b^hitr_Λ* *pok=ti*,
 stay_live=2SG.IMP.INTR inside get.inside=SEQ1
b^hena=i *sat=te=teja*.
 brother.in.law=ERG kill=2=IRR
 ‘Having gotten inside, stay there, (otherwise) brother-in-law will kill you.’
 (CH_CTW_MKC_BRB_040522_1_Cing_Lan_Archive)

References

- Chafe, Wallace L. 1995. The Realis-Irrealis Distinction in Caddo, the Northern Iroquoian Languages, and English. In. <https://api.semanticscholar.org/CorpusID:148063848>.
- DeLancey, Scott. 2018. Deictic and sociopragmatic effects in Tibeto-Burman SAP indexation. In *Typological hierarchies in synchrony and diachrony*, 345–375. Amsterdam: John Benjamins.
- Elliott, Jennifer R. 2000. Realis and irrealis: Forms and concepts of the grammaticalisation of reality. *Linguistic Typology* 4(1). 55–90. <https://doi.org/doi:10.1515/lity.2000.4.1.55>.
- Haan, Ferdinand De. 2012. Irrealis: fact or fiction? *Language Sciences* 34(2). 107–130. <https://doi.org/10.1016/j.langsci.2011.06.016>.
- Mauri, Caterina & Andrea Sansò. 2012. The reality status of directives and its coding across languages. *Language Sciences* 34(2). 147–170. <https://doi.org/10.1016/j.langsci.2011.08.002>.
- Mithun, Marianne. 1999. *The Languages of Native North America*. Cambridge: Cambridge University Press.
- Van Der Auwera, Johan & Maud Devos. 2012. Irrealis in positive imperatives and in prohibitives. *Language Sciences* 34(2). 171–183. <https://doi.org/10.1016/j.langsci.2011.08.003>.

On different ways of changing: a typological study of semantic shifts and constructions with ‘way/road’

This study focuses on various types of linguistic change involving nouns meaning ‘way, road’. Our conventional sample is skewed towards the European languages, but it includes languages from Central America, Africa, and Southeastern Asia. The source-focused perspective here differs from the target-focused methodology, more common in typological studies. Even the fundamental work by Heine and Kuteva (2002) does not cover many patterns of linguistic change attested for particular semantic concepts and is also limited by the targets that are treated as grammatical.

The most obvious type of linguistic change is exemplified by widespread metaphorically motivated polysemy patterns (WAY = MANNER or WAY = MEANS) with the more abstract meaning derived from the more concrete one. Such patterns can be investigated using existing cross-linguistic databases, e.g., CLICS (Rzymiski, Tresoldi 2019), Concepticon (List et al. 2025) or EvoSem (François et al. 2025). Another interesting, but overlooked pattern is WAY = TIME QUANT, typical for the Balkan linguistic area, e.g., Bulgarian *păt(i)* ‘way(s) → time(s)’, cf. also Kakabe *kílá* (Mande; see Pozdniakov 2020) or Mandarin *dào* (Jiang 2004).

The nouns in focus are often involved in a more complicated linguistic change, which often results in the appearance of new larger constructions with the original noun or its derivative occupying a fixed slot. The semantic shift WAY = MEANS often involves certain fossilization and lexicalization of particular forms of the noun, exemplified by simple or complex adpositions, such as Basque *bidez* ‘by means of’ (Jendraschek 2020: 379) or Spanish *por vía de* ‘by way of, through, via’.

Departing from spatial semantics – prolative, perlative, or more unusual ambulative (as in Totonacan, see Beck 2004; Garcia-Vega 2022), – a wide range of meanings is also attained as result of constructionalization of WAY/ROAD, including ‘towards’ (not necessarily spatial), cause and even such functions as nominalizers and nominal classifiers.

For instance, the progressive construction ‘to be on one’s way to’ is attested not only in English (Petré, Davidse & Van Rompaey 2012) or French (*en voie de*, see de Mulder 2019), but also in Korean, where a special mermaid construction is used:

[Na=nun	unhayng=ey	ka-nun]	kil=i-ta.
I=TOP	bank=DAT/LOC	go-ADN.NPST	road=COP-DECL
‘I am on my way to the bank.’ (lit. ‘[I go to the bank] the road is.’) (Kim 2020: 284)			

We embrace the notion of constructionalization (Traugott, Trousdale 2013) to describe the processes yielding such a wide range of resulting linguistic units as complex adpositions, particle verbs, aspectual multiword constructions, etc. Pragmaticalization (Diewald 2011) is important for the development of the discourse markers such as *anyway* and *by the way*, cf. also Polish *swoją drogą* ‘by the way’, lit. ‘by its own road’.

Constructions with ‘way/road’ have been previously studied mainly in the Germanic languages and within the framework of Construction Grammar (Goldberg 1996; Gallez 2021; Mortenmals & Smirnova 2020). We are going to show that more and better structured typological generalizations can be made using more diverse data, as well as to emphasize the vague borderline between the lexical and the grammatical exemplified in our data.

References

- Beck, David. (2004). *Upper Necaxa Totonac*. LINCOM: Europa.
- Diwald, G. (2011). Pragmaticalization (defined) as grammaticalization of discourse functions. *Linguistics*, Vol. 49 (2), pp. 365-390.
- François, Alexandre, Siva Kalyan, Mathieu Dehouck, Martial Pastor & David Kletz. (2025) *EvoSem: A database of dialexification across language families*. Online database. CNRS—LaTTiCe, Paris. <https://tiny.cc/EvoSem>
- Gallez, Françoise. (2021). Particle verbs with weg- in German: a constructional analysis. *Lexique* 28, 35-62.
- Goldberg, Adele E. (1996): Making one's way through the data. In: Shibatani, Masayoshi & Thompson, Sandra A. (eds.): *Grammatical constructions: Their form and their meaning*. Oxford: Clarendon Press. 29–53.
- Heine, Bernhard & Kuteva, Tania. (2002). *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- Jendraschek, G. (2020). Case marking and complex adpositions in Basque. In B. Fagard, J. Pinto de Lima, D. Stosic & E. Smirnova (Ed.), *Complex Adpositions in European Languages: A Micro-Typological Approach to Complex Nominal Relators*. Berlin, Boston: De Gruyter Mouton. Pp. 367-402.
- Kim, Joungmin. (2020). Korean. In Tasaku Tsunoda (ed.), *Mermaid construction: A compound-predicate construction with biclausal appearance*, 283–331. Berlin & Boston: De Gruyter Mouton.
- List, Johann Mattis & Tjuka, Annika & Blum, Frederic & Kučerová, Alžběta & Ugarte, Carlos Barrientos & Rzymiski, Christoph & Greenhill, Simon & Forkel, Robert (eds.) (2025). *CLLD Concepticon 3.3.0* [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.14622303>
- Mortelmans, Tanja and Smirnova, Elena. (2020). Analogues of the way-construction in German and Dutch: another Germanic sandwich?. In: Gunther Vogelaer, Dietha Koster and Torsten Leuschner (eds.), *German and Dutch in Contrast: Synchronic, Diachronic and Psycholinguistic Perspectives*. Berlin, Boston: De Gruyter. Pp. 47-76.
- De Mulder, Walter. (2019). En voie de: du trajet spatial à l'aspect. *Revue Romane. Langue et littérature. International Journal of Romance Languages and Literatures*, 54(1): 39-61.
- Perek Florent. (2018). Recent change in the productivity and schematicity of the way-construction: A distributional semantic analysis. *Corpus Linguistics and Linguistic Theory*, 14(1): 65-97.
- Petré, P., K. Davidse & T. Van Rompaey. (2012). On ways of being on the way. Lexical, complex preposition and aspect marker uses. *International Journal of Corpus Linguistics* 17(2): 229–258.
- Pozdnjakov K. I. (2020). Notes on regular polysemy and homonymy (Mande languages) *Language in Africa* 1(4): 69–84.
- Rzymiski, Christoph and Tresoldi, Tiago et al. (2019). *The Database of Cross-Linguistic Colexifications, reproducible analysis of cross- linguistic polysemies*. <https://clics.clld.org/>
- Traugott, Elizabeth Closs, and Graeme Trousdale. (2013). *Constructionalization and Constructional Changes*, Oxford Studies in Diachronic and Historical Linguistics. Oxford: Oxford University Press.

Diachronic typology of ‘yes’ and ‘no’: preliminary observations

Niyaz Kireyev

(École normale supérieure — PSL, Paris)

Keywords: diachronic typology, etymology, *yes-no* markers, answer particles, lexical databases.

In the domain of polar question answers, the most studied topics so far are:

- answering systems like bipartite *yes* vs. *no*, tripartite *oui* vs. *non* vs. *si*: starting from (Sadock & Zwicky, 1985), with both general (Moser, 2018) and areal studies (Da Milano, 2004);
- short and full answers (Enfield et al., 2019);
- syntactic (Holmberg, 2015) and semantic aspects (Asatiani, 1999; Farkas & Bruce, 2010) of answers.

As concerns the diachrony, however, we can mention only a few case studies on ‘yes’ and ‘no’, both of internal evolution (Mosegaard Hansen, 2020; Huber 2023) and of borrowing (Johanson, 2004), as well as several papers on the origin of polar question markers (Bencini, 2003; Lusini, 2013; Metslang et al., 2017). The only typological work dedicated specifically to the response particles I’m aware of is (Khudina 2021; on ‘yes’).

One can note an asymmetry in Indo-European answer particles: in the majority of languages, the standard words for ‘no’ seem to be etymologically related (*ne*, *nein*, *non* etc.), while ‘yes’ markers are rather heterogeneous (*jī*, *ja*, *da*, *tak* etc.). A lot of other language families exhibit similar asymmetry: cf. common Turkic *yok/yaq* ‘no’ vs. *evet*, *äye*, *bəli* and many other words for ‘yes’ found throughout Turkic languages. To my knowledge, this phenomenon was never explicitly researched.

I argue that this stability of ‘no’ compared to ‘yes’ is not a mere coincidence. This may be because ‘no’ tends to be related to the verbal negation (highly grammaticalised), while there are many possible sources for ‘yes’.

The preliminary check of this hypothesis is possible with the database CLICS³ (Rzymiski 2019). Amongst 81 links to the concept NO, 49 items (60 %) are colexifications with NOT. Amongst 52 links to YES, the most common colexification is with CORRECT, comprising only 18 instances (35 %).

It seems also that ‘yes’ is more likely to be borrowed. E.g., many unrelated languages influenced by Russian adapted *da* for ‘yes’, but maintained their own words for ‘no’ (Romanian, Mari, Komi-Zyryan etc.).

As to the research instruments, it has been noted that the grammars often don’t give the information on how to answer polar questions (Holmberg, 2015, p. 11). I find the verse of Matthew 5:37 useful for researching the *yes-no* typology because it makes it easy to find the relevant particles using multilingual Bible corpora, though translated sources should be used with caution.

This research aims to build a typological database reflecting the origin of the ‘yes’ and ‘no’ items in a convenience sample of languages. In the talk, I will give an overview of the most important sources,

including those that are not present in CLICS³ (e.g. ‘thus, like this’ → ‘yes’: Romance *si*, Polish *tak*, Lithuanian *taip* etc.), as well as general reflections on the diachronic asymmetry between ‘yes’ and ‘no’.

References

- Asatiani, R. (1999). The semantics and typology of yes/no particles: A cross linguistic study. *Open Society Institute*.
<https://www.researchgate.net/profile/Rusudan-Asatiani/publication/266214967>
- Bencini, G. (2003). Toward a Diachronic Typology of Yes/no Question Constructions with Particles. *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, 39(1), 604–621.
- Da Milano, F. (2004). Le domande sì/no nelle lingue del Mediterraneo. *Archivio Glottologico Italiano*, 1, 3–40.
- Enfield, N. J., Stivers, T., Brown, P., Englert, C., Harjunpää, K., Hayashi, M., Heinemann, T., Hoymann, G., Keisanen, T., Rauniomaa, M., Raymond, C. W., Rossano, F., Yoon, K.-E., Zwitserlood, I., & Levinson, S. C. (2019). Polar answers. *Journal of Linguistics*, 55(2), 277–304.
<https://doi.org/10.1017/S0022226718000336>
- Farkas, D. F., & Bruce, K. B. (2010). On Reacting to Assertions and Polar Questions. *Journal of Semantics*, 27(1), 81–118. <https://doi.org/10.1093/jos/ffp010>
- Holmberg, A. (2015). *The Syntax of Yes and No*. Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780198701859.001.0001>
- Huber, Judith (2023). Positive Response Particles in Early Modern Language Teaching Manuals: 1573–1625, in: Renate Bauer, Christine Elswiler, Ulrike Krischke & Kerstin Majewski (eds.), *Travelling texts—texts travelling. Gedenkschrift for Hans Sauer*, München: Utz, 183–204.
- Johanson, L. (2004). On the Turkic Origin of Hungarian igen ‘yes’. *Acta Orientalia Academiae Scientiarum Hungaricae*, 57(1), 93–104.
- Khudina 2021 — Худина, Я. Ю. (2021). Показатели согласия и подтверждения: типологическое исследование. Курсовая работа. Москва: Высшая школа экономики.
- Lusini, S. (2013). *Yes/no question-marking in Italian dialects: A typological, theoretical and experimental approach*. LOT.
<https://scholarlypublications.universiteitleiden.nl/access/item%3A2964540/download>
- Metslang, H., Habicht, K., & Pajusalu, K. (2017). Where do polar question markers come from? *STUF - Language Typology and Universals*, 70(3), 489–521. <https://doi.org/10.1515/stuf-2017-0022>
- Mosegaard Hansen, M.-B. (2020). The role of (historical) pragmatics in the use of response particles: The case of French. *Functions of Language*, 27(3), 307–339.
<https://doi.org/10.1075/fol.18024.mos>
- Moser, E. V. (2018). *Answers to Polarity Questions: A Typological Study* [Stockholm University].
<https://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-157363>
- Rzymiski, Christoph and Tresoldi, Tiago et al. 2019. The Database of Cross-Linguistic Colexifications, reproducible analysis of cross-linguistic polysemies. doi.org/10.1038/s41597-019-0341-x
- Sadock, J. M., & Zwicky, A. M. (1985). Speech Acts Distinctions in Syntax. In *Language Typology and Syntactic Description* (pp. 155–196). Cambridge University Press.

Zero vs. overt marking: The coding of the relation between two nouns from a cross-linguistic perspective

Noora Ahola
(University of Helsinki)

Keywords: zero marking, overt marking, noun juxtaposition, ambiguity, efficiency

This paper examines the interaction of zero and overt marking in coding the relation between two nouns. Sequences formed by two nouns (henceforth, N+N-sequences), when zero-marked, can express various functions across languages: adnominal possession, nominal predication, arguments of (di)transitive verbs, apposition, conjunction, and compounding. Each function can also be overtly marked. Examples (1) and (2) illustrate both marking strategies for possession and conjunction.

1. Possession

- a. Murui Huitoto (Wojtylak 2017: 240)

Lusio yoe-fai

Lucio metal-CLF:SHORT.THICKER

‘Lucio’s machete’

- b. Duhumbi (Bodt 2020: 281)

gepu-waʔ tɕʰiliŋ

king-GEN drum

‘the king’s drum’

2. Conjunction

- a. Baniva de Maroa (Aikhenvald 1998: 246)

tsiɽuɽa kamitsa

trousers shirt

‘trousers and shirt’

- b. Crow (Graczyk 2007: 191)

úuxa-lak iichilikaashi-lak

deer-and elk-and

‘deer and elk’

A genealogically and areally balanced pilot sample of 40 languages reveals that languages commonly express two or more of these functions by simply juxtaposing the nouns. The average number of functions per language is 3, ranging from 0 to all 6. Thus, languages with multiple zero-marked functions are potentially ambiguous, as the forms do not distinguish the function. Piantadosi et al. (2012) suggest that such ambiguity can be linked to the coding efficiency of language.

Examining the overtly marked sequences for each function sheds more light on the relation between zero-marked and overtly-marked sequences. The zero-marked sequence is not always the sole strategy to express a function, and many languages have an additional, overtly marked strategy. These two strategies may be conditioned (e.g., alienability distinction for adpossession constructions) or in free variation (e.g., optionality of overt coordinator). The results indicate that especially for possession and conjunction, only a fraction of languages express them solely as a zero-marked sequence: while the total number of zero-marked strategies for both functions are 13 and 26, respectively, the number of languages with only a zero-marked strategy is 1 and 5, respectively. However, overall, the distribution of solely zero-marked functions strengthens the fact that they are commonly attested both within and across languages. Out of the 40 languages, 22 languages have two or more functions as solely zero-marked. The average number of solely zero-marked functions per language is 2.

The zero-marked N+N-sequences are arguably efficient, as the lack of overt marking makes them easy to produce and process, potentially leading to ambiguity. The initial results suggest that ambiguity is not consistently avoided, as previous research has also indicated (e.g., Wasow 2015). However, the use of overtly marked sequences to express some of the functions to a greater extent suggests a preference for disambiguation by morphological means in certain contexts. By examining the distribution of zero-marked and overtly marked N+N-sequences the study enhances our understanding of language by shedding more light on questions of coding efficiency, processing, and transparency. As this study is part of an ongoing PhD project, the final results will be based on a 70-language sample, better reflecting the behavior of the N+N-sequences cross-linguistically.

References

- Aikhenvald, Alexandra Y. 1998. Warekena. In Desmond C. Derbyshire and Geoffrey K. Pullum (eds.), *Handbook of Amazonian Languages*, 225–440. Berlin & New York: Mouton de Gruyter.
- Bodt, Timotheus Adrianus. 2020. *Grammar of Duhumbi (Chugpa)*. Leiden: Brill.
- Graczyk, Randolph. 2007. *A grammar of Crow, Apsáalooke Aliláau*. Lincoln: University of Nebraska Press.
- Piantadosi, Steven T., Harry Tily & Edward Gibson. 2012. The communicative function of ambiguity in language. *Cognition* 122(3), 280–291.
- Wasow, Thomas. 2015. Ambiguity avoidance is overrated. In Susanne Winkler (ed.), *Ambiguity: Language and Communication*, 29–48. Berlin, München & Boston: Mouton de Gruyter.
- Wojtylak, Katarzyna I. 2017. *A Grammar of Murui (Bue): A Witotoan language of Northwest Amazonia*. Cairns: James Cook University dissertation.

Ideophones as a Strategy for Vividness
A Study of two Creole Adaptations of “Les Fables de la Fontaine”

Patrice Kanndèl Edouard and Jean Odelin Petit-Frère

Ideophones have attracted significant scholarly attention in recent years (Akita & Pardeshi, 2019; Dingemanse, 2019; Hinton, Nichols, & Ohala, 1994; Lahaussois, Marsault, & Treis, 2024; Voeltz & Kilian-Hatz, 2001). However, studies reveal that Western languages like English, Spanish, Dutch, and French exhibit relatively limited ideophone usage, often restricted to phenomena like phonaesthemes or onomatopoeia (Bakker, 2008). By contrast, ideophones are essential in African and Asian languages, vividly conveying sensory experiences and emotions (Dingemanse, 2019). Caribbean Creole languages—including Jamaican Patois, Martinican Creole, Guadeloupean Creole, and Haitian Creole—are similarly recognized for their ideophonic richness (Rougé & Véronique, 2024; Parkvall, 2000). Despite this, little research explores how these languages utilize ideophones to enhance expressivity or how they diverge from their European lexifiers in this regard (Bartens, 1999, 2013; Edouard, 2024a-b; Prou, 1999).

This study examines two early Creole adaptations of *Les Fables de La Fontaine*—*Les Bambous* (Marbot, 1869) in Martinican Creole and *Cric Crac* (Sylvain, 1901) in Haitian Creole—through a typological and comparative lens. Drawing on frameworks by Akita & Pardeshi (2019), Dingemanse (2019), and Lahaussois et al. (2024), it identifies ideophones within these texts and categorizes them by their expressive and sensory functions, shedding light on the African linguistic influences in Caribbean Creoles.

For example, in the text *Le pot de terre et le pot de fer* (Sylvain, 1901), ideophonic expressions can be found, as seen in examples 1 and 2. Similarly, Marbot’s fable *Le loup et l’agneau* (Marbot, 1869) introduces the ideophone *fioupe* to represent a swallowing sound, as shown in example 3.

1. Running

Haitian Creole (Sylvain, 1901):

Couri : tip ! tip ! bô baryè.

run-3SG IDEO IDEO LOC barrier

He runs: tip! tip! toward the barrier.

2. Collision Sounds

Haitian Creole (Sylvain, 1901):

Li : zouîng ! li : pim ! li gnan !

3SG IDEO 3SG IDEO 3SG IDEO

It goes: zouîng! Pim! Gnan!

3. Swallowing Sounds

Martinican Creole (Marbot, 18.):

valé : fioupe ! com yon gombo.

swallow IDEO like 3SG callaloo

It swallows fioupe like a callaloo.

The central research questions are:

1. How do ideophones enhance the vividness and sensory engagement of narratives in Martinican and Haitian Creole adaptations of La Fontaine's fables?
2. In what ways do these ideophonic expressions diverge from their European lexifier, French?
3. What cultural and linguistic markers are revealed through their use?

Methodologically, the study analyzes ideophones in primary texts, examining their frequency, contextual usage, and narrative impact. This is supplemented by secondary literature on Caribbean Creole languages and studies on ideophones.

Results indicate that Haitian and Martinican Creoles exhibit a higher density and variety of ideophones than French, showcasing their superior capacity for sensory engagement. These findings illuminate the role of ideophones in narrative construction, emphasizing their cultural and linguistic significance.

This research deepens our understanding of Creole adaptations of classic literature, highlights the expressive potential of Caribbean Creoles, and underscores the enduring influence of African linguistic heritage on these languages.

References

- Akita, K., & Pardeshi, P. (Eds.). (2019). Ideophones, mimetics and expressives. *Iconicity in Language and Literature*, Vol. 16. Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/ill.16>
- Bartens, A. (1999). Ideophones and sound symbolism in Atlantic Creoles. *Suomalaisen Tiedeakatemia Toimituksia. Sarja Humaniora*, 304.
- Bartens, A. (2013). Creole languages. In Bakker, P. & Matras, Y. (Eds.), *Contact languages: A comprehensive guide* (pp. 65–158). Berlin: De Gruyter Mouton.
- Bakker, P. (2008). Pidgins versus creoles and pidgincreoles. In Kouwenberg, S., & Singler, J. (Eds.), *The handbook of pidgin and creole studies* (pp. 130–157). Oxford: Wiley-Blackwell.
- Dingemanse, M. (2019). Ideophone as a comparative concept. In Akita, K., & Pardeshi, P. (Eds.), *Ideophones, mimetics, and expressives* (pp. 13–33). Amsterdam: John Benjamins.
- Dingemanse, M. (2023). Ideophones. In Van Lier, E. (Ed.), *The Oxford handbook of word classes* (pp. 466–476). Oxford: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198852889.013.15>
- Edouard, P. K. (2024a). *Idéophonie et catégorisation en créole haïtien : une analyse linguistique des idéophones* (Bachelor's thesis). Université d'État d'Haïti, Faculté de Linguistique Appliquée.
- Edouard, P. K. (2024b). *Haitian Creole Ideophones: Typology of a word class*. Manuscript in preparation.

- Lahaussais, A., Marsault, J., & Treis, Y. (2024). Ideophones: honing in on a descriptive and typological concept. *Linguistic Typology at the Crossroads*, 4(1).
- Marbot, L. (1869). *Les Bambous: Fables de La Fontaine travesties en patois créole par un vieux commandeur*. Paris: Lachaud.
- Parkvall, M. (2000). *Out of Africa: African influences in Atlantic Creoles*. London: Battlebridge.
- Prou, M. E. (1999). Haitian Creole ideophones: An exploratory analysis. *Journal of Haitian Studies*, 5/6, 96–112.
- Rougé, J.-L., & Véronique, G. D. (2024). Word classes in Romance-related Creoles. In Ledgeway, A., & Maiden, M. (Eds.), *The Oxford guide to Romance languages* (pp. 689–706). Oxford: Oxford University Press.
- Sylvain, G. (1901). *Cric Crac: Fables de La Fontaine en créole haïtien*. Port-au-Prince: Imprimerie de l'État.
- Voeltz, E. F. K., & Kilian-Hatz, C. (Eds.). (2001). *Ideophones*. Amsterdam: John Benjamins.

Human nouns dispersal: Gender assignment of human nouns in Bantu languages

Rita Popova
(Saarland University)

Keywords: Bantu languages, gender assignment, typology of gender systems

In this talk, I refine some widespread assumptions on gender assignment of human nouns in Bantu languages (a sub-branch of the Atlantic-Congo family comprising around 550 varieties).

Crosslinguistically, gender systems demonstrate different degrees of semantic motivation behind gender assignment of nouns. In some languages, assignment rules are more semantically transparent while in others, assignment principles are largely opaque, or are mostly defined by formal assignment rules (Corbett 1991). Importantly, all gender systems are believed to have a core semantic rule of gender assignment (Corbett 1991, Dahl 2000). Even in the languages where assignment is predominantly opaque or formal, one finds a straightforward semantic principle which governs gender assignment of human (and, sometimes, higher animals) nouns. In sex-based systems, this core rule associates males with the masculine and females with the feminine gender. In animacy (or humanness) based systems, most animals (or humans) belong to a single animate (or human) gender. One potential counterexample to the semantic core universal is Uduk (Koman), as analyzed by Killian (2019).

In Bantu gender systems, the primary semantic contrast is between humans and non-humans. Typically, most nouns referring to humans are assigned to a single 'human' gender value (traditionally labelled as *gender 1/2* in Bantuist notation). In contrast, non-human nouns are distributed across several other gender values, often according to principles that are highly opaque (Katamba 2003). Occasionally, nouns denoting humans with unusual characteristics (e.g. *blind*) are found in gender values other than 1/2 (Van de Velde 2019). These generalizations are derived from observations on well-studied Bantu languages such as Swahili and are rarely based on the systematic examinations of lexicons.

I aim to expand the empirical basis of this line of study by investigating Bantu lexicons available in the RefLex database (Seegerer & Flavie, 2011-2023). My study of 45 Bantu varieties demonstrates that gender assignment of human nouns is a parameter of intra-Bantu variation. While some Bantu languages assign most human nouns to gender 1/2, some others have more than 50% of the human-denoting lexical items assigned elsewhere.

Assignment of human nouns outside gender 1/2 is characteristic for languages spoken in the northwestern region, a zone traditionally recognized as the most diverse within the otherwise relatively homogenous Bantu-speaking world (Nurse & Philippson 2003, p. 165). In my talk, I will demonstrate that in the most divergent cases, the dispersal of human nouns can be partially explained by minor assignment rules that have received little attention in Bantu studies (e.g. assignment of deverbal agent derivations to the 'non-human' gender 3/4). Finally, the Bantu data will be interpreted within a broader typological framework of gender assignment.

References:

Corbett, Greville. 1991. *Gender*. Cambridge: Cambridge University Press.

Dahl, Östen. 2000. Animacy and the notion of semantic gender. In Barbara Unterbeck (ed.), *Gender in grammar and cognition. Vol. 1: Animacy and the notion of semantic gender: Approaches to gender*, 99–115. Berlin: Mouton de Gruyter.

Katamba, Francis. 2003. Bantu nominal morphology. In Derek Nurse & Gerard Philippson (eds.), *The Bantu languages*, 103–120. London: Routledge.

Killian, Don. 2019. Gender in Uduk. In Di Garbo, Francesca, Bruno Olsson & Bernhard Wälchli (eds.), *Grammatical gender and linguistic complexity: Volume I: General issues and specific studies*, 147–168. Berlin: Language Science Press.

Nurse, Derek & Gerard Philippson. 2003. Towards a historical classification of the Bantu languages. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 164–182. London: Routledge.

Segerer, Guillaume & Sébastien Flavien. 2011-2023. *RefLex: Reference Lexicon*, Version 2.2. Paris, Lyon. <https://reflex.cnrs.fr/>

Van de Velde, Mark. 2019. Nominal morphology and syntax. In Van de Velde, Mark, Koen Bostoen, Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 2nd ed., 237–269. London & New York: Routledge.

Reduplication in the Pacific Northwest: A typological and areal account

Simone Mattioli
(University of Pavia)

Keywords: Pacific Northwest, Reduplication, Typology, Areal linguistics.

The Pacific Northwest (PNW) is a very diverse linguistic area with approximately 20 language families plus several isolates. This has favored intense contact which led to one of the most famous *Sprachbünde* (see Thomason 2000:311). Among the linguistic properties of the Northwest Coast Sprachbund (NWCS), reduplication (“the repetition of part or all of one linguistic constituent to form a new constituent with a different function” Inkelas 2014:169) is frequently cited both in general-typological literature (e.g., Rubino 2005: 22) and studies specifically dedicated to the NWCS (Thompson & Kinkade 1990; Campbell 1997; Mithun 1999; Beck 2000). See (1)-(3).

(1) Clallam (Salishan; Montler 2015: 179, 180, 284):

- a. Collective plural: *ləmətú* ‘sheep’ > *ləmləmətú* ‘bunch of sheep’;
- b. Additive plural: *ŋáqsən* ‘nose’ > *ŋiŋáqsən* ‘noses’;
- c. Diminutive: *sqáxaʔ* ‘dog’ > *sqəʔqáʔxaʔ* ‘little dog, puppy’, *štán* ‘walk’ > *šaʔštán* ‘walk a little’.

(2) Nuu-chah-nulth (Wakashan; Davidson 2002: 31, 59, 62, 240)

- a. Additive plural: *ta:yi* ‘older brother’ > *ta:ta:yi* ‘older brothers’, *nu:k* ‘song’ > *nu:knu:k* ‘songs’;
- b. Distributive plural:

<i>ʔa-ʔiʔ</i>	>	<i>ʔa-ʔa-ʔiʔ</i>
stick_like.object.stands-in.house		PL-stick_like.object.stands-in.house
‘(stick-like object) standing in the house’		‘(stick-like objects) standing here and there in the house’

- c. Iterative: *mitxw* ‘turn’ > *mitxmitxš* ‘turn at intervals’

(3) Gitksan (Tsimshian; Rigsby 1986: 54, 97)

- a. Additive plural: *bil’ust* ‘star’ > *bixbil’ust* ‘stars’, *do’o* ‘cheek’ > *dido’o* ‘cheeks’;
- b. Durative:

<i>q’ay</i>	<i>pi-pax=ʔ</i>	<i>qot-t</i>
still	DUR-run=3SG	heart-3SG
‘His heart is still beating .’		

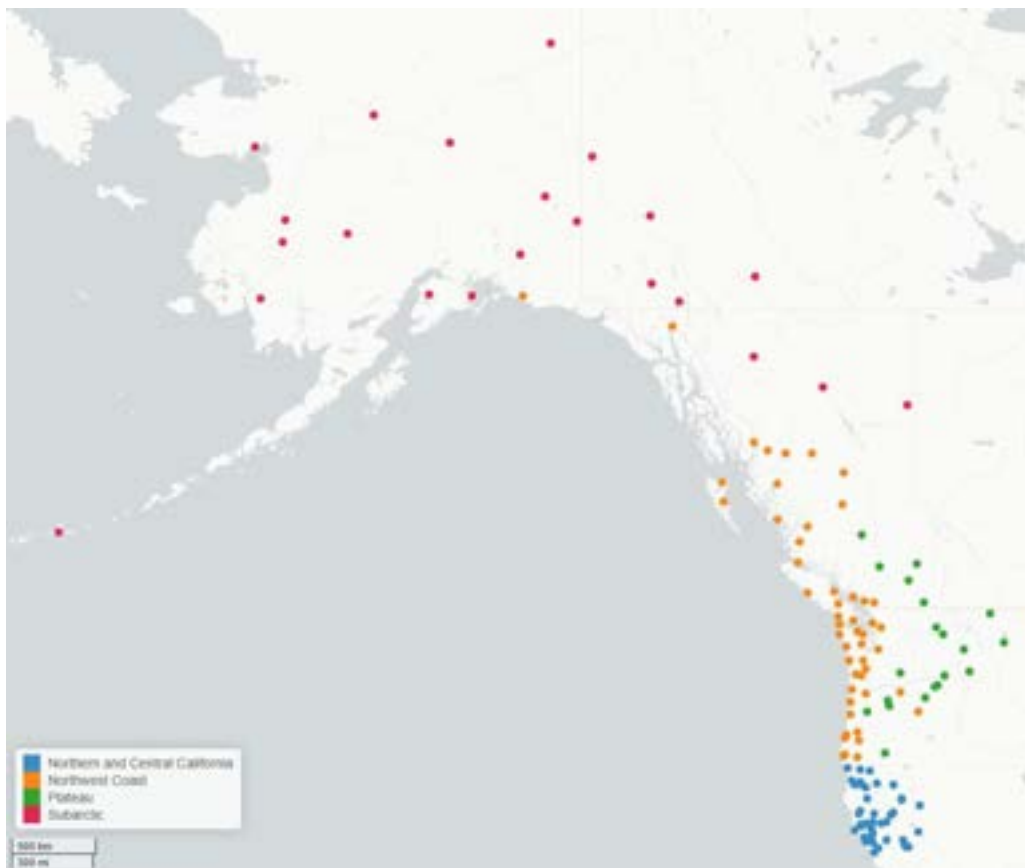
Despite this, a comprehensive cross-linguistic and areal description of reduplicative patterns in the NWCS is still lacking, thus leaving it a *desideratum*. The present contribution aims at describing the formal and functional properties of reduplicative patterns in the NWCS paying particular attention to their areal distribution. To do so, a sample of 50 languages belonging to the NWCS has been designed. Then, an example taken from different sources (grammatical descriptions, dedicated papers, and texts) for each reduplicative construction (pairings of form and function) identified has been extracted and tagged according to the following parameters:

- Formal:
 - type of reduplication (partial, full, etc.);

- position of the reduplicant (initial, internal, final);
- formal nature of the reduplicant (syllable, morpheme, etc.);
- formal schema of the reduplicant (phonetic/prosodic structure)
- lexical category of the base/output (noun, verb, ideophone, etc.);
- Functional:
 - type of function (lexical, grammatical, textual/pragmatic);
 - specific function (plural, pluractional, distributive, etc.).

In addition, to test the areal distributions of constructions, the study has been extended to several neighboring languages (Northern and Central California: 41 languages; Plateau: 20 languages; Subarctic: 22 languages; see Map 1). Each construction has been tagged according to the geographic location of the relative language. The overall sample counts 133 languages.

The present contribution represents the first typological and areal description of reduplicative patterns in the NWCS and, more in general, in the whole PNW (in its widest sense) representing an important contribution to the discussion on NWCS, but also for areal linguistics and for typology of reduplication.



Map 1. The languages of the PNW classified according to their areal/geographic classification.

References

- Beck, David (2000), Grammatical Convergence and the Genesis of Diversity in the Northwest Coast Sprachbund, *Anthropological Linguistics* 42(2), 147–213.
- Campbell, Lyle (1997), *American Indian Languages: The Historical Linguistics of Native America*, Oxford/New York: Oxford University Press.
- Davidson, Matthew (2002), *Studies in Wakashan (Nootkan) Grammar*, Buffalo, NY: State University of New York at Buffalo dissertation.

- Inkelas, Sharon (2014), Non-concatenative derivation: Reduplication, in R. Lieber & P. Štekauer (eds), *The Oxford Handbook of derivational morphology*, Oxford: Oxford University Press, 169–189.
- Kinkade, M. Dale, William W. Elmendorf, Bruce Rigsby & Haruo Aoki (1998), Languages, in D. E. Walker JR. (ed), *Handbook of North American Indians, Vol. 12, Plateau*, Washington, D.C.: Smithsonian Institution, 47–72.
- Mithun, Marianne (1999), *The Languages of Native North America*, Cambridge: Cambridge University Press.
- Montler, Timothy (2015), *Klallam Grammar*, Seattle: University of Washington Press.
- Rigsby, Bruce (1986), *Gitksan Grammar*, Australia: University of Queensland.
- Rubino, Carl (2005), Reduplication: Form, function and distribution, in B. Hurch (ed), *Studies on reduplication*, Berlin: Mouton de Gruyter, 11–29.
- Thomason, Sarah G. (2000), Linguistic areas and language history, in D. Gilbers, J. Nerbonne & J. Schaeken (eds), *Languages in Contact*, Amsterdam: Rodopi, 311–327.
- Thompson, Laurence C. & M. Dale Kinkade (1990), Languages, in Wayne Suttles (ed), *Handbook of North American Indians, Vol. 7, Northwest Coast*, Washington, D.C.: Smithsonian Institution, 30–51.

Gender assignment in Uduk

Stefan Savić & Don Killian
(University of Zurich & University of Helsinki)

Keywords: <gender, Uduk, number, mass noun>

Uduk is a Koman language with two noun classes. Its gender assignment is unique in that it seems to violate Dahl (2000)'s presupposition that semantic principles lie at the foundation of gender assignment in nouns cross-linguistically (Killian 2015, 2019). Even kinship terms are distributed across both classes, and do not exhibit sex, age or relation-based tendencies, cf. (1-4). On the other hand, Killian (2019) notes that gender assignment correlates with borrowed lexicon, certain parts of speech, temporal-spatial semantic concepts, nominalized verbs and agentive nouns. For example, personal pronouns and demonstratives belong to class 1, whereas proper nouns denoting human names are all in class 2, cf. (5-7). It is also noteworthy that although Uduk does not regularly exhibit grammatical expression of number in nouns, plural nouns as well as demonstratives and pronouns largely cluster in class 1, cf. (4, 8). A closer look into the distribution of animacy, quantity-specifications, and identifiability across the two noun classes may identify semantic patterns underlying Uduk noun class organization.

- (1) *'bwā?*
'daughter', class 1
- (2) *à='bóm*
'woman, wife' class 2
- (3) *yà?*
'son', class 1
- (4) *à=kām*
'brother', class 2
- (5) *à=Răbì*
Rabi (personal name), class 2
- (6) *á'dī*
3SG personal pronoun, class 1
- (7) *mún*
locative anaphoric pronoun, class 1
- (8) *ī=kām*
'brothers' (plural), class 1

In the present study we investigate various semantic concepts within a list of approximately 1400 Uduk nouns partly based on Beam and Cridland (1970). Human referents, animals and inanimate objects are largely evenly distributed across both noun classes, confirming Killian (2019). However, some other semantic concepts do seem to cluster together. Class nouns 1 include nouns that denote multiple similar referents, such as groups of people including ethnic identities, bundles of specific crops, forests or fields of a certain plant, cf. (9-12), suggesting all quantification specifications belong to this class. By contrast, class 2 includes most mass nouns (excluding liquids), cf. (13-15).

- (9) *má'dīny*
 'flock', class 1
- (10) *Gàlé?*
 'Ethiopian, Galla', class 1
- (11) *kōth*
 'bundle (of pumpkins, sesame, etc.)', class 1
- (12) *bùntō*
 'field (of sorghum, sesame, etc.)', class 1
- (13) *à='cesh*
 'soil', class 2
- (14) *à='dōṅkōrō?*
 'salt', class 2
- (15) *à=pí?*
 'soot, black ash', class 2

The study suggests that contrary to the crosslinguistic tendencies, the gender assignment in Uduk is not based on animacy, biological gender, or physical properties. Instead, it incorporates quantification distinction (count plural nouns vs. mass) as well as uniqueness and discourse-based referent identifiability strategies.

References

- Beam, Mary S. and Cridland, Elizabeth A. (1970), *Uduk-English Dictionary*, Khartoum: Sudan Research Unit, University of Khartoum.
- Killian, Don (2019), Gender in Uduk. In: F. Di Garbo, B. Olsson, and B. Wälchli (eds), *Grammatical gender and linguistic complexity, vol.1: General issues and specific studies*. Berlin: Language Science Press.
- Killian, Don (2015), Topics in Uduk Phonology and Morphosyntax. PhD thesis: University of Helsinki.

A typological study of applicative uses of AM markers

Timofey Mukhin
(University of Liège)

Keywords: applicatives, associated motion, typology, valency, grammaticalization

This paper investigates the valency-increasing potential of Associated Motion (AM) markers from a cross-linguistic perspective. Applicative use of AM markers was first observed by Pakendorf & Stojnova (2021) in Tungusic languages. Our previous findings show that this functional extension of AM markers is attested in languages across different linguistic areas (Mukhin et al. 2024). However, the main types of variation remain unknown. This paper investigates this issue in a convenience sample of 30 languages with markers showing both applicative and AM uses.

First, we investigate the **polyfunctionality** of a marker: whether it serves both functions in the same context (1), or has either applicative or AM use depending on the verb, as in (2a-b) respectively.

- (1) Murui (Witotoan; Colombia, Peru; Wojtylak 2020: 344)
[Alexis jo-fo-mona] Fransiska=di-no-moloc gui-**zaibi**-t-epred
[Alexis house-CLF-ABL] Francisca=at-CLF-LOC eat-**VENT**-LK-3
'From the house of Alexis (she) came to eat at Francisca's.'
- (2) Nivacle (Matacoan; Argentina, Bolivia, Paraguay; Fabre 2013: 11–13)
a. a-pitej-[yi]-**c'oya**
2s-be.tall-[1]-**AM.ANT.VENT**
'You are taller than me.'
b. j-ovalh-**c'oya**
1A(>3P)-watch-**AM.ANT.VENT**
'I watch(ed), waiting for him/her/they to come.'

Second, we consider parameters relevant to applicatives. In terms of the **semantic role of the applied phrase**, AM markers are found to introduce both spatial and non-spatial participants. In (1), the AM suffix *-zaibi* adds the spatial participant 'house' as a source to the non-motion verb 'eat'. By contrast, in (2a), the AM suffix *-c'oya* introduces the non-spatial participant 'me' taking up the role of standard of comparison – without adding AM. As for the **syntactic status of the applied phrase**, AM markers are not only found in prototypical "*P*-applicative constructions", where the applied phrase gets core argument coding, as in (3b) or (2a) above, but also in so-called "*X*-applicatives", where it is syntactically treated as a non-core argument, as in (1) (cf. Zúñiga & Creissels, 2024: 19). The third parameter is the **status of a non-applicative equivalent**, viz. whether applicatives are *obligatory*, lacking an equivalent with the same participant expressed by other means (1), (2a), or *optional*, cf. (3a-b).

- (3) Paunaka (Arawakan; Bolivia; Terhart 2024: 394–395)
a. pero pi-yunu pi-sane-yae
but 2SG-go 2SG-field-LOC
'But did you go to your field?'
b. kuina Jose ti-yunu-**pu** [uneku]
NEG José 3-go-**DLOC** [town]
'José isn't here, he went to town.'

Third, we investigate whether the availability of an applicative function for AM markers correlates with any parameters of AM, such as the **argument role of the moving figure**, the **temporal relation between the motion and verb event**, and the **direction of the motion** (Guillaume & Koch, 2021: 9). For example, in (4), the marker *-pa* (irrealis variant of *-pu* (3b)) denotes motion of the subject and involves prior motion and itive direction, while in (2b), the AM marker indicates motion of the object and involves subsequent motion and ventive direction. In our data, the applicative extension most commonly appears with prior-motion subject-related ventive (1) or itive (4) AM.

- (4) Paunaka (Terhart 2024: 399)
kuina puero-ina-bu ti-isu-pa-bu
 NEG can-IRR.NV-DSC 3-weed-DLOC.IRR-DSC
 ‘He cannot go to weed anymore.’

Harmonized abbreviations

A	agent of bivalent verb	LK	linker
ABL	ablative	LOC	locative
AM.ANT.VENT	anticipated ventive	NEG	negation
CLF	classifier	P	patient of bivalent verb
DLOC	dislocative	S	subject of monovalent verb
DSC	discontinuous	SG	singular
IRR	irrealis	VENT	ventive
IRR.NV	non-verbal irrealis	1, 3	first, third person

Acknowledgments

Work on this article was funded by grant ARC 23/27-14 – SPACEGRAM of the Research Council of the University of Liège.

References

- Fabre, Alain. 2013. *Applicatives and associated motion suffixes in the expression of spatial relations: A view from Nivacle (Mataguayo family, Paraguayan Chaco)*. Unpublished manuscript. <https://www.academia.edu/12443688/> (accessed 5 May 2025).
- Guillaume, Antoine & Harold Koch (eds.) (2021), *Associated Motion*, Berlin: Mouton de Gruyter.
- Mukhin, Timofey, An Van linden & Dana Louagie. 2024. *A typological study of applicative uses of spatial markers*. Presented at the 57th Annual Meeting of the Societas Linguistica Europaea (SLE), August 21–24.
- Pakendorf, Brigitte & Natalia Stoyanova (2021), *Associated motion in Tungusic languages: a case of mixed argument structure*, in A. Guillaume & H. Koch (eds.), (2021), *Associated Motion*, 855–898, Berlin: Mouton de Gruyter.
- Terhart, Lena (2024), *A grammar of Paunaka*, Berlin: Language Science Press.
- Zúñiga, Fernando & Denis Creissels (2024), *Applicative Constructions in the World’s Languages*, Berlin, Boston: Mouton de Gruyter.
- Wojtylak, Katarzyna Izabela (2020), *A Grammar of Murui (Bue): A Witotoan language of Northwest Amazonia*. Leiden: Brill.