# Complex predicates with multiple auxiliary verbs in Kazakh

Birsel Karakoç & Uldanay Jumabay
(Uppsala University & Goethe University Frankfurt)

Keywords: complex predicates, multiverbal constructions, Kazakh, grammaticalization, actionality, viewpoint aspect, semantics

This paper investigates complex predicates with multiple auxiliary verbs in Kazakh. Similar to some other Turkic languages, Kazakh has a highly elaborate verb system with different types of multiverbal constructions which are based on sequences of verbs forming a single predicate with one argument (see Johanson 2000, 2004, 2021, Karakoç 2005, 2019, Csató et. al. 2019a, 2019b, Jumabay et. al. 2019, Karakoç & Jumabay 2023). One type of such constructions, called postverbial constructions in Turcological literature, consists of a lexical verb ($V^1$) with a converb marker in {-(ǀ)p} or {-A//-y} and a subsequent auxiliary verb ($V^2$), which takes on a predicate suffix. As grammatical operators, postverbial constructions carry a high functional load marking

- actionality, in the sense that the actional content of ($V^1$) is recategorized, e.g. *Oḳï-p šïḳ-tï* ⟨read-B.CONV GO OUT.POSTV-TERM3⟩ 'X read completely' ← *oḳï-* 'to read'.
- viewpoint aspect, e.g. *Oḳï-p žat-ïr-* ⟨read-B.CONV LIE.POSTV-AOR⟩ 'X is reading' ← *oḳï-* 'to read'.
- potentiality, e.g. *Oḳï-p al-dï* ⟨read-B.CONV TAKE.POSTV-TERM3⟩ 'X was able to read' ← *al-* 'to take'.

These constructions are usually ambiguous between monopredicative and pluripredicative readings, distinguishable with distinct prosodic patterns in spoken language (e.g. Karakoç 2005, 2019, Jumabay et. al. 2019). Furthermore, a pluripredicative construction may allow different readings indicating a modifying adverbial clause or a clause chain.

Based on the spoken and written data collected from different Kazakh varieties (see Jumabay 2022), this study focuses on a more complex type of postverbial constructions consisting of more than one auxiliary verb. In most cases, the lexical ($V^1$) is followed by two auxiliary verbs ($V^2 + V^3$), see (1). In rare cases, three consecutive auxiliary verbs ($V^2 + V^3 + V^4$) occur after ($V^1$), see (2).

(1) *Bay bol-ïp šïɣ-a kel-dï* ⟨rich become-B.CONV GO.OUT.AUX-A.CONV COME.AUX-TERM3⟩ 'X suddenly became rich'.

(2) *Ḳupïya-sï-n iš-ïn-e saḳta-p žür-e al-may kel-e-di* ⟨secret-POSS3-ACC inside-POSS3-DAT keep-B.CONV MOVE.AUX-A.CONV TAKE.AUX-NEG.CONV COME AUX-INTRA-3SG⟩ 'X has not been able to keep her secret.'

The requirements and constraints of complex predicates with multiple auxiliary verbs have not yet been investigated, and still await careful scrutiny. Based on a functional approach, and relating to the issues put forward by Kraußе & Aplonova & Czerwinski (2024), we specifically seek answers to the following questions: (1) What type of lexical and auxiliary verbs typically appear in the complex predicates with multiple auxiliaries? (2) What do the limitations on the positioning of the consecutive verb segments depend on? (3) What are the limitations on the choice of converb markers {-(I)p} and {-A//-y}? (4) What are the specific functional contributions of individual segments to the meaning of the complex predicate? (5) What is crucial to consider for the relations and combinational constraints of actional, viewpoint aspectual, and modal notions in these constructions?

An obvious result pertains to the positioning of the actional segments before the segments marking viewpoint aspect. Examples with multiple auxiliary segments expressing different actional contents seem to be particularly challenging. Thus, the paper analyses the relations of various actional segments, and their possible morphosyntactic and semantic constraints.

**References**


Csató, Éva Á., Lars Johanson and Birsel Karakoç (eds.) (2019a), *Ambiguous verb sequences in Transeurasian languages and beyond*, Wiesbaden: Harrassowitz.

Csató, Éva Á., Lars Johanson and Birsel Karakoç (2019b), Introduction, in: Csató, Johanson and Karakoç (2019a): 1–12.

Jumabay, Uldanay (2022), *Clause combining strategies in Kazakh as spoken in China in comparative perspectives*. Unpublished PhD thesis, Johann Wolfgang Goethe-Universität.

Jumabay, Uldanay, Irina Nevskaya, Nida Ocak and Saule Tazhibayeva (2019), *Ambiguous verb sequences in Kazakh in a comparative perspective*, in: Csató, Johanson and Karakoç (2019a): 145–174.

Johanson, Lars (2000), Viewpoint operators in European languages, in Östen Dahl (ed.) (2000), *Tense and aspect in the languages of Europe*, Berlin / New York: De Gruyter Mouton, 27–187.

Johanson, Lars (2004), On Turkic transformativizers and nontransformativizers, *Turkic Languages* 8, 180–190.

Johanson, Lars (2021), *Turkic*. Cambridge: Cambridge University Press

Karakoç, Birsel (2005), *Das finite Verbalsystem im Nogaischen*. Wiesbaden: Harrassowitz.

Karakoç, Birsel (2019), Levels of potential ambiguity in Noghay verb sequences, in: Csató, Johanson and Karakoç (2019a): 175–198.

Karakoç, Birsel & Uldanay Jumabay (2023), Semantic functions of the transformativizing postverbial constructions in Kazakh, *Turkic languages* 27, 160–184.

Kraußе, Daniel & Aplonova, Katya & Czerwinski, Patryk (2024), Exploring the limits of complex predicates. Workshop proposal, 57th Annual Meeting of the Societas Linguistica Europaea (SLE), University of Helsinki.

# Complex predicate as analytical voice marking of antipassivization

Katarzyna Janic, Krzysztof Stroński & Mohammad Tavakoli
(Adam Mickiewicz University, Adam Mickiewicz University & Adam Mickiewicz University)

Keywords: antipassive, analytical voice marking, complex predicates, analytical antipassive marking, multivariate typology

Aim: Although recent studies extended the voice definition from an inflectional category to synthetic (affixation) and analytic marking (analytic verb forms, i.e., complex predicates CPs) (Zúñiga & Kittilä 2019; Creissels 2024), a noticeable disparity remains in how analytical voice marking is treated in voice domain. While CPs received much attention in valency-increasing alternations like causatives (Creissels 2024), there is little research on their detransitivization function. And when compared to passives, antipassives remain under-investigated (Zúñiga & Kittilä 2019: 114). This study aims to bridge this gap in the existing literature, seeking to demonstrate that analytical antipassive marking is more prevalent than previously assumed.

1. Mopan (Mayan; Hofling 2011, cited in Zúñiga & Kittilä 2019: 113-114)

a. *Walak-u-loch-ik-ech*.
INCOMPL-3.I-bend-INCOMPL.TR-2SG.II
'S/he bends you (sg).'

b. *Uch-i* *u-loch*.
**happen**-3SG.I[COMPL.ITR] 3.I-bend
'S/he bent (something).'

Approach, data, method, and research question: We conducted our research in a functional-typological approach based on a sample of 56 unrelated languages from six macroareas. We collected data through grammar mining and analyzed it in multivariate typology (Bickel 2011). Each antipassive construction was thus decomposed into formal variables, with attention to voice marking and its synthetic vs. analytical opposition. The investigation was guided by two research questions: What is the grammatical nature and evolution of the analytic voice form?

Results: Our analysis showed that CPs cover light verbs in the auxiliary function. In Mopan, Hindi, and Worrorra, light verbs serve a functional purpose, coding tense, aspect and/or person. In addition to the aspectual input (aktionsart), some (Mopan 'happen' [1], Hindi 'fall' [2], Worrorra 'be' [3]) act as analytic voice markers, (semi-)detransitivizing transitive construction. They yield antipassives and antipassive-lookalikes, where the ergative (A) argument becomes absolutive (Hindi), or the P argument loses a verb index (Mopan, Worrorra).

2. Hindi (Indo-European; Butt 1995: 107)

a. *us-ne* *gānā-∅* *gāyā*
s/he-ERG song.M.SG sing.PST.M.SG.
'S/he sang a song.'

b. *vo* *gānā-∅* *gā* *paṛ-ā*
he.NOM song.M.NOM sing **fall**-PST.M.SG
'He fell to singing (burst out into song).'

3. Worrorra (Worrorran; Clendon 2014: 331)

a. *minjarl* *nyiN-Ø-mnya=bwu-na*
eat 3F-3-DD=hit-PST
'he ate her.'

b. *mangarri* *minjarl-minjarl* *kaarr=nu-na*
veg.food eat-eat 3PL=**be**-PST
'They used to eat vegetable food.'

Certain light verbs deriving antipassives manifest a well-established evolution. In French, *faire* 'do/make', which is not distinct from tense auxiliaries (Abeillé et al. 1998), can form causative (*faire acheter les chaussures* 'make [someone] buy the shoes') and antipassive periphrases (*faire des achats* 'do [some] shopping'), the latter taking a deverbal event noun in an object role (Sansò 2017). In Makalero, the light verb *kini* 'do' grammaticalized into a synthetic voice marker *-ini* when deriving antipassives. Finally, in Soninke, the antipassive *-ndì* and causative *-ndí* result from the grammaticalization of **tin* 'do' in causative and antipassive periphrases (Creissels 2012). The same holds for *-aɢan* in Mocoví (Juárez & González 2021).

By analogy with synthetic/analytic causative marking and its correlation with the direct vs. indirect causation meaning (Haspelmath 2018), we will check whether similar correlations also exist between the formal properties of antipassive marking and specific aspects of meaning.

Acknowledgments
This research is part of the project No. 2021/43/P/HS2/01395 cofunded 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, headed by Katarzyna Janic as Principle Investigator.

Abbreviations

| | | | |
|---|---|---|---|
| COMPL.ITR | completive intransitive | M | masculine |
| DD | discourse diectic | NOM | nominative |
| ERG | ergative | PL | plural |
| F | feminin | PST | past |
| INCOMPL | incompletive | SG | singular |
| INCOMPL.TR | incompletive transitive | | |

**References**

Abeillé, Anne, Danièle Godard, and Ivan A. Sag. 1998. 'Two Kinds of Composition in French'. In *Complex Predicates in Nonderivational Syntax*, edited by Erhard Hinrichs, Andreas Kathol, and Tsuneko Nakazawa, 1–41. San Diego: Academic Press.

Bickel, Balthasar. 2011. 'Multivariate Typology and Field Linguistics: A Case Study on Detransitivization in Kiranti (Sino-Tibetan)'. In *Proceedings of Conference on Language Documentation and Linguistic Theory 3*, edited by Peter Austin K., Oliver Bond, David Nathan, and Lutz Marten, 3:3–11. University of London: SOAS.

Butt, Miriam. 1995. *The Structure of Complex Predicates in Urdu*. Stanford (CA): CSLI Publications.

Clendon, Mark. 2014. *Worrorra: A Language of the North-West Kimberley Coast*. Adelaide: University of Adelaide Press.

Creissels, Denis. 2012. 'The Origin of Antipassive Markers in West Mande Languages'. Presented at the 45th Annual Meeting of the Societas Linguistica Europaea, Stockholm, September 29.

Creissels, Denis. 2024. *Transitivity, Valency, and Voice*. Oxford (UK): Oxford University Press.

Haspelmath, Martin. 2018. 'Universals of Causative and Anticausative Verb Formation and the Spontaneity Scale'. Edited by Katarzyna Janic and Nicole Nau. *Lingua Posnaniensis* 58 (2): 33–64.

Hofling, Charles A. 2011. *Mopan Maya-Spanish-English Dictionary*. Salt Lake City: The University of Utah Press.

Juárez, Cristian, and Albert Álvarez González. 2021. 'Explaining the Antipassive-Causative Syncretism in Mocoví (Guaycuruan)'. In *Antipassive: Typology, Diachrony, and Related Constructions*, edited by Katarzyna Janic and Alena Witzlack-Makarevich, 315–47. Amsterdam: John Benjamins.

Sansò, Andrea. 2017. 'Where Do Antipassive Constructions Come from? A Study in Diachronic Typology'. *Diachronica* 34: 175–218.

Zúñiga, Fernando, and Seppo Kittilä. 2019. *Grammatical Voice*. Cambridge: Cambridge University Press.

# Circumscribing complex predicate argument structures: Japanese V-V Compounds and argument synthesis

Kazuhiko Fukushima
(Kansai Gaidai University)

Keywords: V-V compounds, argument structure, argument synthesis, thematic proto-riles, Japanese

ARGUMENT SYNTHESIS (AS) in Japanese lexical V1-V2 compounds exhibits various and sometimes unexpected possibilities regarding argument matching/inheritance between/from the component verbal ARGUMENT-STRUCTUREs (ARG-ST) (Kageyama 1988/1993, Matsumoto 1996, Nishiyama 1998, Fukushima 2005/2021, Yumoto 2005, inter alia). However, there is a definite limit regarding what is (im)possible, demonstrating one aspect of limitation in complex predicate formation and offering a potential yardstick for cross-linguistic comparison.

The following four AS types are found. **Pattern 1** (transparent): In transitive Hanako-ga sara-o ***tataki-kowasi-ta***. 'Hanako-NOM plate-ACC hit-destroy-PAST' (V1 ARG-ST<NPi-ga, NPj-o>, V2 ARG-ST<NPi-ga, NPj-o>), the subject PROTO-AGENTs (PA) and direct object PROTO-PATIENTs (PP) of non-head V1 and head V2 are matched, respectively, resulting in: λxλy.(**VIA´(hit**´(x)(y))(**destroy**´ (x)(y))). [N.B.: The function **VIA**´ takes a proposition to return an adverbial propositional modifier for manner, means, cause, etc. modulo contextual plausibility. Also assumed here are Dowty's (1991) PROTO-ROLEs based on two sets of verbal entailments for (i) PA: volition, sentience, causer, movement and (ii) PP: undergoing-change-of-state, incremental-theme, affectee, being-stationary- relative-to-mover.]

**Pattern 2** (additive): *Taroo-ga kodomo-o **ture-sat**-ta*. 'Taroo-NOM child-ACC snatch-leave -PAST' (V1 ARG-ST<NPi-ga, NPj-o>, V2 ARG-ST<NPi-ga> → V1-V2 ARG-ST<NPi-ga, NPj-o>) shows inheritance of a PA (affectee) argument solely from transitive V1 ture and translated as: λxλy.(VIA´(**snatch**´(x)(y))(**leave**´(y))).

**Pattern 3** (subtractive): With no compatible PROTO-ROLE entailment vis-à-vis intransitive V2, the subject ('wearer', PA) of transitive V1 ki is dismissed, while two PPs (affectee, change of state) are matched in *Zyaketto-ga **ki-kuzure**-ta*. 'jacket-NOM wear-deform-PAST' (V1 ARG-ST<~~NP-ga,~~ NPi-o>, V2 ARG-ST<NPi-ga> → V1-V2 ARG-ST<NPi-ga>), yielding (with existential quantification for dismissal): λx∃y.(**VIA´(wear**´(x)(y))(**deform**´(x)))—no wearer is expressed but entailed to exist.

**Pattern 4** (mirror-image): In ditransitive Hanako-ga hon-o Taroo-kara ***yuziri-uke**-ta*. 'Hanako-NOM book-ACC Taroo-from yield-receive-PAST', AS is not: V1 ARG-ST<NPi-ga, NPj-o, NPk-ni>, V2 <NPi-ga, NPj-o, NPk-*kara*>, which is expected and straightforward but leads to a semantic/pragmatic anomaly—the yielder would be the receiver! Instead, AS in this case is mirror-image: V1 <NPk-ga, NPj-o, NPi-ni>, V2 <NPi-ga, NPj-o, NPk-*kara*> → V1-V2 ARG-ST<NPi-ga, NPj-o, NPk-*kara*>. As Dowty tells us, the subjects and indirect objects of the sell/buy type verbs (likewise yield/receive) are PAs (with sentience and volition, i.e. an equal number of PA entailments for each), rendering mirror-image AS possible.

Yet, not every imaginable AS is possible. Constraints on AS: (i) the head verb domination (V2 in the examples above but not total control) with two sub-constraints—(i-a) no obligatory head argument is dismissed; (i-b) V2 determines case-markers (e.g. patterns 3-4). (ii) Argument dismissal is possible before a match is found (pattern 3 vs. pattern 2). (iii) Mirror-image AS is not random and allowed under convergent conditions (PROTO-ROLE equivalence between co-arguments and semantic/pragmatic anomaly avoidance in pattern 4). It is unavailable for regular transitive verbs like *tatak* 'hit' and *kowas* 'destroy' above, where PROTO-ROLE equivalence between co-arguments does not hold. (iv) The more complex AS is, the less productive V-V compounds are—the order descends from patterns 1 to 4. Pattern 2 is ahead of pattern 3, since argument dismissal in the latter requires existential quantification, an extra semantic step otherwise unnecessary.

**References**


Dowty, D. 1991. Thematic proto-roles and argument selection. *Language* 67, 547-619.

Fukushima, K. 2005. Lexical V-V compounds in Japanese: lexicon vs. syntax, *Language* 81, 568-612.

Fukushima, K. 2021. Getting a head with Japanese lexical V-V compounds, in Hae-Sung Jeon et al. (eds.) *Japanese/Korean Linguistics* 28, Stanford: CSLI, 229-243.

Kageyama, T. 1988. Word formation in Japanese, *Lingua* 57, 215-258.

Kageyama, T. 1993. *Bunpoo-to gokeisei* (Grammar and word-formation). Tokyo: Hituzi.

Matsumoto, Y. 1996. *Complex predicates in Japanese: a syntactic and semantic study of the notion 'word',* Stanford: CSLI.

Nishiyama, K. 1998. V-V compounds as serialization, *Journal of East Asian Linguistics* 7, 175-217.

Yumoto, Y. 2005. *Fukugodoshi/haseidoshi-no imi-to togo: mojuru keitairon-kara mita nichieigo-no doshi keisei* (The semantics and syntax of compound verbs/derived verbs: verb-formation in Japanese and English viewed from a modular morphological perspective), Tokyo: Hituzi.

# Object Selection in Japanese resultative V-V compounds

Hideki Kishimoto
(Kobe University)

Keywords: resultative V-V compound, resultative construction, non-subcategorized object, classifier, Japanese

In languages like English. resultative constructions are constructed by combing main verbs with non-verbal predicates, and allow no-subcategorized objects (Levin and Rappaport Hovav 1995). In Japanese, resultative V-V compounds are formed by verb compounding and allow no-subcategorized objects of V1 (see e.g. Matsumoto 1988, Fukushima 2005).

(1) a. Mary cried her eyes out.
b. Mari-ga me-o naki-harasi-ta.
Mari-Nom eye-Acc cry-swell-Past
'Mary cried her eyes swollen.'

In (1b), V1 *naku* 'cry' does not take an object, so the object *me* 'eye' must be the argument of V2. Furthermore, some resultative V-V compounds can have an argument of V2 while suppressing the object of V1, in a way similar to English (cf. Hoekstra 1988).

(2) a. Mary washed {her eyes/the soap out of her eyes}.
b. Mari-ga {me-o arat-ta/sekken-o me-kara arai-nagasi-ta}.
Mari-Nom {eye-Acc wash-Past/soap-Acc eye-from wash-flow-Past}
'Mari washed {her eyes/the soap out of her eyes}.

At first sight, the examples give us the impression that the V-V compounds involve the placement of a small clause in object position. Nevertheless, this view faces difficulty, since the compound *nomi-hosu* 'drink-dry' allows either the object of V1 or V2 to appear in object position.

(3) a. Mary drank {*the beer/the mug} empty.
b. Mari-ga {biiru-o/zyokki-o} nomi-hosi-ta.
Mari-Nom {beer-Acc/mug-Acc} drink-dry-Past
'Mari drank {the beer/the mug} empty(dry).'

Note that if a small clause appears in the object position, just like English, the object of V1 is not expected to occur in the V-V compound construction.

In the resultative V-V compounds, a causative meaning is expressed, while V1, which is most typically activity predicate, represents a causing part, and a caused event by V2, which is most typically an accomplishment/achievement verb. I propose that the selection of the object in (3b) is determined via a derived causative meaning represented as "[x drinks y] causes [[y to become consumed] & [z to become dried]]", where x (=agent), y (=affected theme), z (=location/container) for *noim-hosu*. The lexical meaning includes two caused events, one is provided by *nomu* 'drink' and the other by *hosu* 'dry'.

I suggest that with *nomi-hosu*, when the caused event expressing the consumption of x is prominent, the object of V1 realized, but that the object of V2 is materialized if the caused event representing a change of the container. Even though the presence of y and z is implicated by the lexical meaning of the V-V compound, the object position can be filled by only one argument syntactically. This leads to the expectation that y and z will not be realized simultaneously. I argue that this expectation is verified by making use of floating classifiers; e.g. *iti-rittoru* 'one liter', which identifies a liquid, and *ik-ko* 'one piece, which identifies an inanimate solid entity. Overall, the discussion shows that Japanese utilizes a lexical strategy of verb compounding to fuse the lexical meanings of the component verbs to form a resultative construction, while English resorts to a syntactic operation of placing a small clause constituent in object position.

**References**

Fukushima, Kazuhiko (2005), Lexical V-V compounds in Japanese. *Language* 81, 568– 612.

Hoekstra, Teun (1998), Small clauses results. *Lingua* 74, 101–139.

Levin, Beth & Malka Rappaport Hovav (1995), Unaccusativity: *At the Syntax-Lexical Semantics Interface*. Cambridge, MA: MIT Press.

Matsumoto, Yo (1988), Nihongo-no goiteki fukugoo dooshi-niokeru dooshi-no kumiawase [Verbs combinations in lexical compounds in Japanese]. *Gengo Kenkyu* 114, 37–83.

# Grammaticalization paths, sources, and targets of complex predicates in Beja (North-Cushitic)

Martine Vanhove
(Llacan; CNRS-INaLCO-EPHE)

Keywords: Cushitic, Beja, complex predicates, grammaticalization, TAM

Beja, the sole language of the North-Cushitic branch (Afroasiatic phylum), shows an interesting array of grammaticalization paths of verbal complex predicates, most of them unrecorded so far (e.g. Kuteva et al. 2019; the ComPLETE project https://complete.huma-num.fr/data/ANR-DFG_2021_ComPLETE_Project_Main.pdf). In a bottom-up corpus-based approach, this presentation investigates and discusses the fifteen paths of grammaticalizations attested in this language in the light of their morpho-syntactic properties, which include auxiliary constructions, converb constructions, light verb constructions, and serial verb constructions look-alikes. It will be shown that (i) only four auxiliaries developed from the list of sources of the ComPLETE questionnaire to the proposed targets: BE AT > continuous; GIVE > benefactive; REFUSE > negative intensifier; SAY > volitive; (ii) seven source concepts in this list (as well as in those mentioned in Kuteva et al. 2019) developed into other target concepts than the one mentioned: BE, BECOME > perfect; COME > potential; DO > intensification; LIE (posture) > completive; LOVE / WANT > potential (with negation); SAY > future; SAY > epistemic; SAY > manners (light verb); TAKE > intensification; (iii) four source concepts are not mentioned in the ComPLETE list (but three of the target concepts are), nor in Kuteva et al.: BE DOWN > directional; BE UP > directional; MAKE SEIZE/TIE > intensification; TAKE OUT > completive. A focus will be placed on the (relative) wealth of target concepts that developed from the quotative verb 'say': volitive (converb construction), future and epistemic (auxiliary constructions), and manners (light verb constructions, mainly verbal but also ideophonic). It will be shown that the formal properties of both the lexical verb and the auxiliary contribute to the meaning of the VCPs: 'say' is volitive only when the lexical verb is a Manner converb and the auxiliary is in the Imperfective (or a non-finite form licensed by the syntax) (ex. 1); it marks a future when the lexical verb is a semi-frozen form of the Aorist paradigm and the auxiliary is in the Imperfective (or, again, a non-finite form) (ex. 2), and epistemic values when the auxiliary is in the Aorist paradigm (ex. 3); it expresses manner in a light verb construction only when the lexical verb corresponds to the bare stem (ex. 4, 5). From a diachronic point of view, it is interesting to note that the suffixed conjugation morphemes of the most recent verb class (verb class 2) stemmed from the latter type of construction, and that there are traces of a similar construction in the singular paradigm of verbs of class 1, the oldest one, conjugated with prefixes or infixes (Cohen 1973, Zaborski 1975).

(1) *baro*:*k* *tam-a*:*=ho*:*k* *e*:*-d-na*
2SG.M.ACC eat-CVB.MNR=OBJ.2SG 3-say\IPFV-PL
'You, they want to eat you.'

(2) *tam-i* *i-jad-na*
eat-AOR.SG 3-say\IPFV-PL
'They will eat it.'

(3) *aː=nda* *suːr-ni=heːb* *iː-d-n=eːb* *hiːs-an*
DEF.PL.M.NOM=man\PL precede-AOR.PL=OBJ.1SG 3-say\AOR-PL=REL.M think-PFV.1SG
'I thought that the men would have overpassed me.'

(4) *eːn* *i=takt$^{ʔ}$i* *lawwaːw* *i-ni*
PROX.PL.M.ACC DEF.M=scarecrow prowl 3SG.M-say\PFV
'He prowled swiftly around these scarecrows.

(5) *k$^{w}$iɖ~k$^{w}$iɖ* *iː-d*
disappear~PLAC 3SG.M-say\AOR
'He was astonished for a very short while.'

Acknowledgments: This study is part of the ANR-DFG funded project ComPLETE whose financial support is gratefully acknowledged.

**References**

Cohen, David. 1973. Le renouvellement de l'inaccompli en bédawiye. *Comptes Rendus du GLECS* 14, 69-78.

Kuteva, Tania, Bernd Heine, Bo Hong, Haiping Long, Heiko Narrog, and Seongha Rhee (eds). 2019 (2nd edn.). *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.

Zaborski, Andrzej. 1975. *The Verb in Cushitic*. Krakow: Nakladem Uniwersytetu Jagiellonskiego.

# Simultaneous vs. sequential caused change-of-state events in a corpus of German Sign Language

Cornelia Loos
(Göttingen University)

Background: While sign languages exhibit parallels to spoken languages in all key areas of linguistic description (Sandler & Lillo-Martin, 2006), they also have characteristics unique to the visual-gestural modality. This study builds on one such characteristic, namely the availability of two paired manual articulators which can move independently. The coordination of the two hands allows encoding information about the simultaneous occurrence of two events, which in spoken languages has to be expressed by sequential predicates. In simultaneous event encodings in signed languages, each hand represents one event participant, and the simultaneous presence of both hands represents the parallel existence of these referents in time and space (Perniss, 2007). Simultaneous encoding of information is a much-researched hallmark of sign languages (Vermeerbergen et al., 2007), yet potential limitations on simultaneity have been explored to a much lesser extent.

Aims: The present study examines linguistic and non-linguistic constraints on the simultaneous encoding of complex events in which two subevents occur at the same time. Using data from German Sign Language (DGS), I look at the expression of events that involve an externally caused change of state. An example of such an event might be *Mary hammered the spoon flat*, since Mary's hammering caused a change in the degree of flatness of the spoon. Hammering and becoming flatter go 'hand in hand' here, a fact that a visual-gestural language can potentially express iconically. For example, one hand might represent hammering while the other depicts the spoon flattening over time. Since depictions of the visual properties of an object are typically found in classifier constructions (CCs) in DGS and other sign languages (e.g. Schembri, 2003), CCs form the focus of this study. Here I present the results of a corpus study that identifies simultaneous and sequential CCs used to encode an externally caused change of state.

Method: The data set was obtained from the DGS Corpus (Konrad et al. 2020), which contains over 600.000 annotated tokens from 330 signers from all over Germany performing a variety of elicitation tasks. DGS sentences are glossed and have German and English translations. I first compiled a list of 75 German and 61 English change-of-state verbs on the basis of Levin (1993) and Schumacher (1986), then checked the list against English and German translation tags aligned with DGS sentences in the corpus that contained a CC. The resulting 388 candidates for sequential and simultaneous CCs denoting a caused change of state were annotated for the a) presence of temporal overlap between cause and change of state, b) simultaneous vs. sequential production of predicates, c) meaning, d) type of movement.

Results & Workshop relevance: Only 5/388 CCs depict the simultaneous nature of causing action and change of state simultaneously. I discuss motoric, phonological, and grammatical reasons for this unexpected paucity of simultaneous constructions. The proposed study aligns with workshop 6: It examines limitations on simultaneous complex predicates in sign languages (6) using a corpus-based approach (7).

**References**

Konrad, R., Hanke, T., Langer, G., Blanck, D., Bleicken, J., Hofmann, I., Jeziorski, O., König, L., König, S., Nishio, R., Regen, A., Salden, U., Wagner, S., Worseck, S., Böse, O., Jahn, E., Schulder, M. 2020. MEINE DGS – annotiert. Öffentliches Korpus der Deutschen Gebärdensprache, 3. Release / MY DGS – annotated. Public Corpus of German Sign Language, 3rd release [Dataset]. Universität Hamburg. https://doi.org/10.25592/dgs.corpus-3.0

Levin, Beth. 1993. English Verb Classes and Alternations.: A Preliminary Investigation. Chicago, Ill.: University of Chicago Press.

Perniss, P. 2007. Space and iconicity in German Sign Language (DGS). Ph.D. Dissertation, Radboud University.

Sandler, W. & D. Lillo-Martin. 2006. *Sign language and linguistic universals*. Cambridge: CUP.

Schembri, A. 2003. Rethinking 'classifiers' in signed languages. In K. Emmorey (ed.), *Perspectives on classifier constructions in sign languages,* 13-44. Psychology Press.

Schumacher, Helmut. 1986. Verben in Feldern: Valenzwörterbuch zur Syntax und Semantik deutscher Verben. Berlin: de Gruyter.

Vermeerbergen, M., Leeson, L., & O. Crasborn (eds.). 2007. *Simultaneity in sign language: Form and function*. Amsterdam: John Benjamins.

# Semi-auxiliary negative verbs in Enets

Andrey Shluinsky
(Institute of Linguistics, Russian Academy of Sciences)

My presentation will deal with semi-auxiliary negative verbs Enets (< Samoyedic < Uralic; Taymyr, North Asia; Forest Enets examples are used in the abstract) based on data of text corpora. Expressing clausal negation with an inflected negative verb and a non-finite form (so-called connegative) of a lexical verb, as in (1), is a well-known feature of Uralic (cf. Miestamo et al. (eds.) 2015). Some Samoyedic languages also feature negative verbs with a more complex meaning, which are also used with the connegative form, as in (2). In Forest Enets, these are *bunʲi-* 'after all not', *ketʃi-* 'almost' and *(i)iŋe-* 'of course'. While in the standard negative construction (1), the negative verb is an instance of an auxiliary with a a clearly grammatical function, constructions like that in (2) are semantically and functionally more complex. Such constructions are a non-canonical instance of a complex predicate and challenge the borderline between different types of them, namely auxiliaries and light verbs.

I will focus on the distinctions between negative semi-auxiliaries and negative auxiliary, especially those relevant for a general framework of complex predicates.

First, in contrast to the main negative auxiliary, which has an almost full verbal paradigm, negative semi-auxiliaries are drastically morphologically defective and allow very limited tense-modal inflection. This goes against an expectation that the more an item is grammaticalized, the more should it be morphologically reduced, but can be explained by the fact that meaning of semi-auxiliaries is compatible with only some temporal or modal contexts. Second, while in the standard negative construction, a lexical verb is usually linearly adjacent to the auxiliary, constructions in focus allow that complements or adjuncts of a lexical verb are placed between it and a semi-auxiliary, as in (3). Thus, semi-auxiliaries have a less
tight syntactic relation with lexical verbs.

Third, there can be a pause between a negative semi-auxiliary and a lexical verb, as in (4), and false starts are possible where a auxiliar remains without a lexical verb complement at all, as in (5). These facts also evidence in favor of a greater syntactic autonomy of items in the semi-auxiliary construction.

In sum, Enets constructions with semi-auxiliary negative verbs predictably behave as a less tight type of complex predicates than negative auxiliary constructions based on syntactic evidence, but at the same time morphologically are in a more advanced stage of grammaticalization. This makes them a non-canonical case of complex predicates relevant for elaborating comparative concepts in this domain.

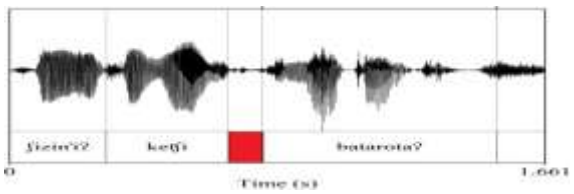
Examples

(1) *bi-t* *n'e-zʔ* *kɔma-ʔ*
water-LAT.SG NEG-1SG want-CNG
'I don't want water.'

(2) *iŋe-zʔ* *kɔma-ʔ* *te-d*
of.course-1SG want-CNG reindeer-LAT.SG
‘Of course, I want reindeer.’

(3) *iŋe-d* *mɛz* *meɔn* *mɔzara-ʔ*
of.course-2SG tent.OBL.SG along work-CNG
‘Of course, you would work at home.’

(4) *ʃizin’iʔ* *ketʃi* *batarota-ʔ*
1DU.ACC almost.3SG overturn-CNG
‘He almost overturned us.’




(5) *nɛɛkuju-r* *kečі…* *d’uzida* *entʃeʔ* *ɛ-zaraxa-bi*
other-2SG almost.3SG clairvoyant person be-SIMIL-PRF.3SG
‘One of them almost… It seems, he was a clairvoyant man.’

**References**

Miestamo, Matti, et al. (eds.). 2015. *Negation in Uralic languages*. Amsterdam: John Benjamins.

# WS7 Lexical affixes

# Ποιέομαι / -ποιέομαι in classical and post-classical Greek: The bound and unbound forms of a support verb

Victoria Beatrix Fendel
(University of Oxford)

Support-verb constructions consist of a verb and a noun that fill the predicate slot of a sentence, e.g. *to make a suggestion* in *I made the suggestion that she join*. Opinions differ as to whether support verbs grammaticalize (Ittzés 2022; Anderson 2006; Slade 2013) or only lexicalise (Butt 2010; Rosén 2020). The support verb 'to do' appears in bound and unbound forms in classical and postclassical Greek. The bound form has been seen in the context of noun incorporation which however does not suit the semantic make-up of the underlying support-verb construction, in that the support verb is discursively secondary as patterns of anaphora show (Asraf 2021; Pompei 2006; Boye 2023). Alternatively, univerbation allots equal weight to the verb and noun and can result from grammaticalisation or lexicalisation (Lehmann 2020; Giomi 2023).

Support verbs, bound and unbound, are not straightforwardly grammatical, like auxiliaries, as they participate in the event structure (Butt 2010) rather than just providing TAM information (Croft 2022), nor are they fully productive – the unbound form imposes collocational restrictions on the range of nouns it combines with (Mel'čuk 2023); the bound form seems to combine with o-stems only (βλάβην ποιέομαι / βλαβοποιέομαι 'to hurt' (intransitive)). Support verbs, bound and unbound, are not fully lexical as they modify the event construal (e.g. Wittenberg & Levy 2017).

The paper looks at *How do lexical affixes differ functionally from grammatical affixes and from non-bound morphemes?* regarding the bound and unbound forms of support verbs. It shows that the two forms discursively differ (e.g. Savary et al. 2018), that 'to do' as a semi- lexical suffix in Greek never becomes fully productive but undergoes domain specialisation (μαρτυροποιέομαι 'to give a witness statement') (cf. Schutzeichel 2014), and that support-verb constructions form an internally heterogenous group of constructions in which 'to do' is an outlier.

**References**

Anderson, Gregory. 2006. *Auxiliary verb constructions* (Oxford Studies in Typology and Linguistic Theory).
Oxford: Oxford University Press.

Asraf, Nadav. 2021. The mechanism of noun incoporation in Ancient Greek. *Glotta* 97. 36–72.

Boye, Kasper. 2023. Grammaticalization as Conventionalization of Discursively Secondary Status: Deconstructing the Lexical–Grammatical Continuum. *Transactions of the Philological Society* 121(2). 270–292. https://doi.org/10.1111/1467-968X.12265.

Butt, Miriam. 2010. The Light Verb Jungle: Still Hacking Away. In Mengistu Amberger, Brett Baker & Mark Harvey (eds.), *Complex Predicates: Cross-Linguistic Perspectives on Event Structure*, 48–78. Cambridge: Cambridge University Press.

Croft, William. 2022. *Morphosyntax: constructions of the world's languages*. Cambridge: Cambridge University Press.

Giomi, Riccardo. 2023. *A functional discourse grammar theory of grammaticalization* (Brill's Studies in Historical Linguistics). Vol. 19. Leiden: Brill.

Ittzés, Máté. 2022. Light Verb, Auxiliary, Grammaticalization: The Case of the Vedic Periphrastic Perfect. *Die Sprache* 54. 95–129.

Lehmann, Christian. 2020. Univerbation. *Folia Linguistica Historica* 41(1). 205–252. https://doi.org/10.1515/flih-2020-0007.

Mel'čuk, Igor. 2023. *General Phraseology*. Amsterdam: John Benjamins.

Pompei, Anna. 2006. Tracce di incorporazione in greco antico. In Pierluigi Cuzzolin & Maria Napoli (eds.), *Fonologia e tipologia lessicale nella storia della lingua greca. Atti del Incontro Internazionale di Linguistica Greca, Bergamo, 15–17 September 2005*, 216–237. Milan: FrancoAngeli.

Rosén, Hannah. 2020. Composite predicates in the layers of Latin. *Journal of Latin Linguistics* 19(2). 231–279. https://doi.org/10.1515/joll-2020-2009.

Savary, Agata, Marie Candito, Verginica Mititelu, Eduard Bejček, Fabienne Cap, Slavomír Čéplö, Silvio Cordeiro, et al. 2018. PARSEME multilingual corpus of verbal multiword expressions. In Stella Markantonatou, Carlos Ramisch, Agata Savary & Veronika Vincze (eds.), *Multiword expressions at length and in depth: Extended papers from the MWE 2017 workshop*, 87–147. Berlin: Language Science Press.

Schutzeichel, Marc. 2014. *Indogermanische Funktionsverbgefüge* (Wissenschaftliche Schriften der WWU Münster 12). Vol. 13. Münster: Monsenstein und Vannerdat OHG.

Slade, Benjamin. 2013. Diachrony of light and auxiliary verbs in Indo-Aryan. *Diachronica* 30(4). 531–578.

Wittenberg, Eva & Roger Levy. 2017. If you want a quick kiss, make it count: How choice of syntactic construction affects event construal. *Journal of Memory and Language* 94. 254–271.

# Lexical affixes in Wao Terero depend on context for properties associated with the lexical-grammatical dichotomy

Noah Diewald
(Ohio State University)

Keywords: lexical suffixes, morphology, Wao Terero, realizational, classifiers, pragmatics

Lexical affixes (glossed LS) are central in discussions of the synchronic lexical-grammatical dichotomy. I consider Wao Terero lexical suffix fieldwork data in light of Boye and Harder (2012), who argue that the grammatical is conventionally discourse secondary. The Wao Terero system is interesting because lexical suffixes occur on many parts of speech and have a variety of uses. I leverage this variety to provide evidence that lexical-grammatical meaning properties are contextual (syntactic, semantic, pragmatic) and are not properties of the affixes themselves, which are only licensed forms. This explains grammaticality clines in function/meaning since affixes have no rigid properties and depend on context for such associations. This may be contrasted with a theory such as Distributed Morphology (DM) (Halle and Marantz 1993) where a morpheme is either a feature bundle (grammatical) or a √root (lexical).

I link the notion of discourse primary and secondary in Boye and Harder (2012) with the concept of proffered and non-proffered in pragmatics (Roberts, 2012). Non-proffered (secondary) content is not expected to be questioned in discourse. A reaction of That's false! to He ran. may target the action asserted, running, but not the non-proffered tense. I probe this distinction in my fieldwork using negation diagnostics.

(1) a. wii yẽdẽ-ka eibe ĩ-dabãĩ ĩ-pa
NEG big-LS.stone above COP-NEG COP-DECL
'The big one (referring to stone) isn't above.'

b. wii di-ka ĩ-dabãĩ ĩ-pa
NEG Ø-LS.stone COP-NEG COP-DECL
'It isn't a stone.'

In (1a), it may be demonstrated that only the adjectival meaning of 'big' is proffered (primary) and canceled under negation. The stone reference is not canceled and speakers may felicitously continue referring to the antecedent. The lexical suffix meaning is conventionally non-proffered (secondary), thus grammatical. An extreme contrast is the nominal example (1b). The root is essentially meaningless, and negation cancels only the meaning associated with the affix, making it proffered (primary), thus lexical. This is unexpected if the lexical affixes are intrinsically lexical or grammatical. Their behavior should be uniform.

One may question whether the distinction proposed by Boye and Harder (2012) provides a sound measure of grammaticality but their proposal is a reaction to serious flaws in previous testsfor the category. My claim is that morphological forms do not directly encode grammaticality. It is not weakened, nor is a DM-like theory supported, if a discrete distinction cannot be measured at all.

In addition, the paper provides a variety of examples and deeper description of the Wao Terero system. I also explore other meaning dimensions relevant to the dichotomy, such as concreteness, and discuss them in terms of contextual factors.

**References**

Boye, Kasper and Peter Harder (2012), A usage-based theory of grammatical status and grammaticalization, *Language* 88(1), 1–44.

Halle, Morris and Alec Marantz (1993), Distributed morphology and the pieces of inflection, in K. L. Hale and S. J. Keyser (eds), (1993), *The View from Building 20,* Cambridge, MA: MIT Press, 111–176.

Roberts, Craige (2012), Information structure in discourse. Towards an integrated formal theory of pragmatics, *Semantics & Pragmatics,* 5(6), 1–69.

# From noun incorporation to lexical affixation in Northwest Caucasian (with focus on Abaza)

Peter Arkadiev
(University of Freiburg)

Keywords: Northwest Caucasian languages, lexical affixes, applicatives, incorporation, grammaticalisation

The polysynthetic Northwest Caucasian languages (Arkadiev, Lander 2020) possess rich systems of verbal affixes expressing spatial relations, the so-called locative preverbs, ranging from several dozen in Circassian to more than a hundred in Abkhaz and Abaza. Most of them function as applicatives introducing an indirect object (Arkadiev et al. 2024) and semantically specifying its spatial relation to the situation expressed by the verb. Preverbs show considerable variation in terms of the degree of abstractness of their meanings and their formal and semantic relations to lexical elements. While some preverbs express very broad spatial relations and have no obvious counterparts among lexical roots, others have clear nominal cognates. Focusing on Abaza and drawing primarily on Klychev (1994, 1995), I show that the Abaza locative preverbs form a cline from incorporated nouns retaining their lexical meaning (body-parts or salient landmarks/artefacts, e.g. 'yard' or 'cradle') (1)–(2) to elements, which, even if synchronically coinciding with nouns, combine with verbs only in abstract spatial meanings (e.g. 'face' > 'vertical surface', 'belly' > 'inside') (3).

(1) *a-ʒə* *jə-**ḳʷa**-ç̌-əw-n*
DEF-water 3SG.M.IO-**LOC:bosom**-pour.liquid-IPF-PST
'Water was pouring into his bosom.' (Klychev 1995: 138)

(2) *a-sabəj* *d-**gara**-l-gʷa-n*
DEF-child 3SG.H.ABS-**LOC:cradle**-3SG.F.ERG-put-PST
'She put the child in the cradle.' (Klychev 1995: 67)

(3) *a-sabəj* *j-qəza* *d-a-**mgʷa**-š'š'-χ-ṭ*
DEF-child 3SG.M.PR-blanket 3SG.H.ABS-3SG.N.IO-**LOC:inside**-crawl-RE-DCL
'The child crawled inside his blanket.' (Klychev 1995: 162)

Between these two extremes are preverbs that occur both in the original and in the spatial meanings, e.g. *ʒa-* 'cheek; lateral surface' (4).

(4) a. *də-j-**ʒa**-kš-əw-n*
3SG.H.ABS-3SG.M.IO-**LOC:cheek**-hit-IPF-PST
'S/he was beating him on the cheek.' (Klychev 1995: 98)

b. *a-kdə* ***ʒa**-r-bəl-wa-n*
DEF-stump **LOC:side**-3PL.ERG-burn-IPF-PST
'They were burning the stump on one side.' (Klychev 1995: 99)

The preverbs differ not only in their productivity (according to Klychev 1995, some of them combine with hundreds of verbs, while others occur with just a handful), but also in the type frequency of verbs with which their original meanings are retained. For instance, the preverb *lakta-* 'face; in front of' retains its lexical meaning with 14 out of 17 verbs it combines with (Klychev 1995: 153–154), while the preverb *qa-* 'head; above' shows an opposite pattern with just three out of more than 200 verbs showing the body-part meaning of the preverb (Klychev 1995: 250–260).

Even with apparently lexical preverbs such signs of grammaticalisation can be observed as co-occurrence with synonymous or hyponymous external noun phrases, whereby preverbs function as verbal classifiers (5), or the morphological distinction between introversive and extraversive forms (6) lacking in nouns, as well as loss of the lexical cognate (as e.g. the preverb *čḳara-* 'courtyard').

(5) *a-**klub*** *jə-**ʕʷna**-n.χ-əj-ṭ*
DEF-club 3PL.ABS-**LOC:house**-work-PRS-DCL
'They work in the club.' (Klychev 1995: 88)

(6) a. *aʕʷ* ***čʕʷa**-l-ga-ṭ*
DEF+trough **LOC:hearth**-3SG.F.ERG-carry-DCL
'She brought the trough to the hearth.' (Klychev 1995: 211)

b. *d-**čʕʷə**-r-ga-χ-ṭ*
3SG.H.ABS-**LOC:hearth.ELAT**-3PL.ERG-carry-RE-DCL
'They carried him out of the hearth.' (Klychev 1995: 218)

All this shows that the system of locative preverbs in Abaza has been shaped by recurrent processes of incorporation and morphologisation and that the distinction between incorporated nouns and affixes is gradual rather than binary.

**References**


Arkadiev, Peter, and Yury Lander (2020), The Northwest Caucasian languages, in M. Polinsky (ed.), (2020), *The Oxford Handbook of the Languages of the Caucasus*, Oxford: Oxford University Press, 369–446.

Arkadiev, Peter, Yury Lander, and Irina Bagirokova (2024), Applicative constructions in the Northwest Caucasian languages, in Fernando Zúñiga, and Denis Creissels (eds.), (2024), *Applicative Constructions in the World's Languages*, Berlin / Boston: De Gruyter Mouton, 871–913.

Klychev, Rauf N. (1994), *Lokal'no-preverbnoe obrazovanie glagolov abazinskogo jazyka* [The locative preverbial derivation of verbs in Abaza], Cherkessk: Adž'pa.

Klychev, Rauf N. (1995), *Slovar' sočetaemosti lokal'nyx preverbov s suffiksoidami i glagol'nymi kornjami v abazinskom jazyke* [The collocational dictionary of locative preverbs with suffixoids and verbal roots in Abaza], Cherkessk: Karačaevo-čerkesskoe knižnoe izdatel'stvo.

# Auxiliaries and the grammatical-lexical distinction

Ronald P. Schaefer & Francis O. Egbokhare
(Southern Illinois University Edwardsville & University of Ibadan)

Keywords: auxiliary type, secondary grammatical status, quantitative modifications, emphatic structures

Boye (2023) has recast grammaticalization as conventionalization of discursively secondary grammatical status. He thereby emphasizes grammaticalization as a process of meaning change undergirded by attentional prominence differentially assigned to some linguistic units relative to others. Accompanying this split in conceptual structure are differences in syntagmatic discourse prominence. Quite naturally this has led to renewed interest in linguistic coding of complex messages and the lexical-grammatical (L-G) distinction.

One domain where this recast can be assessed is that of derivational affixes, which have a variable nature in languages where they occur. Over the years such affixes have become a target of closer scrutiny (Hopper and Traugott 2003 for English, Kinkade 1963 a-b, Mithun 1997, Mattisen 2017). For languages without derivation, analytic/isolating languages of the type found in West Africa, relations among other syntagmatic units gain attention.

For this paper we concentrate on auxiliaries or preverbs in West Benue Congo, Edoid in particular. In Emai for example there are a range of "auxiliaries" that precede the verb and reflect meanings either assigned to traditional tense and modality or to clausal modifications that are, according to Nuyts (2005, 2006, 2016), qualitative (aspectuals, boulomaic modality) or quantitative (dynamic modality, temporal rate, quantum values). In other languages such elements are identified as to auxiliary subtype (primary vs. secondary in Andersen 2007) or preverb subtype (restricted vs. unrestricted in Bamgbose 1966). Regardless of characterization, auxiliaries in Emai present a L-G problem analogous to derivational affixes.

Traditionally, auxiliaries (or preverbs) are assumed to be grammatical units, not lexical (van Gelderen 2010). There is reason to believe that this assumption does not hold for the entire set of Emai auxiliaries. Quantitative modifications associated with dynamic modality (*dábɔ̀* deliberately,' *dègbè* 'carefully'), intensification (*dòbɔ́* REFLEXIVE), and temporal rate/distance (*tùà* 'hurriedly') participate in emphatic structures that call attention to their meanings in traditional oral narrative discourse. Emphasis is conveyed by repetition of a qualitative auxiliary whose character is either additive (*gbò* ADD 'too, also') or distal manner deictic (*ìyɔ́* DMD 'that way, like that').

(1) a. *ɔ́lí ɔ́mɔ́hé ˡíyɔ́ dábɔ̀ íyɔ́ nwú ɔ̀ì fí à.*
ART man:DST PST:DMD deliberately DMD take 3SG propel ITV
'The man deliberately like this tossed it away.'

b. *ɔ́lí ɔ́mɔ́hé ˡgbó tùà gbó é émàè.*
ART man:DST PST:ADD hurriedly ADD eat food
'The man also ate the food in a real hurry.'

c. *ójé ꜜzá íyɔ́ gbò dégbè gbó dóbɔ́ ɔ́ì gbó é émàè.*
Oje:DST PST:RES DMD ADD carefully ADD REFL 3SG ADD eat food
'As a result, Oje that way by himself real carefully ate the food.'

Auxiliaries of the type associated with TAM, including relative tense, epistemic, and deontic modality, do not participate in emphatic structures. Nor do they occur in imperatives or correspond to information questions, as do some quantitative modifications.

**References**


Andersen, Torben (2007), Auxiliary Verbs in Dinka, *Studies in Language* 31(1), 89-116.

Bamgbose, Ayo (1966), *A grammar of Yoruba*, Cambridge: Cambridge University Press.

Boye, Kasper (2023), Grammaticalization as Conventionalization of Discursively Secondary Status: Deconstructing the Lexical-Grammatical Continuum, *Transactions of the Philological Society* 121(2), 270-292.

Hopper, Paul and Elizabeth C. Traugott (2003), *Grammaticalization*, 2nd edition, Cambridge: Cambridge University Press.

Kinkade, M.D. (1963a), Phonology and Morphology of Upper Chehalis I, *International Journal of American Linguistics* 29(3), 181-195.

Kinkade, M.D. (1963b), Phonology and Morphology of Upper Chehalis II, *International Journal of American Linguistics* 29(4), 345-356.

Mattissen, J. (2017), Sub-types of polysynthesis, in M.D. Fortescue, M. Mithun and N. Evans (eds), (2017), *The Oxford Handbook of Polysynthesis*, Oxford: Oxford University Press, 70-98.

Mithun, M. (1997), Lexical affixes and morphological typology, in J.L. Bybee, J. Haiman and S. A. Thompson (eds), (1997), *Essays on Language Function and Language Type,* Amsterdam: John Benjamins, 357-371.

Nuyts, Jan (2006), Modality: Overview and linguistic issues, in William Frawley (ed), (2006), *The Expression of Modality*, Berlin: De Gruyter, 1-26.

Nuyts, Jan (2016), Analyses of modal meanings, in Jan Nuyts and Johan van der Auwera (eds), (2016), *The Oxford Handbook of Modality and Mood*, New York: Oxford University Press, 31-49.

van Gelderen, Elly (2010), *An introduction to the grammar of English: Revised edition,* Amsterdam: John Benjamins.

# Compounds, affixoids and discursive prominence: Two case studies

Muriel Norde
(Humboldt-Universität zu Berlin)

Keywords: morphology, discursive prominence, Dutch, Swedish

The aim of this paper is to test Boye & Harder's model of lexical and grammatical status (Boye & Harder 2012; Boye 2023) against a specific kind of compounding constructions in the Germanic languages. Adopting the view that complex morphological constructions, including compounds, are "syntagms of cooperating morphemes" (Juul Nielsen 2016: 290), it appears that compounding members are lexical by definition, since they can always be "discursively primary", i.e. "the attentional main point of a syntagm" (Boye 2023: 274). This can be illustrated by the Dutch compound *schoolmeisje* 'schoolgirl': even though the right-headedness of Germanic (determinative) compounds implies that the left-hand member modifies the meaning of the head, both compounding members can be focused, as in (1a-b).

(1a) *Het is een MEISJESschool, geen JONGENSschool* 'it is a girls' school, not a boys' school'

(1b) *Het is geen meisjesSCHOOL maar een meisjesGEVANGENIS* 'it is not a girls' school, but a girls' prison'

Things become less straightforward when compounding members change into affixoids, i.e. "morphemes which look like parts of compounds, and do occur as lexemes, but have a specific and more restricted meaning when used as part of a compound" (Booij 2009: 208). Affixoids have been the subject of several (empirical) studies into their collocational and morphosyntactic behaviour (e.g. Ascoop & Leuschner 2006; Booij & Hüning 2014; Battefeld, Leuschner & Rawoens 2018; Norde & Van Goethem 2018), but the issue of their discursive status has not yet been addressed. At the same time, suffixoids (Stevens 2005), have received far less attention than prefixoids. In my talk, I aim to fill these two gaps by focusing on two suffixoids: Dutch *-boer* and Swedish *-jävel(n)*.

The study is based on two annotated data sets, of 1000 tokens each, which are sampled from the TenTen corpora at SketchEngine (Kilgarriff et al. 2014). These form the basis for an analysis of the discursive prominence of the two suffixoids, for which we use the criteria outlined in Boye (2023: 275). In this model, elements are discursively secondary when:

(i) They cannot be focused,
(ii) They cannot be addressed in subsequent discourse,
(iii) They cannot be elaborated through modification, and
(iv) They cannot stand alone in an utterance.

A first analysis of the data suggests different results for the suffixoids in this study. Dutch *-boer* (originally meaning 'farmer'), which means 'seller' in *telefoonboer* 'phone seller', fails to meet criterion (iii), because it can be modified by the first member of the compound. However, it does meet the other criteria: it cannot be focused, and because the suffixoid meaning becomes overridden in all contexts other than compounds by the meaning of the free lexeme, it cannot be addressed, nor can it stand alone, as shown in example (2).

(2) Q: *Wat voor boer ben jij?* 'What kind of farmer / *seller are you?'

A: *Aardappelboer* / **Telefoonboer* 'potato farmer / *phone seller')

For the evaluative suffixoid that is Swedish *-jävel(n)*, originally meaning 'devil' but translating into 'bloody' when used as a suffixoid (*bussjäveln* 'bloody bus'), there is an interesting mismatch because the suffixoid functions as the head morphologically, but not semantically (a *bussjävel* is a bus, not a devil). This means that *-jävel(n)* is discursively secondary according to all criteria: it cannot be stressed or modified, and example (3) shows that it cannot be addressed or used independently in its evaluative meaning.

(3) Q: **Med vilken jävel kom du hit?* 'with which bloody thing did you get here'?

A: *Med bussjäveln* 'with the bloody bus'

**References**


Ascoop, Kristin & Torsten Leuschner (2006), "Affixoidhungrig? Skitbra!" Comparing affixoids in German and Swedish, *Language Typology and Universals* 59(3), 241–252.

Battefeld, Malte, Torsten Leuschner & Gudrun Rawoens (2018), Evaluative morphology in German, Dutch and Swedish, in Kristel Van Goethem, Muriel Norde, Evie Coussé & Gudrun Vanderbauwhede (eds), *Category Change from a Constructional Perspective* , Amsterdam / Philadelphia: John Benjamins, 229–262.

Booij, Geert E. (2009), Compounding and Construction Morphology, in Rochelle Lieber & Pavol Štekauer (eds), *The Oxford Handbook of Compounding*, Oxford: Oxford University Press, 201–216.

Booij, Geert & Matthias Hüning (2014), Affixoids and constructional idioms, in Ronny Boogaart, Timothy Colleman & Gijsbert Rutten (eds), *Extending the Scope of Construction Grammar*, Berlin / Boston: De Gruyter Mouton, 77–106.

Boye, Kasper (2023), Grammaticalization as Conventionalization of Discursively Secondary Status: Deconstructing the Lexical–Grammatical Continuum, *Transactions of the Philological Society* 121(2), 270–292.

Boye, Kasper & Peter Harder (2012), A Usage-Based Theory of Grammatical Status and Grammaticalization, *Language* 88(1), 1–44.

Juul Nielsen, Peter (2016), *Functional Structure in Morphology and the Case of Nonfinite Verbs*, Leiden: Brill.

Kilgarriff, Adam, Vít Baisa, Jan Bušta, Miloš Jakubíček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý & Vít Suchomel (2014), The Sketch Engine: ten years on, *Lexicography* 1(1), 7–36.

Norde, Muriel & Kristel Van Goethem (2018), Debonding and Clipping of Prefixoids in Germanic: Constructionalization or Constructional Change?, in Geert Booij (ed.), *The Construction of Words: Advances in Construction Morphology* (Studies in Morphology), Cham: Springer International Publishing, 475–518.

Stevens, Christopher M. (2005), Revisiting the Affixoid Debate: On the Grammaticalization of the Word, in Torsten Leuschner, Tanja Mortelmans & Sarah De Groodt (eds), *Grammatikalisierung im Deutschen*, Berlin / New York: De Gruyter, 71–84.

# Affixal verbs in Siberia: another case of lexical affixes as an areal feature?

Alexey Vinyar & Daria Ignatenko
(National Research University Higher School of Economics & National Research University Higher School of Economics; )

Keywords: lexical affixes; areal typology; languages of Siberia; denominal verbalizers; polysynthesis
The present study is a part of larger areal-typological project conducted by a Laboratory of which one Author is a member.

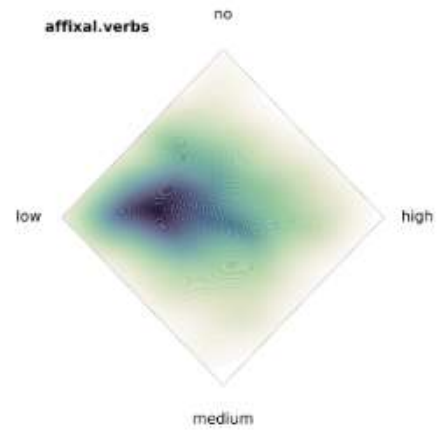
Our areal-typological study deals with a subtype of lexical affixes: denominal verbalizers (see Gerdts & Marlett (2008) conveying root-like semantics, which we label as affixal verbs. Instead of denoting abstract meanings which either largely depend on the verbalized noun's semantics or consist of primitives of event decomposition (like proposed for Inuktitut by Johns (2007)), affixal verbs specify event's manner or (change of) state, a property of $\sqrt{\text{ROOT}}$ in event decomposition studies, see (Harley 2005), (Beavers & Koontz-Garboden 2020) inter alia. Affixal verbs can be primary elements of discourse (see Boye & Harder 2012) because of their status of a single predicative element of a clause.

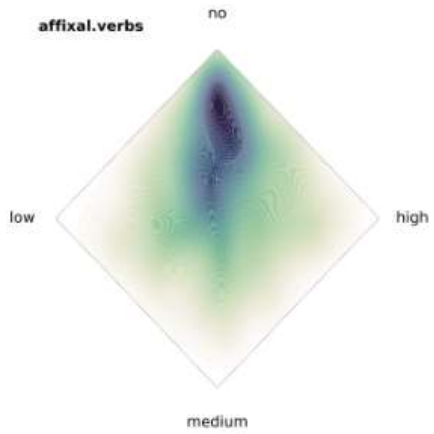
As an example, verbalizer -h denoting hunting, gathering and consumption in Tundra Nenets (Samoyedic < Uralic, see Nikolaeva 2014: 47) does not count as an affixal verb because its host-dependent «getting» meaning derives just from primitives cause (to) have. Its counterpart -ð in Forest Enets (< Samoyedic) derives only verbs of hunting thus specifying manner of «getting» and is an affixal verb.

Our study shows that Siberia may be another «area of lexical affixes» (besides America's Northwest identified by Mithun (1997): affixal verbs are found in Yupik, Samoyedic, Tungusic and Chukotian languages. We have conducted and analyzed a sample of 69 languages of Siberia and neighboring areas (in line with Bickel's (2015) suggestions) and compared it with a 100-language worldwide variety sample. The quantitative analysis based on Ranacher et al. (2021) shows that the presence of affixal verbs in Siberia is best attributed to contact areal factors (see Appendix 1)1. Qualitatively, semantic similarity of affixal verbs in unrelated languages also points toward areal explanation of their origin (see Appendix 2). In our paper we discuss such possible explanations.

Appendix 1. Density plots visualizing posterior distribution of affixal verbs in different areas, based on Ranacher et al. (2021)

Fig. 1. Density plot for Siberia Fig. 2. Density plot for neighboring areas Fig. 3. Density plot for World





Appendix 2. Frequently occurring semantics of affixal verbs in selected unrelated Siberian languages

| | consume | hunt | gather | smell | play |
|---|---|---|---|---|---|
| Central Siberian Yupik (Eskimo-Aleut) | + | + | + | + | + |
| Chukchi (Chukotko-Kamchatkan) | + | + | + | | |
| Even (Tungusic) | + | + | + | + | + |
| Nganasan (Uralic) | + | + | | + | + |

**References**

Badten, L. W., Kaneshiro, V. O., Oovi, M., Koonooka, C., Jacobson, S. A. (2008), St. Lawrence Island / Siberian Yupik Eskimo dictionary. Fairbanks: Alaska Native Language Center, University of Alaska.

Beavers, John & Koontz-Garboden, Andrew. 2020. The roots of verbal meaning. Oxford University Press. Pp. 288.

Bickel, Balthasar (2015), Distributional Typology: Statistical Inquiries into the Dynamics of Linguistic Diversity. in Heine B., and Narrog H. (eds), The Oxford Handbook of Linguistic Analysis, 2nd ed

Boye, Kasper & Peter Harder. 2012. A usage-based theory of grammatical status and grammaticalization. Language 88.1. 1–44.

Gerdts, Donna B. & Marlett, Stephen A. (2008), Introduction: The form and function of denominal verb constructions. International Journal of American Linguistics 74(4). 409-422.

Harley, Heidi (2005), How do verbs get their names? Denominal verbs, manner incorporation, and the ontology of verb roots in English. In N. Erteschik-Shir and T. Rapoport, eds. The Syntax of Aspect, pages 42–64. Oxford: Oxford University Press.

Johns, Alana (2007), Restricting noun incorporation: root movement. Natural Language & Linguistic Theory. 25. 535–576.

Mithun, Marianne (1997), Lexical affixes and morphological typology. In Joan L. Bybee, John Haiman & Sandra A. Thompson (eds.), Essays on language function and language type, 357–371. Amsterdam/Philadelphia: John Benjamins.

Nikolaeva, Irina (2014), A grammar of Tundra Nenets. Berlin: Mouton de Gruyter.

Ranacher, P., Neureiter, N., van Gijn, R., Sonnenhauser, B., Escher, A., Weibel, R., Muysken P., Bickel, B. (2021), Contact-tracing in cultural evolution: A bayesian mixture model to detect geographic areas of language contact. J R Soc Interface.

Tereshenko, N. M. (1966), Enetskij jazyk [Enets language]. In Jazyki narodov SSSR [Languages of USSR people]. Vol. 3, pp. 438-457. Moskow.

Wagner-Nagy, Beáta. 2018. A grammar of Nganasan. Brill (vol. 6).

# “Lexical affixes” of associated motion and associated posture in Tacana and beyond

Antoine Guillaume
(Laboratoire Dynamique du Langage; CNRS; Université Lumière Lyon 2)

Keywords: motion, posture, imperfective aspect, grammaticalization theory, Amazonian languages

In this presentation, my goal is to discuss the spatial verbal categories of associated motion (AM) and associated posture (AP) from the perspectives of the literature on “lexical affixes” and grammaticalization theory.

AM and AP refer to morphemes that are associated with the verb and that have among their possible functions the coding of translational motion in the case of AM (Guillaume & Koch 2021) and postural meanings in the case of AP (Guillaume 2024). In some languages, AM and AP are expressed through forms that qualify, according to formal criteria, for a very advanced stage of grammaticalization. In Tacana, for instance, they are integral components of an elaborate inflectional paradigm featuring nine imperfective circumfixes, as illustrated in (1a) and (1b).

(1) Tacana (Takanan, Lowlands Bolivia, Guillaume 2024)

a. Imperfective & AM

| | |
|---|---|
| ***e-****tsiatsia-**'u*** | ‘**be** shout**ing while going**’ |
| ***e-****tsiatsia-**siu*** | ‘**be** shout**ing while coming**’ |
| ***e-****tsiatsia-**buyu*** | ‘**be** shout**ing while going back**’ |
| ***e-****tsiatsia-**beyu*** | ‘**be** shout**ing while coming back**’ |
| ***e-****tsiatsia-**niuneti*** | ‘**be** shout**ing while wandering**’ |

b. Imperfective & AP

| | |
|---|---|
| ***e-****tsiatsia-**neti*** | ‘**be** shout**ing while standing**’ |
| ***e-****tsiatsia-**sa*** | ‘**be** shout**ing while lying/bending**’ |
| ***e-****tsiatsia-**bade*** | ‘**be** shout**ing while hanging**’ |

c. Imperfective only

| | |
|---|---|
| ***e-****tsiatsia-**(a)ni*** | ‘**be** shout**ing**’ |

AM and AP expressions, akin to those typically examined under the umbrella of “lexical affixes”, pose an intriguing puzzle for traditional grammaticalization theory. They question the assumption, found in numerous works on grammaticalization, that (1) grammaticalizable concepts belong to a limited range of notional domains; (2) motion and posture are excluded from this range; (3) therefore, motion and posture are necessarily lost ("bleached") in the processes of grammaticalization.

In this talk, I will elaborate on earlier presentations of this puzzle in the case of AM by Nicolle (2002; 2007), Guillaume (2006) and Accattoli & Todaro (2017). And I will extend the discussion to include the category of AP.

**References**


Accattoli, Matilde & Giuseppina Todaro. 2017. Verbes de mouvement et grammaticalisation : le cas du sicilien vaffazzu. In Malinka Velinova (ed.), *Normes et grammaticalisation: le cas des langues romanes*, 187–210. Sofia: CU Romanistika.

Guillaume, Antoine. 2006. La catégorie du "mouvement associé" en cavineña : Apport à une typologie de l'encodage du mouvement et de la trajectoire [The category of "associated motion" in Cavineña: Contribution to a typology of motion and path encoding]. *Bulletin de la Société de Linguistique de Paris* 101(1). 415–436.

Guillaume, Antoine. 2024. Associated motion, associated posture and imperfective aspect in Tacana (Amazonian Bolivia). *Studies in Language* [advanced online].

Guillaume, Antoine & Harold Koch. 2021. Introduction: Associated motion as a grammatical category in linguistic typology. In Antoine Guillaume & Harold Koch (eds.), *Associated motion* (Empirical Approaches to Language Typology 64), 3–30. Berlin: De Gruyter Mouton.

Nicolle, Steve. 2002. The Grammaticalisation of Movement Verbs in Digo and English. *Revue de Sémantique et Pragmatique* (11). 47–67.

Nicolle, Steve. 2007. The grammaticalization of tense markers : A pragmatic reanalysis. *Cahiers Chronos* 17. 47–65.

# Identifying lexical affixes in Murrinhpatha

Rachel Nordlinger
(University of Melbourne)

Keywords: lexical affixes, polysynthesis, bipartite stem, complex verb, Australian languages

In this paper I discuss a range of lexical affixes in Murrinhpatha, a polysynthetic non-Pama-Nyungan language of northern Australia. Murrinhpatha is spoken by approximately 2500 people in the remote community of Wadeye in Australia's Northern Territory, and is one of only a small number of Australian Indigenous languages that is still being acquired by children and used as the primary language of the community. There is as yet no comprehensive grammatical description of Murrinhpatha, but it has been the subject of descriptive work by a number of researchers on various aspects of the grammatical structure (e.g. Walsh 1976, Blythe 2009, Nordlinger (2010, 2015), Mansfield 2019, Nordlinger and Mansfield 2021).

Murrinhpatha verbs have a complex templatic structure (Nordlinger 2010) by which single verbs can convey entire complex utterances (1).

(1) ***bena****-ngarru-ngkalip-kathu-dha-ngime-yu*
**3sgS.SEE:RR.PST:IMPV**-1plexclOBL-look.under.arm-HITHER-PST-PC-DM
'He was sneakily watching us by putting his arm across his eyes but looking at us from underneath.'

In addition to the many inflectional and derivational morphs in the Murrinhpatha verb, there are a number of lexical affixes in the sense of Mattissen (2017: 72) or Mithun (1997). These include adverbial affixes (expressing direction, degree, scale), such as *-kathu* in (1); and incorporated body parts which can be used to express a body part of the direct object (2), or metaphorically to modify the predicate (3).

(2) ***dirran****-ngi-**me**-thith=dim*
**3sgS.WATCH.NFUT**-1sgO-**foot**-stare
'She stared at my feet.'

(3) ***yungan****-**marda**-pirt=wurran*
**3sgS.PULL.NFUT**-**stomach**-remove
'She took something out of (the vehicle).'

Of particular interest in this paper is the fact that the verbal predicate consists of a bipartite stem (cf. Klamath, DeLancey 1999) made up of two bound stem elements: the classifier stem (bold in (1-5)) and the lexical stem (underlined in (1-5)). These stem types fall between standard definitions of bound roots and lexical affixes. Each stem contributes lexical semantics and argument structure to the overall verbal predicate, as shown in (4-5) where the classifier stem (4) and the lexical stem (5) are varied, resulting in different predicate semantics.

(4a) ***bangam***-*rtal*
**3SGS.BASH.NFUT**-chop
'He chopped it (e.g. with an axe).'

(4b) ***pan***-*rtal*
**3SGS.SLASH.NFUT**-chop
'He sliced it (e.g. with a knife).'

(5a) ***bangam***-*warnta*
**3SGS.BASH.NFUT**-split.open
'He smashed it open.'

(5b) ***bangam***-*let*
**3SGS.BASH.NFUT**-stick
'He stuck it together.'

Lexical stems constitute a large, closed class and never appear alone without a classifier stem. Classifier stems form 38 paradigms encoding subject person and number as well as tense, aspect and mood. Of these, 27 paradigms *only ever* occur in combination with a lexical stem (Nordlinger 2015, Mansfield 2019). Thus, these two stem elements overwhelmingly depend on the presence of each other in the verbal word. They do not fit standard definitions of roots, since they can't occur alone as the base of a word (e.g. Mattissen 2017: 73); yet, they don't clearly fit definitions of lexical affixes either – for example their meanings are not general or discourse secondary (e.g. Mithun 1997). In this paper I discuss in detail the characteristics and functions of these verbal morphs in Murrinhpatha and their implications for our understanding of lexical affixes and morphological typology.

**References**


Blythe, Joe. 2009. *Doing Referring in Murrinh-Patha Conversation*. PhD dissertation, University of Sydney.

DeLancey, Scott. 1999. Lexical prefixes and the bipartite stem construction in Klamath. *International Journal of American Linguistics*, 65(1): 56-83.

Mansfield, John. 2019. *Murrinhpatha morphology and phonology*. Berlin: de Gruyter Mouton.

Mattissen, Johanna. 2017. Sub-types of polysynthesis. In Michael D. Fortescue, Marianne Mithun & Nick Evans (eds.), *The Oxford handbook of polysynthesis*, 70–98. Oxford: Oxford University Press.

Mithun, Marianne. 1997. Lexical affixes and morphological typology. In Joan L. Bybee, John Haiman & Sandra A. Thompson (eds.), *Essays on language function and language type*, 357–371. Amsterdam/Philadelphia: John Benjamins.

Nordlinger, Rachel. 2010. Verbal morphology in Murrinh-Patha: evidence for templates. *Morphology* 20(2): 321-341.

Nordlinger, Rachel. 2015. Inflection in Murrinh-Patha. Matthew Baerman (ed.) *The Oxford handbook of inflection*, 491-519. Oxford: Oxford University Press.

Nordlinger, Rachel and John Mansfield. 2021. Positional dependency in Murrinhpatha: expanding the typology of non-canonical morphotactics. *Linguistics Vanguard* 7(1): 2020079. https://doi.org/10.1515/lingvan-2020-0079.

Walsh, Michael. 1976. *The Murinypata Language of North West Australia*. PhD dissertation, Australian National University.

# A typology of morphologically bound complementation (with special reference to Siberian Yupik)

Anastasia Panova
(Stockholm University)

Keywords: complementation, lexical affixes, morphosyntax, typology, Yupik

The aim of this paper is to present a typology of affixes which are functionally equivalent to matrix verbs. An example of this type of affix is shown in (1): the suffix *-čuk'* 'check' in Agul constitutes a single wordform with the head of its sentential complement, however, both parts of the wordform retain their syntactic and semantic independence (for example, both predicates have their own subjects). In Maisak's (2016: 838) description of the Agul suffix *-čuk'*, this type of construction is called *morphologically bound complementation*.

(1) Agul, East Caucasian (Maisak 2016: 826)
*gada.ji* *ruš* *quš.u-naj-**čuk'.u**-ne*
boy(ERG) [girl go_back.PF-PRF]-**VERIF.PF**-AOR
'The boy checked whether the girl has gone away.'

While in some languages the only affix involved in morphologically bound complementation is the causative (see discussions on biclausality of wordforms with causative affixes in Japanese (Shibatani 1990: 307-317), Karachay-Balkar (Ljutikova et al. 2006: 131-136)), other languages have whole sets of affixes functionally equivalent to matrix verbs. For example, Yaqui (Uto-Aztecan) has the suffixes 'cause', 'want', 'ask', 'teach', 'see', 'hear' and 'say' (Escalante 1990: 60-80).

Probably the best examples of morphologically bound complementation come from the Eskimo-Aleut languages (Mithun 2000; Sadock 2003: 53; Beach 2011: 187-196). Based on first-hand data, I will describe four productive suffixes *-st* 'cause', *-sq* 'ask', *-nəχsiʁ* 'expect' and *-niq* 'say' involved in morphologically bound complementation in Siberian Yupik and focus specifically on tests confirming their lexical status. The tests are designed to check, for example, the possibility of the independent modification of two parts of the wordform by negation and tense-aspect suffixes (2) or by temporal adverbs (3).

(2) Siberian (Chaplinski) Yupik, Eskimo-Aleut
*taɣi-na-nʁi-**n̥iq**-a*:
go-IMM_FUT-NEG-**SAY**-IND.TR+3SG.3SG
'Someone thinks he is (was) not going to come.'

(3) Siberian (Chaplinski) Yupik, Eskimo-Aleut

| | | | |
|---|---|---|---|
| *matən* | *anana-m* | *muʁus-ima-**niq**-a-tən* | *iʁivɣaq* |
| today | Anana-ERG.SG | hunt-PST-**SAY**-IND.TR-3.2SG | yesterday |

'Today Anana says you went hunting yesterday.'

Not surprisingly, wordforms with the suffixes *-st* 'cause', *-sq* 'ask', *-nəχsiʁ* 'expect' and *-niq* 'say' appear to differ with respect to the degree of event integration. The results of the tests for Siberian Yupik,

coupled with the cross-linguistic evidence, generally conform to the "binding hierarchy" (Givón 1980), which predicts the correlation between syntactic and semantic integration of clauses in standard complementation, but there are some notable exceptions.

Finally, I will consider the so-called echo phenomena in morphologically bound complementation, cf. in (4) the suffix *-sq* 'ask' in a non-finite form is doubled by the matrix verb *piːq* with similar semantics.

(4) Siberian (Chaplinski) Yupik, Eskimo-Aleut

| | | |
|---|---|---|
| *na-m* | ***piːq-aː*** | |
| mother-ERG.SG | **say/do**-IND.TR+3SG.3SG | |
| *siləqa* | *nəʁə-**sqə**-ɬu-ku* | *maŋtaɣ-məŋ* |
| S.ABS.SG | eat-**ASK**-SUB-3SG.DS | mantak-INS.SG |

'Mother asked Seljaka to eat muktuk (whale skin and blubber).'

The possibility of echo constructions is one of the parameters in the typology of morphologically bound complementation that is also relevant to associated motion marking (Pakendorf & Stoynova 2021). I will discuss similarities between these two types of constructions as well how they relate to morphologically bound serial verb constructions.

**References**


Beach, M. D. (2012). Studies in Inuktitut grammar. Ph.D. dissertation. State University of New York at Buffalo.

Escalante F. (1990). Voice and argument structure in Yaqui. Ph.D. dissertation. The University of Arizona.

Givón, T. 1980. The binding hierarchy and the typology of complements. *Studies in Language* 4:333–77.

Ljutikova, E. A., Tatevosov S. G., Ivanov M. Yu., Pazel'skaya A. G., Shluinskiy A. B. (2006). *Struktura sobytija i semantika glagola v karachaevo-balkarskom jazyke* [Event structure and verb semantics in Karachay-Balkar]. Moscow: Gorky Institute of World Literature.

Maisak, T. A. (2016). Morphological fusion without syntactic fusion: The case of the "verificative" in Agul. *Linguistics*, 54(4), 815–870.

Mithun, M. (2000). Valency-changing derivation in Central Alaskan Yup'ik. R. M.W. Dixon, A. Aikhenvald (eds.) *Changing valency, Case studies in transitivity*, 84-114. Cambridge: Cambridge University Press.

Pakendorf, B. & Stoynova, N. (2021). Associated motion in Tungusic languages: a case of mixed argument structure. In A. Guillaume & H. Koch (eds.), *Associated Motion*, 855–898. Berlin, Boston: De Gruyter Mouton.

Sadock, J. M. (2003). *A grammar of Kalaallisut*. Munich: LINCOM.

Shibatani, M. (1990). *The languages of Japan*. Cambridge: Cambridge University Press.

# Affixal verbs in Chukchi and Even: different origins affect morphosyntax

Tatiana Kazakova, Alexey Vinyar & Anna Postnikova
(HSE University, HSE University & independent researcher)

Keywords: affixal verbs, morphosyntax, diachrony, Chucotko-Kamchatkan, Tungusic

Our study deals with syntax and semantics of bound denominal verbalizers (see Gerdts & Marlett 2008) in Chukchi (< Chukotko-Kamchatkan, 17 affixes) and Even (< Tungusic, 18 affixes). Semantically, these verbalizers form a continuum from abstract ones to those expressing events commonly encoded by verbal roots (like 'hunt', 'smell' inter alia). To differentiate the latter from other types of lexical affixes (e.g. describing objects, see Mithun 1997; Matissen 2017) we label them as *affixal verbs* (AV further). In this paper we analyze AV semantics and syntax refining previous descriptions (e.g. Kurebito 2001 for Chukchi). The data was gathered during our fieldwork on Chukchi (Amguema variety, Amguema village, 2016-2023) and Even (Bystraja dialect, Esso and Anavgaj villages, 2021-2023). We used both elicitation data and text analysis (written and oral).

Semantically, there are similarities (AV for hunting, gathering and eating) as well as differences (e.g. detachment coded by AV in Chukchi and olfactory AV in Even) between two systems. Syntactically, affixal verbs differ in two systems. In Chukchi all nominal bases of AV constructions cannot be modified externally (1) but can attach incorporated modifiers (2). Thus, in this construction the free nominal stem has more grammatical status as compared to bound AV which is a single predicational element (see Keizer 2007; Boye & Harder 2012). In Even, nominal bases of some AV can receive external modification: modification is possible for AV constructions which allow discourse reference to nominal bases, compare (3) and (4) with (5) and (6). Consider also the role of affixal verbs in referential and modificational possibilities of their bases in Kalaallisut (< Eskimo-Aleut) described by Van Geenhoven 1998.

On the basis of previous comparative work (e.g. Fortescue 2005 for Chukchi; Cincius 1975, 1977 for Even) and providing our own reconstructions we argue that the syntactic differences come from the sources of AVs. Chukchi denominal verbalizers function in semantic frames established by noun incorporation and develop either from verbal heads of noun incorporation construction or from reanalysis of deverbal derivations (better understood by constructional approach (Hüning & Booij 2014). Thus, the backgrounded status of a verbalized noun is unsurprising. As for Even, incorporation did not shape AV's semantic frames and the number of AVs was extended by analogy (including apophony processes).

Appendix
Examples from Chukchi (1)-(2) and from Even (3)-(6)

(1) **nə-mejəŋ-qin* *ɣe-kupre-nŋe-ɬin*
ST-big-ST.3SG PF-net-GET-PF.3SG
Intended: 'He bought a big net'.

(2) *ɣe-mejŋə-kupre-nŋe-ɬin*
PF-big-net-GET-PF.3SG
'He bought a big net'.

(3) *edge-m* *olla-mi-ri-n*
big-ACC fish-HUNT-PST-3SG
'He caught a big fish'.

(4) *l'oša* *olla-mi-ri-n,* *ŋan* *stud'enta-l* *ere-l-bu* *žip-ti-tən*
l'osha fish-HUNT-PST-3SG and student-PL DEM.PROX-PL-ACC eat-PST-3PL
'L'osha caught a big fish and students ate it'.

(5) a. *budele-m-ne-n*
leg-ACHE-NFUT-3SG
'Her/his leg hurts'.

b. **angida-č* *budele-m-ne-n*
right-INS leg-ACHE-NFUT-3SG
Intended 'Her/his right leg hurts'.

(6) **etiken* *budele-m-ne-n* *temi* *tara-w* *budele-w* *imu-d-də-n*
old.man leg-ACHE-NFUT-3SG thus DEM.DIST-ACC leg-ACC smear-PROG-NFUT-3SG
Intended: 'Old man's leg hurts so he smears it with ointment'.

Abbreviations
3 = third person; ACC = accusative; ACHE = ache; DEM.DIST = distant demonstrative; DEM.PROX = proximal demonstrative; GET = get; HUNT = hunt; NFUT = not-future tense; PF = perfect; PL = pluralis; PROG = progressive; PST = past tense; SG = singular; ST = stative

Acknowledgments
This paper was prepared within the framework of the HSE University Basic Research Program in 2024.

**References**

Boye, Kasper & Peter Harder. 2012. *A usage-based theory of grammatical status and grammaticalization. Language* 88.1. 1–44.

Cincius, V. I. 1975, 1977. *Sravnitelnyi slovar; tunguso-manchzhurskikh yazykov. Materialy k ehtimologicheskomu slovary*, I-II. [Comparative Dictionary of the Tungusic Languages. Materials for the etymological dictionary, I-II]. Leningrad: Nauka.

Fortescue M. 2005. *Comparative Chukotko-kamchatkan dictionary.* Berlin, Boston: Walter de Gruyter. V. 23.

Hüning, M., & Booij, G. 2014. From compounding to derivation the emergence of derivational affixes through "constructionalization". *Folia linguistica*, 48(2). 579-604.

Keizer, Evelien. 2007. The lexical-grammatical dichotomy in Functional Discourse Grammar. *Alfa – Revista de Lingüística* 51.2 [Special issue ed. By Marize M.D. Hatthner & Kees Hengeveld: Advances in Functional Discourse Grammar]. 35–56.

Kurebito T. 2001. On lexical affixes in Chukchi. In O. Miyaoka, F. Endo (eds.), *Languages of the North Pacific Rim*. V. 6. Japan. 65–84.

Mattissen, Johanna. 2017. Sub-types of polysynthesis. In Michael D. Fortescue, Marianne Mithun & Nick Evans (eds.), *The Oxford handbook of polysynthesis*, 70–98. Oxford: Oxford University Press.

Mithun, Marianne. 1997. Lexical affixes and morphological typology. In Joan L. Bybee, John Haiman & Sandra A. Thompson (eds.), *Essays on language function and language type*, 357–371. Amsterdam, Philadelphia: John Benjamins.

Van Geenhoven, V. 1998. *Semantic incorporation and indefinite descriptions: semantic and syntactic aspects of noun incorporation in West Greenlandic.* Stanford: CSLI Publications.

# Where are all the lexical affixes?

Adam J. R. Tallman & Anthony C. Woodbury
(Friedrich Schiller Universität & University of Texasn)

Keywords: Chácabo; Yupik-Inuit; lexical affixes; polysynthesis; discourse prominence

We address the Lexical Affixes Workshop theme by discussing cases in two languages of the Americas: Chácobo (Pano) and Central Alaskan Yupik (CAY; Yupik-Inuit-Unangan). Both languages have been described as polysynthetic (Payne 1991; Loos 1999; Boas 1911: 74; Sapir 1911: 254; Miyaoka 2012: 18-20; Fortescue 2017), and as having robust lexical affixes (Fleck 2006 for Pano; Mithun 1996 for CAY, cf. de Reuse 2009, Fortescue 2015 for related languages). Following Boye & Harder 2012 (B&H) and Boye 2023, even relatively "lexical" affixes should still be discursively SECONDARY, that is, they must always be backgrounded, should not be addressable, should not be focused, and should not be modified. We show that in both languages, there are markers that are clearly affixes but fail to be discursively secondary in one or more of the posited senses.

In Chácabo, associated motion markers are conventionally backgrounded, but they can be modified by markers of speed as in *tápi* in (1). The referents of body-part prefixes can be addressable in later discourse as in (2) with the relationship between *ta-* 'foot' and *mɨkɨ́* 'hand, hoof, paw'.

(1) *tsaya-**ʙaya**-**tápi**=ki*
see-**do&go**-**FAST**=DECL:PST
'He'll look at it (for some time) and then **leave quickly**.' PTCP OBSV

(2)

| | | | | | |
|---|---|---|---|---|---|
| *hawɨ́* | *wíko* | *pɨ* | ***tá**-nɨ̨ş=kɨ* | ... | *hɨnɨ=́* |
| 3SG:GEN | lower.leg | ANX | **foot/hoof**-tie-DECL:PST | ... | water=SPAT |
| *ak=ái=ka* | ***mɨkɨ́*** | *tsi* | *atʃ-a=ş́ona* | | |
| take=NMLZ=REL | **hand/hoof** | LNK | grab=PRIOR:SA | | |

'It (the anaconda) tied itself around the lower leg and hoof (of the cow), taking it into the water after it grabbed its hoof.' ELAR 624:27

In Cup'ik CAY, candidate lexical affixes can carry modifiable, focusable, addressable content as in a conversational snippet in (3), showing all three of these properties. There, several such markers suffix to a lexical noun base meaning 'axe'. In (3)(a) Speaker 1 discusses the lack of an axe at a particular location and is told in (3)(b) by Speaker 2 that he did *'have a crummy little* axe' (focal information modifying and predicating over 'axe'); and then adds in (3)(c) that another person had brought a '*huge* axe' (more focal information addressing the previous focal information 'crummy little').

(3)(a) Speaker 1:

| | | | |
|---|---|---|---|
| *Qalqapa-i-pagg!* | *Qalqapag-taite-llini-ami=ll'* | *kiw-na,* | *...* |
| Axe-lack-what_V! | Axe-lack_there-EVID-When.3sgS=& | inside-ABS.sg | |

'What lack of an axe! And when it appeared there was no axe (in) that (place) inside…'

(3)(b) Speaker 2: (interrupting)
*Qalqapa-**cua-llera-ngqerr-saaq-ua!***
axe-**little-crummy-have-alas!-IND.3sg**
'**I do have a crummy little** axe (but…)'

(3)(c) Speaker 2: (continuing)

| | | | |
|---|---|---|---|
| *<Name>-m* | *im'-um* | *taite-llru-yaaq-aa* | *qalqapa-**paarrluk!*** |
| <name>-ERGsg | that-ERGsg | bring-PAST-alas!-3sg>3sg | axe-**huge.ABS.sg** |

But that guy <name>, he actually brought a huge axe! (but alas, we'll learn it had no handle!) (*Chevak Qaygiq*, 1978-11-26, C-T10A)

Based on a review of the discursive properties of candidate lexical affixes we conclude that these entities are not strictly grammatical in the sense of B&H and Boye (2023) despite their clear status as affixes. We suggest that a better approach does not posit a discrete distinction between grammatical and lexical forms, instead cataloging the properties of linguistic forms in the whole context of the lexico-grammatical systems in which they occur.

**References**


Auderset, S., & al., e. (in press). Constituency and convergence in the Americas - results and discussion. In *Constituency and convergence in the Americas.* Berlin: Language Science Press. Available from https://sites.google.com/view/constituencytypology/papers?authuser=2

Boas, F. (1911). Introduction. En *Handbook of American Indian Languages* (Vols. Bulletin 40, Part 1, pp. 1- 83). Washington D.C.: Bureau of American Ethnology.

Boye, K. (2023). Grammaticalization as conventionalization of discourse secondary status: Deconstructing the lexical-grammatical continuum. *Transactions of the Philological Society, 121*(2), 270-293.

Boye, K., & Harder, P. (2012). A usage-based theory of grammatical status and grammaticalization. *Language, 88*(1), 1-44.

de Reuse, W. J. (2009). Polysynthesis as a Typological Feature: An Attempt at a characterization from Eskimo and Athabaskan perspectives. In M.-A. Mathieu, & N. Tersis (Edits.), *Variations in polysynthesis: the Eskaleut languages (Typological studies in language 86)* (pp. 19-34). Philadelphia: John Benjamins.

Fleck, D. W. (2006). Body-Part Prefies in Matses: Derivation or Noun Incorporation? *International Journal of American Linguistics, 72*(1), 59-96.

Fortescue, M. (2015). The sources of Eskimo-Aleut verbalizing affixes. *unpublished manuscript*. Fortescue, M. D. (2009). Lexical Suffixes and the Position of Proto-Wakashan within Northwest Coast

Linguistic Area. *The Canadian Journal of Linguistics / La revue canadienne de linguistique, 54*(1), 37-41.

Loos, E. E. (1999). Pano. En D. R. W, & A. Aikhenvald (Edits.), *The Amazonian Languages* (págs. 227-251).

Cambridge: Cambridge University Press.

Mithun, M. (1997). Lexical affixes and morphological typology. In J. L. Bybee, J. Haiman, & S. A. Thompson (Edits.), *Essays on language function and language type* (pp. 357-371). Amsterdam: John Benjamins.

Miyaoka, O. (2012). *A Grammar of Central Alaskan Yupik (CAY).* Berlin: Walter de Gruyter.

Payne, D. L. (1990). Morphological Characteristics of Lowland South American Languages. In *Amazonian Linguistics: Studies in Lowland South American Languages* (págs. 213-243). Austin: University of Texas Press.

Sapir, E. (1911). The problem of noun incorporation in American languages. *American Anthropologist, 13*, 250-282. Reprinted in The Collected Works of Edward Sapir, ed. W. Bright, Vol. V, pp. 27–59. Berlin: Mouton de Gruyter, 1990.

Sapir, E. (1921). *Language.* New York: Harcourt, Brace & World.

Sapir, E., & Swadesh, M. (1939). *Nootka texts: Tales and ethnological narrative with grammatical notes and lexical materials.* Philadelphia: Linguistic Society of America.

Spencer, A., & Luís, A. (2012). The canonical clitic. En D. Brown, M. Chumakina, & G. G. Corbett (Edits.), *Canonical morphology and syntax* (pp. 123-150). Oxford: Oxford University Press.

Tallman, A. J. (2021a). Constituency and coincidence in Chácobo (Pano). *Studies in Language, 45*(2), 321-382.
Available at https://sites.google.com/view/constituencytypology/papers?authuser=2

Woodbury, A. C. (in press). Constituency in Cup'ik and the problem of holophrasis (Chapter 2). In A. J. Tallman, S. Auderset, & H. Uchihara (Edits.), *Constituency and convergence in the Americas.* Berlin: Language Science Press. Available at https://sites.google.com/view/constituencytypology/papers?authuser=2

# How to identify a lexical affix

Mads Nielsen, Naja Blytmann Trondhjem & Kasper Boye
(University of Copenhagen, University of Copenhagen & University of Copenhagen)

Keywords: affix, lexical, grammatical, polysynthesis, Kalaallisut (West Greenlandic)

There is little agreement on what lexical affixes are (see e.g., Mithun 1997, Gerdts 2003, Mattissen 2017). Some definitions are specific to one language or a set of related languages and not useful in language comparison. Others take Standard Average European languages as a golden standard to determine lexical status.

In this paper, we first propose a cross-linguistic definition of lexical affixes and illustrate how the definition can be implemented empirically to identify lexical affixes in Kalaallisut. We discuss implications of a theoretically anchored definition of lexical affixes for the distinction between phrases and words, and for the classification of polysynthesis.

Our definition takes its point of departure in the fact that identification of a lexical affix requires two things: identification of an affix and identification of lexical status. Regarding identification of affixes, we follow Haspelmath's (2023) definition according to which "[a]n affix is a bound morph that is not a root, that must occur on a root, and that cannot occur on roots of different root classes" (p. 286), where a root is defined as "a contentful morph (i.e., a morph denoting an action, an object or a property) that can occur as part of a free form without another contentful morph" (p. 287). To our knowledge this definition is the only cross-linguistically applicable definition of affixes.

As for lexical status, we follow Boye and Harder's (2012) definition of lexical elements as potentially discourse primary. We argue, in accordance with Keizer (2007) and Boye (2023), that lexical status can be identified on the basis of modifiability. The idea is that because lexical elements, unlike grammatical elements, can attract attention, they can be modified.

Thus, we suggest that lexical affixes are defined as bound, non-root, root-class dependent morphemes that are potentially discourse primary. In the empirical part of our paper, we identify lexical affixes in Kalaallisut guided by native speaker evaluations, and discuss the distribution of the affixes and the extent to which the distribution reflects general associations between grammaticalization, scope increase and position (cf. Hengeveld 2017). A modifiable and hence lexical Kalaallisut affix is presented in (1):

(1) *anger-pallag-ngaar-ami*
come.home-quickly-very-4SG.CAUS
'because they came home very quickly'

We regard *-pallag* 'quickly' as lexical since its meaning is modified by *-ngaar* 'very'.

The existence of a theoretically anchored notion of lexical affixes challenges the prevailing assumption that words and phrases are different. Traditionally, a word is thought of as consisting of a head surrounded by grammatical material (affixes), and a phrase is thought of as consisting of a head surrounded by lexical material (words). But if some affixes are lexical (and if some words are grammatical), the distinction does not hold.

Besides being crosslinguistically applicable, the proposed definition has the advantage that attestation of lexical affixes becomes an empirical question, in fact one that can be addressed psycholinguistically. To the extent that lexical affixes are a characteristic of polysynthetic languages

(cf. Mattissen 2017), the definition also has the potential to diagnose polysynthesis in a transparent manner.

**References**


Boye, Kasper and Peter Harder (2012), A usage-based theory of grammatical status and grammaticalization, *Language* 88(1), 1-44.

Boye, Kasper (2023), Grammaticalization as conventionalization of discursively secondary status: Deconstructing the lexical-grammatical continuum, *Transactions of the Philological Society* 121(2), 270-292.

Gerdts, Donna (2003), The morphosyntax of Halkomelem lexical suffixes, *International Journal of American Linguistics* 69(4), 345-356.

Haspelmath, Martin (2023), Defining the word, *WORD* 69(3), 283-297.

Hengeveld, Kees (2017), A hierarchical approach to grammaticalization, in K. Hengeveld, H. Narrog and H. Olbertz (eds), *The grammaticalization of tense, aspect, modality and evidentiality*, Berlin / Boston: de Gruyter, 13-37.

Keizer, Evelien (2007), The grammatical-lexical dichotomy in Functional Discourse Grammar, *Alfa – Revista de Lingüística* 51(2), 35–56.

Mattissen, Johanna (2017), Sub-types of polysynthesis, in M. Fortescue, M. Mithun and N. Evans (eds), *The Oxford handbook of polysynthesis*, Oxford: Oxford University Press, 70-98.

Mithun, Marianne (1997), Lexical affixes and morphological typology, in J. Bybee, J. Haiman and S. A. Thompson (eds), *Essays on language function and language type*, Amsterdam: John Benjamins, 357-371.

# (No) lexical affixation in Functional Discourse Grammar

Felix Berner & Evelien Keizer
(University of Vienna & University of Vienna)

Keywords: Functional Discourse Grammar, lexical/grammatical, affixation, stems/roots, polysynthesis

Lexical affixation is often regarded as a necessary feature of polysynthesis (e.g. Fortescue 2016, 2017, Mattissen 2017, Drossdard 2002, De Reuse 2009). From the perspective of Functional Discourse Grammar (FDG; Hengeveld & Mackenzie 2008), however, the notion of lexical affixes is problematic; currently, the theory does not recognize such a category. This is largely due to how FDG defines the notions "affix" and "lexical".

As for affixation, FDG differentiates between stems (free lexical elements), roots (bound lexical elements) and affixes (bound grammatical elements) (Hengeveld & Mackenzie 2008: 404, Giomi 2023: 358-369). Additionally, FDG differentiates between two types of derivation: syntactic derivation (where meaning is predictable, and which takes place in the Grammar, e.g. *buyer*) and lexical derivation (which involves unpredictable meaning elements and takes place in the Lexicon, e.g. *fineliner*) (e.g. Hengeveld & Mackenzie 2008: 217, 229, Keizer 2018). In FDG, the dependent elements involved in both these processes are regarded as grammatical elements.

As for the notion "lexical", FDG makes a distinction between lexical elements (Heads and Modifiers corresponding to Lexemes), grammatical elements (expressing Operators and Functions), and semi-lexical elements (expressing Lexical Operators/Functions). The main criteria employed to distinguish between these three categories are (i) focalizability and (ii) the possibility of an element to have its own modifiers and operators (cf. Keizer 2007, Hengeveld 2017, Giomi 2023).

Therefore, in FDG the notion of lexical affix is a *contradictio in terminis*. If lexical affixes are lexical elements that attach to other lexical elements, they would not be affixes, but stems (free morphemes participating in a process of compounding), roots (dependent lexical element that can only occur in conjunction with another lexical element), or derivational affixes (grammatical elements participating in word formation processes in the Lexicon). Thus, what are often regarded as lexical affixes in polysynthetic languages would in FDG be regarded as roots taking part in processes of incorporation or serialization (see also Genee 2018).

Finally, it will be clear that the criteria used in FDG (and other theoretical approaches; e.g. Boye & Harder 2012, 2020, Boye 2023) to distinguish between lexical and grammatical elements cannot be applied to affixes: if an element is focalizable and modifiable, it would be a (fully) lexical element and not an affix. This raises a number of additional questions, such as (i) How can we determine the lexical status of a bound morpheme (i.e. how do we distinguish roots from affixes)?; and (ii) even if we are able to find such criteria (see e.g. Giomi 2023: 358-369), is it still plausible to maintain that these indicate a difference between lexical and grammatical status? .

This paper presents the FDG perspective on the distinction between lexical and grammatical elements in general, and between lexical and grammatical bound morphemes (roots and affixes) in particular, and discusses the problems involved in the operationalization of the established criteria for lexicality when applied to bound morphemes.

**References**


Boye, Kasper and Peter Harder (2012), A usage-based theory of grammatical status and grammaticalization, L*anguage* 88(1), 1-44.

Boye, Kasper and Peter Harder (2020), Dual processing in a functional-cognitive theory of grammar and its neurocognitive basis, in A. Haselow and G. Kaltenböck (eds), *Grammar and Cognition: Dualistic Models of Language Structure and Language Processing*, Amsterdam: Benjamins, 133-155.

Boye, Kasper (2023), Grammaticalization as conventionalization of discursively secondary status: deconstructing the lexical-grammatical continuum, *Transactions of the Philological Society* 12(2), 270-292.

Drossard, Werner (2002), Ket as a polysynthetic language, with special reference to complex verbs, in N. Evans and H.-J. Sasse (Eds.), *Problems of Polysynthesis*, Berlin: Akademie Verlag, 223-256.

De Reuse, Willem. J. (2009), Polysynthesis as a typological feature: An attempt at a characterization from Eskimo and Athabaskan perspectives, in M.-A. Mahieu and N. Tersis (Eds.), *Typological Studies in Language* (Vol. 86), Amsterdam: John Benjamins, 19-34.

Fortescue, Michael (2016), Polysynthesis: A Diachronic and Typological Perspective. In *Oxford Research Encyclopedia of Linguistics*, Oxford: Oxford University Press.

Fortescue, Michael (2017), What are the Limits of Polysynthesis? In M. Fortescue, M. Mithun and N. Evans (Eds.), *The Oxford Handbook of Polysynthesis*, Oxford: Oxford University Press, 115-134.

Genee, Inge (2018), Measuring polysynthesis: A Functional Discourse Grammar approach, in E. Keizer and H. Olbertz (Eds.), *Studies in Language Companion Series* (Vol. 205), Amsterdam: John Benjamins, 234-272.

Giomi, Riccardo (2023), *A Functional Discourse Grammar Theory of Grammaticalization: Volume 1: Functional Change*, Leiden: Brill.

Hengeveld, Kees (2017), A hierarchical approach to grammaticalization, in K. Hengeveld, H. Narrog and H. Olbertz (Eds.), *The Grammaticalization of Tense, Aspect, Modality and Evidentiality*, Berlin De Gruyter, 13-38.

Hengeveld, Kees and J. Lachlan Mackenzie, (2008), *Functional Discourse Grammar*. Oxford: Oxford University Press.

Keizer, Evelien (2007), The lexical-grammatical dichotomy in Functional Discourse Grammar. *Alfa - Revista de Lingüistica*, *51*(2), 35-56.

Keizer, Evelien (2018), Derivation in Functional Discourse Grammar: Some challenges and implications, *Word Structure* 11(1), 36-66.

Mattissen, Johanna (2017), Sub-Types of Polysynthesis, in M. Fortescue, M. Mithun, & N. Evans (Eds.), *The Oxford Handbook of Polysynthesis* (1st ed.). Oxford: Oxford University Press. 70-98.

# Change detection of lexical and grammatical affixes in West Greenlandic (Kalaallisut)

Solvej W. Kjær, Malu A. Hendriksen, Naja B. Trondhjem, Johanne S.K. Nedergaard, Mads Nielsen & Kasper Boye
(University of Copenhagen, University of Copenhagen, University of Copenhagen, University of Copenhagen, University of Copenhagen & University of Copenhagen)

Keywords: typology, polysynthesis, change detection, psycholinguistics, lexical affixes

According to Boye and Harder (2012) and Boye (2023), the distinction between lexical and grammatical elements has to do with discourse prominence: while lexical elements have the potential to convey discourse primary (i.e. foreground) information, grammatical elements are by convention discourse secondary (i.e. background). Psycholinguistic studies have shown that classifications of elements as lexical or grammatical based on the criteria in Boye (2023) are reflected in the prioritization of processing resources (Boye & Bastiaanse, 2018; Christensen et al., 2021; Lange et al., 2017, 2018). For example, Christensen et al. (2021) found that when people read near-identical sentences, they notice changes to lexical words more frequently than changes to grammatical words.

Kalaallisut (West Greenlandic) offers an interesting testing ground for the difference in attention allocated to lexical versus grammatical items. The language is polysynthetic, meaning that the morphology is unusually complex with extensive systems for both inflectional and derivational affixes. The derivational affixes are especially relevant in this context since it has been claimed that some of them function as lexical items (Boye et al., 2023).

In the present study, we employ a change blindness paradigm, as in Christensen et al. (2021), with self-paced reading in order to test this claim, that is, whether Kalaallisut affixes classified as grammatical based on Boye (2023) (e.g., *-nikuu* 'perfective') are more prone to change blindness than affixes classified as lexical (e.g., *-niar* 'intend to'). In the experiment, participants are shown two presentations of a target sentence. In the second version of the target sentence, the target affix may be either removed or substituted with another affix of its own category (grammatical or lexical). For example:

> Ane aperi**nikuu**voq kammalaatini atuarfimmiit illoqarfiliaqatigisinnaanerlugit.
> 'Ane **has** asked whether she may go out with her friends from school'
>
> Ane aperi**umaar**poq kammalaatini atuarfimmiit illoqarfiliaqatigisinnaanerlugit.
> 'Ane **will** ask whether she may go out with her friends from school'

Our classification of Kalaallisut affixes as lexical or grammatical is specifically based on the modification criterion (see also Keizer, 2007), which states that lexical elements can be modified while grammatical elements cannot (except in corrective contexts). Other diagnostic criteria, such as focus constructions and addressability tests, cannot be used in Kalaallisut as the affixes occur word-internally and there is no way of focusing bound morphemes individually. As an example of using the modification criterion, the affix *-niar* can be modified by means of the affix *-suaq* 'big/very':

*aalisar-niar-suar-poq*
fish-intend-very-DECL.3SG
'He very much intends to fish'

while the affix *-nikuu* cannot be modified:

*aalisar-nikuu-suar-poq*
fish-PERF-very-DECL.3SG
*'He very much has fished'

By using psycholinguistic methods, our study complements existing methods focusing on examples from language use and tests the claim that Kalaallisut has lexical affixes.

**References**


Boye, Kasper (2023), Grammaticalization as Conventionalization of Discursively Secondary Status: Deconstructing the Lexical–Grammatical Continuum, *Transactions of the Philological Society* 121(2), 270–292. https://doi.org/10.1111/1467-968X.12265

Boye, Kasper, & Bastiaanse, Roelien (2018), Grammatical versus lexical words in theory and aphasia: Integrating linguistics and neurolinguistics, *Glossa: A Journal of General Linguistics* 3(1), Article 1. https://doi.org/10.5334/gjgl.436

Boye, Kasper, Bastiaanse, Roelien, Harder, Peter, & Martínez-Ferreiro, Silvia (2023), Agrammatism in a usage-based theory of grammatical status: Impaired combinatorics, compensatory prioritization, or both? *Journal of Neurolinguistics* 65, 101108. https://doi.org/10.1016/j.jneuroling.2022.101108

Boye, Kasper, & Harder, Peter (2012), A Usage-Based Theory of Grammatical Status and Grammaticalization. *Language* 88(1), 1–44.

Christensen, Marie Herget, Kristensen, Line Burholt, Vinther, Nicoline Munck, & Boye, Kasper (2021), Grammar is background in sentence processing. *Language and Cognition* 13(1), 128–153. https://doi.org/10.1017/langcog.2020.30

Keizer, Evelien (2007), The lexical-grammatical dichotomy in Functional Discourse Grammar. *ALFA: Revista de Linguística* 51(2), Article 2. https://periodicos.fclar.unesp.br/alfa/article/view/1436

Lange, Violaine Michel, Messerschmidt, Maria, & Boye, Kasper (2018), Contrasting Grammatical and Lexical Determiners, *Journal of Psycholinguistic Research* 47(3), 467–482. https://doi.org/10.1007/s10936-017-9531-x

Lange, Violaine Michel, Messerschmidt, Maria, Harder, Peter, Siebner, Hartwig Roman, & Boye, Kasper (2017), Planning and production of grammatical and lexical verbs in multi-word messages, *PLOS ONE* 12(11), e0186685. https://doi.org/10.1371/journal.pone.0186685

# WS8 Linguistic typology and the cognitive science of non-WEIRD languages

# Artificial language learning across populations

Jennifer Culbertson
(University of Edinburgh)

A fundamental question in the cognitive science of language is how the human mind, together with the communicative and social functions of language, has shaped language typology. There are a number of methods for exploring this link, including traditional typological, historical and theoretical analysis of languages, as well as newer, more sophisticated quantitative methods for analysis of language data. However, these methods cannot provide evidence for a causal link between a specific aspect of human behavior and language structure. This kind of evidence instead has come from experiments, and in particular artificial language learning (ALL) experiments now play a growing role in the field. ALL experiments were first used to uncover the mechanisms of language acquisition (e.g., Reber, 1967 Reber 1967, Saffran et al., 1996), but over the past 15 years, ALL experiments have also produced results suggesting that *biases* active during language learning and use align with well-known typological patterns. For example, experiments investigating perception of non-word sequences in English speakers suggest a bias favoring suffixing over prefixing (e.g., Hupp et al., 2009; St. Clair et al., 2009), parallel to the typological tendency for suffixes. Similarly, experiments on the ordering of nouns and nominal modifiers like adjectives and demonstratives has revealed that English speakers have a preference for placing the adjective closest to the noun (e.g., Culbertson & Adger, 2014; Martin et al., 2020), the pattern found most commonly across the world's languages. However, the goal of this kind of research is to identify cognitive or psychological *universals*---biases or preferences that are shared by all humans, either independently of, or prior to their experience with a particular language or culture. It is only such biases that offer a potential explanation for typology. Nevertheless, as with the vast majority of work in experimental psychology and cognitive science, participant populations in this area tend to be WEIRD (Henrich et al., 2010). In this talk, I survey recent results that aim to better achieve this goal by expanding the populations of participants tested in artificial language learning experiments. In some cases, this means targeted populations whose language experience actively goes against a typological trend--for example, speakers of a predominantly prefix language. This provides a particularly stringent test of the universality of the hypothesized perceptual bias. In other cases, this means simply expanding our sample to cover different kinds of populations, whose individual language experiences might impact their behavior in different ways. This allows us to explore how experience with different languages or language types affects behavior in a particular task. The experiments I discuss all involve both WEIRD and non-WEIRD speakers, in line with goals of the workshop, and in line with prior work underscoring the need for diversity when studying universality in cognitive science (Blasi et al., 2022; Henrich et al., 2010; Majid, 2023).

**References**

Blasi, D. E., Henrich, J., Adamou, E., Kemmerer, D., & Majid, A. (2022). Over-reliance on English hinders cognitive science. *Trends in Cognitive Sciences*.

Culbertson, J., & Adger, D. (2014). Language learners privilege structured meaning over surface frequency. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(16), 5842–5847. https://doi.org/doi:10.1073/pnas.1320525111

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83.

Hupp, J. M., Sloutsky, V. M., & Culicover, P. W. (2009). Evidence for a domain-general mechanism underlying the suffixation preference in language. *Language and Cognitive Processes*, *24*(6), 876–909.

Majid, A. (2023). *Establishing psychological universals | Nature Reviews Psychology*. https://www.nature.com/articles/s44159-023-00169-w

Martin, A., Holtz, A., Abels, K., Adger, D., & Culbertson, J. (2020). Experimental evidence for the influence of structure and meaning on linear order in the noun phrase. *Glossa-an International Journal of Linguistics*, *5*(1), 1–21. https://doi.org/10.5334/gjgl.1085

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *6*(6), 855–863.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science (New York, N.Y.)*, *274*(5294), 1926–1928.

St. Clair, M. C., Monaghan, P., & Ramscar, M. (2009). Relationships between language structure and language learning: The suffixing preference and grammatical categorization. *Cognitive Science*, *33*, 1317–1329.

# Phonemes as cognitive tools

Caleb Everett, Axel Ekström, Claudio Tennie and Steven Moran
(University of Miami, KTH Royal Institute of Technology, University of Tübingen & University of Neuchâtel)

While the number and characteristics of vowels and consonants are highly variable across the world's spoken languages, all make consistent and deliberate use of a relatively narrow set of contrastive phonemes. Here, we advance the view that phonemes (and systems of phonemes) work as cognitive tools, i.e., shared systems of classified-and-organized cultural inventions that facilitate, support, guide, and extend speaker cognitive capacities. We make several claims toward this point and offer support of these claims primarily from data in non-WEIRD languages. These include data from the PHOIBLE database, which details the phoneme inventories of 2,186 languages worldwide (Moran & McCloy 2019). We underline how speech sounds –individually and as a system–can be conceptualized as extensions of cognitive-communicative capacities, while firmly anchored in physical-acoustic properties of speech sounds themselves, as well as biological constraints on speech and hearing.

While "vowel-like" phenomena are widespread in animal communication - from the *meows* of domestic cats to the *hoots* of chimpanzees - humans make unique use of the tongue in vocal communication. Compared to the vocal communicative systems of nonhuman animals, where the predominant mode of altering such vowel-like sounds is through changing the position of the jaw and lips, human populations use tongue-modulated vowel sounds that enable rapid information transmission rates through syllabic speech. This communicative shift in human history unlocked a universal combinatorial potential (Studdert-Kennedy, 2005; Lieberman, 2012), exploitable by developing human ancestors and, once achieved, became subject to cultural diversification.

In the ontogenetic view, native-language input provides developing human infants with acoustic-perceptual goals. These acoustic-perceptual goals are non-randomly distributed within the phoneme space, which enables their more effective use as cognitive tools. For example, languages adhere to probabilistic implications in the structuring of their vowel systems: Some vowel sounds that are easy to produce but make for less effective acoustic-perceptual goals are wielded only once primary candidates (e.g., point vowels) have been exhausted, with relatively few exceptions cross-linguistically. We observe that the vast majority of languages in PHOIBLE rely on /i/ (92% of inventories) and /u/ (88% of inventories), but only languages with large phoneme and vowel inventories rely on the rounded counterpart of /i/, /y/ (6% of inventories), and the unrounded counterpart of /u/, /ɯ/, (6% of inventories) (Figure 1). Thus, perceptual and articulatory contrast between speech sounds serve double purposes, both facilitating the maintenance of systems of speech (e.g., Liljencrants & Lindblom, 1972; Stevens, 1989) and the learning thereof by "new" speakers. This facilitation of sound-acquisition in turn accelerates the acquisition of semantic concepts, suggesting phonemes are used as cognitive tools to scaffold cognition more broadly.

Our view complements cognitive linguistic perspectives on human perception and consciousness by emphasizing the use of sounds as cognitive tools, across the world's languages.

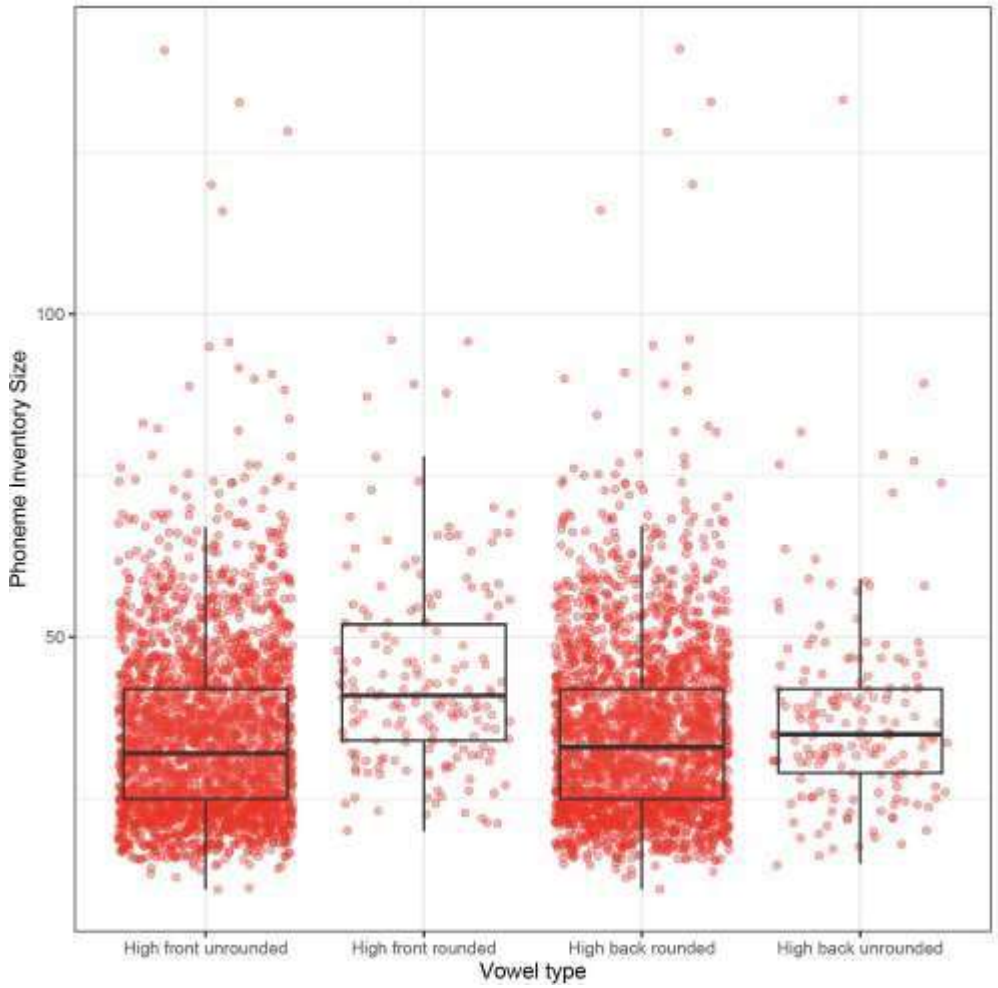



**Figure 1**. High front rounded vowels tend to occur in languages with larger phoneme inventories, when compared to those with high front unrounded vowels. The converse pattern holds for high back vowels, though more weakly. This figure is based on 2,186 languages in PHOIBLE. (Moran and McCloy 2019) Within each box, languages are separated along the x-axis via a randomizing function in R.

**References**

Lieberman, P. (2012). Vocal tract anatomy and the neural bases of talking. *Journal of Phonetics*, *40*(4), 608-622.

Liljencrants, J., & Lindblom, B. (1972). Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language*, 839-862.

Moran, Steven & McCloy, Daniel (eds.) 2019. PHOIBLE 2.0. Jena: Max Planck Institute for the Science of Human History.

Stevens, K. N. (1989). On the quantal nature of speech. *Journal of phonetics*, *17*(1), 3-45. Studdert-Kennedy, M. (2005). How did language go discrete. In M. Tallerman (Ed.), *Language Origins: Perspectives on Evolution*, 48-67. Oxford University Press.

# Maximum diversity sampling reveals universal patterns in language acquisition

Sabine Stoll
(University of Zurich)

Human language is the only communication system acquired to an overwhelming extent through learning, and it is the only communication system that comes in so many extremely diverse variants. Thus, the task for the learning infant is extraordinary but infants acquire any of these systems in the first years of their life, if they are provided with the appropriate input. This ability to learn such diverse systems requires an enormous flexibility and plasticity in learning. To understand this plasticity and flexibility, we need to put comparative crosslinguistic language acquisition research center-stage. But the existence of over 7,000 languages creates a dilemma. Which languages should we study to make universal claims? Randomly selecting individual languages poses the risk of introducing bias in our results, particularly considering the extensive variation inherent in the languages of the world. Further, bias might come from specifics of the randomly picked languages because evidence from acquisition research suggests that language structure can influence learning both with respect to when a grammatical feature is learned and how it is learned. Here we argue that strategic sampling is key for making claims about language acquisition on the species level. We present a "Maximum Diversity" approach which allows us to effectively simulate the global linguistic diversity that children encounter. In the Maximum Diversity approach languages are clustered into structurally maximally diverse groups based on grammatical variables (morphological and syntactic) that are both important for language typology and understanding language acquisition. In other words, this approach is built on the insight that if we find the same strategies and processes in maximally diverse languages, we then can make robust generalizations about universal acquisition mechanisms. This then supports a comprehensive understanding of the acquisition process. The approach has been implemented into a database of longitudinal acquisition studies and it has been effectively applied to various questions of language acquisition. Here I present evidence for some acquisition and input universals, in striking contrast to the lack of evidence for categorical universals in linguistic structure. Two case studies showcase the maximum diversity approach. Case study 1 presents results on the acquisition of grammatical case representing the acquisition of hierarchical structure. Despite the extreme linguistic variation in case marking observed in our language sample we find comparable rates of acquisition. This suggests similar universal acquisition strategies in the acquisition of inflectional morphology, here exemplified with case marking. Case Study 2 presents results on the impact of turn-taking for the acquisition process. We know that turn-taking is potentially more important than the amount of child-directed speech a child receives but we know little about how utterances in turns are structured. Here we compare contingent and non-contingent child-directed speech with respect to grammatical complexity. We find that child-directed speech which is contingent to a child's utterance to be structurally simpler than non-contingent child-directed speech. This supports universal input strategies which help explaining what features of child-directed speech are universal and potentially most relevant for acquisition.

# Is it a man or a flower? Depends on how WEIRD your language is

David Gil
(Max Planck Institute for Evolutionary Anthropology)

Keywords: WEIRD, categorization, schematological hybrid, socio-political complexity, syntactic complexity

This paper adds to an increasing body of empirical evidence pointing to systematic differences between WEIRD and non-WEIRD societies with respect to cognitive processes and the grammatical structures in which they are encoded.

This paper is concerned with a specific phenomenological domain, *schematological hybrids*, in which two different entities are combined into into a single composite entity, as in the following example:

man-flower

A series of studies investigates the ways in which we conceptualize hybrids, focusing on the ways in which they are categorized (Gil and Shen 2019, Mashal et al. 2014, Shen and Gil 2017, 2023). For example, the above image may be conceptualized as a man with his head replaced by a flower, or, alternatively, as a flower whose stem is that of a man. The choice between such alternative categorizations is governed by the interplay of three principles, making reference to the overall schematological gestalt of the image and the relative positions of the two parent entities on schematological and ontological hierarchies.

The conceptualization of hybrids was examined by means of a *Forced-Choice Description Experiment*. Participants were presented with 32 images, such as the one above. Beneath each image, two descriptions were provided, representing its two alternative conceptualizations, and participants were asked to choose the description that they preferred. For example, for the above image, participants had a choice between the following two descriptions:

(1) (a) A man whose head is a flower
(b) A flower whose lower part is that of a man

In the case at hand, all three principles predict that participants will prefer the description in (1a) to that in (1b).

The experiment was run on 8 geographically and sociolinguistically diverse languages, associated with varying degrees of WEIRDness: English, Hebrew, Standard Arabic, Standard Indonesian, Rural Palestinian Arabic, Jakarta Indonesian, Minangkabau, and Yali. The main finding to emerge from the experiment is that the effect of the three principles governing the conceptualization of hybrids is more prominent in languages of WEIRD societies than in those of their non-WEIRD counterparts. For example, for the above image, while in WEIRD English, 89% of participants preferred the description in (1a), in less WEIRD Jakarta Indonesian, a significantly lower 69% of participants preferred the corresponding description in their language.

The differential results of the experiment are argued to represent a particular case of the positive correlation between socio-political and grammatical, and, more specfiically, syntactic complexity (Chen

et al 2024). While the three cognitive principles governing the conceptualization of hybrids appear to be universal, their manifestation in individual languages is mediated by the different grammatical structures of the individual languages. More specifically, languages spoken in communities of lower socio-political complexity, which are typically non-WEIRD, are associated with flatter and less hierarchic syntactic structures, in which the differential grammatical coding of the hybrid's two parents is less pronounced. Thus, for example, while in (1a), 'man' clearly occupies the privileged head position of a relative clause construction, in its Jakarta Indonesian counterpart, the grammatical relationship between 'man' and 'flower' is more symmetric.

**References:**


Chen, Sihan, David Gil, Sergey Gaponov, Jana Reifegeste, Tessa Yuditha, Tatiana Tatarinova, Ljiljana Progovac and Antonio Benitez-Burraco (2024) "Linguistic correlates of societal variation: A quantitative analysis", *PLOS One* 19(4).

Gil, David and Yeshayahu Shen (2019) "How Grammar Introduces Asymmetry into Cognitive Structures: Compositional Semantics, Metaphors and Schematological Hybrids", *Frontiers in Psychology - Language Sciences*.

Mashal, Nira, Yeshayahu Shen, Karine Jospe and David Gil (2014) "Language Effects on the Conceptualization of Hybrids", *Language and Cognition*, 217-241.

Shen, Yeshayahu and David Gil (2017) "How Language Influences the Way We Categorize Hybrids", in H. Cohen and C. Lefebvre eds, *Handbook of Categorization in Cognitive Science*, Second Edition, Elsevier, Amsterdam, 1177-1200.

Shen, Yeshayahu and David Gil (2023) "How to Build a Hybrid, The Structure of Imagination", in M. Gal, D. Gil and Y. Shen eds., *Visual Hybrids*, *Poetics Today* 44.3.

# Trade-offs between information and speech rate in naturalistic speech from 49 non-WEIRD languages

Christophe Coupé, Yoon Mi Oh, Dan Dediu, Frank Seifart & François Pellegrino
(University of Hong Kong, Ajou University, University of Barcelona, Leibniz Zentrum Allgemeine Sprachwissenschaft-ZAS & CNRS)

The regulation of information in language use involves trade-offs between the amount of information conveyed and the speech rate. This is captured in strong proposals such as the “smooth signal redundancy hypothesis” (Aylett & Turk, 2004, Pimentel et al., 2021), the “Uniform Information Density” hypothesis (Levy & Jaeger, 2007, Meister et al., 2021), and the notion of information rate attractors within (Cohen Priva, 2017) and across languages (Coupé et al., 2019). However, these proposals are based on either read speech produced by few speakers and/or data from only a few “big” languages. Do these hypotheses ‘survive’ in natural speech and also generalize to non-WEIRD (Western, Educated, Industrialized, Rich, and Democratic; see Henrich et al. 2010) languages?

Here, we investigate trade-offs within speakers (across different speech segments for an individual) and trade-offs across languages, using a unique collection of natural speech recently made available through the DoReCo corpus (Seifart, Paschen & Stave, 2022, see "references to language corpora" below). The DoReCo core set contains recordings and time-aligned transcriptions for more than 100 hours of natural speech in 51 typologically diverse and mostly non-WEIRD languages (plus English and French). To compute our target variables – average syllabic rate (SR) and average syllabic information density (ID) –, the recordings of all 378 speakers in the dataset were split into segments of ~5s(econds). For each segment, we determined SR as the number of syllables per second. For computing ID we developed a method that relies on the English translations of the transcriptions: for each segment, the total amount of information (to be divided by the number of syllables to compute the ID) was estimated as the surprisal of the translation with a gpt2 model (Radford et al., 2019), fine-tuned on the extended DoReCo corpus.

For instance, a 5s-Gorwaa segment that translates as "they made something to draw people, people were drawn, they agreed this thing continued when it was enough, it was thus" (Harvey 2022: doreco_goro1270_20151202e) has an ID of 3.15 and a SR of 6.89. This contrasts with a 5s-Northern-Alta segment (Garcia-Laguia 2022: doreco_nort2875_nalta093) that translates as "this pineapple however this, if you cutting weeds and the plant is already big" with a higher ID of 3.80, but a slower SR of 5.14, as expected.

Figure 1 summarizes our results. We observed first that across a speaker’s speech segments, the higher the ID is, the lower the SR tends to be. While there is variability, most speakers exhibit a negative correlation between ID and SR, with a median value equal to -.29. Second, after averaging speakers’ SR and ID for each language, we again found a negative correlation ($\rho$ = -.35, $p$ = .012): the faster a language, the less information its syllables tend to carry. It thus appears that the trade-offs previously observed hold against a diverse and naturalistic dataset, suggesting universal and multi-level regulation of the flow of information in language. We also show how recent large language models can be leveraged for research on small and under-researched languages.

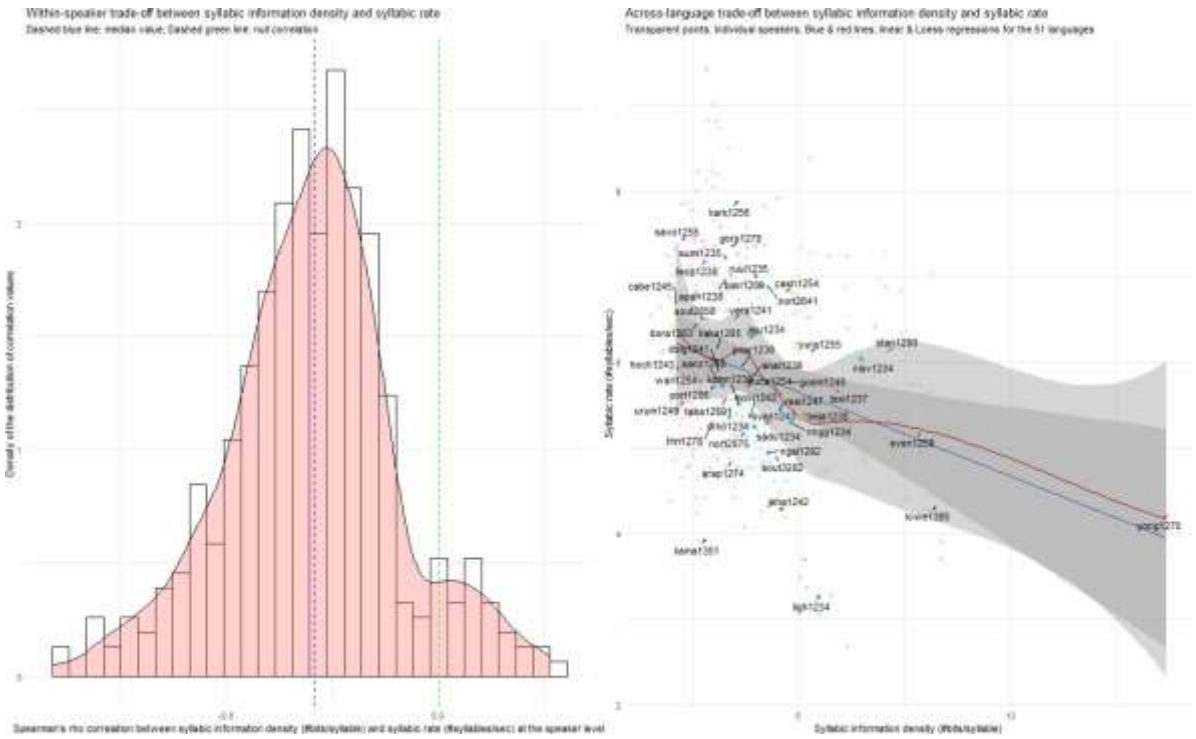



Figure 1: Within-speaker and across-language trade-offs between ID and SR in DoReCo. Left panel: Histogram of the distribution of the 378 speakers' correlation coefficients between ID and SR across 5s speech segments; Right panel: Average ID and SR for the 51 languages (based on speakers' average values).

**References**


Aylett, M., & Turk, A. (2004). The Smooth Signal Redundancy Hypothesis: A Functional Explanation for Relationships between Redundancy, Prosodic Prominence, and Duration in Spontaneous Speech. *Language and Speech*, *47*(1), 31–56. https://doi.org/10.1177/00238309040470010201

Cohen Priva, U. (2017). Not so fast: Fast speech correlates with lower lexical and structural information. *Cognition*, *160*, 27–34. https://doi.org/10.1016/j.cognition.2016.12.002

Coupé, C., Oh, Y. M., Dediu, D., & Pellegrino, F. (2019). Different languages, similar encoding efficiency: Comparable information rates across the human communicative niche. *Science Advances*, *5*(9), eaaw2594. https://doi.org/10.1126/sciadv.aaw2594

Levy, R., & Jaeger, T. F. (2007). Speakers optimize information density through syntactic reduction. In B. Schölkopf, J. Platt, & T. Hofmann (eds.), *Advances in Neural Information Processing Systems 19* (pp. 849–856). The MIT Press. https://doi.org/10.7551/mitpress/7503.003.0111

Meister, C., Pimentel, T., Haller, P., Jäger, L., Cotterell, R., & Levy, R. (2021). Revisiting the Uniform Information Density Hypothesis. *arXiv:2109.11635 [Cs]*. http://arxiv.org/abs/2109.11635

Pimentel, T., Meister, C., Salesky, E., Teufel, S., Blasi, D., & Cotterell, R. (2021). A surprisal–duration trade-off across and within the world's languages. *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, 949–962. https://doi.org/10.18653/v1/2021.emnlp-main.73

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D. & Sutskever, I. (2019). Language models are unsupervised multitask learners. Technical report, OpenAi.

Seifart, F., Paschen, L., & Stave, M. (eds.). (2022). Language Documentation Reference Corpus (DoReCo) 1.2. Berlin & Lyon: Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). https://doi.org/10.34847/nkl.7cbfq779 https://doreco.huma-num.fr/

**References language corpora**


Avanzi, M., Béguelin, M.-J., Corminboeuf, G., Diémoz, F., Johnsen, L.A. (2022). French (Swiss) DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.3520l685

Aznar, J. (2022). Nisvai DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.2801565f

Bartels, H., Szczepański, M. (2022). Lower Sorbian DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.6c6e4e9k

Bogomolova, N., Ganenkov, D., Schiborr, N.N. (2022). Tabasaran DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.ad7f97xr

Burenhult, N. (2022). Jahai DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.6a71xp0p

Cobbinah, A.Y. (2022). Baïnounk Gubëeher DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.a332abw8

Cowell, A. (2022). Arapaho DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.36f5r1b6

Däbritz, C.L., Kudryakova, N., Stapert, E., Arkhipov, A. (2022). Dolgan DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.f09eikq3

Döhler, C. (2022). Komnzo DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.c5e6dudv

Forker, D., Schiborr, N.N. (2022). Sanzhi Dargwa DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.81934177

Franjieh, M. (2022). Fanbyak DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.02084446

Garcia-Laguia, A. (2022). Northern Alta DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.efea0b36

Gipper, S., Ballivián Torrico, J. (2022). Yurakaré DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.7ca412wg

Gippert, J. (2022). Svan DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.9ba054c3

Griscom, R. (2022). Asimjeeg Datooga DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.f77c7m72

Güldemann, T., Ernszt, M., Siegmund, S., Witzlack-Makarevich, A. (2022). N‖ng DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.f6c37fi0

Gusev, V., Klooster, T., Wagner-Nagy, B., Arkhipov, A. (2022). Kamas DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.cdd8177b

Haig, G., Vollmer, M., Thiele, H. (2022). Northern Kurdish (Kurmanji) DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.ca10ez5t

Hartmann, I. (2022). Hoocąk DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.b57f5065

Harvey, A. (2022). Gorwaa DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.a4b4ijj2

Haude, K. (2022). Movima DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.da42xf67

Hellwig, B. (2022). Goemai DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.b93664ml

Hellwig, B., Schneider-Blum, G., Ismail, K.B.K. (2022). Tabaq (Karko) DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.eea8144j

Kazakevich, O., Klyachko, E. (2022). Evenki DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.5e0d27cu

Kim, S.-U. (2022). Jejuan DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.06ebrk38

Krifka, M. (2022). Daakie DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.efeav5l9

Michaud, A. (2022). Yongning Na DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.abe65p95

Mosel, U. (2022). Teop DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.9322sdf2

O'Shannessy, C. (2022a). Light Warlpiri DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.7452803q

O'Shannessy, C. (2022b). Warlpiri DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.042dv614

Ozerov, P. (2022). Anal DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.0dbazp8m

Ponsonnet, M. (2022). Dalabon DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.fae299ug

Quesada, J.D., Skopeteas, S., Pasamonik, C., Brokmann, C., Fischer, F. (2022). Cabécar DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.ebc4ra22

Reiter, S. (2022). Cashinahua DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.a8f9q2f1

Riesberg, S. (2022). Yali (Apahapsili) DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.9d91nkq2

Ring, H. (2022). Pnar DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.5ba1062k

Rose, F. (2022). Mojeño Trinitario DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.cbc3b4xr

Schiborr, N.N. (2022). English (Southern England) DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.9c271u5g

Schnell, S. (2022). Vera'a DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.3e2cu8c4

Seifart, F. (2022a). Bora DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.6eaf5laq

Seifart, F. (2022b). Resígaro DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.ffb96lo8

Skopeteas, S. (2022). Yucatec Maya DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.9cbb3619

Skopeteas, S., Moisidi, V., Tsetereli, N., Lorenz, J., Schröter, S. (2022). Urum DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.ac166n10

Teo, A. (2022). Sümi DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.5ad4t01p

Thieberger, N. (2022). Nafsan (South Efate) DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.ba4f760l

Vanhove, M. (2022). Beja DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.edd011t1

Vydrina, A. (2022). Kakabe DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.d5aeu9t6

Wegener, C. (2022). Savosavo DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.b74d1b33

Wichmann, S. (2022). Texistepec Popoluca DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.c50ck58f

Witzlack-Makarevich, A., Namyalo, S., Kiriggwajjo, A., Molochieva, Z. (2022). Ruuli DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.fde4pp1u

Xu, X., Bai, B. (2022). Sadu DoReCo dataset, in: Seifart, F., Paschen, L., Stave, M. (eds.), *Language Documentation Reference Corpus (DoReCo) 1.2*. Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), Berlin & Lyon. https://doi.org/10.34847/nkl.3db4u59d

# Typology meets cognitive science: moving from WEIRD to non-WEIRD languages

Antonio Benítez-Burraco
(University of Seville)

This contribution will serve as an introductory paper to the workshop. In the first part, we will provide a brief state-of-the-art of the research on the cognitive science of language. Over the time, we have identified a core cognitive device (or more properly, a set of interconnected devices) that enables humans to acquire and use a language (our faculty of language). However, variation pervades language facts, from the diversity of language uses, to the diversity of human languages, to the diversity of varieties of each individual language, and ultimately, to the differences between users when speaking the same variety of the same language for fulfilling the same function. Cognitively, the successful management of all this diversity is challenging and typically entails the differential recruitment of additional resources. In the second part of the talk, we will focus on the cognitive science of language diversity, which is the topic of the workshop. Linguistic typology has identified many of the aspects in which languages diverge, including infrequent or rare phenomena. Nonetheless, our knowledge of how human cognition deals with this diversity mostly derives from the study of a limited set of WEIRD (Western Educated, Industrialized Rich Democratic) languages (Blasi et al., 2022). Besides stressing this circumstance, we will also highlight the benefits of doing research on non-WEIRD languages, not only for the cognitive science of language, but also for linguistic typology, since cognitive biases are one key source of language universals through their impact on language change. We will also mention the challenges for a non-WEIRD cognitive science of language, including the design of ad-hoc experiments aimed to examine language-specific phenomena, particularly, exceptional phenomena, which are expected to impose cognitive loads and demands different to those of WEIRD-languages. The talk will finish with a brief overview of the contributions to the workshop.

**References**

Blasi, D. E., Henrich, J., Adamou, E., Kemmerer, D., & Majid, A. (2022). Over-reliance on English hinders cognitive science. Trends in cognitive sciences, 26(12), 1153–1170. https://doi.org/10.1016/j.tics.2022.09.015

# Codability of faces across diverse languages

Ewelina Wnuk & Jan Wodowski
(University of Warsaw & University of Warsaw)

Keywords: codability, ineffability, lexical semantics, semantic typology, descriptions of faces

The question of what meanings can and cannot be coded in words is central to our understanding of human language and cognition (Levinson & Majid 2014). Yet, predictions about inexpressible (or ineffable) domains often rely on English and related languages and, as recent research in the field of smell has shown (Wnuk & Majid 2014, Majid 2021), they do not always hold when considering a more diverse language sample. The present study addresses ineffability in another area, i.e., the human face (Mangini & Biederman 2004). While there is extensive research in psychology feeding into hypotheses on why faces might be difficult to verbalize (Meissner et al. 2008), there is little systematic work in linguistics investigating how speakers of different languages describe faces. In fact, this domain has been singled out as neglected in semantic typology (Evans 2011), despite its potentially high significance for semantic theory and our understanding of the limits of language. The present study addresses ineffability of faces through a series of stimulus naming tasks targeting Maniq, an Austroasiatic language of Southern Thailand, known for considerable semantic specificity in some of its basic vocabulary (Wnuk 2016). Maniq is compared to a representative of an Indo-European family – Polish. The investigation focuses on codability of faces, operationalized as naming agreement, as well as semantic and morphosyntactic analysis of responses. We consider facial features, facial movements, and emotions on the face. The results suggest Maniq and Polish display some commonalities; for instance, in both languages descriptions of facial features frequently reference shape, magnitude, color, and other surface properties. Nevertheless, upon direct comparison, Maniq and Polish differ significantly in codability of faces and facial lexicons, with Maniq displaying higher naming agreement for some aspects of faces and having more vocabulary dedicated to the face. This suggests Maniq is better suited to talking about faces and—although linguistically coding faces appears generally difficult—the constraint on their codability is not strict.

**References**

Evans, Nicholas (2011), Semantic typology. In J. J. Song (ed.), *The Oxford handbook of linguistic typology*, 504–533. Oxford: Oxford University Press.

Levinson, Stephen C. & Asifa Majid (2014), Differential ineffability and the senses, *Mind & Language* 29, 407–427.

Majid, Asifa (2021), Human olfaction at the intersection of language, culture, and biology, *Trends in Cognitive Sciences* 25(2), 111–123.

Mangini, Michael C. & Irving Biederman (2004), Making the ineffable explicit: estimating the information employed for face classifications, *Cognitive Science* 28(2), 209–226.

Meissner, Christian A., Siegfried L. Sporer & Kyle J. Susa (2008), A theoretical review and meta-analysis of the description-identification relationship in memory for faces. *European Journal of Cognitive Psychology* 20(3), 414–455.

Wnuk, Ewelina (2016), *Semantic specificity of perception verbs in Maniq*. Nijmegen: Radboud University PhD dissertation.

Wnuk, Ewelina & Asifa Majid (2014), Revisiting the limits of language: The odor lexicon of Maniq, *Cognition* 131(1), 125–138.

# The global recurrence and variability of kinship terminology structure

Sam Passmore
(Australian National University)

The extent to which kinship terminology varies between linguistic groups is a long-debated but unresolved social and linguistic puzzle (Jones, 2010). Whilst contemporary research has shown the common assumption that kinship terminology can be collapsed into a typology of six robust types is overly simplistic (Godelier et al., 1998), no alternative typologies have reached broad acceptance. This paper takes a data-driven solution to this problem.

Using data from Kinbank (Passmore et al., 2023), a global database of 1,156 kinship terminology, I quantitatively review the global structural diversity of kinship terminology and derive a new typology of kinship terminology that better captures global diversity. In a two-part analysis, I first show that more than six types are needed to represent the variability in cousin kinship terminology structure. While the analysis finds broad support for the existing types (e.g. lineal, bifurcate-merging, etc.), there is sufficient variation with these groups to justify further types. Specifically, there is evidence to consider types that, in addition to the bifurcate and lineal distinctions made within the existing typology (Lowie, 1928), distinguish kin based on relative age, relative gender, and gender of speaker – as originally proposed by Kroeber (1909).

Secondly, while typology can assume within-group homogeneity, variance within different typological categories is not homogenous. Some types of kinship terminology show identical structure across languages, but others may share a single structural feature but no other commonalities. Broadly speaking, I observe that terminology which fall into the broad "lineal" type show high levels of structural homogeny (i.e., all languages are more-or-less structured identically), whereas terminology that fall into the broad "bifurcate-merging" category shows low levels of structural homogeny (i.e., languages share a bifurcate-merging feature, but have little else in common). Some variability can be attributed to the global distribution of types: types that are found within small geographical bounds tend to be more similar, likely as a result of common inheritance, whereas types that are found globally exhibit more variability due to regional particularities. However this is not always the case, and I consider the possibility of universal constraints on kinship terminology diversity.

Finally, I also show that different subsets of kin contain varying levels of structural homogeneity. In smaller subsets of genealogical kin, such as siblings or grandparents, there is more structural homogeneity, whereas subsets involving more kin (siblings and cousins, niblings) tend to show more variability. In part, this can be attributed to the combinatorial possibilities imposed by the quantitative analysis, rather than a reflection of a real-world constraint, but some analysis of the patterning of these relationships is suggestive of additional cognitive constraints, such as the disinclination towards disjunctive categories.

Through this data-driven approach, I demonstrate that we can identify further division within this semantic domain than the existing typology currently describes. Further, a data approach allows the exploration of within-type variability, which furthers our understanding of type validity and opens up

the discussion for the usefulness of typology more generally, and whether we can opt for a more dynamic approach to understanding diversity.

**References**


Godelier, M., Trautmann, T. R., & Tjon Sie Fat, F. E. (1998). *Transformations of Kinship*. Smithsonian Institute.

Jones, D. (2010). Human kinship, from conceptual structure to grammar. *Behavioral and Brain Sciences*, *33*(5), 367–381. https://doi.org/10.1017/S0140525X10000890

Kroeber, A. L. (1909). Classificatory Systems of Relationship. *The Journal of the Royal Anthropological Institute of Great Britain and Ireland*, *39*, 77–84. https://doi.org/10.2307/2843284

Lowie, R. H. (1928). *A Note on Relationship Terminologies*. JSTOR. https://www.jstor.org/stable/661336

Passmore, S., Barth, W., Greenhill, S. J., Quinn, K., Sheard, C., Argyriou, P., Birchall, J., Bowern, C., Calladine, J., Deb, A., Diederen, A., Metsäranta, N. P., Araujo, L. H., Schembri, R., Hickey-Hall, J., Honkola, T., Mitchell, A., Poole, L., Rácz, P. M., … Jordan, F. M. (2023). Kinbank: A global database of kinship terminology. *PLOS ONE*, *18*(5), e0283218. https://doi.org/10.1371/journal.pone.0283218

# Neural evidence on the role of ergative case marking in sentence planning

Sebastian Sauppe
(University of Zurich)

Keywords: ergative, case marking, language processing, Basque, Hindi

Speakers (can) adapt the time course of sentence planning with respect to what linguistic information and which words planned when, e.g., in response to extralinguistic factors such as the visual salience of referents (Gleitman et al., 2007; Myachykov et al., 2011). It remains largely unknown, however, how planning is adapted to different grammars and whether adaptations to grammatical affordances are the same across languages—mainly due to a lack of comparative cross-linguistic production research (Jaeger & Norcliffe, 2009; Norcliffe et al., 2015; Nordlinger et al., 2022; Sarvasy et al., 2022).

We present two EEG studies that explore how early sentence planning processes are shaped by case marking. Some languages assign *ergative* case to the subjects of agentive verbs, distinguishing them from most other arguments. Other languages assign the same case (nominative or unmarked) to both transitive verb's agents and most other arguments (Bickel et al., 2015). We hypothesized that planning ergative sentences requires more extensive and earlier relational and structural encoding (Griffin & Bock, 2000) because verb agentivity is signaled on the sentence-initial noun through case, while nominative case does not reveal transitivity early on.

In the presented studies (Egurtzegi et al., 2022; Sauppe et al., 2021), participants described drawings of intransitive and transitive events while neural oscillatory activity in different EEG frequency bands was recorded. In the first study, a comparison of Basque (all agents ergative-marked) and Swiss German (nominative case marking pattern) revealed that Basque speakers quickly encoded the event-relational structure as well as the sentence-final verb's argument structure to decide on case marking at an early time point. These processes were reflected in event-related synchronisation (ERS) in the theta (~4-8 Hz) and alpha (~8-12 Hz) bands in the first 600 ms of the planning process (Egurtzegi et al., 2022). In the second study, a within-language comparison in Hindi built on the split nature of the language's case marking system (agents in perfective aspect carry ergative, but nominative/unmarked case in imperfective aspect) and found different patterns of neural activity. The planning of sentences with unmarked agents elicited increased theta-band ERS and alpha-band event-related desynchronization (ERD) in Hindi (Sauppe et al., 2021). We argue that these effects likely reflect the concurrent planning of possible alternative structures since sentences with unmarked initial NPs may continue in many different ways, whereas commitment to an ergative sentence structure early on eliminates alternative plans.

These results show that a language's case marking system profoundly shapes the earliest phases of sentence planning beyond, while language-specific characteristics of case marking determine the timing and nature of steps that speakers take before speaking (e.g., likely accessing verbal argument structure earlier in Basque than in Hindi). In addition, the current studies demonstrate that EEG is a viable method to reveal the neural underpinnings of sentence planning before speech onset and a versatile future tool for gaining fine-grained insights into cross-linguistic differences in language processing.

**References**


Bickel, B., Witzlack-Makarevich, A., Choudhary, K. K., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2015). The neurophysiology of language processing shapes the evolution of grammar: evidence from case marking. PloS One, 10(8), e0132819.

Egurtzegi, A., Blasi, D. E., Bornkessel-Schlesewsky, I., Laka, I., Meyer, M., Bickel, B., & Sauppe, S. (2022). Cross-linguistic differences in case marking shape neural power dynamics and gaze behavior during sentence planning. Brain and Language, 230, 105127.

Gleitman, L. R., January, D., Nappa, R., & Trueswell, J. C. (2007). On the give and take between event apprehension and utterance formulation. Journal of Memory and Language, 57(4), 544–569.

Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. Psychological Science, 11(4), 274–279.

Jaeger, T. F., & Norcliffe, E. J. (2009). The Cross-linguistic study of sentence production. Language and Linguistics Compass, 3(4), 866–887.

Myachykov, A., Thompson, D., Garrod, S., & Scheepers, C. (2011). Referential and visual cues to structural choice in visually situated sentence production. Frontiers in Psychology, 2, 396.

Norcliffe, E., Harris, A. C., & Jaeger, T. F. (2015). Cross-linguistic psycholinguistics and its critical role in theory development: early beginnings and recent advances. Language, Cognition and Neuroscience, 30(9), 1009–1032.

Nordlinger, R., Garrido Rodrigues, G., & Kidd, E. (2022). Sentence planning and production in Murrinhpatha, an Australian “free word order” language. Language, 98(2), 197–220.

Sarvasy, H. S., Morgan, A. M., Yu, J., Ferreira, V. S., & Momma, S. (2022). Cross-clause planning in Nungon (Papua New Guinea): Eye-tracking evidence. Memory & Cognition. https://doi.org/10.3758/s13421-021-01253-3

Sauppe, S., Choudhary, K. K., Giroud, N., Blasi, D. E., Norcliffe, E., Bhattamishra, S., Gulati, M., Egurtzegi, A., Bornkessel-Schlesewsky, I., Meyer, M., & Bickel, B. (2021). Neural signatures of syntactic variation in speech planning. PLoS Biology, 19(1), e3001038.

# An information-theoretic formulation of language complexity

Sihan Chen
(Department of Brain and Cognitive Sciences, MIT)

Keywords: language complexity, information theory, linguistic typology, communication

One of the big questions in language research is whether a language is more complex than an- other. However, current studies seem to have few agreement not only on whether some languages are more complex, but also in the methodology to quantify complexity (cf. Ehret et al., 2021, for review). On one hand, some researchers argue that all languages are relatively similar in complexity (e.g. Hockett, 1958), claiming that if a language is more complex than some other language in some aspect, there should be another aspect where the language is less complex than that other language (e.g. Fenk and Fenk-Oczlon, 1993; Juola, 1998; Sinnemäki, 2008, 2014; Ehret and Szmrecsanyi, 2016; Cotterell et al., 2019). On the other hand, there are also studies showing that some languages are either "grammatically less complex" than other languages (e.g. Everett, 2005; Gil and Shen, 2019; Gil, 2017; Dahl, 2008; Everett, 2018) or on the process of simplification or complexification over time (e.g. McWhorter, 2007; Kusters, 2003; Deutscher, 2000). Studies also differ in the formulation of complexity: *absolute complexity* and *relative complexity*, as noted in Miestamo (2004). The first approach measures the minimal length to describe the rules in a language, often in information- theoretic terms, as was followed by studies such as McWhorter (2001); Miestamo (2006); Ackerman and Malouf (2013); Çöltekin and Rama (2018); Cotterell et al. (2019); the other focuses on how difficult the language is processed or learned, as was adopted by Kusters (2003); Hawkins (2009). Scholars have also pointed out issues studies attempting to measure a type of complexity, such as conflating concepts generated when describing properties inherent to one language with those used when comparing properties across multiple languages (Haspelmath, 2018), using metrics for absolute complexity to measure relative complexity (Ehret et al., 2021), or ignoring factors other than grammar at play (Riddle, 2008).

Here we plan to assess language complexity from a more practical perspective, considering the relation between communication and language, by formulating complexity as the mutual information between meaning and utterances, which has been used to evaluate a communication system's complexity in various semantic domains (e.g. Zaslavsky et al., 2018, 2019; Chen et al., 2023).

Specifically, if all the meanings in a language form a set $M$ with a distribution $p(m)$, and all the utterrances in a language form a set $S$ with a distribution $p(s)$, then the complexity of a language, defined as $I[S;M]$, can be computed by Equation 1:

$$I[S;M] = H(M) - H(M \mid S) \tag{1}$$

The first term $H(M)$ is the entropy of meaning, which measures the diversity of meanings: the more meaning a language distinguishes, and the more uniformly these meanings are conveyed, the higher this term is. The second term $H(M \mid S)$ is the conditional entropy of meaning, conditioned on the utterance, which measures the ambiguity of language. From this equation, a language is complex if 1) it has a high diversity in meaning and 2) it is less ambiguous.

This one single formulation offers a quantitative explanation of previous proposals on complexity. For example, (McWhorter, 2001) state that less complex language relies more on context more to disambiguate utterance. (Gil, 1994, 2001) argue that less complex languages make less distinctions in thematic roles. In (Kusters, 2003), more complex languages are said to be more transparent, with one form corresponding to one and only one meaning. We hope this paper will foster more discussion on language complexity from this more quantitative, communication-centered perspective, including generating experimental methods to evaluate complexity of languages under this framework.

**References**

Ackerman, F. and Malouf, R. (2013). Morphological organization: The low conditional entropy conjecture. *Language*, pages 429–464.

Chen, S., Futrell, R., and Mahowald, K. (2023). An information-theoretic approach to the typology of spatial demonstratives. *Cognition*, 240:105505.

Çöltekin, Ç. and Rama, T. (2018). Exploiting universal dependencies treebanks for measuring morphosyntactic complexity. *Proceedings of the First Shared Task on Measuring Language Complexity*, pages 1–8.

Cotterell, R., Kirov, C., Hulden, M., and Eisner, J. (2019). On the complexity and typology of inflectional morphological systems. *Transactions of the Association for Computational Linguistics*, 7:327–342.

Dahl, Ö. (2008). Grammatical resources and linguistic complexity. *Studies in Language Companion Series (SLCS)*.

Deutscher, G. (2000). *Syntactic change in Akkadian: the evolution of sentential complementation*. OUP Oxford.

Ehret, K., Blumenthal-Dramé, A., Bentz, C., and Berdicevskis, A. (2021). Meaning and measures: interpreting and evaluating complexity metrics. *Frontiers in Communication*, 6:61.

Ehret, K. and Szmrecsanyi, B. (2016). An informationtheoretic approach to assess linguistic complexity. *Complexity, isolation, and variation*, 57:71.

Everett, C. (2018). Contrasting phonetic complexity across languages: Two approaches. Proceedings of the first shared task on measuring language complexity.

Everett, D. (2005). Cultural constraints on grammar and cognition in pirahã: Another look at the design features of human language. *Current anthropology*, 46(4):621–646.

Fenk, A. and Fenk-Oczlon, G. (1993). Menzerath's law and the constant flow of linguistic information. In *Contributions to quantitative linguistics*, pages 11–31. Springer.

Gil, D. (1994). the structure of riau indonesian. *Nordic Journal of Linguistics*, 17(2):179–200.

Gil, D. (2001). Escaping eurocentrism: Fieldwork as a process of unlearning. *Linguistic fieldwork*, pages 102–132.

Gil, D. (2017). Isolating-monocategorial-associational language. In *Handbook of categorization in cognitive science*, pages 471–510. Elsevier.

Gil, D. and Shen, Y. (2019). How grammar introduces asymmetry into cognitive structures: compositional semantics, metaphors, and schematological hybrids. *Frontiers in Psychology*, page 2275.

Haspelmath, M. (2018). How comparative concepts and descriptive linguistic categories are different. *Aspects of linguistic variation*, pages 83–114.

Hawkins, J. A. (2009). *An efficiency theory of complexity and related phenomena*. na. Hockett, C.

F. (1958). A course in modern linguistics.

Juola, P. (1998). Measuring linguistic complexity: The morphological tier. *Journal of Quantitative Linguistics*, 5(3):206–213.

Kusters, W. (2003). *Linguistic complexity*. Netherlands Graduate School of Linguistics.

McWhorter, J. (2007). *Language interrupted: Signs of non-native acquisition in standard language grammars*. Oxford University Press on Demand.

McWhorter, J. H. (2001). The worlds simplest grammars are creole grammars.

Miestamo, M. (2004). On the feasibility of complexity metrics. In *FinEst linguistics, proceedings of the annual Finnish and Estonian conference of linguistics, Tallinn*, pages 11–26.

Miestamo, M. (2006). On the complexity of standard negation. *A man of measure: Festschrift in Honour of Fred Karlsson on His 60th Birthday*, pages 345–356.

Riddle, E. M. (2008). Complexity in isolating languages: Lexical elaboration versus grammatical economy. *Language complexity: Typology, contact, change*, pages 131–151.

Sinnemäki, K. (2008). Complexity trade-offs in core argument marking. *Language complexity: Typology, contact, change*, 67:88.

Sinnemäki, K. (2014). Complexity trade-offs: A case study. In *Measuring grammatical complexity*, pages 179–201. Oxford University Press.

Zaslavsky, N., Kemp, C., Regier, T., and Tishby, N. (2018). Efficient compression in color naming and its evolution. *Proceedings of the National Academy of Sciences*, 115(31):7937–7942.

Zaslavsky, N., Kemp, C., Tishby, N., and Regier, T. (2019). Color naming reflects both perceptual structure and communicative need. *Topics in cognitive science*, 11(1):207–219.

# WS9 Marginal phonemes

# Are phonetic rara marginal phonemes?

Tihomir Rangelov
(Max Planck Institute for Evolutionary Anthropology, Leipzig)

Keywords: phonetic rara, marginal phonemes, phonetics/phonology interface, phonological typology, functional load

Cross-linguistically rare sounds are often characterized by some phonetic constraints: they may be considered difficult to articulate and/or integrate into connected speech, or they may require very precise articulatory settings that easily fail. In this context, phonetic rara are more likely to persist in a language if they become contrastive, i.e. when they have a higher functional load. Establishing phonemic status is, however, not always straightforward (c.f. Hall 2013), and some oppositions involve marginal contrasts. In this study, I investigate how often, and in what circumstances, phonetic rara tend to be marginal phonemes.

I compile data on several phonetic rara, including clicks, bilabial trills, linguolabials, labial-velars, voiceless sonorants, the *th*-sound [θ] (Maddieson 2013), the Swedish *sj*-sound [ɧ], and the Czech *ř*-sound [r̝]. To determine whether they are marginal phonemes, I estimate their functional load. I interpret here marginal phonemes as phonemes with relatively low functional load using primarily the criteria set out by Labov (1994): they display low lexical opposition (few minimal pairs) and high lexical predictability (skewed phonological distribution).

The preliminary results show that phonetic rara are marginal phonemes in at least some cases. Examples are linguolabials (Lynch 2020), bilabial trills in most languages (Maddieson 1989, Rangelov et al. 2023), clicks in some Bantu languages (Sands 2020), and voiceless sonorants in Icelandic (Árnason 2005). Other rara, such as labial-velars, the Czech *ř*-sound, the Swedish *sj*-sound (Riad 2014), and the *th*-sound are much less commonly marginal phonemes. There is, however, a tendency for phonetic rara to be less marginal in linguistic areas where they are most commonly found, for example, clicks in languages spoken in the Kalahari Basin Area, and voiceless sonorants in Southeast Asian languages. This tendency overlaps with existing typologies of rara, according to which they can be either absolute (found rarely across language families) or relative (globally rare, but common in a geographic area or a language family) (c.f. Cysouw and Wohlgemuth 2010).

These tendencies can be explained by at least two factors. Firstly, in languages where phonetic rara have been (relatively recently) borrowed, these tend to be more marginal. Secondly, the persistence of rare sounds is likely dependent on extralinguistic factors, such as in-group identity attachment (cf. Rangelov et al. 2023, Guérin 2011). This is in line with Ladefoged & Everett's (1996) observation that phonetic rarities play a role that goes beyond "widespread phonological processes."

With this paper, I aim to contribute towards a typology of the phonological relationships of phonetic rara by discussing the above tendencies in detail. I elaborate on the behavior of the sounds mentioned above in different languages. I also explore cases where phonetic rara are not phonemic, or their phonemic status is difficult to establish.

**References:**


Árnason, Kristján. 2005. *Íslensk tunga. Handbók um hljóðfræði og hljóðkerfisfræði / Bd 1, Hljóð*. Reykjavík : Almenna bókafélagið

Cysouw, Michael & Jan Wohlgemuth. 2010. The other end of universals: Theory and typology of rara. In Jan Wohlgemuth & Michael Cysouw (eds.), *Rethinking universals: How rarities affect linguistic theory*. Berlin, New York: De Gruyter Mouton.

Guérin, Valérie. 2011. *A grammar of Mavea: An Oceanic language of Vanuatu*. Honolulu: University of Hawai'i Press.

Hall, Kathleen Currie. 2013. A typology of intermediate phonological relationships. *The Linguistic Review* 30(2). 215-275. https://doi.org/10.1515/tlr-2013-0008

Labov, William. 1994. *Principles of linguistic change. Volume 1: Internal factors*. Oxford: Blackwell.

Ladefoged, Peter & Daniel Everett. 1996. The status of phonetic rarities. *Language* 72(4). 794-800. https://doi.org/10.2307/416103

Lynch, John. 2020. The Bilabial-to-Linguolabial Shift in Southern Oceanic: A Subgrouping Diagnostic? *Oceanic Linguistics* 58(2). 292-323

Maddieson, Ian. 1989. Aerodynamic constraints on sound change: The case of bilabial trills. *UCLA Working Papers in Phonetics* 72. 91-115

Maddieson, Ian. 2013. Presence of Uncommon Consonants. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) *WALS Online* (v2020.3) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.7385533 (Available online at http://wals.info/chapter/19, Accessed on 2023-11-09.)

Rangelov, Tihomir, Mary Walworth & Julie Barbour. 2023. A multifaceted approach to understanding unexpected sound change: The bilabial trills of Vanuatu's Malekula Island. *Diachronica* 40(3). 384–432 https://doi.org/10.1075/dia.21051.ran

Riad, Tomas. 2014. *The phonology of Swedish*. Oxford University Press.

Sands, Bonny. 2020. Click Consonants: an Introduction. In Bonny Sands (ed.), *Click Consonants*. Leiden, Boston: Brill.

# The emergence of a marginal phoneme as a harmonic and contact-assisted phenomenon: The case of the Neo-Aramaic [œ]

Lidia Napiorkowska
(University of Warsaw)

Keywords: Diyana Neo-Aramaic, vowel harmony, vowel fronting, language contact, Semitic typology

Language contact has been recognised as an important source of marginal phonemes (for example in the case study of Romanian in Renwick 2014). The main factor in this process is lexical borrowing and its "non-adaptation" strategy (Martin, van Heugten, Kager, and Peperkamp 2022:4). Even though the distribution of marginal phonemes is highly restricted (Evans 2022), their functional load rather low, recognising minimal contrasts benefits the typological and diachronic studies of sound change (Renwick 2024:20). In this paper, I will discuss the contact-induced emergence of a marginal phoneme in Neo-Aramaic, viewed, however, from a different angle. I will also demonstrate the importance of representing marginal phonemes in documentation and descriptive work as segments indicative not only of the typological status of a language, but also of the self-identity of the language community.

My analysis pertains to suprasegmental features and concerns the contact between languages from different families in East Anatolia (Haig 2001). I will argue that the phone [œ] in a Neo-Aramaic dialect of Diyana (a Semitic language) should not be classified as an allophone of /*o*/, but rather a marginal phoneme. I will demonstrate that /*ö*/ occurred due to a harmonic process of vowel fronting, triggered by regressive assimilation from the newly emerged vowel /*ü*/. Although contrast between /*o*-*ö*/ in Diyana has not been attested, we may compare the same word in two dialects, one with /*u*/ vowel and one with /*ü*/:

(1a) beθoxun [beθo:xun]

(Jewish Betanure, Mutzafi 2008:42)

(1b) betöxün [bet$^{h}$œ:xʏn]

(Neo-Aramaic of Diyana, fieldwork)

'your (PL) house'

The above phenomenon of fronting harmony is unusual in Semitic languages, as are fronted rounded vowels (Kogan 2011:119-120). The harmony attested in modern times in Modern Hebrew loans is not only a rightward (i.e. progressive) spreading, but also a non-rounding harmony (Cohen 2013). The case of the Diyana Neo-Aramaic [œ] is therefore of a different nature. Since the appearance of /*ü*/ in the Neo-Aramaic of Diyana was probably contact-assisted to begin with (Napiorkowska 2015:39-40), the emergence of /*ö*/ and the harmonising process may be attributed to the spreading of features in the area of East Anatolia, cf. similar vowel qualities in Kurmanji Kurdish (Iranian, Öpengin 2012) and Turkish (Haig and Öpengin 2018:194). Though the influence of the neighbouring languages on Neo-Aramaic has been widely discussed (see Kapeliuk 2001, Khan 2008:7-8, Haig 2001), structural similarities have mainly been underlined. By contrast, the emergence of the fronted marginal phoneme /*ö*/, next to /*ü*/ should be considered a novelty in Semitic. Thus, the emergence of /*ö*/ indicates a typological development in the Neo-Aramaic phonology to converge with the phonological systems characteristic

of the varieties of East Anatolia (Donabedian and Sitaridou 412). This typological advancement in Neo-Aramaic could easily be overlooked if we subsume the vowel quality [œ] under /*o*/, rather than recognise the existence of a marginal phoneme.

Finally, as a field researcher, I wish to adduce the speakers' perspective and their perception of [œ] in Diyana Neo-Aramaic as the characteristic trait of the dialect. Linguistic self-identity is likely to play a role in sustaining rare sounds in a language, regardless of their functional load (Rangelov at al. 2023:406). Respecting the expression of the community's identity further indicates the need to recognise the existence of marginal phonemes (against more categorical approaches, cf. Martin, van Heugten, Kager, and Peperkamp 2022:30) and the benefits thereof to the organisation of the phonological phenomena.

**References**


Cohen, Evan-Gary. 'The Emergence of the Unmarked: Vowel Harmony in Hebrew Loanword Adaptation.' *Lingua* 131 (2013), 66-79.

Donabedian, Anaïd and Ioanna Sitaridou. 'Anatolia'. *The Routledge Handbook of Language Contact*, Evangelia Adamou and Yaron Matras (eds.), 404-433, 2020.

Evans, Nicholas. 'Pushing the Boundaries: Marginal Phonemes and Dialogic Interaction.' *Russian Journal of Linguistics* 2022, 26 (4): 995-1011.

Haig, Geoffrey. 'Linguistic Diffusion in Present-Day East Anatolia: From Top to Bottom.' *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics Areal Diffusion and Genetic Inheritance*, Alexandra Aikhenvald and R. Dixon (eds.), 195-224. Oxford: Oxford University Press, 2001.

Haig, Geoffrey and Ergin Öpengin. 'Kurmanji Kurdish in Turkey: Structure, Varieties and Status.' *Linguistic Minorities in Turkey and Turkic-speaking Minorities of the Periphery*, Christiane Bulut (ed.), 157–230. Wiesbaden: Harrassowitz, 2018.

Kapeliuk, Olga. 'Language Contact between Aramaic Dialects and Iranian.' *The Semitic Languages: An International Handbook*, Stefan Weninger et al. (eds.), 738-747. Muton: De Gruyter, 2011.

Khan, Geoffrey. *The Jewish Neo-Aramaic Dialect of Urmi*. Piscataway, New Jersey: Georgias Press, 2008b.

Kogan, Leonid. 'Proto-Semitic Phonetics and Phonology.' *The Semitic Languages: An International Handbook*, Stefan Weninger et al. (eds.), 54-151. Muton: De Gruyter, 2011.

Martin, Alexander, Marieke van Heugten, René Kager, and Sharon Peperkamp. 'Marginal Contrast in Loanword Phonology: Production and Perception'. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 2022, 13 (1): 1–36.

Mutzafi, Hezy. *The Jewish Neo-Aramaic Dialect of Betanure (Province Dihok).* Wiesbaden: Harrassovitz, 2008.

Napiorkowska, Lidia. *A Grammar of the Christian Neo-Aramaic Dialect of Diyana-Zariwaw*. Leiden: Brill, 2015.

Öpengin, Ergin. 'Dialectal Variation in Kurmanji: Isoglosses in Phonology and Morphosyntax.' Lacito CNRS, Paris 3 – Bamberg University, Handout for 2nd International Conference of Kurdish Studies, Exeter, Sept 6, 2012.

Rangelov, Tihomir, Mary Walworth and Julie Barbour. 'A Multifaceted Approach to Understanding Unexpected Sound Change: The Bilabial Trills of Vanuatu's Malekula Island.' *Diachronica* 2023, 40 (3): 384–432.

Renwick, Margaret E. L. *The Phonetics and Phonology of Contrast: The Case of the Romanian Vowel System*. Mouton: De Gruyter, 2014.

—. 'Robustness and Complexity in Italian Mid Vowel Contrasts.' *Languages* 2024, 9 (4): 1-24. https://www.mdpi.com/2226-471X/9/4/150.

# Voiced Sibilant Affricates in Bulgarian: Evolution and Phonemics

Denitsa Dimitrova
(University of Sofia)

Keywords: voiced sibilant affricates, marginal phonemes, Bulgarian phonology, phonemic analysis, dialectal geography

Affricates, representing a characteristic convergence of stop and fricative articulations, are a widespread and varied feature in the world's languages, as comprehensively detailed in Berns' (2014) typological survey. In her study, Berns found that 66.96% of languages in the UPSID database feature at least one affricate, with a notable dominance of voiceless over voiced affricates (74.7% vs. 25.3%). Sibilant affricates, especially the voiceless palatoalveolar /tʃ/ and dental/alveolar /ts/, are particularly favoured. In Bulgarian phonology, both voiced and voiceless alveolar and postalveolar affricates are present; however, there is no scientific consensus regarding the status of the voiced alveolar affricate.

Voiced sibilant affricates, while common in various languages, including Slavic, Romance, and Germanic, often face marginalization due to their articulatory and aerodynamic complexity (Żygis 2008; Żygis, Fuchs & Koenig 2012). This complexity frequently leads to their devoicing, decomposition, or transformation into other sounds. Mirroring this global trend of marginalization, voiced affricates, particularly [d͡z], hold an ambiguous status in Bulgarian phonology. This ambiguity mainly arises from its limited phonologic functions; while some linguists, such as Scatton (1993), argue for its exclusion from the phonemic inventory due to its rarity in both native and foreign words, others, including Ternes and Vladimirova-Buhtz in the Handbook of IPA (1999), recognize it as part of the phonological system.

The affricate [d͡z] traces its roots to Old Bulgarian, emerging from the early palatalization of /k, g, x/ in front of front vowels and /j/. Conversely, [d͡ʒ] also originates from Old Bulgarian, but its prevalence is supported by the significant influence of Turkish loanwords up to the 20th century, many of which are now considered archaic. Studies by Velcheva (1999) and Boiadzhiev (2004) suggest that in some Bulgarian dialects, these sounds evolve from corresponding sibilants in positions following a sonorant consonant in a process of secondary affricatization. For example:

инжекция [in'ʒɛkt͡sijɐ]
injection

инджекция [inˈd͡ʒɛkt͡sijɐ]
'injection'

Additionally, there are instances of morphologization in suffixes, where these affricates alternate with /g/, and lexicalization in specific words. These findings indicate that the voiced sibilant affricates in Bulgarian exhibit unique developmental patterns, reflecting both historical influences and contemporary phonological processes.

This research project aims to investigate the origins, distribution, combinatorial properties, and phonemic roles of the voiced sibilant affricates [d͡z] and [d͡ʒ] in Bulgarian, providing insights into their current usage and potential future trends. Utilizing a corpus-based approach, the study will analyze

their occurrences within the Bulgarian National Corpus (Bulgarian Academy of Sciences 2009) and the Online Dictionary of Bulgarian (Bulgarian Academy of Sciences 2014), specifically examining the frequency and distribution of these affricates in word-initial, word-medial, and word-final positions.

Furthermore, the project will employ a linguo-geographical approach using resources such as the Bulgarian Dialectal Atlas (Antonova-Vasileva et al. 2001) and the online spoken corpus of Bulgarian dialects (Alexander & Zhobov 2016). The outcome will include a detailed map depicting the geographical distribution of [d͡z] and [d͡ʒ] across Bulgaria.

**References**


Alexander, R. & Zhobov, V. (2016). Bulgarian Dialectology as Living Tradition. Retrieved January 11, 2024, from http://www.bulgariandialectology.org

Antonova-Vasileva, L., Vasileva, V., Vitanova, M., Valcheva, E., Keremidchieva, S., Kostova, T., … Vakarelska, D. (2001). Balgarski dialekten atlas. Obobshtavasht tom. Sofia: Trud.

Berns, J. (2014). A typological sketch of affricates. Linguistic Typology, 18(3), 369–411. DOI: 10.1515/lingty-2014-0017.

Boiadzhiev, T. (2004). Proizhod i distributsia na zvuchnite afrikati v balgarskite dialekti. Balgarska rech, 10(3), 42-48.

Bulgarian Academy of Sciences (2009). Bulgarian National Corpus. Retrieved January 11, 2024, from http://dcl.bas.bg/bulnc/

Bulgarian Academy of Sciences (2014). Online Dictionary of Bulgarian. Retrieved January 11, 2024, from https://ibl.bas.bg/rbe/

Scatton, E. (1993). “Bulgarian”. In: The Slavonic Languages. London: Routledge. pp. 188-249.

Ternes, E. & Vladimirova-Buhtz, T. (1999). "Bulgarian". Handbook of the International Phonetic Association. Cambridge University Press. pp. 55–57.

Velcheva, B. (1999). Yavleniya, svyrzani s [dz] i [dž] v balgarskite govori. In Sbornik “Dialektologiya i lingvistichna geografiya”. Sofia: Izdanie na Sofiyski universitet. pp. 39-45.

Żygis, M. (2008). On the avoidance of voiced sibilant affricates. ZAS Papers in Linguistics, 49, 23–45. http://dx.doi.org/10.21248/zaspil.49.2008.362

Żygis, M., Fuchs, S. & Koenig, L. L. (2012). Phonetic explanations for the infrequency of voiced sibilant affricates across languages. Journal of the Association for Laboratory Phonology, 3, 299-336. DOI: 10.1515/lp-2012-0016.

# How do marginal phonemes emerge? Case studies from minor Finnic varieties

Fedor Rozhanskiy & Elena Markus
(University of Tartu & University of Tartu)

Keywords: phonology, marginal phonemes, Finnic languages, moribund languages, language contact

Marginal phonemes present a major challenge for linguists compiling a phonological description of a language (Scobbie & Stuart-Smith 2008, Hall 2013, Evans 2022, and Martin et al. 2022). In the current paper, we discuss how and why such phonemes emerge and evolve using material from several moribund Finnic varieties: Vaipooli Votic, Soikkola and Lower Luga Ingrian, and Siberian Ingrian Finnish. The data come mostly from our own fieldwork and also from recent publications by our colleagues. Most of these data have not been previously analysed from a theoretical phonological perspective.

In the mentioned varieties, marginal phonemes are quite numerous. The goal of our analysis is to illustrate the mechanisms that lead to their emergence. Among the reasons and processes behind the development of marginal phonemes we address the following phenomena.

1. Changes in the phonetic environment of segments that used to occur in a complementary distribution but can no longer be considered allophones, e.g. *l* vs *ɲ* in Jõgõperä Votic, cf. *iɲm* 'weather' and *silm* 'eye'.

2. The loss of final vowels and the transition of their phonological properties to the preceding consonants, e.g. the emergence of palatalized and/or labialized consonants in Siberian Ingrian Finnish, e.g. *tükk'$^{o}$* 'smouldering piece of wood' < *tükkü*, *pehk$^{o}$* 'bush' < *pehko* (Sidorkevič 2014).

3. Non-standard behaviour of segments in consonant clusters (e.g. no expected change *s* > *š* in the cluster *ts* in Soikkola, cf. *okša* < *oksa* 'branch' but *otsa* 'edge') promotes their possible interpretation as indivisible segments (here, an affricate, see Rožanskij & Markus 2020).

4. A large number of borrowings changes the status of phonemes in a language, e.g. palatalized consonants in Vaipooli Votic were originally attested only in marginal lexicon such as onomatopoeia or specific pejorative/diminutive words (*n'ookkima* 'to hup', *n'aukkuma* 'to meow', *mut'u* 'genitals of a boy') but became frequent because of Russian loanwords, e.g. *varen'n'ə̂* 'home-made jam', *ut'ugə̂* 'iron (device)'.

5. Prosodic changes in a language (e.g. replacing quantitative oppositions with qualitative ones) cause restructuring of the vocalic system and the emergence of marginal phonemes. For instance, in Luuditsa Votic the contrast of the type *ā*/*ā̆* vs *a*/*ä* (in all syllables) changed into *ā*/*ā̆* (initial syllable) vs *a*/*ä* (all syllables) vs reduced *ə̂*/*ə* (final syllables), e.g. *r**aa**k**aa*** 'branch, brushwood.PART' > *r**aa**k**a***, *j**a**ɲk**a*** 'foot' > *j**a**ɲk**ə̂***. Similarly, short final vowels developed into voiceless vowels that contrast with final voiced (originally long) vowels in Lower Luga Ingrian (see details in Kuznetsova 2015).

In our data, there are two main groups of factors that trigger the emergence of marginal phonemes. First, there are internal factors, namely the changes in the phonetic system that distort the original phonetic contexts. Second, there are external factors: contact induced changes that restructure the phonological system, and borrowings from contact languages. The latter may well contain new or rare

segments not common to the original phonological system of a given language. Often, internal and external factors work simultaneously, increasing the probability for marginal phonemes to emerge.

**References**


Evans, Nicholas (2022), Pushing the boundaries: Marginal phonemes and dialogic interaction. *Russian Journal of Linguistics* 26(4), 995–1011.

Hall, Kathleen Currie (2013), A typology of intermediate phonological relationships. *The Linguistic Review* 30 (2), 215–275.

Kuznecova Natalia V. (2015), Dve fonologičeskije redkoski v dialektax ižorskogo jazyka. *Acta Linguistica Petropolitana* 11(2), 201–257.

Martin, Alexander, Marieke van Heugten, René Kager, and Sharon Peperkamp (2022), Marginal contrast in loanword phonology: Production and perception. *Laboratory Phonology* 13(1), 1–36.

Rožanskij, Fedor I., and Elena B. Markus (2020), O fonologičeskom statuse affrikaty ts v sojkinskom dialekte ižorskogo jazyka. *Acta Linguistica Petropolitana* 16 (3), 440–479.

Scobbie, James M., and Jane Stuart-Smith (2008), Quasi-phonemic contrast and the fuzzy inventory: Examples from Scottish English. In Peter Avery, B. Elan Dresher, and Keren Rice (eds.), (2008), *Contrast in phonology: theory, perception, acquisition*. Berlin/New York: Mouton de Gruyter, 85–114.

Sidorkevič Dar'ja V. (2014), Jazyk ingermanlandskix pereselencev v Sibiri (struktura, dialektnye osobennosti, kontaktnye javlenija). Unpublished Ph.D. thesis. St-Petersburg: Institute for Linguistic Studies.

# Proto-Mordvin *ć: a marginal phoneme or an artifact of reconstruction?

Sampsa Holopainen
(University of Helsinki)

Keywords: Phonological reconstruction, Mordvin, Uralic languages, comparative method

In this presentation I will discuss the status of the phoneme *ć (a voiceless palatal affricate) in Proto-Mordvin (the reconstructed predecessor of the Mordvin languages Erzya and Moksha, part of the Uralic language-family) and investigate whether this phoneme is a marginal phoneme or simply a result of erroneous reconstruction. The purpose is both to highlight problems and challenges of diachronic research on marginal phonemes and to contribute to a more precise reconstruction of Mordvin historical phonology.

It is customary to reconstruct *ć to Proto-Mordvin phoneme system (Keresztes 1987: 78, Bereczki 1988: 316, Bartens 1999: 33 and Zhivlov 2023: 129) but already Keresztes (1987: 78) notes that this phoneme occurs only in few words, some of which are probably of onomatopoetic origin. In the etymological material of Keresztes, there are 16 examples showing Proto-Mordvin *ć, 7 of which are labeled as problematic etymologies or show an alternative reconstruction. Moreover, contrary to most Proto-Mordvin consonant phonemes, *ć does not have a clear Proto-Uralic source: although it has been argued that Proto-Mordvin *ć in some cases reflects Proto-Uralic *ć (Bereczki 1988: 318), this is not accepted by Bartens (1999: 42) or Zhivlov (2023: 129), according to whom Proto-Uralic *ć regularly becomes Proto-Mordvin *ś (*ź between vowels). The alleged Uralic or Finno-Ugric etymologies with *ć in the material of Keresztes (1987) and the etymological dictionary UEW often manifest other phonological problems or a limited distribution within Uralic which points to erroneous etymologies or parallel loanwords (for example, Erzya *ćoćana* ‘top’ ~ Hungarian *csúcs* id. is likely an erroneous etymology due to problematic sound correspondences and Proto-Mordvin (?) *ćil’gə is rather a Turkic borrowing as are its alleged cognates, such as Hungarian *süly* ‘Skorbut’ [WOT: 749]). Furthermore, many etymologies with Proto-Mordvin *ć are based on disputed Erzya and Moksha evidence, as the dialectal variation of the alleged reflexes of *ć is great: for example, the reflexes of Proto-Mordvin (?) *purćəs ‘piglet’ (> Erzya dialectal *pursuz*, *purcos* etc.) and Proto-Mordvin (?) *ćil’gə (> Erzya *śil’ge*, Moksha *ćil’gä* ‘wart’) show aberrant correspondences that are difficult to interpret diachronically. It seems plausible to assume that ć developed independently in Erzya and Moksha after the split of Proto-Mordvin, possibly with the influx of loanwords. In my presentation I will discuss the dialect data in more detail.

The evidence for reconstructing *ć as a separate phoneme is rather meagre. Issues with marginal phonemes have been encountered in the reconstruction of protolanguages with a more distant past (see Pronk 2019 regarding Proto-Indo-European), and Proto-Mordvin *ć is an interesting case study, as this is a protolanguage with more shallow time-depth (split sometime during the first millenium CE; cf. Keresztes 1987: 42–43) and its phonology can be reconstructed in detail. If it turns out that reconstructing *ć is, in the end, justified, this will be a useful example for scholars working with marginal phonemes in reconstructed languages and interesting from the point of view of the methodology of phonological reconstruction.

**References**


Bartens, Raija. 1999. *Mordvalaiskielten rakenne ja kehitys*. Helsinki: Finno-Ugrian Society.

Bereczki, Gábor. 1988. *Geschichte der wolgafinnischen Sprachen*. In Sinor, Denis (ed.), *The Uralic languages: description, history and foreign influences*, 314–350. Leiden – Boston: Brill.

Keresztes, László. 1987. *Geschichte des mordwinischen Konsonantismus* I. Szeged: JATE Finnugor tanszék.

Pronk, Tijmen. 2019. Proto-Indo-European *a. *Indo-European linguistics* 7. 122–163.

UEW = Rédei, Károly. 1988. *Uralisches Etymologisches Wörterbuch*. Wiesbaden: Harrassowitz.

WOT = Róna-Tas, András & Berta, Árpád. 2011. *West Old Turkic. Turkic loanwords in Hungarian*. Wiesbaden: Harrassowitz.

Zhivlov, Mikhail 2023. Reconstruction of Proto-Uralic. In Abondolo, Daniel & Valijärvi, Riitta-Liisa (eds.), *The Uralic languages*, 117–175. 2nd edn. London & New York: Routledge.

# Contact-influenced transitory widening of consonantal periphery in vanishing Lower Luga Ingrian (Finnic)

Natalia Kuznetsova
(Università Cattolica del Sacro Cuore)

Keywords: Ingrian, Votic, marginal phonemes, unfinished sound change, phonological rarities

Lower Luga Ingrian is the most idiosyncratic dialect of the vanishing Ingrian language (Finnic) spoken at the Russian-Estonian border. This dialect has been influenced by several closely related varieties within the *Lower Luga Sprachbund* (language union): Votic, Ingrian Finnish, and Estonian (Kuznetsova, Markus & Muslimov 2015). In particular, the combination of linguistic, archaeological, and genetic data (Kuznetsova et al. 2022) indicates that Lower Luga Ingrian speakers represent a mix of original Votic and Ingrian population. The dialect itself is a convergent variety between Ingrian and Lower Luga Votic.

The data used in this study was collected by the author and colleagues in 2006-2015 in Lower Luga, Soikkola Ingrian, and Lower Luga Votic villages (in total, ~80 speakers), see Kuznetsova (2009). The consonantal system of Lower Luga Ingrian is very different from the systems of core Ingrian dialects, e.g. of Soikkola Ingrian (cf. Figures 1-2), due to several innovations:

(1) transformation of rare ternary consonantal quantity into more common binary (through the merger of the intermediate and the longest category into a single type of long consonants);

(2) split of the lenis/fortis contrast of *p-, t-, k-, s-types* into two separate contrasts, in length and in voicing;

(3) split of *s-types* into two separate types of phonemes: *s-types* and *ʃ-types*;

(4) formation of phonemic *t͡s-types*;

(5) expansion of the inventory of palatalised phonemes (as a result of the syncope and apocope of reduced vowels).

These innovations have brought the Lower Luga consonantal inventory close to that of Lower Luga Votic (cf. Markus & Rozhanskiy 2017: 2:16). Importantly, they have significantly widened the inventory of marginal consonants, i.e. those with a limited distribution (marked as non-bold in Figures 1-2).

The crucial difference between the Soikkola and Lower Luga marginal consonants, however, is not only in their number, but also in the status of such phonemes. In Soikkola, all marginal consonants are almost doubtless independent phonemes. In Lower Luga Ingrian, the status of many of them is uncertain (marked in italic in Figure 2), because several sound changes are still ongoing (the final vowel loss, the formation of *t͡s-types*). Note that also Lower Luga Votic has a smaller inventory of the marginal consonants with unclear status than Lower Luga Ingrian.

Such an contact-induced inventory with an extremely inflated periphery of unclear marginal phonemes might be just a brief transitory period in the life of a vital language. However, the Ingrian language is nearly obsolete, so there are high chances that exactly this stage will remain in the descriptions of Lower Luga Ingrian and used for typological comparisons (e.g. as a phonological rarity). Typology, therefore, should take sociolinguistic factors related to language loss into account more closely. Otherwise, such factors may create certain unwanted biases in data, which may lead to the controversial statements like “small indigenous languages concentrate more linguistic rarities” (viz. Anderson, Kuznetsova & Easterday in press).

**References**


Anderson, Cormac, Kuznetsova, Natalia & Easterday, Shelece. in press. Rarities in phonetics and phonology. In *Rarities in phonetics and phonology: structural, typological, evolutionary, and social dimensions* [Topics in Phonological Diversity]. Berlin: Language Science Press, <https://langsci-press.org/catalog/book/415>

Kuznecova, Natal'ja Viktorovna. 2009. Fonologičeskije sistemy ižorskih dialektov [Phonological systems of Ingrian dialects]. Diss. ...kand. filol. nauk [Candidate thesis], Sankt-Peterburg: Institut lingvističeskih issledovanij RAN [Institute for Linguistic Studies, Russian Academy of Sciences].

Kuznetsova, Natalia, Balanovsky, Oleg, Muslimov, Mehmed & Kuleshov, Viacheslav. 2022. Contact of Ingrians and Votes in the light of linguistic, archaeological and genetic data. Oral presentation at the Congressus XIII Intenationalis Fenno-Ugristarum, Aug 21-26, 2022, Vienna, August 24, <https://cifu13.univie.ac.at/programme/symposia/e2-the-interdisciplinary/>

Kuznetsova, Natalia, Markus, Elena & Muslimov, Mehmed. 2015. Finnic minorities of Ingria: the current sociolinguistic situation and its background. In *Cultural and linguistic minorities in the Russian Federation and the European Union: Comparative studies on equality and diversity* [Multilingual Education 13], Heiko F. Marten, Michael Rießler, Janne Saarikivi & Reetta Toivanen (eds), 127-67. Berlin: Springer, https://doi.org/10.1007/978-3-319-10455-3_6

Markus, Elena & Rozhanskiy, Fedor. 2017. *Sovremennyj Vodskij Jazyk. Teksty i Grammatičeskij Očerk. 2-e Izdanije, Ispravlennoje i Dopolnennoje*. Vol. 2. Sankt-Peterburg: Nestor-Istorija.


Figure 1. Soikkola Ingrian consonants

| manner \ place | | labial | dental | palatal(ised) | velar |
|---|---|---|---|---|---|
| stops | | **p p**ˑ **p**: | **t t**ˑ **t**: | tʲ tʲˑ tʲ: | **k k**ˑ **k**: |
| fricatives | | f fˑ f: | **s s**ˑ **s**: | sʲ sʲˑ sʲ: | **h h**ˑ **h**: |
| sono-rants | nasals | **m m**ˑ **m**: | **n n**ˑ **n**: | nʲ nʲˑ | |
| | laterals | | **l l**ˑ **l**: | lʲ lʲˑ | |
| | thrills | | **r r**ˑ **r**: | rʲ rʲˑ | |
| affricates | | | | t͡ʃʲ t͡ʃʲ: | |
| glides | | **v v**ˑ **v**: | | **j j**ˑ **j**: | |

Figure 2. Lower Luga Ingrian consonants

| manner \ place | | labial | | dental | | post-alveolar | velar | |
|---|---|---|---|---|---|---|---|---|
| | | plain | palatalised | plain | palatalised | | plain | palatalised |
| stops | voiceless | **p p**: | *pʲ pʲ:* | **t t:** | tʲ tʲ: | | **k k**: | *kʲ kʲ:* |
| | voiced | **b** b: | | **d** d: | dʲ dʲ: | | **g** | |
| fricatives | voiceless | f f: | *(fʲ) fʲ:* | **s s**: | sʲ sʲ: | ʃ ʃ: | **h h**: | *hʲ (hʲ:)* |
| | voiced | | | **z** z: | *(zʲ) zʲ:* | ʒ ʒ: | | |
| sono-rants | nasals | **m m**: | *mʲ mʲ:* | **n n**: | nʲ nʲ: | | | |
| | laterals | | | **l l**: | lʲ lʲ: | | | |
| | thrills | | | **r r**: | rʲ *rʲ:* | | | |
| affricates | | | | *(t͡s) t͡s:* | t͡sʲ *t͡sʲ:* | t͡ʃʲ t͡ʃʲ: | | |
| glides | | **v v**: | *vʲ vʲ:* | | **j j**: | | | |

Notes to fonts used in Figures 1-2: **bold non-italic** — cardinal phonemes; non-bold non-italic — marginal but doubtless phonemes; *italic without parentheses* — attested marginal phonemes with unclear status; *(italic in parentheses)* — marginal phonemes with unclear status, which are potentially possible but unattested in actual data.

# Marginal vowels: The Romani sample

Márton A. Baló
(HUN-REN Research Institute for Linguistics; Eötvös Loránd University)

Keywords: Romani, language contact, unrounding, vowel length, phonemic status

Romani is a New Indo-Aryan language spoken primarily in Europe (Baló 2022), but also elsewhere in the world, such as the Americas or Australia (cf. e.g. Hancock 2013 and Sayer 2017, respectively). Its core vowel inventory consists of front unrounded /i/ and /e/, open central /a/ and back rounded /u/ and /o/. However, additional vowels have emerged in various Romani dialects, which can be divided into two groups: vowels coming from contact languages and appearing in loanwords only, and vowels emerging through various contact-induced modification processes and replacing or complementing core vowels in the inherited, Indo-Aryan component of the lexicon (Baló 2020).

The first group of vowels, appearing in loanwords only, varies from dialect to dialect, with the most common vowels including /y/ and /ø/, as well as /œ/, cf. e.g. Bodnárová (2015, p. 92) for Vend, a Central Romani variety. Naturally, other varieties, where the vowel system of the contact language includes these vowels, also borrow them, like köprüs 'bridge' ~ [ˈkøprys] from Turkish köprü in Xoraxani, šöro 'beer' ~ [ˈʃøro] from Hungarian sör in Romungro, or tjüteti 'town' ~ [tʲyˈteti] from Albanian qytet in Mečkari. As we see in Vekerdi (1985), and as Elšík et al. (1999, p. 309) also remark, a strategy of unrounding to adapt these vowels is sometimes used; therefore, Hungarian felhő 'cloud' and büntet 'punish' may become [felˈheːvo] and [bintetisar-ˈ], alongside [felˈhøːvo] and [ˈbyntetin-].

Examples of the the second group include centralised /ɨ/and /ə/ in Kalderaš, a Vlax Romani dialect (Boretzky 1991). These stem from Romanian, and besides appearing in Romanian loans, they are also transferred onto inherited words, replacing /i/ and /e/ under certain conditions. Determining whether they can be considered distinct phonemes is difficult and hinges on such diverse factors as the classification of the lexicon and the uncertainties of phonetic realisation.

While no phonemic status is attributed to vowel length in Romani in general, and the emergence of long or lengthened vowels is not a historically inherited feature (see already Miklosich 1872–1880, pp. ix, 24; also cited in Matras 2002, p. 59), vowel lengthening does take place in several varieties of Romani, often under the influence of contact languages. Thus, for instance, Finnish Romani abounds in long vowels, particularly in the first vocalic position of the word, often accompanied by a shift of stress to the first syllable, e.g. pani 'water' ~ [ˈpaːni] instead of [paˈni]. Recent studies (Elšík 2022; Bodnárová & Wiedner 2022) also show that some varieties have developed phonemic vowel length.

In this paper, I would like to give an overview of some of the most prominent features of the marginal vowels that are found in Romani with special regard to varieties in contact with Hungarian, such as Southern Central Romani and Vlax Romani dialects, while also trying to explore the continuum along which the marginality of these phonemes lie and address the question whether and how phonemic status itself may be considered a gradual concept.

**References**

Baló, Márton A. (2020), Romani Phonology, in Y. Matras, and A. Tenser (eds), (2020), *The Palgrave Handbook of Romani language and linguistics*, Cham: Palgrave Macmillan, 119–153.

Baló, Márton A. (2022), The Role of the Romani Language in the Preservation of Romani Identity in Hungary, in P. Ronan and E. Ziegler (eds), (2022), *Language and Identity in Migration Contexts*, Oxford: Peter Lang, 141–162.

Bodnárová, Zuzana (2015), *Vend Romani: Grammatical description and sociolinguistic situation of the so-called Vend dialects of Romani*, PhD dissertation, Prague: Charles University.

Bodnárová, Zuzana and Jakob Wiedner (2022), Development of phonemic length alternation in adjectives and possessive pronouns in South Central Romani, *Romani Studies* 32(2), 221–245.

Boretzky, Norbert (1991), Contact-induced sound change, *Diachronica* 8, 1–15.

Elšík, Viktor (2022), Vowel length in Selice Romani: Phonology, morphophonology, and diachrony, *Romani Studies* 32(2), 185–219.

Elšík, Viktor, Milena Hübschmannová, and Hana Šebková (1999). The Southern Central (ahi-imperfect) Romani dialects of Slovakia and Northern Hungary. In D. W. Halwachs and F. Menz (eds), (1999), *Die Sprache der Roma: Perspektiven der Romani-Forschung in Österreich im interdisziplinären und internazionalen Kontext*, Klagenfurt: Drava, 277–390.

Hancock, Ian (2013), The schooling of Romani Americans. An overview, In M. Miskovic (ed), *Roma education in Europe: Practices, policies and politics*, New York: Routledge, 87–100.

Sayer, Mandy (2017), *Australian Gypsies, Their Secret History*, Sydney: NewSouth Publishing.

Vekerdi, József (1985), *Cigány nyelvjárási népmesék*, Debrecen: Kossuth Lajos Tudományegyetem.

Matras, Yaron (2002), *Romani: A linguistic introduction*, Cambridge: Cambridge University Press.

Miklosich, Franz (1872–1880), *Über die Mundarten und Wanderungen der Zigeuner Europas*, Vienna: Karl Gerold's Sohn.

# WS10 Mass nouns in a typological perspective

# Functions of plural forms for mass nouns in Evenki

Elena Rudnitskaya & Elena Klyachko
(Institute of Oriental Studies RAS & HSE)

Pluralization tends to have special functions for mass nouns denoting abundance, sort recategorization or distributive plurality (Corbett 2000; Nicolas 2008). In this paper, we will use text corpora to study the effects of pluralization on the mass nouns of Evenki, an endangered Tungusic language spoken in Russia and China.

In Evenki, mass nouns like *mu:* 'water' or *ullə*/*uldə* 'meat' can be used as singulars ((1), (2)):

(1) *ə:gda-t* *ə-ŋnə-Ø-nnə* ***mu:-jə*** *əmə-w-rə*
what-INSTR NEG-HAB-NFUT-2SG water-PART come-TR-CVB.NEG
'Why don't you bring water?'

(2) *hila-t-na* *dʲəwu-fki:* *tar* ***uldə-jə*** *əmə-w-rəki-n*
barbecue-DUR-CVB eat-PTCP.HAB that meat-PART come-TR-CVB-3SG
'(The monster) used to eat barbecuing that **meat** when he brought (it)'

However, plural forms are also possible. For instance, *burduki-l* (flour-PL) can be used when talking of multiple bags with flour (distributive plurality) or about large amounts of flour (abundance plurality) (elicitation), whereas, *ari:-l* 'oil-PL' can mean 'different types of oil'. Other effects of pluralization on mass nouns in Evenki include similative plurality or imprecise reference.

Nevertheless, some examples cannot be interpreted as belonging to one of the types mentioned above:

(3) *magazin-du:* *zdaj-tʃa:-wun* *ollo-l-wo* ***ullə-l-wə*** *tʃipka:-r-wə*
shop-DAT deliver.R-PST-1PL.EXCL fish-PL-ACC meat-PL-ACC sable-PL-ACC
'We delivered the fish, the **meat**, and the sables to the shop'

(4) ***imu:rə-ri-n*** *nurgi-dʲa-ra-Ø*
fat-PL-3SG.POS drip-IPFV-NFUT-3PL
'**Her fat** is dripping'

(5) *e:da* ***ullə-l*** *mana-w-dʲə-rə-Ø*
why meat-PL finish-PASS-IPFV-NFUT-3PL
'Why are we running out of meat (lit. Why is **meat** ending) ?'

These cases are most probably related to the category of plural mass nouns (Alexiadou 2011, Corbett 2019: 82). Does Evenki have non-obligatory plural mass nouns?

The table below shows the frequencies of several Evenki nouns forms in the corpora (Kazakevich et al. 2023, Däbritz & Gusev 2021, Shakhovtsev et al. 2017): the first number is singular and the second one is plural. The results show that the nouns are at different points along the plurality scale, with none of them being either pluralia or singularia tantum.

| | Northern | Southern |
|---|---|---|
| Meat | 48/21 | 18/11 |
| Fish | 38/38 | 8/18 |
| Food | 45/14 | 3/6 |
| Berry | 1/9 | 2/6 |

| | | |
|---|---|---|
| Flour | 6/4 | 8/1 |
| Resin | 14/3 | 0/1 |
| larch resin | 6/3 | 0/0 |
| Blood | 9/3 | 4/3 |

Some contexts endorse singular or plural use. In most caritive contexts, a singular form is used ((6)):

(6) *arkạ* *taptịra*ː*-nə-n* *hojo-wị*
Arka wipe-PTCP.PF-3SG forehead-ACC.RFL
*e*ː*-matʃ͡* ***ʃə*ː*kʃə-jə*** *a*ː*tʃ͡ịn* *hụjạ* *a*ː*tʃ͡ịn* *o*ː*-da-n*
what=FOC **blood-PART** NEG.CAR wound NEG.CAR become-NFUT-3SG
'Arka wiped his forehead, there is not blood, there is no wound' (Southern dialect, INEL)

On the contrary, a form with the attenuative marker or an intensifier is often pluralized ((7), (8)):

(7) *tadu:* ***oldo-ko-r-ə*** *d'awa-ŋnə-rə-w*
there **fish-ATTEN-PL-ACC** catch-HAB-NFUT-1PL.EXCL
'We used to catch fish there'

(8) *kələ…* ***nʲutə-maːtə-l-wa*** *namaru-dʲa-kaldu*
INTJ… resin-INTS-PL-ACC stick-IPFV-IMPER.2PL
'Stick the resin!'

According to the definition in (Plungian 2003: 107), number is a non-obligatory grammatical category in Evenki; for instance, both singular and plural forms are possible with numerals (Boldyrev 2007: 409-410). We suggest that we still face plural mass nouns in Evenki, whereas the variation of plural and singular forms of mass nouns is related to the non-obligatory character of the number category.

Abbreviations: ATTEN – attenuative, CVB - converb, DAT - dative, DUR - durative, EXCL - exclusive, FOC - focus, HAB - habitual, IMPER - imperative, INSTR - instrumental, INTS - intensifier, IPFV - imperfective, NEG - negation, NFUT – non-future, PART - partitive, PASS - passive, PF - perfective, POSS - possessive, PST - past, PTCP - participle, R - Russian, RFL - reflexive, SG - singular, TR – transitivizer.

**References**


Alexiadou, Artemis (2011), *Plural mass nouns and the morpho-syntax of number*. In M. Byram Washburn, K. McKinney-Bock, E. Varis, A. Sawyer, & B. Tomaszewicz (Eds.), *Proceedings of the 28th West Coast Conference on Formal Linguistics* (pp. 33–41), Somerville: Cascadilla Proceedings Project, 33-41.

Boldyrev, Boris (2007), *Morfologiya evenkiyskogo yazyka [Evenki morphology]*, Novosibirsk: Nauka.

Däbritz, Chris Lasse & Gusev, Valentin (2021), *INEL Evenki Corpus* (≈47 000 words). Version 1.0. Publication date 2021-12-31. Archived at University of Hamburg. https://hdl.handle.net/11022/0000-0007-F43C-3 . In: The INEL corpora of indigenous Northern Eurasian languages. https://hdl.handle.net/11022/0000-0007-F45A-1

Corbett, Greville G (2000), *Number*, Cambridge University Press.

Corbett, Greville G (2019), *Pluralia tantum nouns and the theory of features: a typology of nouns with noncanonical number properties*, Morphology, 29 (2019): 51-108.

Kazakevich Olga et al (2023), *Evenki corpus* (≈35 000 words). Publication date 2023. At https://gisly.net/corpus

Nicolas, David (2008), *Mass nouns and plural logic*, Linguistics and philosophy 31 (2008): 211-244.

Plungian, Vladimir (2003), *Obshchaya morfologiya. Vvedenie v problematiku [General morphology. Introduction]*, Moscow: Editorial URSS.

Shakhovtsev, Kirill, Funk, Dmitry, Mamontova, Nadezhda, & Terekhina, Alexandra (2017), *Korpusy IEA RAS (Corpora of IEA RAS)* (≈ 121 000 words). Publication date 2017. At https://corpora.iea.ras.ru/corpora/

# Is there a unique category of transnumeral nouns in Foodo (Guang)?

Fiedler, Ines
(Humboldt-Universität zu Berlin)

Keywords: gender, transnumerals, mass nouns, Niger-Congo languages, morphology

This paper aims at analyzing the morphosyntactic features of transnumeral nouns in Foodo, a Guang language spoken in Benin. Foodo has a very productive, elaborated system of nominal classification, consisting of 10 nominal form classes and 10 agreement classes (Plunkett 2009). Nearly all of these classes can host transnumeral nouns (Table 1). The investigation is based on the lexical database compiled by Gray Plunkett from 1994 to 2007. It consists of about 2.300 entries, with around 1.500 nominal concepts, and will be accompanied by field notes of the author.

| **AGR class** | | **Default class number value** | **NF class** | **Number of nouns** | **Example** | **Gloss** |
|---|---|---|---|---|---|---|
| class 1 | o | SG, TN | O-.-O | 10 | *ò-yúl-ó* | shadow |
| | | SG, TN | Ø | 61 | *káwá* | beauty |
| class 2 | ba | PL, TN | A-.-SUF | 3 | *à-nyɔ́-ánà* | twin(s) |
| class 3 | kU | SG, TN | KU-.-U | 24 | *kʊ́-cé-ʊ̀* | rainy season |
| class 4 | a | PL, TN | A-.-SUF | 25 | *à-dá-ḿ* | blood |
| class 5 | dU | SG, TN | DI-.-DI | 24 | *dì-cíígbé-lí* | filth |
| class 6 | yU | PL, TN | I-.-SUF | 19 | *ì-ŋmá-m* | hair |
| class 7 | ka | SG, TN | KA-.-A | 35 | *ká-néέy↓á-á* | poverty |
| class 8 | sU | PL, TN | N-.-SUF | 6 | *m̀-bɔ́-ḿ* | money |
| class 9 | bU | TN | | 31 | *ǹ-cɔ́-ḿ* | water, rain |
| | | | BU-.-SUF | 400+ | *bʊ̀-kʊ́lá-á* | gathering |
| class 10 | tU | TN | Ø, A-.-SUF | 7 | *ʊ̀kùtú-tɔ́-ɔ́* | boisterousness |

Table 1: Distribution of transnumeral nouns over agreement classes in Foodo (based on Plunkett (2007)

The paper will first analyze how transnumeral nouns, i.e. nouns without inherent number differentiation, fit morphosyntactically into the nominal classification system of the language. Second, it will investigate how the great distribution over all classes can be explained, and whether there is a common semantic denominator unifying nouns differentiated for number and transnumeral nouns, so that both types of nouns can be assigned to one and the same class. This is especially interesting for class 5 nouns, as this class contains a good number of transnumeral nouns, but serves at the same time as class unitizing masses, as in (1), thus resulting in nominal triplets:

(1) *ì-ŋmá-m* ‘hair’ (mass noun)
dì-ŋmáḿ.bí-lí / à-ŋmáḿ.bé-é ‘single hair / single hairs’

Further, on basis of Grimm’s (2018) ‘scale of individuation’ it will be explored how the transnumeral nouns in Foodo can be semantically classified, e.g. as mass nouns, collectives, abstract nouns, verbal nouns etc., and whether there is a direct reflex of this classification in their morphosyntactic marking.

This in-depth analysis will allow for a better understanding of similar observations in other Guang languages as well as languages of the wider Niger-Congo family, and will shed light on the original semantics of noun classes and the historical developments that might underlie the assignment of nouns.

**References**

Grimm, Scott (2018), Grammatical Number and the Scale of Individuation?, *Language* 94, 527- 574.

Plunkett, Gray C. (2007), Foodo: Toolbox lexicon project, (Ms.).

Plunkett, Gray C. (2009), An overview of Foodo, a linguistic island in Benin, *Journal of West African languages* 36, 107-138.

# One fish, two fish, less fish, more fish:
# An experimental investigation of Arabic count and collective nouns

Connor Caston Rouillier
(The Ohio State University)

Keywords: Collectives, Pluratives, Number, Derivation, Arabic

Count nouns (see Table 1) and mass nouns (see Table 2) are often treated as a rigid morphological dichotomy (Grimm, 2018). However, in Arabic, collective nouns (see Table 3) are discussed as morphologically distinct from the other two noun types. These collective nouns have multiple plural forms: the base form, the plurative, and the plural of kind. This paper investigates the relationship between these collective nouns and count nouns. I conducted an experiment exploring how the number of entities seen by participants affects their selection of morphological form. Based on the results, I argue that the selection of the plurative form is not strictly based on number like the singular or dual and that the same cognitive pressures affect the selection of word form in the collective nouns as in the count nouns.

| Singular | Dual | Plural |
|---|---|---|
| zawdʒ | zawdʒ-aani | ʔazwaadʒ |
| husband.M.SG | husband.M-DU | husband.M.PL |

Table 1: Three forms of a count noun.

| Singular |
|---|
| θaldʒ |
| snow.M.SG |

Table 2: The form of a mass noun.

| Base | Singulative | Dual | Plurative | Plural of Kind |
|---|---|---|---|---|
| samak | samak-ah | samak-at-aani | samak-aat | ʔasmaak |
| fish.M.PL | fish-F.SG | fish-F-DU | fish-F.PL | fish.M.PL |

Table 3: Five forms of a collective noun.

The morphological relationship between the forms of the collective paradigm is unclear. Dali and Mathieu (2021) and Jaradat and Jarrah (2022) argue that these forms are inflectionally related to each other. These approaches claim that the singulative arises from a regular, productive process and acts like a classifier. Under this view, the collective paradigm has five forms contrasting with the count paradigm. However, Acquaviva (2008) and Kramer (2015) argue that the singulative form is derived from the base form. Acquaviva argues that a gender shift marks a derivational process, while Kramer's arguments question whether the singulative is productive as might be expected of inflection. Through derivation, the collective paradigm splits into two paradigms of three forms, aligning more closely with the paradigm of the count nouns.

In a forced-choice task, participants were asked to select between two morphological forms (base and plurative for collectives; singular and plural for count) when shown pictures with different numbers of entities. I found that, when shown fewer entities, participants were more likely to select the plurative and plural forms, aligning the plural and plurative with individuation. Singular and base forms were selected more when participants saw higher numbers of entities, aligning these forms with a mass interpretation.

Further, there seems to be an association between masculine singularity and mass interpretation, which can be seen in the base form of collectives and the form of mass nouns which are both masculine singular.

Taken together, these results question the dichotomy set up between these noun types, following Grimm (2018) and others. Additionally, these results suggest a parallel in the cognitive mechanisms, such as subitization, for how Arabic speakers select the morphological form of count and collective nouns. Further, this parallel between count nouns and collective nouns supports the derivation of the singulative form from the base form consistent with Acquaviva (2008) and Kramer (2015).

**References**


Acquaviva, Paolo. (2008). *Lexical plurals: A morphosemantic approach.* Oxford: Oxford University Press.

Dali, Myriam and Eric Mathieu. (2021). "Singulative systems." In P. C. Hofherr and J. Doetjes (eds), *The Oxford handbook of grammatical number.* Oxford: Oxford University Press, 275-290.

Grimm, Scott. (2018). "Grammatical number and the scale of individuation." *Language* 94(3), 527-574.

Jaradat, Abdulazeez and Marwan Jarrah. (2022). "The syntax of plurals of collective and mass nouns: Views from Jordanian Arabic." *Poznan Studies in Contemporary Linguistics* 58(3), 509-539.

Kramer, Ruth. (2015). *The morphosyntax of gender.* Oxford: Oxford University Press.

# Why the mass/count distinction is not morphologically marked: Evidence from Francoprovençal

Tabea Ihsane & Elisabeth Stark
(University of Geneva & University of Zurich)

Keywords: mass/count, morphosyntactic number, semantic number, partitive article, Francoprovençal

According to Gil's (1987) typology, Romance languages display a grammaticalized distinction of countability (Type 3). In this paper, we focus on Francoprovençal (FrPr), a highly endangered Gallo-Romance language, spoken by 110'000 to 160'000 speakers, in Italy, Switzerland, and France (Zulato, Kasstan & Nagy 2018:13), and demonstrate that, at least in Romance, the mass-count distinction is a *byproduct* of two oppositions related to number that are marked morphosyntactically: plural/singular (morphosyntactic number, usually by inflection) and cumulative/atomic reference (semantic number, in the sense of a semilattice interpretation, cf. Link 1983, Heycock & Zamparelli 2005, usually by nominal determination). Our analysis relies on indefinite examples like (1), which can be either mass singular or indefinite plural: *də buli* thus means 'mushroom' or 'mushrooms'.

(1) *də* *buli* FrPr
DE mushroom
'mushroom' or 'mushrooms'

In examples like (1), the noun phrase is completely unmarked for the mass/count interpretation or for morphosyntactic number (no marking on D, nor on N). This absence of marking characterizes most nominal arguments of a majority of FrPr varieties (Ihsane, Winistörfer & Stark 2023) and contrasts with languages like French, which use, with mass nouns and indefinite plurals, a partitive article (PA), historically a combination of DE and the Latin ILLE, as illustrated in (2).

(2) a. *du* *pain* b. *des* *pommes* French
of.the.M.SG bread of.the.PL apples
'bread' 'apples'

In French, the mass/count opposition thus seems to be overtly marked on the determiner (in addition to morphological number), as *du* is restricted to mass nouns and *des* to count nouns. What we argue, however, is that these determiners encode morphosyntactic and semantic number and that the mass/count interpretation is only a byproduct of the combination of these two number features.

More precisely, we propose that the invariable DE in FrPr (*də* in (1)) and the DE component of the PA in French (cf. (2)) encode *semantic* number (cumulative vs. atomic reference, parallel to bare arguments in Spanish). This element combines with the definite article, which encodes morphosyntactic number (*du* = *de* 'of' +*le* 'the.M.SG'; *des* = *de* 'of' +*les* 'the.PL'). Therefore, Romance languages mark the threefold opposition 'count plural' – 'mass singular' – 'count singular' differently: they mark at least either the *morphological* category of number (e.g., sigmatic *-s* on nouns in Spanish) or the *semantic plural* (e.g., invariable DE in FrPr) or both (e.g., PAs in French). In other words, *du* 'of.the.M.SG' in (2a) marks the morphosyntactic singular and the semantic plural (via DE) and *des*

'of.the.PL' the morphosyntactic plural and the semantic plural (via DE). The mass interpretation thus results from the combination of the features [cumulative, singular], and the count interpretation from the combinations [cumulative, plural] (as in *des pommes* in (2b)) and [atomic, singular] (as in *une pomme* 'an apple').

Countability is therefore not grammaticalized in Romance, and there is no explicit marking for "mass". This suggests that some aspects of Gil's (1987) typology for Romance languages, in particular the grammaticalized distinction of countability (Type 3), needs to be revised.

**References**


Gil, David (1987), Definiteness, Noun Phrase Configurationality, and the Count-Mass distinction, in E. Reuland, and A. G. B ter Meulen (eds), (1987), *The Representation of (In)definiteness*, Cambridge: MIT Press, 254–269.

Heycock, Caroline and Roberto Zamparelli (2005), Friends and colleagues: Coordination, plurality, and the structure of DP, *Natural Language Semantics* 13, 201–270.

Ihsane, Tabea, Olivier Winistörfer and Elisabeth Stark (2023), Francoprovençal: a spatial analysis of 'partitive articles' and potential correlates in Swiss and Italian varieties, *Isogloss. Open Journal of Romance Linguistics* 9(1)/4, 1–30.

Link, Godehard (1983), The logical analysis of plurals and mass terms, in R. Bäuerle, C. Schwarze, and A. von Stechow (eds), (1983), *Meaning, use and interpretation of language*, Berlin: Mouton de Gruyter, 302–323.

Zulato, Alessia, Jonathan Kasstan and Naomi Nagy (2018), An overview of Francoprovençal vitality in Europe and Northern America, *International Journal of the Sociology of Language* 2018(249), 11–29.

# Agreement of mass nouns in Dargwa

Michael Daniel & Nina Sumbatova
(Collegium de Lyon & Russian State University for the Humanities )

The paper discusses agreement patterns associated with mass nouns in Dargwa, a branch of East Caucasian. In Dargwa, many – though not all – mass and abstract nouns are morphologically singular but treated as plural for the purposes of agreement. They are thus not *inflectional* pluralia tantum – because they do form morphological plurals – but rather *concordial* pluralia tantum (cf. Corbett 2019).

In Table 1a from Mehweb Dargwa, the singular of 'sour cream', *burt*, controls *b*-agreement, as expected of a non-human noun; but the noun *ni?* 'milk', also morphologically singular, controls *d*-agreement, typical for non-human plurals. At first glance, nouns like 'milk' can be seen as constituting a special gender. However, 'milk' and similar nouns also control plural agreement on demonstratives (non-cumulative with gender). This excludes the interpretation of *d*- as an exponent of singular. As a result, gender assignment is modeled as working separately for morphologically singular and plural forms of the same noun, placing both in the same non-human plural category, cf. Table 1b.

Principles of the distribution of Dargwa mass nouns between those treated as concordial plurals (e.g. 'blood', 'flour', '(hoar)frost' 'milk', 'rain', 'smell') versus concordial singulars (e.g. 'bile', 'sugar', 'snow', 'sour cream', 'light') are not understood (cf. singular and plural mass nouns in English in McCawley 1975); and some nouns may follow either of the two patterns. One aim of the study is to investigate the diachronic stability of the split, based on the data currently available from different Dargwa varieties (Mehweb, Kadar, Tanti and Itsari).

In addition to internal stability, data from Dargwa also shows an interesting diachronic connection. Cf. the gender paradigm in Standard Lak, Dargwa closest sister branch (Table 2). Gender 4 in Lak has the same exponence as concordial pluralia tantum in Dargwa (*d*- in both singular and plural), and contains most mass nouns and nomina abstracta (Zhirkov 1955: 22). But Gender 4 also contains other nouns, e.g. utensils and body parts (Kibrik and Kodzasov 1990). In Dargwa, the approach in Table 2b is primarily possible because, from the cross-linguistic typology of mass nouns, we know that these may indeed behave as plural nouns; we can expect them to be inherently plural for the purposes of agreement. This argument does not work for decomposing Gender 4 in Lak, because not all its members are names of substances or abstract nouns. On the other hand, as Zhirkov (1955: 72) notes, mass nouns (but again not all members of Gender 4) control plural agreement on demonstrative pronouns, which make them similar to mass nouns in Dargwa.

The two systems may thus be evolutionarily related one to the other. The challenge is to model the direction of the evolution between the two systems. Siemund (2006) suggests reanalysis of pronominal gender agreement into mass / count marking in Germanic and Romance. If one wants to link Lak and Dargwa systems, the opposite evolution seems more plausible, Dargwa style concordial pluralia tantum developing into a separate Gender 4 in Lak by attracting other nouns.

| | | | Agreement | | Agreement |
|---|---|---|---|---|---|
| Non-humans | ‘yoke’ | *duk’* (Sg) | *b-* | *duk’ane* (Pl) | *d-* |
| **Mass nouns (sg agreement in Sg)** | ‘sour cream’ | *burt* (Sg) | ***b-*** | *burtane* (Pl) | *d-* |
| **Mass nouns (pl agreement in Sg)** | ‘milk’ | *niʔ* (Sg) | ***d-*** | *niʔane* (Pl) | *d-* |

Table 1a. Patterns of agreement of mass nouns in Mehweb Dargwa

| | | | Agreement | | Agreement |
|---|---|---|---|---|---|
| Non-humans | ‘yoke’ | *duk’* (Sg) | *b-* | *duk’ane* (Pl) | *d-* |
| **Mass nouns (sg agreement in Sg)** | ‘sour cream’ | *burt* (Sg) | *b-* | *burtane* (Pl) | *d-* |
| **Mass nouns (pl agreement in Sg)** | ‘milk’ | | | *niʔ* (Sg)<br>*niʔane* (Pl) | *d-* |

Table 1a. Patterns of agreement of mass nouns in Mehweb Dargwa (reconsidered)

| | Singular | Plural | |
|---|---|---|---|
| 1 (M) | Ø- | b- | Pl |
| 2 (F) | d- | | |
| 3 | b- | | |
| 4 | d- | d- | NPl |

Table 2. Gender paradigm in Lak

**References**


Corbett, Greville G. 2019. Pluralia tantum nouns and the theory of features: a typology of nouns with non-canonical number properties. *Morphology* 29. pp. 51–108. https://doi.org/10.1007/s11525-018-9336-0

Kibrik, Aleksandr, and Sandro Kodzasov. 1990. Sopostavitel’noe izučenie dagestanskix jazykov. Imja. [Comparative study of the languages of Daghestan. The Noun]. Moscow: MSU.

McCawley, James D. 1975. Lexicography and the Count-Mass Distinction. *Proceedings of the First Annual Meeting of the Berkeley Linguistics,* pp. 314-321

Siemund, Peter. 2006. Reanalyzing gender markers as mass/count agreement. In: Elżbieta Górska and Günter Radden (eds.) *Metonymy-Metaphor Collage*. Warsaw: Warsaw University Press, pp. 223-254.

Zhirkov, Lev. 1955. Lakskij jazyk. Fonetika i morfologija [Lak. Phonetics and morphology]. Moscow: AN SSSR.

# Estonian and Ukrainian derived mass nouns and object case

Olena Andrushenko, Lesia Chaika, Natalia Lehka, Anne Tamm & Denys Teptiuk
(Universität Augsburg, Universität Augsburg, Károli Gáspár University of the Hungarian Reformed Church, Károli Gáspár University of the Hungarian Reformed Church & University of Tartu)

Keywords: object case, aspect, verbs, genitive, partitive

The paper discusses scalar features in deadjectival derivation. The focus is object case in Estonian and Ukrainian. We want to demonstrate that deadjectival mass nouns do not form a uniform group within mass nouns, exploring the gradability of the nouns in question.

Several analyses regard gradability not only as a property of adjectives, but of nouns, verbs, adverbs, and prepositions, and the scalar properties of one category may persist in derivation (Hay et al 1999). These analyses underline semantic features beyond quantization, which captures the count/mass and telic/atelic parallels.

Elicitation, introspection, and corpora are used to clarify the data in the previous Ukrainian literature on mass and abstract nouns, aspect, and aspect (Vyxovanec' 1992: 120, Ševčuk 2010: 131, Pljušč 2005: 87, 107, Šypovyč, Ihnatolja, and Dančenko 2020) to highlight the parallels between Ukrainian and Estonian as in the following data, where we see an accusative/genitive alternation that parallels Estonian (in Chaika et al 2024).

(1) Ukrainian

a. *V kimnati bulo xolodno. Ya prynesla dyvan, lampu i prosti štory,
i tak stvoryla zatyšku
… created cosy.feeling.**GEN**
Intended to mean:
‘The room was cold.
I brought a sofa, a lamp and simple curtains and so I created cosyness.’

b. … і так створила затишок
… i tak stvoryla zatyšok
… created cosy.feeling.**ACC**
‘The room was cold.
I brought a sofa, a lamp and simple curtains and so I created cosyness’

c. V kimnati bulo xolodno. Ya prynesla dyvan, lampu i prosti štory,
i tak dodala zatyšku
… added cosy.feeling.**GEN**
‘The room was cold.
I brought a sofa, a lamp and simple curtains and so I added cosyness’
(more of an existing property was added)

d. i tak dodala zatyšok
… added cosy.feeling**.ACC**
‘The room was cold.
I brought a sofa, a lamp and simple curtains and so I added cosyness’
(it was not cosy before, a new property was added)

Relations between aspect and object case are central topics in Ukrainian as well as Estonian linguistics (cf Metslang 2001 a.o. for Estonian). However, a comparison between them, how abstract mass nouns (and plurals) interact with aspect as well as case has not been explored from the angle of recent partitivity literature (Luraghi, Sleeman 2023; Chaika et al 2024).

**References**


Chaika, Lesia., Lehka, Natalia, Tamm, Anne, & Vaiss, Natalia (2024), Ukrainian aspect and object case in ukTenTen. In eds. Glaser, Elvira; Sleeman, Petra; Strobel, Thomas; Tamm, Anne *Partitive constructions and partitive elements within and across language borders in Europe.* LiVVaL. Linguaggio e Variazione 3. Venice: Edizioni Ca' Foscari. Venice University Press. 173-223. https://edizionicafoscari.it//it/edizioni/libri/978-88-6969-795-1/ukrainian-aspect-and-object-case-in-uktenten/, doi 10.30687/978-88-6969-795-1

Hay, Jennifer.; Chris Kennedy; Beth Levin. (1999), Scale structure underlies telicity in 'degree achievements'.

Luraghi, Silvia; Petra Sleeman (eds.), (2023), *Partitives cross-linguistically. Dimensions of variation*.

Metslang, Helle. (2001), On the Developments of the Estonian Aspect: The Verbal Particle *ära*".

Pljušč, M. J. (2005), *Hramatyka ukrajins'koji movy. Morfemika, slovotvir, morfolohija.* [Ukrainian grammar. Morphemics, word formation, morphology], vol. 1.

Ševčuk, N. (2010), Hramatyčni osoblyvosti ob'jekta pry dijeslovax na poznačennja tvorennja. [Grammatical features of the object with verbs denoting creation].

Šypovyč, M. V.; Ihnatolja, A. A.; Dančenko, O. B. (2020), Semantyčni funkciji vidminkiv imennyka v ukrains'kij movi. [Semantic functions of noun cases in Ukrainian].

Vyxovanec', I. R. (1992), Narysy z funkcional'noho syntaksysu ukrajins'koji movy. [Essays on functional syntax in Ukrainian].

# On subtle differences between mass and count nouns in Seri (isolate, Mexico)

Robert Henderson & Jeremy Pasquereau
(University of Arizona & Laboratoire de Linguistique de Nantes)

In Seri, an isolate language spoken in northwest Mexico (Marlett, 2016), (non-possessed) nouns typically mark a sg/pl distinction by adding a lexically-determined plural suffix to the sg/unmarked form. The interpretation of plural forms is similar to that found in English (we leave out plural forms of so-called possessed nouns from our study as they involve a different kind of plurality marking (anonymized reference)).

By contrast, Seri mass nouns (O'Meara and Bohnemeyer, 2011; Marlett, 2016) do not show the same type of sg/pl alternations: they are either singular in form (*hesen* 'iron wood'), pluralia tantum (*xiica coopol* 'coffee'), or the sg/pl distinction some nouns make is not interpreted as it is with count nouns. For instance the word for sugar is either the singular *ziix cooatjö* or the plural x*iica cooatjö:* both forms can refer to an underspecified quantity of sugar, and both can be used with the sg indefinite zo. The (count) noun for 'tortilla' shows a similar alternation *ziix/xiica cxatlc* but the formally plural form *xiica cxatlc* cannot refer to a singular tortilla and be used with the sg indefinite zo.

Another difference with count nouns relates to quantity underspecification and agreement. While it is possible to underspecify quantity for mass nouns when it is not at issue or known using plural determiners (1a; also see Marlett 2016, 560), for count nouns using plural determiners entails a plurality (1b). Mass nouns in Seri are thus pluralia tantum in the sense that, barring a packaging/kind interpretation, they take plural agreement (Corbett, 2019).

(1) a. Mass nouns: plural agreement to underspecify quantity
If your father tells you:
**Xiica** **c-oosot-oj** **pac** c-amjc!
thing.pl sbj.nmlz-be_narrow-pl indf.pl imp-bring
*Bring rice!*
*And you bring him a plate of rice, is that ok?* Judg: Yes, it's fine. [EDSEI3JUL2023DRPM]

b. Count nouns: no way to underspecify quantity
If your father tells you:
**Xiica** **c-xat-lc** **pac** camjc!
thing.pl sbj.nmlz-be_flat-pl indf.pl imp-bring
*Bring me tortillas!*
*And you bring him one tortilla, is that ok?* Judg: No, he said 'pac', there should be more than one tortilla. [EDSEI3JUL2023DRPM]

Finally, while mass nouns are compatible with both the singular indefinite D *zo* and numerals (provided apportionment by kind or packaging is supported by context), numerals tend to impose the stricter requirement that the counted portions be as small as possible. By contrast no such requirement seems to hold of the identifIcation of a portion with the singular indefinite D. For instance, in (2a), *xiica coosotoj quih tazo* 'lit. one rice' is interpreted as 'one grain of rice' whereas in (2b), *xiica coosotoj zo* 'lit. a rice' is interpreted as one package of rice.

(2) a. Context: there's a grain of rice on your fork

**Xiica coosotoj quih tazo** eenim quinláz cap
thing.pl sbj.nmlz.be_narrow.pl det one metal sbj.nmlz.have_fingers def.sg
iti quiij quih hant yaait.
[3.poss]on sbj.nmlz.sit det land rls.yo.fall
*A grain of rice that was on my fork fell.*" [EDSEI2JUL2023DRPM, ELAB]
*And you bring him a plate of rice, is that ok*? Judg: Yes, it's fine. [EDSEI3JUL2023DRPM]

b. If your father tells you:

**Xiica c-oosot-oj zo** c-amjc!
thing.pl sbj.nmlz-be_narrow-pl indf.sg imp-bring
*Bring rice!*
*What would you bring him?* Judg: A bag of rice. [EDSEI3JUL2023DRPM]

The talk will elaborate on these generalizations and others relating to agreement/concord against the background of number marking in Seri and of properties of mass nouns across languages.

**References**

Corbett, G. G. (2019). Pluralia tantum nouns and the theory of features: A typology of nouns with non-canonical number properties. Morphology 29, 51–108.

Marlett, S. A. (2016). Cmiique Iitom: the Seri language. Unpublished grammar (2016 draft).

O'Meara, C. and J. Bohnemeyer (2011). Mass-count coercion with the Universal Sculptor in Seri. In A. Black and M. Louie (Eds.), Papers for WSCLA 16: The Sixteenth Workshop on Structure and Constituency in Languages of the Americas, pp. 121–134. UBC working papers in linguistics.

# Mass Nouns in Old English & the Roots of the Category of Countability in English

Ondrej Tichy & Vojtech Ripl
(Charles University & Charles University)

While the workshop's proposal notes that the research on mass nouns has been "focused on data from English in a clearly disproportionate way", this is only true for Present-day English. The research into the roots and the development of the category described by Quirk as "basic in English" (1985: 247) has been perhaps as limited as its research in non-SAE languages. In 1998 Denison (Denison in Romaine, ed., 1998: 96) noted that "[i]t seems very likely that there is a systematic process of change from noncount to count for some nouns, but full evidence is not yet available". Such evidence is, in fact, missing for most of the history of this category in English all the way to its roots in OE. The few studies that have covered the topic (mostly in a fragmentary way or in passing, e.g. Strang 1970, Görlach 1991, Fischer 1992, Lass 1992, Rozumko 2002, Sinkko-Latvala 2009, Koike 2004, Kharlamenko 2020, Tichý 2022), have sometimes arrived at seemingly widely different conclusions. While Marckwardt (1970) has shown clear similarities in distributional preferences of *much* and *many* between OE and PDE, Ackles (1997) and Toyota (2009) have concluded respectively that OE did not have any mass/count distinction or that OE was a classifier language.

In this paper, we would like to look more specifically at the status of mass nouns in OE and, more generally, into its further development in view of the typological shift in the history of English. The corpus-based probe will use the data of the YCOE to analyze OE nouns that persist as mass, collectives and abstracts into PDE in interactions (informed by typological and comparative perspective) with classifiers, measuring and counting constructions, number congruence and partitive genitive.

Preliminary results e.g. in case distribution suggest that verbs like *on-fōn* (to take, receive) that have so far been simply described as taking objects in acc., gen. and dat. show a clear preference for mass/abstract nouns (e.g. *hlāf* "bread", *wǣt* "fluid, moisture" or *deað* "death"), in gen. but no clear evidence in acc. C.f. the construction with genitive as in:

"[fæt] … þæt ðæs wynsuman wætan_GEN þær onfen" ([a vessel] … that collected the wonderful fluid)[1]

and with accusative as in:

"[he] onfeng Godes word_ACC" ([he] received God's word).

This distribution can be seen as analogical to those in other languages such as Czech "dostal piva/pivo" ([he] received [a] beer) with the mass noun being grammatical both in gen. and acc. while "dostal stůl/!stolu" ([he] received a table) with the count noun is only grammatical in acc.

[1] All examples are excerpted from YCOE.

Since this paper is an initial stage of a larger incipient project mapping the development of countability in HEL, comparisons are drawn to the development of the category in later English (and other languages such as Czech) stressing the dynamic and changeable nature of the phenomenon (c.f. the contemporary usage of *less* and *fewer*).

**References**


Ackles, N. M. (1997). *Historical syntax of the English articles in relation to the count/non-count distinction*. University of Washington, PhD Thesis.

Allan, K. (1980). Nouns and Countability. *Language*, *56*(3), 541–567.

Baldwin, T., & Bond, F. (2003). A plethora of methods for learning English countability. *Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing*, 73–80. Association for Computational Linguistics.

Brinton, L. J. (1998). Aspectuality and countability: a cross-categorial analogy. *English Language and Linguistics*, *2*(1), 37–63.

Bosworth, Joseph. An Anglo-Saxon Dictionary Online, edited by Thomas Northcote Toller, Christ Sean, and Ondřej Tichy. Prague: Faculty of Arts, Charles University, 2014. https://bosworthtoller.com/.

Corbett, G. G. (2000). Number. In *Cambridge Textbooks in Linguistics*. Cambridge: Cambridge University Press.

Downing, A., & Locke, P. (2002). *A University Course in English Grammar*. Routledge. Huddlestone, R., & Pullum, G. K. (2002). *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

Kemenade, A. van, & Vincent, N. (1997). *Parameters of morphosyntactic change*. Cambridge, U.K.; New York, NY: Cambridge University Press.

Kharlamenko, O. (2020). The expression of non-individual in some Old English Nouns. In R. P. Lorido, C. P. Alonso, & P. Rodríguez-Puente (Eds.), Of ye Olde Englisch Langage and Textes: New Perspectives on Old and Middle English Language and Literature (pp. 25–50). Peter Lang Verlag.

Rozumko, A. (2002). *Countable, uncountable and collective nouns in the early eighteenth century English - an overview*. Studia Anglica Posnaniensia: International Review of English Studies, 131–160.

Lass, R. (Ed.). (1992). The Cambridge History of the English Language Volume III: 1476–1776. Cambridge University Press.

Marckwardt, A. H. (1970). *Much and Many: The Historical Development of a Modern English Distributional Pattern*. In J. L. Rosier (Ed.), Philological Essays (pp. 50–54). De Gruyter Mouton.

Murphy, M. L. (2012). *Lexical meaning*. Cambridge: Cambridge University Press.

Fischer, O. (1992). Syntax. In N. Blake (Ed.), The Cambridge History of the English Language. Volume II. 1066–1476 (pp. 207–408).

Goddard, C. (2010) *A Piece of Cheese, a Grain of Sand: The Semantics of Mass Nouns and Unitizers.* In Pelletier, F. J. (2010). *Kinds, things, and stuff: mass terms and generics*. Oxford: Oxford University Press.

Görlach, M. (1991). Introduction to Early Modern English. Cambridge University Press.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. A. (1985). *Comprehensive Grammar of the English Language*.

Romaine, S. (Ed.). (1998). *The Cambridge History of the English Language. Vol. 4, 1776-1997* (5th print.). Cambridge: Cambridge University Press.

Rozumko, A. (2002). *Countable, uncountable and collective nouns in the early eighteenth-century English – an overview*. In *Studia Anglica Posnaniensia: International Review of English Studies*, 131–160.

Sinkko-Latvala, S. (2009). *A Study of the Countability of Some Usually Uncountable Nouns in British English from the 16th Century to the Present Day*. University of Tampere.

Strang, B. M. (1970). A History of English. Methuen.

Svensson, P. (1998). *Number and countability in English nouns : an embodied model*. Umeå University, Umeå.

Taylor, A., A. Warner, S. Pintzuk, and F. Beths. (2003). The York-Toronto-Helsinki Parsed Corpus of Old English Prose (YCOE).

Toyota, J. (2009). *When the Mass Was Counted: English as Classifier and Non-Classifier Language.* Skase Journal of Theoretical Linguistics, 6(1).

Tichý, Ondřej. (2022). *Countability in the History of English: Evidence of Grammars and Dictionaries*. Linguistica Pragensia. 32(1), 39–58.

Wierzbicka, A. (1988). *The semantics of grammar*. Amsterdam; Philadelphia: J. Benjamins Pub. Co.

# Mass nouns in locative and existential clauses in Siberian languages

Chris Lasse Däbritz
(Universität Hamburg)

Both locative and existential clauses prototypically describe the episodic presence or absence of a referent – the figure (a.k.a. theme, pivot) – at a ground (a.k.a. location, coda) (Hengeveld 1992: 119, Sarda & Lena 2023: 21). However, as commonly known, existential clauses serve to (re-)introduce a referent into the discourse, whereas locative clauses serve to predicate the location of an already established discourse referent. This entails the figure belonging to the focus domain in existential clauses, but being topical in locative clauses (ibid.). Thus, it becomes evident that the figure is a concrete and individualizable referent in the latter case (Creissels 2019: 41; Croft 2022: 317): To make the sentence *the books are on the table* pragmatically adequate, it must be clear for both speaker and hearer, to which books the speaker refers.

Given that mass nouns prototypically refer to substances or similar concepts, i.e. non-individualizable referents (Ghomeshi & Massam 2012), the question arises how mass nouns behave in existential and locative clauses. I argue that mass nouns appear without further restriction in existential clauses, but not in locative clauses. Following Goddard's (2010) approach to mass nouns and unitization, I hypothesize that mass noun referents receive a unitized reading in locative clauses, by default. In other words, *there is beer in the fridge* simply tells about the presence of beer as a substance, whereas *the beer is in the fridge* means that a certain perceivable unit (bottles, two liters etc) of the substance is located in the fridge.

This hypothesis is tested on natural language data from sixteen genetically diverse Siberian languages, the database containing 1,221 locative and 4,045 existential clauses. First, I observe that mass nouns appear significantly less frequently in locative clauses (35 clauses; 2.9%) than in existential clauses (453 clauses; 11.2%). Second, they always receive a unitized reading indeed, when occurring in locative clauses. The unitized reading can be acquired by possessive marking (1) (see also Aralova & Pakendorf 2023) or the usage of demonstratives (2), but it can also be formally unmarked, only being derived from the context (3).

(1) […] *da-s'íla-ur* *da-xúš'-kej-ɛʾ.*
3SG.M-power.R-water 3SG.M-birch.bark.tent-LOC-PRED.3.INAN
'[I know], his power water is in his birch-bark tent.'
(Yugh < Yeniseian; Werner 1997: 255)

(2) *gibər* *dĭ* *süt.*
where this milk
[I have a cow. It gives me milk.] 'Where is this milk [= the milk the cow gave that morning]?'
(Kamas < Uralic; Gusev et al. 2019)

(3) *a* *et* *taškïr* *pol-odu.*
and.R meat outside be-PRS.3SG
'[We store pickles in the cellar. We slaughter the livestock.] And the meat is outside.'
(Chulym Turkic < Turkic; ELAR: Filchenko 2016–2019)

These observations contribute to the wider typological discussion of mass nouns since they imply that mass nouns occurring in contexts, which call for reference to a concrete and individualizable entity, receive a unitized reading by default. Additionally, the analyzed data from the given Siberian languages strongly imply that locative clauses are such a context.

**References**


Aralova, Natalia & Pakendorf, Brigitte. 2023. Non-canonical possessive constructions in Negidal and other Tungusic languages: a new analysis of the so-called "alienable possession" suffix. *Linguistics* (ahead of print).

Creissels, Denis. 2019. Inverse-locational predication in typological perspective. *Italian Journal of Linguistics* 31(2). 37–106. (doi:10.26346/1120-2726-138)

Croft, William. 2022. *Morphosyntax: Constructions of the World's Languages*. Cambridge: Cambridge University Press.

Filchenko, Andrey. 2016. *Comprehensive documentation and archiving of Teleut, Eushta-Chat, and Melets Chulym: three areally adjacent critically endangered Turkic languages of Siberia*. (http://hdl.handle.net/2196/00-0000-0000-0010-8981-B)

Ghomeshi, Jila & Massam, Diane. 2012. The count mass distinction: Issues and perspectives. In: Massam, Diane (ed.), *Count and mass across languages*, 1–8. Oxford: Oxford University Press.

Goddard, Cliff. 2010. A piece of cheese, a grain of sand: The semantics of mass nouns and unitizers. In Pelletier, Francis J. (ed.), *Kinds, things, and stuff: Mass terms and generics*, 132–165. Oxford: OUP.

Gusev, Valentin & Klooster, Tiina & Wagner-Nagy, Beáta. 2019. *INEL Kamas Corpus*. Version 1.0. (http://hdl.handle.net/11022/0000-0007-DA6E-9)

Hengeveld, Kees. 1992. Non-verbal predication: Theory, typology, diachrony (Functional Grammar Series 15). Berlin, New York: Mouton de Gruyter.

Sarda, Laure & Lena, Ludovica. 2023. Existential constructions. In search of a definition. In Sarda, Laure & Lena, Ludovica (eds.), *Existential constructions across languages. Forms, meaning and function* (Human Cognitive Processing), vol. 74, 1–32. Amsterdam, Philadelphia: John Benjamins.

Werner, Heinrich. 1997. *Das Jugische (Sym-Ketische) [Yugh (Sym-Ket)]* (Veröffentlichungen der Societas Uralo-Altaica 50). Wiesbaden: Harrassowitz.

# Construing mass nouns in Supyire

Robert Carlson
(Africa International University; SIL)

Although the morphosyntax of the noun class system of Supyire (Senufo, Gur, Niger-Congo) has been extensively described (see Carlson 1994, 2022), mass nouns and related notions such as collectives have not hitherto been dealt with in an adequate manner. This paper describes how speakers use morphosyntactic constructions to construe masses, approaching the analysis from a cognitive-functional perspective (Langacker 2006, 2008, 20017; Harder 2010).

Nouns in Supyire, which has a typical Niger-Congo noun class system of the suffixing Senufo variety, can be viewed as constructions in which the suffix is normally the profile determinant: the noun class suffix can be viewed as symbolizing a highly abstract "thing" with a particular semantics. Thus, in addition to singular-plural pairs forming three "count" genders, as is common in Gur languages (cf. Miehe & Winkelmann 2007, Miehe et al. 2012, and Fiedler et al. 2022) there are two classes specializing in mass nouns: TI, specializing in solid masses, and PU, specializing in liquid and powder masses. The profile determinant role played by the suffix is especially clear in the case of nominalizations, where the suffix imposes its "thing" construal. For example, PU class *fùn-mɔ̀* 'sweat (n)' is from *fûn* 'sweat (v)', *fyere-mɛ* 'urine' is from *fyeere* 'urinate', and TI class *kya-ra* 'meat' is from *kya* 'eat by chewing'. Similarly, the TI class is extensively used to derive "collective" nouns, e.g. TI *ci-re* '(collection of) trees' has the same root as *ci-ge* 'tree' (KU) and *ci-ye* '(countable) trees' (YI).

The masses in the dedicated classes TI and PU do not by any means exhaust the inventory of mass nouns: there are just as many mass nouns in the "singular" classes U and KU, where they have no plural counterparts in the corresponding PI and YI classes. In U most of these are borrowed, such as *kafé* 'coffee' from French and *mɔ̀ɔ̀* 'rice' from Bambara. In KU there are many prototypical mass nouns such as *lwɔ-ʔɔ* 'water', *sìshyèn-gɛ̀* 'blood', and *yɔ̀ɔ̀-gɔ̀* 'mud'.

This paper examines several other constructions where mass status makes a difference. For example, a unitizing function is played by the diminutive suffix *-rV́* (cf. Nurmio 2023), which results in a noun in singular class LI: cyɛn-rɛ́ 'small amount of sauce' from KU *cyɛn-gɛ* sauce, and *kyā-rá* 'little piece of meat' from TI *kya-ra* 'meat'.

The "individualizing" universal quantifier construction N *màʔa* N (e.g. *cige máʔá cige* 'every tree') with mass nouns encourages a subtype construal: *lwɔʔɔ máʔá lwɔʔɔ* 'each type of water' (e.g. from a well, from a stream, from a watertower). Similarly the interrogative determiner construction (e.g. *cigé ŋ̀kìré yɛ?* 'which tree'): *lwɔʔé ŋ̀kìré yɛ?* 'which type of water'. The "indefinite" construction (e.g. *cigé kà* 'a (certain) tree'gives a partitive meaning with masses: *lwɔʔé kà* 'some water'.

These and other constructions are described in detail to show how speakers impose a variety of construals on mass nouns and related nouns such as collectives.

**References**


Carlson, Robert. 1994. *A grammar of Supyire.* Berlin: Mouton de Gruyter.

Carlson, Robert. 2022. Gender and noun classes in Supyire, in Fiedler et al., 15-44.

Fiedler, Ines & Junglas, Jan & Schulze, Michael (eds). 2022. *Gur languages and nominal classification diversity.* Köln: Rüdiger Köppe.

Harder, Peter. 2010. *Meaning in mind and society*: *a functional contribution to the social turn in cognitive linguistics.* Berlin: De Gruyter Mouton.

Langacker, Ronald. 2008. *Cognitive Grammar: a basic introduction.* Oxford: Oxford Univerity Press.

Langacker, Ronald. 2017a. *Ten lectures on the basics of Cognitive Grammar.* Leiden: Brill.

Langacker, Ronald. 2017b. *Ten lectures on the elaboration of Cognitive Grammar.* Leiden: Brill.

Miehe, Gudrun & Winkelmann, Kerstin (eds). 2007. *Noun class systems in Gur languages*: *Southwestern Gur languages (without Gurunsi).* Köln: Rüdiger Köppe.

Miehe, Gudrun & Reinicke, Brigitte & Winkelmann, Kerstin (eds). 2012. *Noun class systems in Gur languages*: *North central Gur languages.* Köln: Rüdiger Köppe.

Nurmio, Silva. 2023. Towards a typology of singulatives: an overview of markers. In Arbes, Deborah (ed), *Number categories: dynamics, contact, typology,* 155-181. Berlin: De Gruyter.

# Singulatives in Bantu languages

Antti Laine, Emma Kurkinen & Silva Nurmio
(University of Helsinki, University of Helsinki & University of Helsinki)

Keywords: Bantu languages, singulatives, mass nouns, collective nouns, noun classes

Bantu languages (Atlantic-Congo) are known for their noun class systems (Van de Velde 2019; Maho 1999): nouns belong to one of up to 19 classes governing agreement in both verbs and nominal modifiers. The classes, numbered according to cognacy for cross-Bantu comparability, form singular-plural pairs so that nouns in one class have their plurals in another (e.g. CL2 generally has plurals of CL1 nouns). However, not all individual pairings neatly follow this schema, and not all classes have a dedicated plural class. Some classes have dedicated derivational functions, such as diminutive, augmentative or locative.

Singulative functions of noun classes, particularly classes 5 (example 1) and 11 (example 2), are occasionally mentioned in descriptions of Bantu languages. We define singulatives as noun forms denoting one individual or one unit of a referent which typically occurs as a mass, collection or plurality; singulatives formed from mass nouns are particularly common typologically (Nurmio 2023; Grimm 2018; Acquaviva 2016). While Grégoire (1994) discusses singulative functions of CL11 as an areal development in certain northwestern Bantu languages, no Bantu-wide typological work on singulatives exists to date. Our paper aims to address this gap and contribute both to the typology of singulatives and mass nouns.

1. Eton (Van de Velde 2008)

| | | | |
|---|---|---|---|
| *mè-ndím* | 'water (6)' | *è-ndím* | 'a drop of water (5)' |

2. Fwe (Gunnink 2022)

| | | | |
|---|---|---|---|
| *bw-ékè* | 'grains (14)' | *rw-ékè* | 'a single grain (11)' |

The Fwe singulative seems to be quite productive (Gunnink 2022), but productivity is rarely discussed in other descriptions, even if singulativity is explicitly mentioned. Most commonly one finds listed examples of CL11 nouns with singulative-like meanings without any discussion of functions.

In our paper, we provide an overview of previous research mentioning singulatives in Bantu. We also present the first systematic cross-Bantu study on the expression of singulatives, based on a geographically balanced sample of 50 languages showing that singulatives are found across the Bantu area, particularly in CL11, but also others. We present a detailed survey of the functions and semantics of CL11 and CL5 and an overview of other classes which form singulatives less commonly.

Our results establish singulatives as a feature of Bantu languages. The bases from which singulatives are formed typically comprise those also found typologically, especially aggregates (e.g. hair, grass, grains, firewood), and liquids (water, blood). There are also examples of packaged meanings for mass referents that lack natural minimal units ('metal' > 'piece of metal'). Some languages show three-part paradigms, with collective, singulative and plural meanings:

3. Swahili (Contini-Morava 1999)

| | | | |
|---|---|---|---|
| *ma-nyasi* | ‘grass (collective, 6)’ | *u-nyasi* | ‘blade of grass (11)’ |
| *nyasi* | ‘blades of grass (10)’ | | |

We also take a diachronic perspective to see what historical inferences can be made about the marking of singulativity in our sample. While Grégoire (1994) argues that CL11 singulatives arise from a wider reorganization of the noun class systems in the northwestern region, we show that singulative functions of CL11 are far more widely attested and not dependent on such reorganization, suggesting that this may already have been a function of CL11 in Proto-Bantu.

**References**


Acquaviva, Paolo. 2016. Singulatives. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *HSK Word-Formation. An international handbook of the languages of Europe*, 1171–1183. Berlin: De Gruyter Mouton.

Contini-Morava. Ellen. 1999. Noun class and number in Swahili. In: Ellen Contini-Morava and Yishai Tobin (eds.), *Between Grammar and the Lexicon*, 3−29. Amsterdam/Philadelphia: Benjamins.

Grégoire, Claire. 1994. A Diachronic Approach to Classes 10 and 11 in Bantu with Special Reference to North-Western Languages. *Proceedings of the Twentieth Annual Meeting of the Berkeley Linguistics Society: Special Session on Historical Issues in African Linguistics.* 21–34.

Grimm, Scott. 2018. Grammatical number and the scale of individuation. *Language* 94(3), 527–574.

Gunnink, Hilde. 2022. *A grammar of Fwe*. Berlin: Language Science Press.

Maho, Jouni Filip. 1999. *A comparative study of Bantu noun classes*. (Orientalia et africana gothoburgensia, 13.) Gothenburg: Acta Universitatis Gothoburgensis.

Nurmio, Silva. 2023. Towards a typology of singulatives: An overview of markers. In: Deborah Arbes (ed.), *Number categories: Dynamics, contact, typology*, 155–181. Berlin: De Gruyter.

Van de Velde, Mark. 2008. *A Grammar of Eton*. Berlin: Mouton de Gruyter.

Van de Velde, Mark. 2019. Nominal morphology and syntax. In Mark Van de Velde, Koen Bostoen, Derek Nurse & Gérard Philippson (eds.), *The Bantu Languages* (2nd edn.), 237–269. London: Routledge.

# WS11 Mismatches in Information Structure

# A focus mismatch: Free vs. subcategorized indirect objects in Spanish

Florian Schäfer & Steffen Heidinger
(Humboldt Universität & University of Graz)

Key words: focus affinity, indirect objects, argument, adjunct, Spanish

Background. Grammatical roles (subject, objects, types of adjuncts) differ with respect to their focus affinity, i.e., their likelihood to be the focus of the sentence. Although this idea has circulated for quite some time (Contreras 1978, Sgall et al. 1986, Firbas 1992, Lambrecht 1994, Drubig 2003), systematic experimental studies have not been conducted until recently (Heidinger 2018, Heidinger & Onea 2021). One generalization from these experiments is that adjuncts exhibit a higher degree of focus affinity than arguments. This paper addresses the question of how robust this generalization is and whether it extends to different types of objects.

A focus mismatch. We present novel evidence from a forced-choice experiment on focus interpretation in Spanish. In this experiment, 48 native speakers were shown sentences combining a direct and an indirect object as (1) and (2). These sentences were presented together with paraphrases identifying either the direct or the indirect object as the narrow focus of negation. The participants' task was to indicate which of the two paraphrases corresponds to their preferred interpretation of the sentence.

The results show that the preferred focus interpretation depends on the syntactic status of the iO (cf. Table 1): In sentences combining a dO and a free iO (as in (1)), dO is interpreted more often as the narrow focus than iO; in sentences combining a dO and subcategorized iO (as in (2)), iO is interpreted more often as the narrow focus than dO. Hence, iOs with adjunct status exhibit a lower degree of focus affinity than subcategorized iOs. This constitutes a focus mismatch in that the ranking w.r.t focus affinity (cf. (3)) does not follow the argument-adjunct (i.e., obligatory-optional) distinction (unlike in Heidinger 2018).

Discussion. Objects taking the form of a lexical XP (as in (1) and in our experimental stimuli) must be legitimated in the face of the economy principle (Grice 1975, Horn 1984). Since focus requires the use of a stressable form, being the focus of a sentence legitimates the use of a lexical XP. Given that lexical iOs are more marked than lexical dOs (Ashby & Bentivoglio 1993, Belloro 2009), interpreting the iO and not the dO as focus minimizes the violation of the economy principle (resulting in the ranking $iO_{subcat} > dO$ in (3a)). Since this holds for subcategorized and free iOs alike, there must be an additional factor at play to account for $dO > iO_{free}$ in (3b). In the light of recent theoretical work on datives, we discuss two potential factors: The non-at-issueness of the *affectedness* meaning typically associated with some free iOs (Bosse et al. 2012), and the reading of sentences with free iOs where the iO escapes the scope of negation (Bar-Asher Siegal & Boneh 2015). A post-hoc semantic classification of our experimental items reveals, however, that neither factor can account for the observed focus mismatch.

(1) No le=rompieron el coche al vecino. (*al vecino* = free iO)
NEG CL=broke the car to.the neighbor
'They did not break the neighbor's car'

(2) No le=dieron el regalo al niño. (*al niño* = subcategorized iO)
NEG CL=gave the gift to.the child
'They did not give the gift to the child'

| | dO | iO | |
|---|---|---|---|
| iO = free | 55% | 45% | 100% |
| iO = subcat | 31% | 69% | 100% |

Table 1: Interpretation as narrow focus

(3) Focus affinity

a. $iO_{subcat}$ > dO
b. dO > $iO_{free}$
c. $iO_{subcat}$ > dO > $iO_{free}$

**References**


Ashby, William J. & Paola Bentivoglio (1993), Preferred argument structure in spoken French and Spanish. *Language Variation and Change*: 5, 61–76.

Bar-Asher Siegal, Elitzur & Nora Boneh (2015), Decomposing Affectedness: Truth-Conditional Non-core Datives in Modern Hebrew. *MIT Working Papers in Linguistics*: 78, 1-22.

Belloro, Valeria A. (2009), Spanish datives: remarks on the information-structure side of the story. In L. Guerrero & S. Ibáñez & V. Belloro (eds.), *Studies in Role and Reference Grammar*, 491–516. México: UNAM.

Bosse, Solveig, Benjamin Bruening & Masahiro Yamada (2012), Affected experiencers. *Natural Language & Linguistic Theory*: 30, 1185–1230.

Contreras, Heles (1978), *El orden de palabras en español*. Madrid: Ediciones Cátedra.

Drubig, Bernhard H. (2003). Toward a typology of focus and focus constructions. In *Linguistics*: 41, 1–50.

Firbas, Jan (1992), *Functional sentence perspective in written and spoken communication*. Cambridge University Press, Cambridge.

Grice, Paul H. (1975), Logic and Conversation. In P. Cole, & J. L. Morgan (eds.), *Syntax and Semantics, Vol. 3, Speech Acts*, 41–58. New York: Academic Press.

Heidinger, Steffen (2018), *Sekundäre Prädikation und Informationsstruktur: Fokus und Informationsstatus bei spanischen Depiktiven*. Berlin: Peter Lang.

Heidinger, Steffen & Edgar Onea (2021), Focus affinity in Spanish: An experimental study. *Journal of Pragmatics*: 186, 100–116.

Horn, Laurence R. (1984), Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. In D. Schiffrin (ed.), *Meaning, Form, and Use in Context: Linguistic Applications*, 11–42. Washington, D.C.: Georgetown University Press.

Lambrecht, Knud (1994), *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge: Cambridge University Press.

Sgall, Petr, Eva Hajičová & Jarmila Panevová (1986), *The meaning of the sentence in its semantic and pragmatic aspects*. Dordrecht: Reidel.

# From semantics to discourse representation alone: The topic of presentational constructions

Delia Bentley, Francesco Maria Ciconte & Alina McLellan
(University of Manchester, University of Chieti-Pescara & University of Manchester)

Keywords: discourse, presentational construction, topic, Gallo-Italian, Kréol Rényoné

Presentational constructions, i.e., constructions which introduce an event into the universe of discourse, raise the question of what it means for an utterance to be entirely new in information structural terms. There is growing consensus that there cannot be a topicless predication (Bianchi 1993, Erteschik–Shir 1997) and that presentational constructions are not completely new, but rather predicate an event about a situational or contextual topic (Klein 2008, among others).

Our talk aims to shed light on the variety of strategies that are attested for the encoding of the topic of presentational constructions and on the diachronic and structural underpinnings of this variation. The Romance languages of Northern Italy are our starting point, in that their presentational constructions often exhibit an etymologically locative clitic in a subject clitic position, which has been analysed as their topic and, at the same time, a syntactic subject (Tortora 1997, Parry 2013). The topic of presentational constructions can also be spelled out as a subject in German (Fuß and Hinterholz 2023).

Availing ourselves of data collected through fieldwork and online, we discuss case studies from Turinese (Gallo-Italian) (Bentley and Ciconte 2024) and Kréol Rényoné (KR) (McLellan 2023), which suggest that the topic of presentational constructions need not be a syntactic subject. The case of Turinese is illustrated in (1), where *je* is an etymologically locative clitic.

(Turinese)

(1) A l'è nassu**je** / A son nassu**je** tante fior.
SCL.3SG AUXCL-be.3SG be.born.PSTP-PRES.CL SCL.3PL be.3PL be.born.PSTP-PRES.CL many flowers
'Many flowers have appeared.'

We rule out a subject analysis of *je* because *je* fails to occur in subject position and is also compatible with V-S agreement (for preverbal *a l* see Regis 2006, Tosco et al. 2023:177-179). We compare the case of Turinese *je* to that of *nana*, which is the possessive and existential form 'have' in KR.

(2) **Nana** in num privé la tèl amwin. (KR)
PRES.COP INDF number private PRF phone 1SG
'A private number rang me.'

*Nana* cooccurs with a noun phrase in subject position and cannot therefore be a subject. We also reject its analysis as a strategy to avoid clause-initial focus (see Lambrecht 2000 for Spoken French *y a*) because this is not ruled out in KR (McLellan 2023).

We propose a parallel architecture (Jackendoff 2002, Van Valin 2023) account of the structures in (1) and (2), in which the grammaticalization of *je* and *nana* results in a semantics-discourse mismatch, with *je* and *nana* being mapped from discourse to syntactic representation alone without having an

argument or predicate correlate in semantic representation. Marking the deixis of the discourse situation, Turinese *je* indexes the discourse situation as the topic of the construction. In turn, *nana*, being an existential copula, i.e., a form that normally occurs in a context-dependent predication (Francez 2007, Bentley et al. 2015), also grammaticalizes into a marker of the discourse situation in which a new event is announced, thus coming to index this situation, albeit via a different grammaticalization route to that of *je*.

Highlighting a mismatch between predicate semantics and discourse, we propose that the topic of presentational constructions can originate in discourse alone and we make a case for discourse as a separate level of linguistic representation.

**References**


Bentley, Delia, Francesco Maria Ciconte and Silvio Cruschina (2015), *Existentials and Locatives in the Romance Dialects of Italy*. Oxford: Oxford Univesity Press.

Bentley, Delia and Francesco Maria Ciconte (2024), Microvariation at the interfaces: The subject of predication of broad focus VS constructions in Turinese and Milanese. *Languages* 9:37. https://doi.org/10.3390/languages9020037.

Bianchi, Valentina (1993), Subject positions and e-positions. *Quaderni del laboratorio di linguistica* 7, 51–69. Pisa: Scuola Normale Superiore.

Erteschik–Shir, Nomi (1997), *The dynamics of focus structure*. Cambridge: Cambridge University Press.

Francez, Itamar (2007), Existential propositions. Ph.D. Thesis. Stanford University, Stanford, California USA.

Fuß, Eric and Roland Hinterhölz (2023), On the historical development of pronouns referring to situations: The rise of pre-finite 'expletives' in German. *Journal of Historical Syntax*, Volume 7, Article 2, 1-54.

Jackendoff, Ray (2002), *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.

Klein, Wolfgang (2008), The topic situation, in B. Ahrenholz, U. Bredel, W. Klein, M. Rost-Roth & R. Skiba (eds), (2008), *Empirische Forschung und Theoriebildung: Beiträge aus Soziolinguistik, Gesprochene-Sprache- und Zweitspracherwerbsforschung: Festschrift für Norbert Dittmar*. Frankfurt am Main: Lang, 287–305.

Lambrecht, Knud (2000), When subjects behave like objects. An analysis of the merging of S and O in Sentence-Focus Constructions across languages. *Studies in Language* 24 (3), 611-682.

McLellan, Alina (2023), Relative and cleft constructions in Kréol Rényoné. PhD thesis. University of Manchester.

Parry, Mair (2013), Variation and change in the presentational constructions of north-western Italo-Romance varieties, in J. Barðdal, Cennamo, M. and E. van Gelderen (eds), (2013), *Argument Structure in Flux: The Naples/Capri Papers*. Amsterdam: Benjamins, 511-548.

Regis, Riccardo (2006), Se i clitici soggetto sono facoltativi: Il caso piemontese. *Studi e Ricerche Quaderni del Dipartimento di Scienze del Linguaggio e Letterature Moderne e Comparate dell'Università di Torino* 1, 191–232.

Tortora, Christina (1997), The syntax and semantics of the weak locative. PhD thesis, University of Delaware, Newark, Delaware USA.

Tosco, Mauro, Emanuele Miola and Nicola Duberti (2023), *A Grammar of Piedmontese. A Minority Language of Northwest Italy*. Leiden: Brill.

Van Valin, Robert D. Jr. (2023), Principles of Role and Reference Grammar, in D. Bentley, R. Mairal Usón, W. Nakamura and R. D. Van Valin, Jr. (eds), (2023), *The Cambridge Handbook of Role and Reference Grammar*. Cambridge: Cambridge University Press. Cambridge, 17-177.

# Optional topic movement in German scrambling

András Bárány & Jutta-Maria Hartmann
(Bielefeld University / Bielefeld)

Keywords: Information structure, Morphosyntax, German, Scrambling, Acceptability judgements

German scrambling in the middle field is influenced by, i.a., scope, prosody, information structure (IS), however, the precise nature of IS effects in particular is disputed. Among others, Frey (2004) argues for an obligatory topic position preceding sentence adverbials, but if available at all, topic movement seems optional (Fanselow 2006). We contribute to the empirical and theoretical understanding of IS-related scrambling in German considering optional movement and tolerance of mismatches with preceding context. Our research questions are, first, *which information-structural concepts are relevant for German* and, second, *how "optional" or information-structurally conditioned movement can be modelled?* We hypothesise that movement to a topic position is not obligatory as IS interpretations can be reached via syntax or pragmatic reasoning; as syntactic marking is explicit, we expect that (i) non-pronominal topical arguments front more than non-topical ones; (iii) topical arguments are continued more frequently as pronouns.

Data**:** We report the results of a pilot study on scrambling that involved a context followed by a forced choice between a scrambled (DAT–ADV–NOM) and an unscrambled (NOM–ADV–DAT) order, followed in turn by a discourse completion task (using jsPsych; de Leeuw et al. 2023), see a translated example in (1). 80 German native speakers, recruited using Prolific, were shown contexts with or without a proper name as topic (topic vs. non-topic conditions) and two sentences which differed in (non-)scrambling of an indirect object (proper name with a dative determiner, e.g. *der*(DAT) *Tina*). Participants chose one of the two sentences in a forced choice task and were then prompted to provide a discourse continuation.

(1) a. Context
Topic condition: *Do you know what I heard* ***about Tina****?*
No-topic condition: *Do you know what I heard?*
b. Choice (i: unscrambled vs. ii: scrambled)
*I was told that …*
i. *selbstverständlich der Veranstalter* ***der Tina*** *einen Rabatt geben wird.*
ii. ***der Tina*** *selbstverständlich der Veranstalter einen Rabatt geben wird.*
'… the host will naturally give **Tina** a discount.'
c. Continuation
*Do you believe that …*
Response: '… ***she****'ll be happy with that?'*

Results and discussion: First, topic vs. no-topic context had no effect on choice of order, i.e., topic movement is optional. In the discourse completion task, participants chose the original dative phrase as the subject of their completion more frequently than the original (nominative) subject, although this was not statistically significant. Second, when the dative became the subject of the completion, it was pronominalised significantly more frequently than the original subject. Both promotion to subject and pronominalisation indicate topic continuation. We propose that topic movement is optional in relation to the preceding discourse as topics can be established via a pragmatic or morphosyntactic mechanisms, resulting in apparent mismatches between context and syntactic marking of topics. Syntactic marking is, however, stronger and leads to an increase in topic interpretation for the following discourse, supported by making a topic a subject and pronominalising it. In a follow-up experiment, we further explore the role of case and grammatical function for topic interpretation.

**References**

Fanselow, Gisbert. 2006. On pure syntax (uncontaminated by information structure). In Patrick Brandt & Eric Fuß (eds.), *Form, structure, and grammar: A festschrift presented to Günther Grewendorf on occasion of his 60th birthday*, 137–158. Berlin: Akademie Verlag. DOI: 10.1524/9783050085555.137.

Frey, Werner. 2004. A medial topic position for German. *Linguistische Berichte* 198, 153–190.

de Leeuw, Joshua R., Gilbert, Rebecca A. & Luchterhandt, Björn. 2023. jsPsych: Enabling an open-source collaborative ecosystem of behavioral experiments. *Journal of Open Source Software* 8(85), 5351. DOI: 10.21105/joss.05351.

# Exploring focus-contrast mapping mismatch in Romance languages: a comparative study of cleft structures and contrastive prosodic patterns in Italian and French

Bianca Maria De Paolis, Cecilia Andorno, Sandra Benazzo
(University of Turin, Italy and Lab. SFL - University of Paris 8 / CNRS, France; University of Turin, Italy; Lab. SFL - University of Paris 8 / CNRS, France)

Keywords: Romance, cleft sentences, prosody, focus, contrast

Our research investigates the mapping of various information-structural notions — givenness, focus, and contrast — onto syntax and prosody. The study concerns two Romance languages, Italian and French, specifically examining a subset of strategies that are shared and commonly employed to mark discourse prominence in both languages: prosodic emphasis conveyed through vowel lengthening or tonal movements (Delais-Roussarie et al., 2015; Gili Fivela et al., 2015) and it-clefts (Roggia, 2008); see example 1.

(1) C'est Marie qui achète le journal.
È Maria che compra il giornale.
*It is Mary who is buying the newspaper.*

We aim to explore how these marking strategies interact in contexts displaying varying degrees of contrastiveness and/or newness: background, broad focus, narrow-identification-focus, narrow-correction focus (Repp, 2016; Cruschina, 2021).

The analysis was conducted on task-elicited speech involving 35 native speakers of both languages (18 French, 17 Italian). The data collection protocol included a picture-constrained question-and-answer task, targeting subject constituents (Marie/Maria) in four distinct information-structural conditions. Specifically, the collected dataset comprises 12 utterances per speaker, with three for each of the four target conditions: background, broad focus, narrow-identification focus, narrow-correction focus. In total, the dataset includes 420 utterances. Syntactic phenomena were coded manually, while prosodic ones were measured and labeled using an automatic transcription tool (Polytonia, Mertens 2014). For the quantitative analysis, data were fitted into generalized linear mixed models with informational condition as independent variable and speaker as random factor.

Results reveal a shared pattern in French and Italian groups, where background subjects undergo prosodic de-accentuation, i.e. reduction in duration and pitch span. This result challenges a conventional belief that Romance languages lack prosodic reduction of given elements (Gussenhoven, 2004, Ladd, 1996). However, notable distinctions emerge between the two languages. Among Italian speakers, prosody takes on a default role in marking both new and focal elements, using intonation and durational cues together. As for syntactic marking, Italian speakers resort to it-cleft structures selectively, primarily in higher contrast situations (corrective focus). Conversely, French speakers default to it-cleft structures as the primary narrow-focus strategy, employing prosody, especially on clefts, for more contrastive contexts. Also, they exhibit a preference for tonal over durational cues in implementing prosodic prominence.

Our study underscores a crosslinguistic alignment in marking new and given elements, while highlighting a mismatch between prosody and syntax in the combinations of focus and contrast. Despite both languages employing increased means for marking contrast, in fact, these devices appear

to serve distinct functions. Additionally, the analysis of separate phonetic measures within prosodic marking reveals a disparity between durational and tonal cues, with Italian speakers favoring the former and French speakers favoring the latter. Moreover, our work highlights a crucial observation: apart from language-specific mismatches, results from Italian and French groups point to an additive relationship between syntactic and prosodic strategies in marking different degrees of prominence at the discourse level. This challenges the previous notion of a syntax/prosody trade-off, as proposed in studies on focus in Romance languages (Lambrecht 2001).

Acknowledgements
The research was partly funded by the Projet Vinci grant, awarded by the Italo-French University (UIF-UFI) in 2022. The authors would like to express their gratitude to Christoph Gabriel for providing the materials used in data collection.

**References**

Cruschina, S. 2021. The Greater the Contrast, the Greater the Potential: On the Effects of Focus in Syntax. *Glossa: A Journal of General Linguistics*, 6 (1), 1-30.

Delais-Roussarie, E., Post, B., Avanzi, M., Buthke, C., Cristo, A. D., Feldhausen, I., Jun, S. A., Martin, P., Meisenburg, T., Rialland, A., Sichel-Bazin, R., & Yoo, H. Y. 2015. "Intonational phonology of French: Developing a ToBI system for French". In S. Frota & P. Prieto, (eds.), *Intonation in Romance*. Oxford: Oxford University Press.

Gili Fivela, B., Avesani, C., Barone, M., Bocci, G., Crocco, C., D'Imperio, M., Giordano, R., Marotta, G., Savino, M., and Sorianello, P. (2015). "Varieties of Italian and their intonational phonology". In S. Frota and P. Prieto (eds.), *Intonation in Romance* (pp. 140–197). Oxford: Oxford University Press.

Gussenhoven, C. 2004. *The phonology of Tone and Intonation*, Cambridge, Cambridge University.

Ladd, R. 1996. *Intonational phonology*, Cambridge, Cambridge University Press.

Lambrecht, K. 2001. "A Framework for the Analysis of Cleft Constructions", *Linguistics*, 39, 463-516.

Mertens, P. 2014. "Polytonia : a system for the automatic transcription of tonal aspects in speech corpora". *Journal of Speech Sciences*, 4(2), 17–57.

Repp, S. 2016. "Contrast : Dissecting an Elusive Information-Structural Notion and Its Role in Grammar". In C. Féry & S. Ishihara (eds.). *The Oxford Handbook of Information Structure*.

Roggia, C. E. 2008. "Frasi Scisse in italiano e in francese orale: evidenze dal C-ORAL-ROM.", *Cuadernos de Filología Italiana*, 15: 21.

# On a mismatch between the definiteness/specificity-hierarchy and topicality: The case of fronted bare quantifiers in Romance

Ion Giurgea & Eva-Maria Remberger
(Romanian Academy Institute "Iorgu Iordan - Alexandru Rosetti & University of Vienna)

Keywords: bare quantifiers, indefinites, topicalization, contrastive topics, verum focus

It is by and large agreed that indefinites can be topicalized only if they are (i) specific, carrying an existential presupposition (Erteschik-Shir 1997), or (ii) in the restriction of a quantifier over situations (Endriss & Hinterwimmer 2008; Endriss 2009) or (iii) if their descriptive part (the NP-property) is given. We discuss a less known type of topicalization, the contrastive topicalization of bare quantifiers (BQ), which shows a complete mismatch between givenness or the definiteness/specificity-hierarchy (e.g. Aissen 2003 a.o.) and topicality:

(1) Sp. [Context: Juan didn't eat anything]
No; **Algo**, Juan sí (*lo) comió, pero no mucho. (Arregi 2003:4)
no something J. yes it.CL ate but not much
'No, he ate something.'

(2) Ro. **Ceva** el ŞTIE. (Giurgea 2015)
something he knows
'He knows something.'

(3) It. **Qualcosa** avrà fatto, nella vita. (Benincà et al. 1988:143)
something will-have.1SG done in-the live
'He will have done something in live.'

Previous analyses of this construction include focalization (Zubizarreta 1998), focus-affected reading (Quer 2002) or just verum focus, with expletive focus on the bare quantifier (Leonetti & Escandell-Vidal 2009). Arregi (2003) was the first to recognize this fronting as a type of contrastive topicalization, which is supported by prosody and by the possibility of inserting a preverbal subject (see (1)-(2)), an analysis extended to other Romance languages in Giurgea & Remberger (2011), Giurgea (2015). Arregi however did not explain why this construction typically contains verum focus (cf. Höhle 1992) or focus on an epistemic necessity expression.

Proposal. We use Büring's (1999) analyses of contrastive topics as introducing sets of sets of focal alternatives, each alternative to the topic being associated to a different choice of the focus. We argue that in BQ topicalization, for a fixed value of the focus, the topic alternatives form a series of propositions in which the proposition at hand is the weakest, the one entailed by all the others – e.g., in (2), 'he knows something', 'he knows *a*', 'he knows *b*' (*a, b* are specific things), 'he knows many things' etc. The generalized quantifier *something* ($\lambda P.\exists x.P(x)$) is associated to certainty (verum focus can be analyzed as focus on an epistemic FOR-SURE, see Romero & Han 2002). For the other choices of the topic – specific entities (which in generalized quantifier terms can be represented as $\lambda P.P(a)$, $\lambda P.P(b)$), or the indefinite *many things* – the modality is set to uncertainty. Evidence comes from negative sentences. In this case, because the entailment relations are reversed, we find *all* and *many* as BQ topics, instead of *something/somebody*:

(4) It. Tutto non lo si può avere. (Floricic 2013, ex. (14b))
everything not CL.3MSG.ACC CL.IMPERS can.3sg have
'One cannot have everything.'

We explain the specialization of BQ topicalization to this type of focus by the fact that a BQ cannot be contrasted to other topic choices in terms of properties. This leaves modality as a possible choice for the focus (the variable part inside the Comment of the topic alternatives).

**References**


Aissen, Judith. 2003. Differential Object Marking: Iconicity vs. Economy. *Natural Language & Linguistic Theory* 21, 435–483.

Arregi, Karlos. 2003. Clitic Left dislocation is contrastive topicalization. In Elsie Kaiser and Sudha Arunachalam (eds.), *Proceedings of the 26th Annual Penn Linguistics Colloquium*, *U. Penn Working Papers in Linguistics 9* (1), 31–44.

Beninca, Paola, Giampaolo Salvi & Lorenza Frison. 1988. L'ordine degli elementi della frase e le costruzioni marcate. In Lorenzo Renzi and Giampaolo Salvi (eds.). *Grande grammatica italiana di consultazione*, 115–225. Bologna: Il Mulino.

Büring, Daniel 1999. Topic. In Peter Bosch and Rob van der Sandt (eds.). *Focus*, 142–165. Cambridge: Cambridge University Press.

Floricic, Franck 2013. Bare quantifiers and topics in Italian. In Johannes Kabatek and Albert Wall (eds.). *New Perspectives on Bare Noun Phrases in Romance and Beyond*, 255–282. Amsterdam/Philadelphia: John Benjamins.

Giurgea, Ion. 2015. Bare quantifier fronting as contrastive topicalization. *Bucharest Working Papers in Linguistics* 17 (2), 23–38

Giurgea, Ion & Eva-Maria Remberger. 2011. Indefinite pronouns as contrastive topics. Talk at the SLE – *44th Annual Meeting of the Societas Linguistica Europaea*, Universidad de la Rioja, Logroño, 11.9.2011.

Endriss, Cornelia. 2009. *Quantificational Topics. A Scopal Treatment of Exceptional Wide Scope Phenomena.* Springer.

Endriss, Cornelia & Stefan Hinterwimmer. 2008. Direct and indirect aboutness topics. *Acta Linguistica Hungarica* 55, 297–307.

Erteschik-Shir, Nomi, 1997. *The dynamics of focus structure.* Cambridge: CUP.

Höhle, Tilman, 1992. Über Verum Fokus in Deutschen. *Linguistische Berichte Sonderheft* 4, 112–141.

Leonetti, Manuel & Vicki Escandell-Vidal. 2009. Fronting and verum focus in Spanish. In Andreas Dufter and Daniel Jacob (eds.). *Focus and Background in Romance Languages*, 155-204. Amsterdam/Philadelphia Benjamins.

Quer, Josep. 2002. Edging quantifiers. On QP-fronting in western Romance. In Claire Beyssade, Reineke Bok-Bennema, Frank Drijkoningen and Paola Monachesi (eds.). *Romance Linguistics and Linguistic Theory* 2000, 254–270. Amsterdam, Philadelphia: John Benjamins.

Romero, Maribel & Chunghye Han. 2002. Verum Focus in negative yes/no questions and Ladd's *p*/¬*p* Ambiguity. In Brendan Jackson (ed.). *Proceedings of the 12th Semantics and Linguistic Theory Conference*, 204–224. Ithaca, NY: CLC Publications.

Zubizarreta, Maria Luisa. 1998. *Prosody, focus and word order. Linguistic Inquiry Monographs* 33, MIT Press.

# Discourse determined morphosyntax in Hehe: A case of subject inversion

Simon Simbahomile Msovela
(University of Gothenburg; Dar es Salaam University College of Education)

Keywords: Bantu, discourse roles, morphosyntax, word order, subject inversion

The current paper presents how discourse roles determine Hehe morphosyntax. Despite the long-lived argument that the linear word ordering in Bantu is based on grammatical roles and that Subject-Verb-Object-Adjunct (SVOA) is a canonical word order (Heine 1976), it is uncertain how far this order is rigidly fixed. The uncertainty has called for some scholars to argue for the fact that discourse roles determine syntax (Kiss 1995; Yoneda 2011; Van der Wal 2015a; Kerr et.al 2023). After studying subject inversion in Hehe (G62), a Bantu language spoken in Iringa, Tanzania, in line with Kerr et.al (2023), I assert that discourse roles (topic and focus) do not only determine word order but also nominal and verbal morphology. While the topic is defined as a referent whose proposition expresses what the sentence is about, the focus is new information and the complement of a topic (Lambrecht 1994; Krifka 2008; Yoneda 2011).

The paper addresses two questions: first, "Which types of subject inversion are attested in Hehe?" Second, "what determines this word order?" With fresh data from the field, I illustrate how topic and focus determine the morphosyntax of Hehe, I present five different types of subject inversion and explain the discourse roles that determine these word orders. First, the presentational focus determines the formal locative inversion whereby an NP marked with the prototypical locative classes 16, 17, or 18 precedes the verb and determines agreement on the verb as in [1]. The contrastive focus determines the semantic locative inversion in [2] whereby the preverbal NP with a noun class prefix other than locative prefixes agreeing with the verb is locative only in meaning but not in form. Also, the simple focus determines the agreeing inversion whose postverbal subject determines the subject marker on the verb and agrees with it in [3]. The contrastive topic determines the instrument inversion in [4]. As its name suggests, the grammatical subject is a thematic instrument, and the verb shows agreement with it. The final type is patient inversion which is also called object-subject reversal. To communicate the all-new information, the direct object bearing a patient thematic role in a transitive construction occurs preverbally and triggers agreement with the verb in the inversion construction [5].

[1] pa-luhengo=po pa-anik-e wusee
16-11.penthouse=16 16SM-dry-PFV 14.flour
'There is flour dried at the penthouse.'

[2] i-li-wodi i-li li-gon-**i**-a **va-**va-nya sumi
AUP-5-ward 5-DEM.PROX 5SM-sleep-**CAUS**-FV **EXH**-2-with 9.tuberculosis
'This ward admits only tuberculosis patients (no other patients).'

Context: who will go to the market?

[3] ku-m-nada a-bit-a Munofumumi
17-3-market 1SM-go-FV 1.Munofumumi
'Munofumumi will go to the market.'

[4] i-ki-mage che-i-kes-**el**-a u-yuvahenge i-mboga
AUP-7-knife 7REL-SM.PRS-slice-**APPL**-FV AUP-1.aunt AUP-9.vegetable
'The knife that aunt chops vegetable with… (not another knife)'

[5] i-li-lasi li-tem-ite u-mw-ana
AUP-5-bamboo tree 5SM-cut-PFV AUP-1-child
'The bamboo tree has cut the child.'

The grammatical subject occupies the topical position in each, and the logical subject is postposed and underspecified. As for agreement, apart from the agreeing inversion whose postverbal NP triggers agreement, all the rest have agreement triggered by the grammatical subjects. This suggests that Hehe has syntactically dedicated positions for both topic and focus (i.e., preverbally, and immediately after verb, respectively). In line with Kerr et.al. (2023), I argue that subject inversion constructions above are a blueprint of discourse i.e. they result from speakers' accommodation of discourse roles. I also argue that, like other Bantu languages, discourse can best explain Hehe morphosyntax.

**References**


Elisabeth J. Kerr, Allen Asiimwe, Patrick Kanampiu, Zhen Li, Ernest Nshemezimana and Jenneke van der Wal, (2023). Bantu word order between discourse and syntactic relations *Linguistique et Langues Africaines* 9(1).

Heine, B. 1976. *A Typology of African Languages Based on the Order of Meaningful Elements*. Berlin: Dietrich Reimer

Kiss, K. É. 1995. *Discourse-configurational languages. Oxford*: Oxford University Press.

Krifka, M. 2008. Basic notions of information structure. Acta Linguistica Hungarica 55(3–4). 243–276.

Lambrecht, K. 1994. *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents.* New York: Cambridge University Press.

van der Wal, J. 2015a. *Bantu syntax*. In Oxford handbooks online. Oxford: Oxford University Press.

Yoneda, N. 2011. Word order in Matengo (N13): Topicality and informational roles. *Lingua 121 (5):*, 754–771.

# Mismatches between information structure and prosody: Contrast-marking in spoken British English

Matthias Mittendorfer
(University of Vienna)

Keywords: parallel structures, contrast-marking, mismatches, prosody, information structure

Contrast as an information-structural notion is notoriously hard to define (see, i.a., Chafe 1976, Molnár 2006, Molnár & Winkler 2010, and Repp 2010, 2016). Semantically, it is often associated with the notion of alternativeness (set membership, exclusion requirement, etc.; see, i.a., Molnár 2002, Repp 2016), while prosodically, Contrast is claimed to be marked by the pitch accents L*+H and L+H* (e.g. Büring 2003, Pierrehumbert & Hirschberg 1990, and Steedman 2000). However, partly due to issues with the reliable coding of information structure (see, i.a., Hedberg 2006), the prosodic realization of Contrast in natural data seems less clear, indicating that mismatches play an important role in Contrast-marking. One case in which the pragmatic presence of Contrast is relatively straightforward is parallel structures, in which two or more assertions are contrasted in at least one syntactic position. As such, they have often been considered "prime examples of contrastiveness" (Repp 2010: 1339).

The present paper will analyse the information structure and prosody of parallel structures taken from the British component of the International Corpus of English (Nelson et al. 2002). In (1) below, a sample analysis following the question-under-discussion approach, an annotation framework offering a relatively strict formalism to deal with parallel structures (QUD; Riester 2019, and Riester et al. 2018) is given:

(1) […] Michael Heseltine has charisma. Douglas Hurd has a safe pair of hands.

$Q_1$: {Who has what?}
> $Q_{1.1}$: {What does Michael Hesltine have?}
>> $A_{1.1}$: [[Michael Heseltine]$_{CT}$ has [charisma]$_{CF}$]~.
> $Q_{1.2}$: {What does Douglas Hurd have?}
>> $A_{1.2}$: [[Douglas Hurd]$_{CT}$ has [a safe pair of hands]$_{CF}$]~.

[ICE-GB; S1B-043 #001]

In this particular example, both the Contrastive Topics and Contrastive Foci are marked by an L+H* pitch accent (in the ToBI framework; see Beckman et al. 2006), thus supporting the above-mentioned association of Contrast with accenting. However, even as strict a framework as the QUD approach cannot completely solve the issue of Contrast-marking in English, as variation in the prosody remains an issue.

Consequently, the proposed paper aims at investigating the nature of these mismatches between Contrastive Topics/Foci and their prosodic realization in parallel structures. More specifically, the following questions will be explored:

RQ1: Can a refined analysis of natural data using QUDs contribute to our understanding of Contrast (-ive Topic/Focus) and its prosodic features?

RQ2: What is the nature of the interaction between Contrast (-ive Topic/Focus) and prosody? Does a contrastive pitch accent exist or is the interface characterized by mismatches?

RQ3: How can the information structure–prosody interface be accounted for in Functional Discourse Grammar (FDG)?

In a final step, the data will be analysed in FDG (Hengeveld & Mackenzie 2008), which, with its top-down architecture, in which pragmatics and semantics govern morphosyntax and phonology, is well-suited for investigations into phenomena at the interfaces of different levels of analysis (see, for instance, Contreras-García & García Velasco 2021). Therefore, the proposed paper will contribute to the understanding of the interaction between the Interpersonal (i.e. pragmatic) and Phonological Levels.

**References**


Beckman, Mary E., Hirschberg, Julia, & Schattuck-Hufnagel, Stefanie (2006), The Original ToBI System and the Evolution of the ToBI Framework, in S.-A. Jun (ed.), *Prosodic typology: The phonology of intonation and phrasing*, Oxford: Oxford University Press, 9–54.

Büring, Daniel (2003), On D-trees, beans, and B-accents, *Linguistics and Philosophy* 26(5), 511–545.

Chafe, Wallace L. (1976), Givenness, contrastiveness, definiteness, subjets, topics, and point of view, in C. N. Li (ed.), (1976), *Subject and topic*, London: Academic Press, 25–55.

Contreras-García, Lucía, & García Velasco, Daniel (eds.), (2021), *Interfaces in Functional Discourse Grammar: Theory and applications*, Berlin: De Gruyter.

Hedberg, Nancy (2006), Topic-focus controversies, in V. Molnár & S. Winkler (eds.), (2006), *The architecture of focus*, Berlin: De Gruyter, 373–397.

Hedberg, Nancy, & Sosa, Juan M. (2007), The prosodic structure of topic and focus in spontaneous English Dialogue, in C. Lee, M. Gordon, and D. Büring (eds.), (2007), *Topic and focus: Cross-linguistic perspectives on meaning and intonation*, Dordrecht: Springer, 101–120.

Hengeveld, Kees, & Mackenzie, Lachlan J. (2008), *Functional Discourse Grammar: A typologically-based theory of language structure*, Oxford: Oxford University Press.

Molnár, Valéria (2002), Contrast—from a contrastive perspective, in H. Hasselgård (ed.), (2002), *Information structure in a cross-linguistic perspective*, Amsterdam: Rodopi, 147–161.

Molnár, Valéria (2006), On different kinds of contrast. In V. Molnár, and S. Winkler (eds.), (2006), *The architecture of focus*, Berlin: De Gruyter, 197–233.

Molnár, Valéria, and Winkler, Susanne (2010), Edges and gaps: Contrast at the interfaces, *Lingua* 120(6), 1392–1415.

Nelson, Gerald; Wallis, Sean, and Aarts, Bas (2002), *Exploring natural language: Working with the British component of the international corpus of English*, Amsterdam: John Benjamins.

Pierrehumbert, Janet B., and Hirschberg, Julia (1990), The meaning of intonational contours in the interpretation of discourse, in P. R. Cohen, J. Morgan, and M. E. Pollack (eds.), (1990), *Intentions in communication*, Cambridge, MA: MIT Press, 365–378.

Repp, Sophie (2010), Defining 'contrast' as an information-structural notion in grammar, *Lingua* 120(6), 1333–1345.

Repp, Sophie (2016), Contrast: Dissecting and elusive information-structural notion and its role in grammar, in C. Féry, and S. Ishihara (eds.), (2016), *The Oxford handbook of information structure*, Oxford: Oxford University Press, 270–289.

Riester, Arndt (2019), Constructing QUD trees, in M. Zimmermann, K. Von Heusinger, and V. E. Onea Gaspar (eds.), (2019), *Questions in Discourse: Volume 2: Pragmatics*, Leiden: Brill, 164–193.

Riester, Arndt, Brunetti, Lisa, and De Kuthy, Kordola (2018), Chapter 14: Annotation guidelines for Questions under Discussion and information structure, in E. Adamou, K. Haude, and M. Vanhove (eds.), (2018), *Studies in Language Companion Series*, Amsterdam: John Benjamins, 403–444.

Steedman, Mark (2000), Information Structure and the Syntax-Phonology Interface, *Linguistic Inquiry* 31(4), 649–689.

# Mismatches in Pintupi-Luritja focus particle placement

James Gray
(Western Sydney University)

Keywords: Australian languages, focus particles, syntax, semantics, fieldwork

It has long been recognised that particles that associate with focus often have a transparent syntactic relation to their focus associate (Jackendoff 1972, Tancredi 1990, a.m.o.). Despite a common transparent relation, recent work has been exploring morphosyntactic mismatches in how particles relate to their focus associates; in particular the phenomenon of *anti-pied-piping* (Branan & Erlewine 2023) where a sub-element contained *within* the semantic focus is morphosyntactically targeted. In this paper I investigate this issue in Pintupi-Luritja, a Western Desert (Pama-Nyungan) language of Central Australia, based on original fieldwork and collected corpora. Despite a general transparent syntactic regularity between focus-sensitive particles like *kutju* `only' with their focus associates, I show that mismatches also arise in certain configurations.

For nominals and non-finite verb clauses, focus particles consistently mark the right edge of their associate, as seen in example 1.

**(1)** **Prompt**: *Only adults can drive a car.*
**Anangu tina-ngku kutju** mutukayi trayip-mila-lpayi.
**person big-ERG only** car drive-LOAN-HABIT
Only [adults] drive cars.

However with many finite verb phrases and medial verb clauses, we see a shifted placement of these particles. Here, focus particles mark a *subset* of the focus material, and do not sit on its right edge. This occurs apparently obligatorily with finite VPs that include a direct object or locational adjunct, but optionally with medial verb clauses. This can be seen in example 2, where the exclusive *kutju* has been shifted to the left within the finite VP associate.

**(2)** **Prompt***: I wanted to sing and play guitar at the concert, but the microphone was broken, so I only played guitar.*
… ngayulu **kiita kutju** pliyi-rri-ngu.
1SG.ERG **guitar only** play-INCH-PST
I only [played guitar]$_F$

Similar distributions can be seen with other focus sensitive elements like *-tarra* 'also', suggesting that this is a general pattern for focus sensitive elements in Pintupi-Luritja. The fact that clause type determines this pattern to differing degrees of optionality shows that this shifting of focus particle placement cannot be due to more general phonological or intonational considerations. This suggests instead that syntactic and clausal considerations are relevant for describing the phenomenon in the language.

In-depth descriptions of these focus mismatches in under-described languages help inform a cross-linguistic understanding of the phenomenon of focus marking strategies and the extent to which they vary.

**References**

Branan, K. & Erlewine, M. Y. (2023). Anti-pied-piping. In *Language* 99(3), 603–653.

Jackendoff, R. (1972). *Semantic interpretation in generative grammar*. Cambridge: Mass.: MIT Press.

Tancredi, C. (1990). *Not only even, but even only*. Unpublished manuscript. Massachusetts Institute of Technology.

# Information Structure of Right Dislocation in Japanese

Satoshi Tomioka & Dursun Altinok
(University of Delaware & University of Delaware)

Postposing/Right Dislocation (RD) in Japanese is arguably sensitive to information structure. More specifically, Kuno (1978) and Tanaka (2001) claim that the postposed constituent must correspond to given material. This claim is problematic for the very analysis of RD that Kuno and Tanaka endorse: the bi-clausal analysis, in which the postposed constituent is a 'left-over' of clausal ellipsis, as illustrated below.

(1) a. Kinoo at-ta-yo, Aya-ni

yesterday meet-past-disc.part, Aya-dat
'I met (her) yesterday, Aya.'

b. Kinoo **e** at-ta-yo, Aya-ni1 ~~kinoo~~ t1 ~~at-ta-yo~~

The ellipsis-based derivation of RD and the given status of the postposed constituent do not mix well, however. According to Rooth's (1992) theory of focus and ellipsis, the remnant of ellipsis should contain a focus. Therefore, a non-focal remnant of an ellipsis is expected to be impossible. Kuno's and Tanaka's claim has also been challenged empirically. Several authors (e.g., Takami 1995, Shimojo 2005) have noted that so-called 'all focus' sentences can be transformed into RD structures. For instance, to the question of 'what's happening?', one can answer with (1).

(2) Kenka shiteru-ndayo, Aya-to Ken-ga.

flight do-prog-disc.part, Aya-and Ken-nom.
'(they) are having a fight, Aya and Ken.'

On the other hand, a narrow focus or a constituent containing a narrow focus cannot be placed in the RD-ed position. For instance, a wh-phrase cannot be postposed in a RD-ed constituent question.

(3) *Motte ki-mashi-ta-ka, nani-o?
Have come-polite-past-q what-acc
'Lit: Did you bring (it), what?'

In this presentation, we will revisit Simon's (1989) I*mportant Information Early* principle for the Japanese RD construction and propose that Simon's directive based on communicativeness encourages the speaker to place the main predicate as early as it is allowed. A predicate in Japanese, an agglutinative language, can be morphologically complex and can carry not only the predicate's lexical meaning but also tense, aspect, negation, evidentiality, modality, clause-type, politeness and other discourse-related information. Thus, the earliness directive singles out a predicate complex as an intrinsically information-loaded expression that deserves to be placed earlier than its canonical sentence-final position. We further argue that the notion of importance applicable to predicates interacts with information structural importance, which accounts not only for the general tendency of givenness of a postposed constituent but also for the puzzling contrast between narrow vs. broad focus described above. If time permits, we will the broader consequences of the proposed analysis, including the strictly conversational nature of RD structures and the 'long-before-short'

preference observed by Yamashita and Chang (2001).

**References**

Kuno, Susumu (1978), Danwa- no Bunpoo, Grammar of Discourse. Tokyo, Taishukan Publishing Company.

Rooth, Mats (1992), "A theory of focus interpretation." *Natural Language Semantics* 1,117–121.

Shimojo, Mitsuaki (2005), *Argument Encoding in Japanese Conversation*. New York: Palgrave Macmillan

Simon, Mutsuko Endo (1989), *An analysis of the postposing construction in Japanese*. Ph.D. thesis. University of Michigan.

Takami, Ken-ichi (1995), *Kinoteki Kobunron ni yoru Nichieigo Hikaku* [A Comparison between Japanese and English in Functional Syntactic Theories]. Tokyo: Kuroshio Shuppan.

Tanaka, Hidekazu (2001), "Right-dislocation as Scrambling." *Journal of Linguistics* 37: 551– 579.

Yamashita, Hiroko, and Franklin Chang (2001). "'Long before short' preference in the production of a head-final language." Cognition 81, no. 2, B45-B55.

# Information structure and pitch accents in British English

Jiseung Kim, Na Hu, Riccardo Orrico, Stella Gryllia & Amalia Arvaniti
(Radboud University, Radboud University, Radboud University, Radboud University & Radboud University)

Keywords: Information structure, pitch accents, f0, pragmatic functions, British English

For American English (AE), a distinction is posited between H* and L+H* accents, such that the former is realized as high pitch and encodes new information while the latter is realized as rising pitch and encodes contrastivity (Pierrehumbert & Hirschberg 1990). This distinction is said to apply in British English (BE) as well, though empirical evidence to this effect is limited (cf. Ladd 2008). We investigated whether putative H*s and L+H*s are phonetically distinct in BE and whether the differences relate to information structure distinctions, as posited for AE. If so, H* and L+H* should be phonetically distinct, and, by and large, H* should be associated with items conveying new information, and L+H* with items conveying contrastive and corrective information, resulting in some version of one-to-one mapping between phonetic realization and pragmatic function.

Pitch accents (N = 2,127) from 1,188 unscripted speech utterances elicited from 8 BE speakers (5 F) from the greater London area were analyzed. The accents were annotated separately for f0 shape and pragmatic function, to avoid one classification scheme being guided by the other and vice versa. The phonetic annotation was based on visual and aural inspection of the audio files: accented items were marked as L+H* if they included a deliberate f0 dip at the onset of the accented syllable, and as H* otherwise. The pragmatic annotation was based solely on the orthographic transcripts: items were marked as *corrective* if they were an explicit correction of a previously mentioned item, and as *contrastive* if they were contrasted to a contextually available or easily inferable set of alternatives. All items marked as H* or L+H* and not labeled as corrective or contrastive were marked as *non-contrastive*. The normalized and smoothed f0 curves of the accented items were analyzed using a Generalized Additive Mixed Model, with phonetics (H*, L+H*) and pragmatics (non-contrastive, contrastive, corrective) as fixed intercepts and smooth terms, and speaker as a factor smooth (random intercept and slope). The fitted smooths in different conditions were compared to identify the intervals of significant difference.

The modelling showed that accents classified as H*s were realized as falls, while those classified as L+H*s were realized as rise-falls. Further, the accents differed in f0 scaling (H* < L+H*) and peak location (earlier for H*s). The information structure-based distinctions were not clear-cut: corrective accents showed a more marked dip (i.e., a rise-fall) relative to non-contrastive accents, but there were no significant differences between contrastive and non-contrastive and corrective accents. In short, the information-structure distinctions investigated here were not uniquely associated with phonetically distinct accents, against the hypothesis and what has been argued for AE. Rather, both H* and L+H* were used to mark corrective and contrastive items. Additionally, L+H* was employed for functions unrelated to information structure, such as to highlight unexpected items in the story-telling task. In conclusion, our findings provide insights into the complex role of H* and L+H* in BE and indicate that the assumption of one-to-one mapping between accent shape and information structure does not hold.

Acknowledgements
This work has been financially supported by the European Science Foundation through grant ERC-ADG-835263 (SPRINT) awarded to Amalia Arvaniti. This support is hereby gratefully acknowledged.

**References**

Ladd, D. R. (2008). Intonational phonology (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511808814

Pierrehumbert, J. B., & Hirschberg, J. B. (1990). The meaning of intonational contours in the interpretation of discourse (pp. 271–311). In P. Cohen, J. Morgan, & M. Pollack (Eds.), Intentions in Communication. https://doi.org/10.7916/D8KD24FP

# Object focus form without the object

Eva-Marie Bloom Ström & Onelisa Slater
(University of Gothenburg & Rhodes University)

In Xhosa, word order interacts with morphological marking and phonological phrasing to convey information structural notions (Jokweni 1995, Bloom Ström 2017, 2020). An important part of this is a paradigmatic distinction between two morphologically distinct verb forms of the same tense-aspect (Jokweni 1995) referred to as conjoint/disjoint forms (CJ/DJ) in Bantu studies (Downing and Hyman 2016, van der Wal 2017). For example, the CJ form of the recent past is used when 'child' is in focus such as in (1a), while the DJ form is used when 'beat' is in focus (1b):

1. *a) ú-bhéth-é* *umntwa*:*na*
   SM.3G-beat-REC.CJ 1.child
   (Who did she beat?) 'She beat the child.'

   *b) ú-m-bhét-i*:*le* *(umntwa*:*na)*
   SM.3G-OM.1-beat-REC.DJ (1.child)
   (What did she do with the child?) 'She beat her (the child)'

That the CJ form is never final is the most pervasive characteristic in languages with this distinction (van der Wal 2017). This talk will present a mismatch in the use of the CJ form in Xhosa; it can in fact be clause-final, and moreover take unexpected object marking (2):

2. *ú-m-bethé*:
   SM.3G-OM.1-beat-REC.CJ
   'She beat her a lot/very hard'

Here, the reading of the verb contains an element of surprise or intensity of the action. This use of the CJ form is previously unnoticed, to the best of our knowledge. Since the verb-form is expected to be followed by something, such as in (1a), the focus in the sentence appears to shift to the verb, which is in the 'wrong' form for verb focus. In this talk, we present an analysis of the restrictions of this use of the CJ form and its implications for the semantics of the phrase. This analysis will based on introspections and interviews, as well as examination of relevant forms in a newly developed corpus of spoken Xhosa (Bloom Ström et al. 2023).

**References**

Bloom Ström, Eva-Marie. 2017. What the giant tells us about agreeing post-verbal subjects in Xhosa. *Stellenbosch papers in linguistics Plus 52*:73-100.

Bloom Ström, Eva-Marie. 2020. The existential in Xhosa in relation to indefiniteness. *Studies in African Linguistics 49*:213-239.

Bloom Ström, Eva-Marie, Onelisa Slater, Aron Zahran, Aleksandrs Berdicevskis and Anne Schumacher. 2023. Preparing a corpus of spoken Xhosa. *Association for Computational Linguistics*:62-67.

Downing, Laura J. and Larry M. Hyman. 2016. Information structure in Bantu. In Caroline Féry and Shinichiro Ishihara (eds): *The Oxford Handbook of Information Structure*. Oxford handbooks online: Oxford University Press.

Jokweni, Mbulelo W. 1995. Aspects of IsiXhosa phrasal phonology. Doctoral disseration. Urbana-Champaign: University of Illinois.

van der Wal, Jenneke. 2017. What is the conjoint/disjoint alternation? Parameters of crosslinguistic variation. In Jenneke Van Der Wal and Larry M. Hyman (eds): *The conjoint/disjoint alternation in Bantu*, 14-60. Berlin: De Gruyter.

# Mismatches between prosody and context: Factors influencing individual differences

Jutta M. Hartmann, Farhat Jabeen & Petra Wagner
(Bielefeld University, Bielefeld University & Bielefeld University)

Keywords: Prosody, focus alternatives, German, picture selection task, mismatches

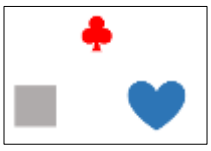
In a felicitous discourse, information-structural (=IS-)marking is expected to match the context. Mismatches between the two are therefore often used to test IS in experiments (where context can be textual or pictorial), see for example for studies on exhaustivity Destruel et al. (2015), Hartmann (2016), Onea & Beaver (2011), Washburn et al. (2013). Responses are however often not as clear-cut as expected. One source of this effect might lie in individual participants' different construals of implicit alternatives (for a more general overview on the representation of alternatives see Gotzner 2017, Gotzner & Spalek 2019 and references therein).

In this paper, we investigate the effect of focus alternatives and prosody on participants' behavior in mismatching contexts. We set up a picture selection task for objects of a specific shape and color. Participants were asked to pick one object, see (1), from a set of three, where (at least) one matched in color and (at least) one in shape, but none in the full description, so that participants were confronted with a situation in which the instructions and the pictorial context did not match. The three objects were manipulated such that two presented separate alternatives {red floret, blue heart} and a distractor, or a set of focus alternatives for objects of a specific shape {blue heart, blue florett} or objects of a specific color {red florett, green florett}, see (2).

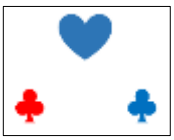
(1) Bitte klicke [$_{PP}$ auf das rote Herz ].
Please click on the red heart

(2) Picture box with three different objects

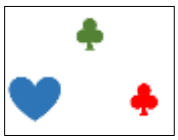
a. General contrast



b. Contrast in Shape



c. Contrast in Color



Prosody varied on the PP between a wide focus prosody (H*L-% on Herz) vs. contrastive focus on color (adj *rot*, L+H*, L-%) vs. on shape (noun *Herz*, L+H*, L-%) (on narrow focus in German see Bauman et al 2015, Féry 1993; on different analyses of realization of wide focus see Grice et al. 2012, Kügler & Féry 2017, Baumann & Schumacher 2020).

The results show that participants select mostly the object that matches in shape, see (3), while some participants consistently clicked on shape independent of prosody. The latter are excluded from the data analysis.

For the other participants, we found that the pictorial context was not a significant predictor. We only found a significant main effect of focus prosody: When color was contrastively focused, there were significantly more clicks on matching color, as compared with contrastive focus on shape. Additionally, there were more clicks on matching color in wide focus than with contrastive focus on shape.

(3) Number of clicks …

a. … per prosody and context

| Prosody | General contrast | | | Shape contrast | | | Color contrast | | |
|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 1 | 2 | 3 | 1 | 2 | 3 |
| | ♥ | ♣ | ■ | ♥ | ♣ | ♣ | ♥ | ♣ | ♣ |
| Color | 80 | 45 | 1 | 83 | 43 | - | 82 | 44 | - |
| Shape | 97 | 29 | - | 92 | 34 | - | 90 | 35 | 1 |
| Wide | 91 | 35 | - | 87 | 39 | - | 81 | 45 | - |

b. … per prosody

| Prosody | ♥ 1 | ♣ 2 | 3 |
|---|---|---|---|
| Color | 245 | 132 | 1 |
| Shape | 279 | 98 | 1 |
| Wide | 259 | 119 | - |

We conclude from these results that while there is a strong preference for shape (syntactic head, prominence of nouns generally), prosody contributes to overriding this preference for a subgroup of participants. Moreover, focus alternatives in the pictorial context did not have the same effect (no interaction). In future work, we plan to further investigate the relevance of mismatches and construal of focus alternatives with linguistic context instead of pictorial context in order to better understand the role of prosody and construal of alternatives in different scenarios.

Acknowledgements: This research is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – CRC-1646, project number 512393437, project A03.

**References**

Baumann, Stefan, Christine T. Röhr and Martine Grice, (2015), Prosodische (De-)Kodierung des Informationsstatus im Deutschen, *Zeitschrift für Sprachwissenschaft* 34, 1–42.

Baumann, Stefan and Petra Schumacher, (2020), The incremental processing of focus, givenness and prosodic prominence, *Glossa: A Journal of General Linguistics* 5(1): 6, https://doi.org/10.5334/gjgl.914 .

Destruel, Émilie., Daniel Velleman, Edgar Onea, Dylan Bumford, and Jingyang Xue (2015), Cross-linguistic study of the non-at-issueness of exhaustive inferences, in F. Schwarz (ed), *Experimental Perspectives on Presuppositions*, Berlin: Springer, 135–156.

Féry, Caroline, (1993), *German Intonational Patterns*, Tübingen: Niemeyer.

Grice, Martine, Stefan Baumann and Ralf Benzmüller, (2012), *German intonation in Autosegmental-Metrical phonology*, in Sun-Ah-Jun (ed.), (2012), *Prosodic Typology: The Phonology of Intonation and Phrasing*, Oxford: Oxford University Press, 55–83.

Gotzner, Nicole, (2017), *Alternative sets in language processing: How focus alternatives are pepresented in the mind,* Cham: Palgrave Macmillan.

Gotzner, Nicole and Katharina Spalek, (2019), The life and times of focus alternatives: Tracing the activation of alternatives to a focused constituent in language comprehension, *Language and Linguistics Compass* 13(2), https://doi.org/10.1111/lnc3.12310 .

Hartmann, Jutta M. (2016), *The Syntax and Focus Structure of Specificational Copular Clauses and Clefts.* Tübingen: Tübingen University, Habilitationsschrift. https://lingbuzz.net/lingbuzz/007384.

Kügler, F & C Féry. 2017. Postfocal downstep in German*, Language and Speech* 60, 260–288.
Onea, Edgar & David Beaver (2011), Hungarian focus is not exhausted, in Ed Cormany, Satoshi Ito and David Lutz (eds), Proceedings of Semantics and Linguistic Theory (SALT) 19, 342–359.
Washburn, Mary B. (2013*), Narrowing the Focus: Experimental Studies on Exhaustivity and Contrast*. Los Angeles, California: University of Southern California Dissertation.

# Focus mismatches in Italian wh-questions

Valentina Bianchi, Giuliano Bocci & Silvio Cruschina
(University of Siena, University of Siena & University of Helsinki)

Keywords: focus, prosody, types of wh-phrases, D-linking, interfaces

In main questions with a bare wh-element, Italian shows a mismatch between prosodic prominence and interpretation: the nuclear pitch accent (NPA) falls on the lexical verb without a focal interpretation, as indicated by the boldface in (1) (cf. Marotta 2001). The situation is however different —and still less clear— with wh-elements other than bare. Bocci, Cruschina & Rizzi (2021), for example, discuss a distinct pattern with a specific type of partitive wh-phrase (cf. 2), where the NPA is significantly more likely to associate with the wh-element rather than with the verb:

(1) Che cosa hai rega**la**to a tua madre per Natale?
what have.2SG given to your mother for Christmas
'What did you give to your mother for Christmas?'

(2) Chi di **voi** ha lavato il divano? who of you have.3SG washed the couch
'Who of you washed the couch?'

In order to shed light on this mismatch and on the factors that are responsible for the variation related to the type of wh-phrase, in this paper we will report the results of a production experiment featuring questions with four types of wh-phrases: (i) partitive *which*-phrase (3), (ii) *which*-phrase following a dislocated partitive PP (4), (iii) *which*-phrase (5), and (iv) *what*-phrase (6):

(3) *Quale di questi libri* leggerai? (partitive *which*-phrase) '*Which of these books* will you read?'

(4) Di questi libri, *quale* leggerai? (*which*-phrase following a dislocated partitive PP)
'Of these books, *which one* will you read?'

(5) *Quale libro* leggerai? (*which*-phrase) '*Which book* will you read?'

(6) *Che libro* leggerai? (*what*-phrase) '*What book* will you read?'

We assume that NPA marks an occurrence of the [focus] feature borne by the wh-phrase (Calabrese (1982), as well as by its copies in the wh-chain (Bocci, Bianchi & Cruschina 2021). We thus use the four conditions illustrated above to test two hypotheses:

**Hypothesis 1**: NPA must fall on the lexical verb whenever the wh-phrase contains no overt lexical material, as with the wh-phrase *what* 'che cosa' in (1). This prosodic account, based on a PF filter that prevents certain functional elements from bearing main prominence, predicts that NPA systematically falls on the verb in (4), while the NPA assignment shows some degree of optionally in (3), (5), (6).

**Hypothesis 2**: NPA must fall on the verb whenever the wh-phrase has a non-D-linked interpretation, as is the case for the bare wh-phrase *che cosa* in (1). When followed by lexical material, as in (6), *what*-phrases differ from *which*-phrases in allowing a non-restricted (hence non-D-linked) interpretation (Heim 1987). By testing *what*-questions in contexts that force a non-D-linked interpretation, the prediction is that NPA systematically falls on the verb in (6), but is subject to optionality in (3)–(5). Under this account, the interpretation (D-linking *vs* non-D-linking) plays a crucial role.

We will discuss the implications of our results as to whether the mismatch between prosody and interpretation that characterizes Italian main wh-questions is determined at the syntax-prosody interface (Hypothesis 1) or through the interplay of the syntactic component with both the interpretive and the phonological component (Hypothesis 2).

**References**


Bocci, Giuliano, Valentina Bianchi & Silvio Cruschina. 2021. Focus in wh-questions. Evidence from Italian. *Natural Language & Linguistic Theory* 39(2): 405–455.

Bocci, Giuliano, Silvio Cruschina & Luigi Rizzi. 2021. On some special properties of *why* in syntax and prosody. In *Why is 'why' Unique? Its syntactic and semantic properties*, Gabriela Soare (ed.), 293–316. Berlin: De Gruyter.

Calabrese, Andrea. 1982. Alcune ipotesi sulla struttura informazionale della frase in italiano e sul suo rapporto con la struttura fonologica. *Rivista di Grammatica Generativa* 13: 489–526.

Heim, Irene. 1987. Where does the definiteness restriction apply? evidence from the definiteness of variables. In *The Representation of (In)definiteness*, Eric Reuland & Alice ter Meulen (eds.), 21–42. Cambridge, MA: MIT Press.

Marotta, Giovanna. 2001. I toni accentuali nelle interrogative aperte (wh-) dell'italiano di Lucca. In *Atti del II congresso di studi dell'Associazione Italiana di Linguistica Applicata*, Camilla Bettoni, Antonio Zampolli, and Daniela Zorzi (eds), 175–194. Perugia: Guerra Edizioni.

# Bilinguals and the uniformity of discourse strategies: A pilot study of Topic Preposing in Puerto Rican codeswitching

Ángel L. Jiménez-Fernández, Melvin González Rivera & M. Carmen Parafita Couto
(University of Seville, University of Puerto Rico & University of Leiden)

Keywords: English-Spanish codeswitching; Puerto Rican bilingual speakers; topic preposing; experimental work; syntax-information structure interface

Studies on the origins of Antillean Caribbean Spanish (ACS) and its grammatical properties have been characterized by the possible influence that other languages, mostly English, have had on this variety (Ortiz López 2016). Code-switching is a widely used strategy in situations where bilingualism is prevalent and languages come into contact. In Puerto Rico this topic has not received much attention, but studies on code-switching in this Caribbean variety from different perspectives are beginning to emerge (Clachar 2016; Guzzardo Tamargo & Vélez Avilés 2017; Balam et al. 2020). These studies on code-swicthing have focused mainly on issues other than those belonging to the syntax-information structure interface.

Our main goal in this work is to examine Topic Preposing in the grammar of Puerto Rican bilingual speakers in situations of English/Spanish code-switching. We analyze (Clitic) Left Dislocation both in matrix and embedded sentences. Our starting hypothesis is that, when English is the matrix language (i.e. English provides the morphosyntactic frame), preposed topics will be less accepted than when this matrix language is Spanish, due to the rigid SVO order in English in contrast with Spanish.

To carry out this (pilot) empirical study, a total of 21 bilingual Puerto Rican speakers completed an acceptability judgement task, based on a 4-point Likert scale (Stadthagen González et al. 2017) and a questionnaire about their linguistic background. Both surveys were built and run by using *Qualtrics*. The test included 35 tokens containing examples where the matrix language was either English or Spanish (see examples 1-5) and examples where there was no code-switching (6-9). It also contained 6 sentences to control the informant's attention to the precise task that they were asked to fulfill.

Regardless of whether the matrix language was English or Spanish, the statistical analysis revealed that the bilingual informants tended towards a slightly acceptable position in regards to the code-switched examples, with means between 2.14 and 3.38.Furthermore, when the matrix language is English and the example in fully in English, the mean lowers to 1.90, which is expected due to the rigid word order in English (as opposed to the Spanish examples, with means between 2.8 and 3.88, also to be expected because of the flexibility of the language; Jiménez-Fernández & Miyagawa 2014).

Our preliminary conclusion, concerning the syntax-discourse interface, is that code-switching is acceptable independently of the matrix language (cf. Stadthagen González et al. 2017 for similar results in the nominal domain), despite the rigidity/flexibility distinction, contrary to our initial hypothesis. This pioneering study opens a research line about the connection between syntax and information structure in code-switching in ACS, and bilingualism in general.

Examples:

1. El arroz con habichuelas Ana hated as a child, but she liked pasta.
   the rice with beans …
2. Antonio said that el arroz con habichuelas he hated as a child.
3. El arroz con habichuelas Antonio said that he hated as a child.
4. El rice and beans Ana lo odiaba de pequeña, pero le gustaba la pasta.
   the CL.ACC hated of small, but CL.DAT liked the pasta
   'Ana hated rice and beans as a child, but she liked pasta.'
5. Antonio dijo que el rice and beans lo detestaba cuando era pequeño.
   Antonio said that the … CL.ACC hated when was small
   'Antonio said that he hated rice and beans when he was a child.'
6. Rice and beans Ana hated as a child, but she liked pasta.
7. The rice and beans Ana hated as a child, but she liked pasta.
8. Antonio dijo que el arroz con habichuelas lo detestaba cuando era pequeño.
9. El arroz con habichuelas Antonio dijo que lo detestaba cuando era pequeño.

Acknowledgments: This research was funded by Spain's Ministry of Science, Innovation, and Universities (MICINN), grant number PID2022-137233NB-I00.

**References**

Balam, Osmer, M. Carmen Parafita Couto & Hans Stadthagen-González (2020), Bilingual Verbs in Three Spanish/English Code-switching communities, *International Journal of Bilingualism*, 24(5-6), 952-967.

Clachar, Arlene (2016), Codeswitching among African-American English, Spanish and Standard English in Computer-Mediated Discourse: The Negotiation of Identities by Puerto Rican Students in Rosa E. Guzzardo Tamargo, Catherine M. Mazak & M. Carmen Parafita Couto (eds.), *Spanish-English Codeswitching in the Caribbean and the US*, Amsterdam/Philadelphia: John Benjamins, 61-80.

Guzzardo Tamargo, Rosa & Jessica Vélez Avilés (2017), La Alternancia de Códigos en Puerto Rico: Preferencias y Actitudes, *Caribbean Studies* 45(1/2), 43-76.

Jiménez–Fernández, Ángel L. & Shigeru Miyagawa (2014), A Feature-Inheritance Approach to Root Phenomena and Parametric Variation, *Lingua* 145, 276-302.

Ortiz López, Luis (2016), Dialectos del Español de América: Caribe Antillano (Morfosintaxis y Pragmática), in Javier Gutiérrez-Rexach (ed.), *Enciclopedia de Lingüística Hispánica*, London: Routledge. https://doi.org/10.4324/9781315713441

Stadthagen-González, Hans, Luis López, M. Carmen Parafita Couto & C. Alejandro Párraga (2017), Using Two Alternative Forced Choice Tasks and Thurstone's Law of Comparative Judgments for Code-Switching Research. *Linguistic Approaches to Bilingualism* 8(1), 67-97.

# WS12 Origin of Discourse Particles: Borrowing and Grammaticalization

# Contact-induced semantic change: The origin of the Sibe particle =*da*

Veronika Zikmundová & Jan Křivan
(Charles University & Charles University)

Keywords: discourse marker, limitation, Sibe, language contact, etymology

The aim of the paper is to elucidate the origin of the Sibe (Manchu, Jurchenic, Tungusic) clitic *=da* surveying its current functions (based on Zikmundová's fieldwork) in light of the data from contact languages of Sibe, taking into account a rich history of its contacts. Sibe speakers have been detached from the main body of Manchu speakers in Manchuria for almost 300 years. At their current location in Xinjiang they have been in contact with Turkic languages (Uyghur, Kazakh and others), two varieties of Mongolian, and Chinese, among others (cf. Janhunen 2013). Prior to their movement, contacts with Khorchin Mongol and Nanaic (Tungusic) are assumed.

The Sibe clitic *=da* has a central position within Sibe discourse markers. It is used in combination with other discourse markers, as well as independently for modelling of information flow. It has been analysed differently as a topic marker with limitation (1) as its basic meaning (Jang and Payne 2012) or as an identificational marker (2) (Kogura 2017).

(1) *muku* *oeme-maqe=da* *jiave-xei.*
water drink-CVB=PRT walk-PST
'He only drank water and went.' (Jang and Payne 2012: 166).

(2) *bi=da* *meNjangge.*
1SG=PRT right
'It's me.' (answering the question "Who is the principal of this school?") (Kogura 2017: 162)

Both analyses generally agree on the crucial importance of *=da* in maintaining narrative coherence (using Chafe's 1976 framework). Limitation semantics of =*da* including emphasis and foregrounding of a portion of speech is very close to the semantics of the Chinese particle *jiu*.

There is no straightforward etymology of the Sibe *=da.* Kogura (personal information) has suggested borrowing from Kazakh. However, the typical additive semantic and syntactic properties of the particle DA in Kazakh, and many non-Jurchenic Tungusic languages, which share this particle, are absent in Sibe.

Screening of Turkic, Tungusic and Mongolic language data has revealed that most Turkic and Tungusic languages employ a presumably cognate *=da* particle, yet with the additive meaning. Subsequently, we examined the data from the contact languages of Sibe in Xinjiang (Modern Chinese, Uyghur, Kazakh, and the Jungarian varieties of Mongolian), and compared it to the contact language data from a pre-Xinjiang period, based on descriptions of Manchurian Tungusic languages (particularly Nanaic).

Data from modern Chinese (Hole 2004) and from Nanaic (Sunik 1985, Avrorin 1961) bring us to the hypothesis that the semantic and syntactic shift of the Tungusic formerly additive particle *=dA* had already begun in Nanaic, and Sibe borrowed the particle already at this time. After the resettlement to Xinjiang, the use of this particle may have been reinforced by the contact with Kazakh,

and, at the same time, the increasing influence of Mandarin and widespread bilingualism in Sibe population led to the increased similarity in the use of *=da* and the Chinese *jiu*.

**References**


Avrorin, Valentin A. (1961), *Grammatika nanajskogo jazyka* (A grammar of Nanai), vol. 2, Leningrad: Nauka.

Chafe, Wallace L. (1976), Givenness, contrastiveness, definiteness, subjects, topics and point of view, in Li, Charles N. (ed) *Subject and topic*, New York: Academic Press, 25-55.

Hole, Daniel (2004), *Focus and Backgroung Marking in Mandarin Chinese: System and Theory behind cai, jiu, dou and ye*, Routledge.

Jang, Taeho, and Payne, Thomas (2012), Topic Marking and the Construction of Narrative in Xibe, in P. Suihkonen, B. Comrie, and V. Solovyev (eds), *Argument Structure and Grammatical Relations. A crosslinguistic typology*, John Benjamins, 151–176.

Janhunen, Juha (2013), The Tungusic languages: A history of contacts, in K. Juwon Kim, and K. Dongho (eds), *Current Trends in Altaic Linguistics: A Festschrift for Professor Emeritus Seong Baeg-in on his 80th Birthday*, Altaic Society of Korea, 17–60.

Kogura, Norikazu (2017), The function of particles =ni' and da in Sibe: Focus and Nominal Reference, in *Proceedings of the 13th Seoul International Altaic Conference*, 151–168.

Sunik, Orest P. (1985), *Ul'čskij jazyk. Issledovanija i materialy* (The Ulcha language: Explorations and materials), Leningrad: Nauka.

# Grammaticalization patterns of Yugan Khanty demonstratives

Schön Zsófia
(LMU Munich)

Keywords: Khanty, nominal demonstrative, manner demonstrative, particle, grammaticalization

In my paper, I study grammaticalization patterns of Yugan Khanty nominal demonstratives in spoken and written discourse (narrative texts and unstructured interviews from monolingual and bilingual speakers collected between 2010–2015 in Siberia).

Yugan Khanty is a local variant of Surgut Khanty, being one of the Eastern Khanty dialects, belonging to the Finno-Ugric language family. It differs from the other variants of Surgut Khanty, by means of phonotactics, lexicon and degree of Russian influence.

The aim of this pilot study is to describe how some demonstratives developed different functions and meanings beyond canonical demonstrative functions, and others not. There are eight Yugan Khanty nominal demonstratives which can be told apart by the following characteristics: a) proximity, b) visibility, c) inflection, and d) adnominal/independent use. Csepregi (2017) claims that four of these demonstratives (namely *tʲi* 'this – proximal, abstract, non-inflectable, adnominal'; *tʲu*ː 'that – distal, abstract, non-inflectable, adnominal' (see the first word in (1)); *tʲit* 'this – proximal, abstract, inflectable, independent'; *tʲuːt* 'that – distal, abstract, inflectable, independent') have developed into particles in order to structure discourse, expressing meanings like 'well' or affirmation. Csepregi observes these new functions only in spoken language (2017: 110–115).

(1) [Two men lay down, one in a deserted house, the other one in a cemetery].

| | | | | | |
|---|---|---|---|---|---|
| ***tʲu***ː | *jiməŋ* | *puːɬ-e* | *oɬint-əm* | *ot* | *məttə* |
| **DEM.DIST** | holy | village-LAT | lay_down-PTCP.PST | thing | it_is_audible |
| *qoːɬ-ɬ-ətəɣ* | *əj* | *teːt-nə*[;] | ***tʲi***ː | | |
| hear-PRS-SG<3SG | one | time-LOC | **PTCL** | | |

'That one who had lain down on the cemetery, it [= the awakening of the death] is audible, one day he [= the man] hears it; **So**!' (OUDB 2017)

However, my data shows that only two of the four demonstratives, *tʲi* and *tʲuː*, have undergone changes, first into manner demonstratives (second to last word in (2)), and then into particles, and this not exclusively in spoken language but in narratives too. Example (2) shows *tʲi* as indicating a new topic, introducing a clause in which the two protagonists leave a place they had visited.

(2)

| | | | | | | |
|---|---|---|---|---|---|---|
| *jeː* | *qoβɬi* | *βeːnɬi* | *βoɬ-m-eɬ* | | *puɯːrnə* | |
| well | long | shortly | live-PTCP.PST-3SG | | after | |
| ***tʲi***ː | *oːs* | *jaq* | *məɣ-i-nem* | *tʲi* | *mən-əɣən* | |
| **PTCL** | again | home | land-PROPR-APP | so | go[PST]-3DU | |

'Well, after they lived long or short, then[?] they went so back again to their homeland.' (OUDB 2019)

In order to find explanations for the similarities and discrepancies, I analyze the examples using Fischer's (2006) theory. I will argue that even if a common tendency can be observed in the use of *tʲi* and *tʲu*ː from changing from a demonstrative into a particle, different patterns can be established.

While *tʲu*: dominantly stands with verbs, and has a stronger manner reading, *tʲi* additionally overtook some functions from the Russian particle *vot* and as such may make an utterance of its own (last word in (1)).

**References**


Чепреги, Марта [Csepregi, Márta] (2017), *Сургутский диалект хантыйского языка*. Ханты-Мансийск: Обско-Угорский Институт Прикладных Исследователя и Разработок.

Fischer, Kerstin (2006), Towards an understanding of the spectrum of approaches to discourse particles: Introduction to the volume. *Approaches to Discourse Particles*. Ed. K. Fischer. (Studies in Pragmatics 1.) Oxford–Amsterdam: Elsevier, 1–20.

OUDB 2017 = OUDB Yugan Khanty (2010–) Corpus. Text ID 1543. http://www.oudb.gwi.uni-muenchen.de/?cit=1543 (Accessed on 2024-05-31)

OUDB 2019 = OUDB Yugan Khanty (2010–) Corpus. Text ID 1648. https://www.oudb.gwi.uni-muenchen.de/?cit=1648 (Accessed on 2024-05-31)

# Language contact and grammaticalization: The phasal adverb 'already' in Romance

Eva-Maria Remberger
(University of Vienna)

Keywords: phasal adverbs, discourse markers, language contact, Romance, left periphery

This paper deals with some common properties of Romance 'already'-elements (AEs), as in Italian *già*, French *déjà* (Hansen & Strudsholm 2008), and in other Romance languages (Bazzanella 2005 Calaresu 2015). AEs originally are phasal adverbs, which Van der Auwera (1998:25) described by the two parameters 'change of state' or 'continuity', including their negation. Similarly, Kroon & Risselada posited three basic ingredients for the Latin AE *iam*, namely, 'phasality', 'polarity', and '(counterpresuppositional) focality'. Thus, in a sentence like *John is already at home*, there is a change of state from one phase to another, which is focused, and sometimes therefore perceived as unexspected. The functions of AEs vary corresponding to the parameters involved and do not always follow pre-established pathways. Particularly relevant to the topic of the workshop are possible interpretations of AEs as discourse markers: as has been observed, a discourse marker can be interpreted as a kind of "paralinguistic", gesture-like element, which is easily borrowed from one language into another, because it exhibits a high degree of "pragmatic detachability" (Matras, 1998, 2009:153f; Stolz, 2008:23); see also the place of discourse particles in the *borrowing scale* proposed by Thomason & Kaufman (1988) or the *scale of adoptability* proposed by Haugen (1950:224). Some interpretations of AEs show minimal variation and change and the different functions of AEs observed often can be interpreted as a result of language contact, when either matter or patterns of use of the AE are "copied" (cf. Johanson 2008) from another language. The cases that will be discussed more thoroughly in this context are particular discourse properties of Sardinian affirmative *ge* (probably ascribable to contact with Spanish), cf. (2), the use of Spanish affirmative *ya* in the Basque country (cf. González 2000), cf. (3), and special usage contexts of Northern Italian *già* (where contact with French seems to be at work, cf. e.g. Fedriani & Miola 2004, Cerruti 2009, Squartini 2014), cf. (4).

(2) ***Gei*** *ddu scit* *ca deu a crésia no ddu andu ...* (Srd.)
AE it=you.know that I to church not there=I.go
'You (surely) know that I don't go to church... ' (Lobina 2004: 286)

(3) A: *y en el pueblo ahora todavía se acuerda de / si tiene que hablar en euskera puede?*
B: *ah! Sì sì* ***ya*** *suelo hablar / con los niños también /* (Sp.)
'And in the town, do you still remember… if you need to use Basque, do you remember?' 'Yes, yes, I often speak Basque. With the kids too, yes.' (González 2000: 314–315)

(4) *Dove vi siete sposati,* ***già****?* (N.-It.)
where you are married AE
'Where did you get married, again?' (Fedriani & Miola 2014: 181)

My proposal is that AEs can either "climb up" in syntactic structure in a process of grammaticalization from the position of a phasal adverb to a left peripheral discourse element, or, being highly detachable and thus noticeable at the discourse interface in the left periphery, they can borrow new functions in a kind of pragmatic calque from another language.

Bazzanella, Carla et al. 2005. Dal latino *iam* agli esiti delle lingue romanze: verso una configurazione pragmática complessiva. *Cuadernos de Filología Italiana* 12, 49–82.

Calaresu, Emilia. 2015. Modalizzazione assertive e funzioni confirmative/asseverative dell'avverbio fasale GIÀ: Usi preverbali di Sardegna (sardo e italiano regionale) vs. usi olofrastici in italiano standard. In Kirsten Kragh and Jan Lindschouw (eds.), *Les variations diasystématiques et leurs interdépendance dans les langues romanes*. Strasbourg: ELiPhi, 113–127.

Cerruti, Massimo. 2009. *Strutture dell'italiano regionale: morfosintassi di una varietà diatopica in prospettiva sociolinguistica*. Frankfurt: Lang.

Fedriani, Chiara and Emanuele Miola. 2014. French *dejà*, Piedmontese Regional Italian *già*: A case of contact induced pragmaticalization. In Chiara Ghezzi,, Piera Molinelli (eds.). *Pragmatic markers from Latin to the Romance languages*. Oxford: Oxford University Press, 166–189.

González, María José. 2000. Emphatic affirmation in the Spanish of the Basque Country: the equivalence between *ya* and *sí (que)*. In Héctor Campos (ed.). *Hispanic Linguistics at the Turn of the Millennium*. Somerville, MA: Cascadilla Press, 308–321.

Hansen, Maj-Britt M. and Erling Strudsholm. 2008. The semantics of particles: advantages of a contrastive and panchronic approach: a study of the polysemy of French *déjà* and Italian *già*. *Linguistics* 46/3, 471–505.

Haugen, Einar. 1950. The analysis of linguistic borrowing. *Language 26*, 210–231.

Koike, Dale A. 1996. Functions of adverbial *ya* in Spanish narrative discourse. *Journal of Pragmatics* 25, 267–279.

Johanson, Lars 2008. Remodeling grammar. Copying, conventionalization, grammaticalization. In Peter Siemund and Naomi Kintana (eds.). *Language contact and contact languages.* Amsterdam/Philadelphia: John Benjamins, 61–79.

Kroon, Caroline and Rodie Risselada. 2002. A feature analysis of the Latin particle *iam*. *Belgian Journal of Linguistics* 16, 65–78.

Matras, Yaron. 1998. Utterance modifiers and universals of grammatical borrowing. *Linguistics 36*, 281–331.

Matras, Yaron. 2009. *Language contact*. Cambridge: Cambridge University Press.

Squartini, Mario. 2013. From TAM to discourse. The role of information status in North-Western Italian *già* 'already'. In Liesbeth Degand, Bernd Cornilie, Paola Pietrandrea (eds.). *Discourse Markers and Modal Particles. Categorization and description*. Amsterdam / Philadelphia: John Benjamins, 163–190.

Squartini, Mario. 2014. The pragmaticalization of 'already' in Romance: From discourse grammar to illocution. In Chiara Ghezzi and Piera Molinelli (eds.). *Discourse and Pragmatic Markers from Latin to the Romance Languages*. Oxford: Oxford University Press, 190–210.

Remberger, Eva-Maria. 2020. Discourse and Pragmatic Markers in the Romance Languages. In *The Oxford Research Encyclopedia of Linguistics*. Oxford University Press.

Thomason, Sarah G. and Terrence Kaufman. 1988. *Language contact, creolization and genetic linguistics*. Berkeley: University of California Press.

Traugott, Elizabeth and John Waterhouse. 1969. *Already* and *yet*: A suppletive set of aspect markers? *Journal of Linguistics* 5, 287–304.

# Polish deconverbal discourse markers with the negation marker

Marzena Stępień
(University of Warsaw)

Keywords: pragmaticalization, discourse markers, negation, converbs, Polish

The paper focuses on the problem of Polish discourse markers or discourse structuring elements (DSE; on the problem of the terms referring to this phenomenon c.f. Fischer 2006) consisting of at least two elements: the exponent of negation *nie* (ang. *not*) and an expression homographic with an adverbial participle (converb - Dryer, Haspelmath 2011; Haspelmath, König 1995) such as e.g. *nie mówiąc o_*(*to say nothing of _, let alone_*), *nie wspominając o_* (*not to mention_*), *nie wytykając_* (*not to mention names*), *nie ukrywając* (*not denying*) etc.

A similar process of the gradual transformation of converbs into separate lexical units was the subject of the analysis in: Kortman & König (1992), Ramat (1992), Brinton & Traugott (2005), Brinton (2012) (mostly in German and Romance languages). In Polish, the problem was addressed by Weiss (2005) who identified and described Polish conjunctions, prepositions and adverbs of converbial origin, as the result of the degrammaticalization and lexicalization processes (van der Auwera (2002).

However, it was Birzer who drew attention to the Russian and Polish DSE of converbial origin (c.f. Birzer 2015: Russian *voobšče govorja* and Polish *ogólnie mówiąc – generally speaking*, Birzer 2017a,b) as the results of pragmaticalization process interpreted as grammaticalization of discourse functions (Diewald 2011). Nevertheless, the development of DSE consisting of negation marker *nie* and expression homographic with a converb has not been scrutinized yet.

The list of expressions being the subject of the further research was excerpt in 3 steps: (1) literature review, i.a. Weiss (1975, 2005), Bogusławski, Wawrzyńczyk (1993), Bogusławski, Danielewiczowa (2005), Moroz (2010), selected dictionaries of contemporary Polish (ISJP, USJP, WSJP); (2) formulation of the hypotheses for the semantic-syntactic features of matrix verbs (regularity in valency patterns cf. valency dictionaries: SGSCzP, Walenty; shared semantic components), (3) verification of the list in the text corpora (NKJP, Spokes) accompanied by manual inspection of the data.

The objective of the research, conducted with corpus-based qualitative-quantitative analysis, was to determine syntactic changes accompanying the process of pragmaticalization of expressions in question along with indicating semantic features of the matrix verbs.

The research allowed to establish the stage of pragmmaticalization of DSE (from construction through expression with generalized conversational implicature to lexical unit; Levinson (1995, 2002), Traugott & Dasher (2002)). Most of them are results of univerbation understood as the loss of the ability to open any valency slots; in other cases, valency pattern of DSE is much reduced in comparison to the matrix verb. Whereas some of DSE have the status of lexical units (idioms, e.g. *nie mówiąc o_*(*to say nothing of _, let alone_*), *nie wspominając o_* (*not to mention_*)) with a component: 'X, saying that he or she will not say about something that p, is actually saying: p', others shall be considered rather as constructions with a generalized conversational implicature which has not been fully coded yet.

**References**


van der Auwera, Johan (2002), More thoughts on degrammaticalization, in I. Wischer, G. Diewald (eds.) (2002), *New Reflections on Grammaticalization [Typological Studies in Language 49]*, Amsterdam/ Philadelphia: John Benjamins, 19–29.

Birzer, Sandra (2015), Generally speaking, connectivity and conversation management combined: the functions of Russian voobšče govorja and Polish ogólnie mówiąc, *Russian Linguistics* 39(1), 81–115.

Birzer, Sandra (2017a), Assessing the role of pattern and matter replication in the development of Polish discourse structuring elements based on non-finite verba dicendi. Part 2 – Historical development, *Zeitschrift für Slavische Philologie* 73(2), 397-440.

Birzer, Sandra (2017b), Assessing the role of pattern and matter replication in the development of Polish discourse structuring elements based on non-finite verba dicendi. Part 1 – Synchronic state, *Zeitschrift für Slavische Philologie* 73(1), 159-185.

Bogusławski, Andrzej & Danielewiczowa, Magdalena (2005), *Verba polona abscondita. Sonda słownikowa III*, Warszawa: Elma Books.

Bogusławski, Andrzej & Wawrzyńczyk Jan (1993), *Polszczyzna, jaką znamy. Nowa sonda słownikowa*, Warszawa: Wydawnictwo Uniwersytetu Warszawskiego.

Brinton, Laurel (2012), "The ghosts of old morphology". Lexicalization or (de)grammaticalization?, in: K. Davidse, T. Mortelmans, L. Brems, T. Breban (eds.), (2012), *Grammaticalization and language change: new reflections*, Amsterdam/Philadelphia: John Benjamins Publishing Company.

Brinton, Laurel, Traugott, Elizabeth Closs (2005), L*exicalization and Language Change*, Cambridge: CUP.

Diewald, Gabrielle (2011), Pragmaticalization (defined) as grammaticalization of discourse functions, *Linguistics* 49(2), 365-390.

Dryer, Matthew S. & Haspelmath, Martin (eds.), 2011, *The world atlas of language structures online*, Munich: Max Planck Digital Library; [online] http://wals.info.

Fischer, K., (ed.), 2006, Approaches to Discourse Particles. Amsterdam et al. Haspelmath, M., K.nig, E. (eds.), (1995), *Converbs in cross-linguistic perspective*, Berlin: Mouton de Gruyter.

Haspelmath, Martin & König, Ekkehard (eds.) (1995), *Converbs in cross-linguistic perspective. Structure and meaning of adverbial verb forms – adverbial participles, gerunds* (*Empirical Approaches to Language Typology; 13)*, Berlin & New York: Mouton de Gruyter.

Kortmann, Bernd, König, Ekkehard (1992), Categorial reanalysis: The case of deverbal prepositions, *Linguistics* 30, 671–697.

Levinson, Stephen C. (1995), Three levels of meaning, in F. R. Palmer (ed.) (1995), *Grammar and meaning*, Cambridge: Cambridge University Press, 90-115.

Levinson, Stephen C. (2000), *Presumptive meanings: the theory of generalized and conversational implicature*, Cambridge: The MIT Press.

Moroz, Andrzej (2010), *Parenteza ze składnikiem czasownikowym we współczesnym języku polskim*, Toruń : Wydawnictwo UMK.

Ramat P., 1992, Thoughts on degrammaticalization, *Linguistics* 30(3), 549–560.

Traugott, Elizabeth Closs, Dasher, Richard B. (2002), *Regularity in semantic change*, (Cambridge Studies in Linguistics 97), Cambridge: Cambridge University Press.

Weiss, Daniel (1977), *Syntax und Semantik polnischer Partizipialkonstruktionen im Rahmen einer generativ-transformationellen Sprachbeschreibung*, Bern-Frankfurt: Peter Lang.

Weiss, Daniel (2005), Nowe przyimki o pochodzeniu imiesłowowym? In M. Grochowski (ed.) (2005), *Przyimki i przysłówki. Studia ze składni i semantyki języka polskiego,* Toruń: Wydawnictwo UMK, 175-205.

Dictionaries:

ISJP: Bańko, Mirosław (ed.) (2000), *Inny słownik języka polskiego*, Warszawa: PWN.

SGSCzP: Polański, Kazimierz (ed.) (1980-1992), *Słownik syntaktyczno-generatywny czasowników polskich*, Wrocław / Kraków : Zakład Narodowy im. Ossolińskich / Instytut Języka Polskiego PAN,.

USJP: Dubisz, Stanisław (ed.) (2003), *Uniwersalny słownik języka polskiego*, Warszawa: PWN. WSJP: Żmigrodzki, Piotr (ed.) (2007-2023), *Wielki słownik języka polskiego*, Kraków (online: http://www.wsjp.pl).

Corpora:

NKJP: Narodowy Korpus Języka Polskiego (National Corpus of Polish: www.nkjp.pl): Przepiórkowski, Adam, Bańko, Mirosław, Górski, Robert, Lewandowska-Tomaszczyk Barbara (2012), *Narodowy Korpus Języka Polskiego*, Warszawa.

SPOKES – Conversational Corpus http://spokes.clarin-pl.eu/: Pęzik, Piotr (2015), Spokes – a Search and Exploration Service for Conversational Corpus Data, in J. Odijk, *Selected papers from the CLARIN 2014 Conference*, October 24–25, 2014, Soesterberg, The Netherlands, Linköping: Linköping University Electronic Press, pp. 99–109.

Walenty: http://zil.ipipan.waw.pl/Walenty: Przepiórkowski Adam, Hajnicz Ewa, Patejuk Agnieszka, Woliński Marcin, Skwarski Filip, Świdziński Marek (2014), Walenty: Towards a comprehensive valence dictionary of Polish, in N. Calzolari et al. (eds.), *Proceedings of the Ninth International Conference on Language Resources and Evaluation, LREC 2014*, Reykjavík, Iceland. European Language Resources Association (ELRA), pp. 2785–2792.

# Interjection? Conjunction? Discourse marker? The history of the Hungarian *nohát ~ nahát* 'well! wow!'

Mónika Varga & Csilla Ilona Dér
(HUN-REN Hungarian Research Centre for Linguistics & HUN-REN Hungarian Research Centre for Linguistics)

Keywords: discourse marker, interjection, conjunction, pragmaticalization, mirativity

The fused form of the interjection/discourse marker *no~na* 'now, so, well, wow' and the discourse marker *hát* 'well, so' can convey a wide range of attitudes and emotions: surprise, appreciation, threat, but it also has discourse structuring roles: it can express summary, conclusion, or topic-shift. The stand-alone usage of it *nohát~nahát* in today's Hungarian primarily denotes mirativity:

> (1) *'– Száznegyvenezer – mondta Cilike olyan hangszínnel, mint amikor egy kártyajátékos kijátssza az adut. – **Nahát!** – csodálkoztunk.'* (MNSz2, #1845137,doc#97, literature)
> '"One hundred and forty thousand," Cilike said in a tone like a card player who plays the trump card. "**Wow!**" – we wondered.''

The earliest occurrences of the two elements (17th c.) next to each other (written separately) in addition to nagging/urging ('let's do it') typically carried inferential meaning ('so, therefore') (ÚESz) (2), which meaning was also a determining stage in the grammaticalization of the discourse marker *hát* (cf. Schirm 2007–2008) and formed the basis of many of its additional functions (e.g. evidentiality, specification):

> (2) *monda az Aszony az Urnák hogy **no hat** tuttad az en kurvasagomat, tudua laktal velem* [Mv, MvLt 291. 146a) (1638, SzT. IX. 4: 685)
> 'the woman said to the lord **well** you knew my whoredom, you lived with me knowing that'

In today's Hungarian, like the element *tehát* 'therefore, so', which has also become a discourse marker, *nohát~nahát* can still express the continuation of what is being said, but it is typically associated with an attention-grabbing function ('this here is interesting/important in the story'), which *tehát* cannot (3):

> (3) *Ehhez tudni kell, hogy Kuncz Aladárt mi, kolozsváriak, csak per Dadi hívtuk és tituláltuk, vagyis gyermekkori becenevén. **Nohát** Dadi, az én kedves tanárom és lakótársam a penzióban, sokat beszélt rólam nagy barátjának, Ady Endrének* (MNSz2, #18243127, doc#363, literature)
> 'For this, it is necessary to know that we, the people of Cluj, only called and titled Kuncz Aladár per Dadi, that is, by his childhood nickname. **Well,** Dadi, my dear teacher and roommate in the boarding house, talked a lot about me to his great friend Endre Ady'

The aim of this presentation is to answer how and in what order *nohát~nahát* acquired its numerous functions that can be observed in today's Hungarian, and which of these were formed after the fusion of the two elements. We will discuss which of the variants (*no hát, na hát, nohát, nahát*) are more grammaticalized/pragmaticalized and richer in functions, for which we will analyze diachronic and synchronic databases (TMK, MTSz, MNSz2), using the discourse variation analysis method of variational pragmatics (Pichler 2013).

The examination also covers the positions (left/right periphery, internal or stand-alone) of both compound forms in relation to their functions, where possible, we also perform statistical tests. Moreover, we will determine the status of *nahát~nohát* in its different development phases and

occurences (interjection like *no/na* 'now, wow', discourse marker like *hát* 'well, so', or conjunction similar to *tehát* 'so, therefore').

Acknowledgments: This research was supported by the Hungarian National Research, Development and Innovation Office of Hungary, projects No. FK 135186.

**References**

Pichler, Heike (2013), *The structure of discourse-pragmatic variation,* Amsterdam / Philadelphia: John Benjamins.

Schirm, Anita (2007–2008), A *hát* diskurzusjelölő története [The history of the discourse marker *hát* 'well, so'], *Nyelvtudomány* III–IV., 185–201.

MNSz2 = *Magyar Nemzeti Szövegtár 2.* [Hungarian Gigaword Corpus 2., 1998–, 1.5 billion words] http://clara.nytud.hu/mnsz2-dev/

MTSz = *Magyar Történeti Szövegtár* [Hungarian Historical Corpus, 1772–2010, 30 million words] http://clara.nytud.hu/mtsz/run.cgi/first_form

SzT = *Erdélyi Magyar Szótörténeti Tár* [Transylvanian Hungarian Etymological Dictionary] IX/4. https://mek.oszk.hu/08300/08370/pdf/emszt_09_04.pdf

TMK = *Történeti Magánéleti Korpusz* [Historical Corpus of Personal Life: Middle Hungarian Court Records, private correspondences, 16–18th cent., 1 million words) http://tmk.nytud.hu/3/

ÚESz = *Új Magyar Etimológiai Szótár* [New Etymological Dictionary of Hungarian], https://uesz.nytud.hu/index.html

# From intensifiers to discourse particles and conjunctions: The case of East Caucasian

Aigul Zakirova & Timur Maisak
(University of Potsdam & HSE University)

In this talk, we discuss the evolution of intensifying pronouns into discourse elements in two distantly related lects: Upper Andi (<Andic, East Caucasian) and Rutul (<Lezgic, East Caucasian). Although these are lects without a long historical record, we will show that their synchronic data point at possible grammaticalization processes yielding elements which can be classified as restrictive particles and adversative conjunctions.

Our data is partly elicited and partly from spoken and written texts. The elicited data come from the authors' fieldwork with speakers of Zilo Andi (Botlikh district, Dagestan, Russia, fieldwork conducted remotely) and Kina Rutul (Rutul district, Dagestan, Russia). The text sources under investigation were Magomedova & Alisultanova (2010) for Upper Andi, Makhmudova (2001) and Alieva (2013) for Rutul.

The Upper Andi pronoun *ʒi<GN>gu*[1] combines reflexive, intensifying and logophoric uses. The same polyfunctionality is found with the Rutul pronoun *<GN>iž* and is overall typical of East Caucasian intensifiers (Testelec & Toldova 1998, Lyutikova 2002).

Additionally, our data show that Upper Andi *ʒi<GN>gu* also has two types of uses which have not been previously described for East Caucasian intensifiers. We labelled the first *restrictive* (1), as it seems to be functionally equivalent to a restrictive particle (cf. English *only*, *just*). The second use (2) we labelled *adversative*, as it is functionally equivalent to conjunctions *but* and *however*.

(1) *muħammadi* *w-aχo-r-sːu,* *he-w* ***ʒi-w=gu*** *helli-r.*
Muhammad M-fight-PROG-NEG this-M **self-M=EMPH** run-PROG
'Muhammad isn't fighting, he is **just** running around'.

(2) *di-j* *q'oroq'oj* *b-iʁi* *rej-ʔo* *w-uʔinn-u,*
I-DAT need N1-stop.AOR yard-LAT M-leave-INF
***ʒi-r=gu*** *ts'a* *r-eɬ:i.*
**self-N2=EMPH** rain N2-go.AOR
'I wanted to go outside **but** it started to rain'.

Rutul *<GN>iž* also features adversative uses (3) but no restrictive uses analogous to (1).

(3) *anuc-a* *jɨw-a* *raǯekim* *lɨž* *i‹w›i-r-i,*
Anuts-ERG seven-MULT time lot <N1>draw.PFV-CVB-COP
***již*** *sada-ki* *iʁi‹b›q'ɨ-r-diš.*
N2-self once-ADD <N1>win.PFV-CVB-COP.NEG
'Anuts drew lots seven times but never won' (Makhmudova 2001: 197).

The restrictive component is present in some uses of intensifiers (Lyutikova 2002), cf. *You yourself are to blame* (= 'no one else is to blame'). However, the restrictive use of *ʒi<GN>gu* in (1) is different from the interpretation described by Lyutikova, as the prejacent of *ʒi<GN>gu* is not an NP but a verb. Besides, *ʒi<GN>gu* allows scalar readings ('just X, not X+1'), whereas intensifiers usually only have complement exclusion readings ('only X, not Y'). We propose that this may be an

[1] <*GN*> is a slot for the gender-number agreement marking. The segment *-gu* goes back to the emphatic / intensifying clitic *=gu*, functionally equivalent to the Avar *=go* (Forker 2015).

indication of *ʒi<GN>gu* developing into a discourse particle. Another indication of this categorial reanalysis may be petrification of gender-number marking in Mukhad Rutul *<GN>iǯ,* which only occurs in this function as *jiǯ*.

**References**


Alieva, F. 2013. Svod pamjatnikov fol'klora narodov Dagestana [The Collection of folklore monuments of the peoples of Daghestan]. Vol 3. Fairy-tales. Moscow.

Forker, D. 2015. Towards a semantic map for intensifying particles: Evidence from Avar. STUF-Language Typology and Universals, vol. 68, pp. 485-513.

Lyutikova, Ye. 2002. Kognitivnaya tipologiya: refleksivy i intensifikatory [Cognitive Typology: Reflexives and Intensifiers], Moscow.

Magomedova, P. A., & Alisultanova, M. A. (Eds.). 2010. *Dunjalla baχunnirsːi χalgiltɬol gwanabol gwanab micˈːillasːol muχal* [Fairy tales from around the world in the Andi language]. Makhachkala.

Testelec, Ja. G. & S. J. Toldova. 1998. Refleksivnye mestoimenija v dagestanskix jazykax i tipologija refleksiva [Reflexive pronouns in Dagestanian languages and the typology of reflexive]. Voprosy Jazykoznanija, pp. 35–57.

Makhmudova, S. M. Morfologija rutul´skogo jazyka [Morphology of the Rutul language]. Moscow, 2001

# The puzzle of the "Ottoman" BRE and its relatives

Vladimir Panov
(HSE University)

A number of languages spoken across the Aegean, Adriatic, and the Black seas and beyond exhibit a variety of semi-bound discourse-level devices ("particles") whose morphosyntactic behavior and functions are similar across multiple idioms genealogically related and not (various branches of Indo-European, Turkic, Semitic). The spread region corresponds roughly to the northern part of the former Ottoman empire. Typically, such particles either precede a vocative noun phrase and thus behave as non-obligatory vocative markers or scope over the whole utterance. In the latter case, the particles are strongly addressee-oriented and may be roughly labeled as allocutive with additional pragmatic flavors such as speaker's attitude. The two modern Greek examples represent the two types:

Standard Modern Greek:

(1) *ja* *su* ***re*** *Janni!*
health your **VOC** Yannis
Hi Yannis!

(2) *ti* *thelis* ***re****?*
what want.PRS.2SG **ALLOC**?
What do you want [dude]?

In a part of the idioms of interest, such particles have clearly spread through direct borrowing and exhibit similar phonetic shapes – *bre*, *bire*, *vre*, *re, ra, be* etc. (further *BRE*). Traditionally, these are derived from the medieval Greek *more* (baby.VOC.M), which itself functions as a discourse particle *more/mori* in different languages (Joseph 1997). Although a nominal origin of *(b)re* is likely, the mainstream "baby" hypothesis is not unproblematic (Vastenius 2011).

At the same time, many idioms of the region exhibit etymologically unrelated particles whose functioning and meaning flavors parallel that of *BRE*, an example being the Turkish *ya*:

Turkish (Adıgüzel 2023: 61):

(3) S1: *niye* *bana* *söyle-me-di-n?*
why I.DAT tell-NEG-PST-2SG
Why didn't you tell me?

S2: *söyle-di-m* *ya!*
tell-PST-1SG YA
I did! [You should remember!]

Although there are numerous language-particular studies of such devices (e.g. Karachaliou & Archakis 2012; Adıgüzel 2023), the regional "big picture" remains unclear. In this talk, I put parts of the puzzle together and treat functions, morphosyntax, language contact patterns, and lexical sources in broad

historical and areal context in order to track the interaction of a variety of factors in the spread of *BRE* and its structural-functional relatives.

Given that the scope of the talk is providing a preliminary "big picture", my data is of a mixed nature. I analyze and compare existing language-specific descriptions of around a dozen idioms of the region, I use data from available spoken corpora and I evoke data acquired from interviews with native speakers.

I conclude that he region presents a very complex mosaic in which matter borrowing, pattern replication, and genealogical inheritance are tightly intertwined. However, the very presence of the structural features of interest – sentence-level allocutive particles and adnominal vocative particles – is strikingly stable across the region and covers an area larger than the one in which the corresponding functions are expressed by *BRE*. Finally, I argue that studying discourse-level devices can teach us more about the history of language contact in an area than more "traditional" grammatical categories.

**References**


Adıgüzel, Fatih M. 2023. The functional spectrum of the Turkish pragmatic marker Ya. *Journal of Pragmatics* 212. 58–71.

Joseph, Brian D. Methodological Issues in the History of the Balkan Lexicon: The Case of Greek vré / ré and Relatives. In Vkitor A. Friedman, M. Belyavski-Frank, M. Pisaro & D. Testen (eds.), *Balkanistica Vol. 10 (Studies Dedicated to the Memory of Zbigniew Golab 19 March 1923 - 24 March 1994)*. 255-277.

Karachaliou, Rania & Argiris Archakis. 2012. Το ρε ως δείκτης του απροσδόκητου. Ανάλυση δεδομένων από συνομιλιακές αφηγήσεις [The Greek particle re as a marker of unexpectedness: Evidence from the analysis of conversational narratives]. *Studies in Greek Linguistics. Proceedings of the 32nd Annual Meeting of the Department of Linguistics, Faculty of Philosophy, Aristotle University of Thessaloniki* 172–183.

Vastenius, Anu. 2011. *Expressive Particles in Serbian, Bulgarian, Greek and Kurdish. BA Thesis in General Linguistics.* Lund: Lunds universitet.

# Interaction and competition between borrowed and native discourse particles in Khoekhoe (Khoe-Kwadi)

Alena Witzlack-Makarevich, Sylvanus Job & Anna Veselovsky
(Hebrew University of Jerusalem, University of Cape Town & Hebrew University of Jerusalem)

Khoekhoe is a Khoe-Kwadi language spoken primarily in Namibia. Since the early days of colonization by European settlers it has been in continuous contact with Afrikaans (Germanic). One of the results of this prolonged language contact is the large amount of Afrikaans discourse particles used in Khoekhoe, which include lexical items familiar from and extensively studies in other Germanic languages, such as *tox* (cf. Dutch *tog* and German *doch*).

In this first study of discourse particles in Khoekhoe, we explore the structural and semantic properties of both borrowed and native discourse particles in Khoekhoe on the basis of a large corpus (around 1 mil. words) of spoken and written Khoekhoe texts of various registers from a wide range of speakers with variable level of proficiency in Afrikaans. For this study, we identified 40 most frequent discourse particles (20 borrowed from Afrikaans and 20 native ones, 2611 and 9393 tokens respectively). We annotated their usage in the corpus for formal properties (e.g. position in the turn, syntactic position in a clause, presence of contiguous pause) and functions (using the taxonomy inspired by Crible 2006 and Zeevat 2006).

The discourse particles borrowed from Afrikaans are much more frequent in spontaneous spoke Khoekhoe (conversations) than in written registers. After surveying the properties of the discourse particles based on the annotation, we proceed with an in-depth discussion of a small number of particularly common particles and use these cases to highlight different outcomes of borrowing on the Khoekhoe system of discourse particles. We will show that the functions of the particle *tox* (< Afrikaans *tog*) in Khoekhoe, as in (1), resemble its counterpart in Afrikaans: it is used as a marker of common ground to indicate propositions that are assumed to be accessible to all communication partners (e.g. Fischer 2021). However, it is also frequently combined with the particle *mut,* as well as the native assertive particle and results in readings unattested in Afrikaans. The particle *mut* (<Afrikaans *moet* 'must') is reshaping the Khoekhoe system of modality: it is used in combination with the native irrealis particle *nî* to express strong necessity, as in (1), a function for which no dedicated native means of expression exist. The particle *nee* (< Afrikaans *nee* 'no') is being recruited as a common strategy to introduce reported discourse, as in (3). In this function it competes with the native particle *hî-î* (also used as the interjection 'no'), this competition occasionally results in double marking. Finally, by providing first corpus-based survey of the Khoekhoe discourse particles *hūga* and *koma*, we also address earlier claims (Conradie 2015) about the contribution of speakers of Khoe languages to the development of Afrikaans discourse particles *hoeka* 'ostensibly' and *kamma* 'supposedly', which were not inherited from Dutch.

Examples

(1) Context: the speaker draws conclusions at the end of the conversation

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *want* | *ǂgui* | *xū-n* | *ge* | ***tox*** | *ra* | *ī* | *sa-da* | *ǂnamipe* |
| because | many | thing-3C.PL | DECL | ***TOX*** | IPFV | happen | 1INCL-1C.PL | around |

'… because many things are happening around us.'

(2) *want* *ǂan=ta* *a* *xū-i* *ǀgui-i-a*
because know=1SG.SBJ STAT.PRS thing-3C.SG only-3C.S.OBL
*=ta* *ge* ***mut*** *nî* *dī*
=1SG.SBJ DECL ***MUT*** IRR do
'because I must do only things that I know.'

(3) Context: "Now why did you chase the boy?" my uncle asked him.
*"**nee*** *axa-b* *ge* *go* *ǀui-s* *ǀkha* *ǂnoa* *te,* *…"*
***NEE*** child-3M.SG DECL RPST stone-3F.SG with throw_at 1SG.OBJ
(He replied) "The child threw at me with sones, …"

**References**


Conradie, C. Jac. 2015. Die gebruik van Afrikaanse modale partikels deur voormalige Khoisprekers. *Stellenbosch Papers in Linguistics PLUS* 47, 35–57.

Crible, Ludivine. 2006. Towards an operational category of discourse markers: A definition and its model. In Fedriani Chiara & Andrea Sansó (eds), *Pragmatic markers, discourse markers and modal particles: New perspectives*, 99–124. Amsterdam: John Benjamins.

Fischer, Kerstin. 2021. Grounding and common ground: Modal particles and their translation equivalents. In Anita Fetzer & Kerstin Fischer (eds), *Lexical markers of common grounds*, 46-66. Leiden: Brill.

Zeevat, Henk. 2006. A dynamic approach to discourse particles. In Kerstin Fischer (ed.), *Approaches to discourse particles*. Leiden: Brill.

# Particle entanglement:
# The parallel grammaticalization of the proximal modal particles in Danish

Lennart Westergaard
(Lund University)

Keywords: Modal particles, grammaticalization, semantic change, paradigm, dialogue

The modal particles *nu* ('now') and *da* ('then') form a subparadigm among the Danish modal particles: the so-called proximal particles. They express 1) that the utterance of the speaker is in conflict with a proposition available in the context, and 2) that the proposition expressed by the speaker is identifiable or mutually manifest (in the case of *da*) or unidentifiable (*nu*) (cp. Andersen 1982: 80; Davidsen-Nielsen 1996: 290; Hansen & Heltoft 2011: 1054-1058):

(1) A: Hvad har du brugt pengene på?
B: Det skal jeg **da/nu** ikke svare på.
'A: How did you spend the money?
B: I don't **DA/NU** have to answer that question
(*da*: as opposed to what you think, and as you should know)
(*nu*: as opposed to what you think, but I don't expect you to know)'
(adapted from Hansen & Heltoft 2011: 1056)

Based on a corpus study covering the period from Middle Danish to Modern Danish, I will present an account of the development of these modal particles.

What is striking about their development is their parallelism. They develop along similar paths and the emergence of new meanings seems to be temporally coinciding with only few decennials between first attestations of the corresponding meanings.

Roughly, both modal particles seem to develop through the following stages:

| **Middle Danish** | | | **Early Mod. Danish** | **Mod. Danish** |
|---|---|---|---|---|
| temporal | > text connective | > modal | contextually motivated questions/ surprise questions | > proximal particle |

According to this scenario, the modal particles develop their conflict meaning through utterances where they express that some contextual element creates a conflict of hypotheses as in the following example from a 17th century play (a satire about the *Briefadel*):

(2) [discussing whether *countess* or *baroness* is a more noble title]
**Grevinden:** (…) vi er dog mere end Friherinder.
**Friherinden:** Hvad **nu**? Ere vi ikke ogsaa Herrestandspersoner?
'**Countess:** We are better than baronesses after all.
**Baroness:** What **NOW**? Aren't we nobles as well?'
(Grevens og Friherrens Komedie, 1675, p. 16)

In such an utterance, *nu* expresses that a new piece of information from the context generates a conflict of hypotheses and thereby makes the speaker ask the question. Similar examples can be provided for *da*. The development of similar interrogative modal particles with similar source meanings have been discussed in e.g. Detges (2008, Germ. *jetzt*), Wegener (2002, Germ. *denn*) and Haselow (2011; 2012, Engl. *then*). From such contexts it spread to declarative clauses where the modal particles primarily express conflicts of belief or similar as in (1).

I will argue that the parallel development of the two modal particles is due to parallel source meanings motivating similar uses in relevant bridging contexts. Similarities in meaning alone cannot explain the temporal coincidence of the changes, however. Arguably, throughout their history, *nu* and *da* form a paradigmatic hyper-construction in the sense of Diewald (2020; cp. Nørgård-Sørensen et al. 2011) where changes in one particle facilitate corresponding changes in the other through analogy.

**References**


Andersen, Torben (1982): Modalpartikler og deres funktion i dansk. *Danske Studier*, 86-95.

Davidsen-Nielsen, Niels (1996): Discourse particles in Danish. In: E. Engberg- Pedersen, M. Fortescue, P. Harder, L. F. Jakobsen, L. Heltoft (eds.), *Content, Expression and Structure. Studies in Danish Functional Grammar.*. Amsterdam/Philadelphia: John Benjamins, 283-314.

Detges, Ulrich (2008): Altfranzösisch or und deutsch jetzt in Fragen. In: E. Stark, R. Schmidt-Riese & E. Stoll (eds.), *Romanische Syntax im Wandel*. Tübingen: Narr, 418-436.

Diewald, Gabriele (2020). Paradigms lost–paradigms regained. In: Sommerer, L., & Smirnova, E. (Eds.), *Nodes and networks in diachronic construction grammar* (Vol. 27). John Benjamins Publishing Company, 277-315.

Hansen, Erik & Lars Heltoft (2011): *Grammatik over det danske sprog*. København & Odense: Det Danske Sprog- og Litteraturselskab & Syddansk Universitetsforlag.

Haselow, Alexander (2011). Discourse marker and modal particle. The functions of utterance-final *then* in spoken English, *Journal of Pragmatics 43* (14), 3603–3623.

Haselow, Alexander (2012). "Discourse Organization and the Rise of Final *then* in the History of English". In: I. Hegedüs and A. Fodor (eds)*. Selected Papers from the Sixteenth International Conference on English Historical Linguistics (ICEHL 16), Pecs, 23-27 August 2010*, John Benjamins Publishing Company, 153-176.

Nørgård-Sørensen, Jens, Lars Heltoft & Lene Schøsler (2011): *Connecting grammaticalisation*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Wegener, Heide (2002). ''The evolution of the German modal particle *denn''*, in G. Diewald and I. Wischer (eds.), *New reflections on grammaticalization.* Amsterdam: Benjamins, 379–394.


Primary sources


*Grevens og Friherrens Komedie*, 1675. Published 1871 by Soph. Birket Smith. Copenhagen: G. S. Wibes Bogtrykkeri.

# Decomposing the Tagalog contrastive particle *naman*

Henrison Hsieh & Michael Yoshitaka Erlewine
(Hong Kong Polytechnic University & National University of Singapore)

Tagalog is known for its rich inventory of second position clitics (Schachter & Otanes 1972), which include pronominal clitics and a closed class of adverbial clitics. In this talk, we discuss the synchronic grammatical status and uses of the contrastive particle *naman* and its likely grammaticalization path.

Schachter & Otanes (1972: 425) and AnderBois (2016) describe *naman* as a discourse particle with several uses, including introducing contrastive topics (1), expressing obviousness ('of course'), expressing "mild reproach" (S&O) in imperatives (2), and emphasizing the relevance of gradable predicate evaluations (3), among others. These prior works however do not substantially discuss the fact that *naman* appears to reflect the combination of *na* (expressing a change of state or temporal earliness, described as an "iamitive" in e.g. Klimenko 2022) and *man*, the (concessive) scalar particle meaning 'even' or 'even if.'

(1) Nag-a~aral si Linda. Nagla~laro naman si Carmen.
AV.IPFV~study NOM Linda AV.IPFV~play *naman* NOM Carmen
'Linda is studying. Carmen, on the other hand, is playing.' (Schacter & Otanes 1972)

(2) Tulung-an mo naman ako. help-
PV(IMP) 2SG.GEN *naman* 1SG.NOM
'Please help me. (Don't just sit there.)' (Schacter & Otanes 1972)

(3) Napaka-rumi naman nito!
EXCL-dirty *naman* this.GEN
'How dirty this is!' (Schacter & Otanes 1972)

We propose that *naman* indeed historically derives from *na* and *man*, but has grammaticalized into a single particle, with some residual behaviors reflecting its historical decomposition. For example, although monosyllabic clitics generally precede disyllabic clitics (see e.g. Kaufman 2010), *naman* can often be linearized early as if it is treated as two monosyllabic clitics; compare (4) and (5). At the same time, *naman* can also cooccur with the particles *na* and *pa* 'still,' even though multiple *na* or *na pa* combinations are otherwise ungrammatical (6); this suggests that *naman* is synchronically distinct from *na* and *man*. In addition, we will show that the semantic contribution made by *naman* cannot be derived transparently from the conventional semantics for *na* and *man*.

(4) Tulung-an mo {naman} din {naman} ako. help-
PV(IMP) 2SG.GEN *naman* also *naman* 1SG.NOM
'Please help me as well. (Don't just sit there.)'

(5) Tulung-an mo {*kasi} din {kasi} ako. help-
PV(IMP) 2SG.GEN because also because 1SG.NOM
'Help me as well. (You didn't before. Now look what's happened.)'

(6) Marami pa {naman /*na} =ng pagkain dito. many still *naman* IAM =LK food here ‘There’s still a lot of food here anyway.’

Finally, we also discuss the contemporary use of *na* ‘already’ and *naman* together to express repetitive ‘again,’ suggesting a next stage in the development of conventionalized discourse particle combinations.

**References**

AnderBois, Scott. 2016. A QUD-based account of the discourse particle *naman* in Tagalog. In *Proceedings of AFLA 23*, ed. Hiroki Nomoto, Takuya Miyauchi, Asako Shiohara, and Takuya Miyauchi, 20–34. Asia-Pacific Linguistics.

Kaufman, Daniel. 2010. *The morphosyntax of Tagalog clitics: A typologically driven approach*. Doctoral Dissertation, Cornell University.

Klimenko, Sergei B. 2022. Conjunctions, clisis, and topicalization in Tagalog. *Acta Lin- guistica Petropolitana* 18:61–117.

Schachter, Paul, and Fe T. Otanes. 1972. *Tagalog reference grammar*. University of California Press.

# From manner adverb to epistemic particle: Estonian *äkki* and *järsku*

Helle Metslang
(University of Tartu)

Keywords: epistemic modality, (inter)subjectivity, mirativity, particles, registers

The words *äkki* and *järsku* exhibit a similar polysemy: both function in Modern Estonian as adverbs meaning 'suddenly, abruptly, unexpectedly' (1) and as epistemic modal particles expressing medium certainty (2). Of these words, *äkki* is more frequent in modern language, and its modal use is more common than its lexical use. The word *järsku* is less frequent and primarily used as a manner adverb.

(1) *Laps hakkas **äkki~järsku** nutma.*
'The child **suddenly/abruptly/unexpectedly** started crying.'

(2) ***Äkki~järsku** on laps haigeks jäänud.*
'**Maybe/perhaps** the child has fallen ill.'

In the presentation, I will examine the usage of these words, seeking answers to the questions as: (a) how has the usage of these words as modal particles evolved? (b) what subjective and intersubjective functions do these words carry in contemporary Estonian language registers?

The approach is centered on usage-based and corpus-driven studies of discourse and is based on the ideas of discourse grammar (see e.g. Heine et al. 2021) and the concepts of subjectivity and intersubjectivity (see e.g. Narrog 2017). The method is corpus-pragmatic qualitative analysis with the use of quantitative data (see e.g. Rühlemann 2019). The data are taken from the corpora of Estonian of previous centuries (VAKK and CELL) and from the balanced corpus of modern Estonian registers "Pragmatics" (Prillop et al. 2021).

The preliminary results show that the modal use of *äkki* has developed through the semantic shifts 'suddenly/abruptly > unexpectedly > unpredictably > maybe, perhaps'. The mirative lexical meaning 'unexpectedly' is the source of the weak epistemic certainty (cf. Squartini 2018). The modal usage of *äkki* developed during the 20th century. Contact language support is also possible: a similar polysemy has been known for the Russian word *vdrug* since the 19th century (see Bonola 2016). In modern usage, subjective epistemicity is accompanied by intersubjectivity (3): various interrogative and directive communicative functions, their softening, mirativity (unexpectedness for the addressee, cf. Hengeveld & Olberts 2012).

(3)
(Library, the workday ends in a few minutes)
A: tere päevast.
'hello'
B: 'tere.
'hello'
A: vabandage, kas **äkki** võiks raamatu veel ära pikendada (Institutional interaction)
'Excuse me, could I ÄKKI renew a book'

In the second half of the 20th century, the word *järsku* also began to be used as an epistemic particle (4). As a developing particle, *järsku* copies the functions that *äkki* has developed step by step. The word *järsku*, however, is ideophonic both etymologically and in terms of phonetic shape (ETY; cf. Mikone 2002) and has likely emerged as an expressive, stronger alternative to *äkki*.

(4) Annika ära 'vannituppa 'mine seal on **järsku** 'märg.=hh (0.3) 'mis ta seal 'lödistas. (Everyday conversation)
'Annika, don't go into the bathroom, it might be (lit. 'is JÄRSKU') wet there. What has s/he splashed there'

The epistemic *järsku* occurs primarily in spoken language, presumably due to its expressive sound and more complicated written form. *Äkki* is common in spoken language as well, but is especially frequent in online discourse, which due to its lack of multimodality uses more particles than spoken language (Hennoste et al. 2021).

Acknowledgement: The study was supported by the Estonian Research Council (grants PRG341 and PRG2184).

**References**

Bonola, Anna (2016), Semantiko-pragmatičeskij potencial slova *vdrug* v russkom jazyke i ego perevod na ital'janskij: kontrastivnyj analiz na osnove parallel'nogo russko-ital'janskogo korpusa [Semantic and pragmatic potential of Russian word *vdrug* and its translations into Italian: contrastive analysis on the material of the parallel Russian-Italian corpus], *Vestnik Novosibirskogo gosudarstvennogo pedagogičeskogo universiteta* 6(2), 24–37. http://dx.doi.org/10.15293/2226-3365.1602.02

CELL = *The Corpus of Estonian Literary Language 1890–1990*. https://keeleressursid.ee/et/151-baaskorpus (31.5.2024)

ETY = *Eesti etümoloogiasõnaraamat* [Estonian etymological dictionary]. https://www.eki.ee/dict/ety/ (31.5.2024)

Heine, Bernd, Gunther Kaltenböck, Tania Kuteva, and Haiping Long (2021), *The Rise of Discourse Markers,* Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108982856

Hengeveld, Kees and Hella Olbertz (2012), Didn't you know? Mirativity does exist! *Linguistic Typology* 16(3), 487–503. https://doi.org/10.1515/lity-2012-0018

Hennoste, Tiit, Helle Metslang, Külli Habicht, and Külli Prillop (2021), Kuue (inter)subjektiivsuspartikli kasutus eesti keele registrites [The use of six (inter)subjectivity particles in Estonian registers], *Emakeele Seltsi aastaraamat* 66, 91−123.

Mikone, Eve (2002), *Deskriptiiviset sanat: määritelmät, muoto ja merkitys* [Descriptive words: definitions, form and meaning], Helsinki: Suomalaisen Kirjallisuuden Seura.

Narrog, Heiko (2017), Three types of subjectivity, three types of intersubjectivity, their dynamicization and a synthesis, in Van Olmen, Daniel, Herbert Cuyckens, and Lobke Ghesquière (eds.), (2017), *Aspects of Grammaticalization: (Inter)subjectification and directionality*, Berlin/Boston: De Gruyter Mouton, 19–46. https://doi.org/10.1515/9783110492347-002

Prillop, Külli, Tiit Hennoste, Külli Habicht, and Helle Metslang (2021), Ei saa me läbi „Pragmaatika“ korpuseta. Korpuspragmaatika ja pragmaatikakorpus [We cannot do without the "Pragmatics" corpus. Corpus pragmatics and corpus of pragmatics], *Mäetagused* 81, 161−176. https://doi.org/10.7592/MT2021.81.pragmaatika

Rühlemann, Christoph (2019), *Corpus linguistics for Pragmatics.* A Guide for Research (Routledge Corpus Linguistics Guides), Abingdon / New York: Routledge.

Squartini, Mario 2018. Mirative extensions in Romance: evidential or epistemic? in Guentchéva, Zlatka (ed.), (2018), *Epistemic Modalities and Evidentiality in Cross-Linguistic Perspective*, Berlin, Boston: De Gruyter Mouton, 196-216. https://doi.org/10.1515/9783110572261-009

VAKK = *Corpus of Old Literary Estonian*. https://vakk.ut.ee/ (31.5.2024)

# Serbian and Croatian *pa*: From 'again' to discourse particle

Mikołaj Dunikowski
(Jagiellonian University)

Keywords: pragmaticalization, historical linguistics, semantic change, etymology, Slavic languages

Modern BCMS (Bosnian-Croatian-Montenegrin-Serbian) *pa* is a multifunctional conjunction and particle. One of its functions is to introduce surprise, disbelief, or summarization. Its roots can be traced back to Old Church Slavonic *pakъ, paky* 'again; furthermore; then' through Shtokavian *pak(e)* 'however', *pače* 'and (corrective)'. The aim of this paper is to retrace the semantic and functional change and construct a semantic map (Cf. Haspelmath 2007; Narrog & Van Der Auwera 2012) allowing to compare the pattern to more typological data.

The studies of Slavic discourse particles are focused on synchrony, but BCMS *pa* has not been a subject of any study so far. Some work has been done for Slovenian *pa* (Žagar 2010), a discourse particle of the same origin, but different function. Available etymological dictionaries do not include entrenched semantic change analysis of particles. Etymological Dictionary of Croatian Language (Matasović et al. 2021) notes, that *pa* comes from earlier *pak* and was first recorded 14th century according to *Monumenta Serbica* (Miklosich 1868). The form, however, appears only twice and is not confirmed elsewhere before 18th century, hence it could be considered a mistake.

Retracing the semantic change of *pak* that led to the modern discourse particle is limited by the availability of texts. We can rely on Old Church Slavonic, Shtokavian vernacular texts since 12th century (Bosnian and Serbian charters, Krajina borderland letters), and then 19th century texts and dictionaries. The fact that discourse particles mostly appear in spoken languages is another limitation.

Old Church Slavonic *pakъ, paky*, probably derived from *opaky* 'back, oppositely' (< PIE $*h_2epo$-$h_3k^w$-; Derksen, 2008), aside from adverbial meaning 'again' was used in contexts indicating adversative meaning: *běaše bo ti lice blědo vidimo* ***nynia že paky*** *světьlo aky slьnce* (*Your face was seen pale,* ***but now*** *it's as bright as the sun*) (SJS 1968). This function appears later in Shtokavian (see also corrective *pače* in Pavlović, 2014). At the beginning of 19th century, *pa* appears in dictionaries as a separate entity described as a counterpart of Latin *vero; tum, post haec* (Karadžić 1818). Nowadays it is commonly used in BCMS as a conjunction with adversative and temporal meanings, and a discourse particle in several meanings: ***pa*** *jest* (***well…*** *it is,* ***I guess***); ***pa*** *se ženi!* (*marry!,* imperative, in reaction to one's bad experience with marriage); ***pa*** *to nije moguće* (*that's impossible*).

The examination of this semantic shift is a contribution to the research of language change dynamics, offering a valuable case study for broader linguistic typology. It contributes to South Slavic historical linguistics and offers a framework for comparative analyses, enriching the discourse on language change globally.

**References**

Derksen, Rick (2008), *Etymological dictionary of the Slavic inherited lexicon*. Leiden / Boston: Brill.

Haspelmath, Martin (2007), Coordination, in T. Shopen (ed.), (2009), *Language Typology and Syntactic Description, Second Edition: Volume II: Complex Constructions*, Cambridge University Press, 1–51.

Karadžić, Vuk (1818), *Srpski rječnik,* Vienna.

Matasović, Ranko, Pronk, Tijmen and Ivšić Majić, Dubravka (2021), *Etimološki rječnik hrvatskoga jezika,* vol. 2, O–Ž, Zagreb: Institut za hrvatski jezik i jezikoslovlje.

Miklosich, Franz (1868), *Monumenta serbica spectantia historiam Serbiae, Bosnae, Ragusii*, Vienna.

Narrog, Heiko and Van Der Auwera, Johan (2012), Grammaticalization and semantic maps, in B. Heine and H. Narrog (eds.), *The Oxford Handbook of Grammaticalization*, 1st ed., Oxford University Press, 318–328.

Pavlović, Slobodan (2014), Suprotne naporednosložene rečenice u srpskom jeziku XII–XV veka, *Južnoslovenski Filolog* 70, 111–136.

SJS = (1968), *Slovník jazyka staroslověnského*. Academia. http://gorazd.org/?q=en/node/21

Žagar, Igor Ž., (2010), Pa, a modifier of connectives. An argumentative analysis, in M. N. Dedaić and M. Mišković-Luković (eds.), *South Slavic Discourse Particles*, vol. 197, Amsterdam: John Benjamins Publishing Company, 133–162.

# Additive and scalar particles and clitics in Finnish: Asymmetries between nouns and verbs

Elsi Kaiser
(University of Southern California)

Keywords: Finnish, additivity, scalarity, discourse particles, clitics

This research investigates Finnish focus-sensitive additive and scalar clitics and particles. While additive and scalar particles (*myös/jopa/edes*) show parallel patterns with different parts-of-speech, the implicatures associated with additive and scalar clitics (*-kin/kAAn*) differ strikingly with verbs vs. nouns. I identify differences between verb- and noun-attached clitics, including historical asymmetries, propose a possible source for these differences, and explore consequences of the verb/noun asymmetry for theories of focus-sensitive particles.

The particles *myös* 'too/also' and $jopa_{pos}$/$edes_{neg}$ 'even' associate with nouns and verbs, showing similar effects with both (Hakulinen et al. 2005). *Edes* is used in the scope of negation; *jopa* elsewhere.

(1) {Myös/Jopa} Pekka {myös/jopa} osti {myös/jopa} auton.
{Too/Even} Pekka {too/even} bought {too/even} a car.

The clitic [*-kin*] occurs at the right word-edge. Its negative allomorph [*-kAAn*] (subject to vowel harmony) occurs under negation. On nouns, [*-kin/-kAAn*] has an additive 'too/also'-meaning and a scalar 'even'-meaning (Hakulinen et al. 2005, Vilkuna 1984), ex.(2). Adjectives and adverbs pattern similarly (*even a [red] car*). I show that these patterns can be captured by extending Lahiri's (1998) work on Hindi.

(2) Pekkakin osti auton.
[Pekka], too, bought a car. / Even $[\text{Pekka}]_{FOC}$ bought a car.

However, verbs differ: I show that [*-kin/kAAn*]-marked verbs (ex.3) can be used when the entire event of Pekka buying a car (not just the clitic-bearing verb) is low-likelihood/unexpected (4a) or high-likelihood/expected (4b, see also Hakulinen et al. 2005, Vilkuna 1984).

(3) Pekka ostikin auton.
Pekka bought-KIN car.

(4a) *Context where the event with [-kin/-kAAn]-marked verb is low-likelihood/unexpected:*
I know that Pekka won the lottery. We are wondering what he will do with the money: travel, buy something? He does not know how to drive and has shown no interest in cars. The next day, someone tells me: "Guess what! Pekka bought-KIN car with his winnings!" (ex.3)

(4b) *Context where the event with [-kin/-kAAn]-marked verb is high-likelihood/expected:*
Pekka needs a new car. During a full day of visiting car dealerships with his friend Lasse, Pekka sees a certain car that catches his eye: Price, size, color -- everything is perfect, he tells Lasse.

The next day, Lasse tells me what happened, describes Pekka telling him about the perfect car and then says “And Pekka bought-KIN that car right on that very same day.” (ex.3)

Historically, verb-attached use emerged later (see VVKS). Agricola (1500s) used nominal [-*kin*], but rarely used [-*kin*] on verbs (5%). However, by the late 1800s, 20% of [-*kin*] use by Juhani Aho was on verbs (Lievonen 1985). [-*kAAn*] is more frequent but shows a similar trajectory. I take this diachronic pattern as evidence of noun- and verb-attached uses being dissociable.

I analyze verb-attached clitics as having two distinctive properties: (a) Elements on the likelihood scale can be events, not just verbs. In contrast, (i) with [-*kin/kAAn*]-marked nouns, the scale’s elements are also nouns and (ii) with the particle ‘even’ *jopa/edes*—when associated with focused verbs—the scale’s elements are verbs. (b) Events can be expected or unexpected. I derive both properties from verbs being the core of an event. If one’s communicative aim is to express additional information about the entire event, the verb is the optimal element target: Although not ideal, the verb is the best choice under the circumstances. This can result in the blocking of the default ‘host-based’ interpretation (which obtains with nouns and adjectives)—and this blocking is indeed what we find with verbs.

In other words, since clitics must attach to something, if we want to express information about events, verbs are the ‘not-ideal-but-best-available’ clitic-hosts for event-level meanings. Diachronic data suggests event-level use grammaticalized after verb/noun/adjective-uses.

Summary: Noun- vs. verb-attached scalar/additive clitics in Finnish have different semantics which correspond to different diachronic trajectories. The noun-verb asymmetry supports approaches emphasizing the discourse management function of additives/scalars.

**References**

Hakulinen, A., M. Vilkuna, R. Korhonen, V. Koivisto, T. R. Heinonen, and I. Alho (2005). *Iso suomen kielioppi.* Helsinki: Suomalaisen Kirjallisuuden Seura.

Lahiri, U. (1998). Focus and negative polarity in Hindi. *Natural Language Semantics*, 6(1), 57-123.

Lievonen, M. (1985) *Mikael Agricolan teosten liitepartikkelit.* Pro gradu, Jyväskylän yliopiston suomen kielen laitos.

Vilkuna, M. (1984). Voiko kin-partikkelia ymmärtää?. *Virittäjä*, 88(4), 393.

Virtuaalinen vanha kirjasuomi (VVSK) https://vvks.it.helsinki.fi/

# Perspectives on the origin of discoursive functions of the Meadow Mari 3SG possessive suffix ŽE

Johannes Hirvonen
(Northwestern University)

Several Uralic languages (Meadow Mari, Udmurt, Northern Mansi, and Tundra Nenets, among others) employ possessive suffixes in functions related to discourse organization, including the marking of identifiability, contrast, or topic shift (e.g. Nikolaeva 2003, Simonenko 2014, Gerland 2014). The origin of these "non-possessive" functions of Uralic possessive suffixes has received internal explanations (Fraurud 2001, Kiss & Tánczos 2018), but several scholars have also suggested language contact as factor in the development of the possessive suffixes' functions in individual Uralic languages like Meadow Mari or Komi-Zyryan (e.g. Serebrennikov 1972: 10, Leinonen 1998: 75, Nikolaeva 2003: 140–
142, Yurayong 2020: 119).

Adding to this body of research, this study explores the origin of the multi-faceted discourse functions of the Meadow Mari 3SG possessive suffix ŽE. The basis for this study is an ongoing survey of the usage of ŽE in both historic data from Mari varieties and contemporary spoken language (Volkova et al. 2023). In contemporary spoken Meadow Mari, the suffix is used in a number of different discoursive functions which include (but are not limited to) marking (1) membership in a previously introduced superset, (2) a topic shift, (3) contrastive topics, and (4) that information is presented as uncontroversial (the so called 'enimitive', cf. Panov 2020):

(1) *Memnan indeš čəve ulo.* ***Iktə-že*** *ala-kušan jomən.*
2PL.GEN nine chicken exist one-ŽE somewhere.LAT get.lost.PST2.3SG
'We have nine chickens. **One of them** has gone off somewhere.'

(2) ***A.že*** *tide podružkəžo mo?*
A.PX3SG this female.friend.PX3SG Q '
**(What about) A.** – is this her friend?'

(3) ***Pij-em-že*** *pəzəm kočkaš jörata,* ***pərəs-em-že*** *koləm*
dog-PX1SG-ŽE meat.ACC eat.INF love.3SG cat-PX1SG-ŽE fish.ACC
*kočkaš jörata.*
eat.INF love.3SG
'(What do your pets like to eat? –) **My dog** likes to eat meat, (and) **my cat** likes to eat fish.'

(4) ***Rəvəž-še*** *čəvənam kočkən.*
fox-ŽE chicken.PX1PL.ACC eat.PST2.3SG
'(I cannot bring you any eggs today.) **The fox** has eaten our chickens, **you see**.'

Both historic data from the beginning of the 20th century (e.g. Beke 1951) and current studies (Georgieva 2022) display an even larger variety of functions, including anaphoric marking and nominalization marking, complicating explanations for the possessive suffix's development. In this study, we juxtapose historic and contemporary language data and attempt to draw conclusions about the history of the discoursive usage of possessive suffixes in Mari, focussing on contact influences.

We suggest that the existence of different discoursive functions reflects different stages of Mari language contact with Turkic and Russian: in addition to functions which Mari varieties share with other Uralic languages (superset, topic shift), the suffix characteristically patterns with the neighboring Turkic languages of the Volga-Kama area (Chuvash, Tatar) in marking contrastive topics (instead of an additive clitic which is typical in this function, cross-linguistically; cf. Russkikh & Oskol'skaya 2021). The enimitive pattern, on the other hand, appears to be the result of recent language contact with Russian – where the Mari possessive suffix took over some functions of the Russian discourse particle *zhe* – as this function is almost absent in texts recorded in the early 20th century, but common in contemporary speech.

**References**


Beke, Ödön (1951). *Volksdichtung und Gebräuche der Tscheremissen (Maris) 1. Band. Gesammelt und herausgegeben von Ödön Beke*. Budapest: Akadémiai kiadó.

Fraurud, Kari (2001). "Possessive with extensive use: A source of definite articles?" In: *Dimensions of possession.* Ed. by Irène Baron, Michael Herslund, and Finn Sørensen. Amsterdam: John Benjamins, 243–267.

Georgieva, Ekaterina (2022a). "Fake possessives, partitives and definites sharing the same D". Handout at GLOW 45, 28.04.2022.

Gerland, Doris (2014). "Definitely Not Possessed? Possessive Suffixes with Definiteness Marking Function". In: *Frames and Concept Types: Applications in Language and Philosophy.* Ed. by Thomas Gamerschlag et al. Springer, 269–292.

Kiss, Katalin É. & Tánczos, Orsolya (2018). "From Possessor Agreement to Object Marking in the Evolution of the Udmurt *-jez* Suffix: A Grammaticalization Approach to Morpheme Syncretism." *Language* 94/4, 733–757.

Leinonen, Marja. (1998). The postpositive particle *-to* of Northern Russian dialects compared with Permic languages (Komi Zyryan). *Studia slavica finlandensia* 25, 74-90.

Nikolaeva, Irina (2003). "Possessive affixes in the pragmatic structuring of the utterance: Evidence from Uralic". In: *International Symposium on Deictic Systems and Quantification in Languages Spoken in Europe and North and Central Asia: Collection of papers*. Ed. by Pirkko Suihkonen and Bernard Comrie. Izhevsk and Leipzig, 130–145.

Panov, Vladimir (2020). "The marking of uncontroversial information in Europe: presenting the enimitive". In: *Acta Linguistica Hafniensia* 52.1, 1–44.

Russkikh & Oskol'skaya (2021) = Русских, Алина Алексеевна & Оскольская, София Алексеевна (2021). "Аддитивная частица в тюркских языках Поволжья." *Oriental Studies* 14/6, 1324–1352.

Serebrennikov (1972) = Серебренников, Борис Александрович. (1972). "О некоторых отличительных признаках волгокамского языкового союза". In: *Языковые контакты в Башкирии. Тематический сборник*. Ed. by Т. М. Гарипов and Н. В. Черемисина. Уфа, 7–17.

Simonenko, Alexandra (2014). "Microvariation in Finno-Ugric possessive markers". In: *Proceedings of the 43th annual meeting of the North East Linguistic Society (NELS 43). Vol. 2*. Ed. by H.-L. Huang et al., 127–140.

Volkova, Anna et al. (2023). *Spoken corpus of Meadow Mari (as spoken in the village of Staryj Toryal, Novyj Torjal district, Mari El Republic, Russia)*. lingconlab.ru/spoken_meadow_mari/

Yurayong, Chingduang (2020). "Postposed Demonstratives in Finnic and North Russian Dialects". Doctoral Dissertation. Helsingin Yliopisto.

# Discourse particles of Latvian-origin in Courland Livonian

Marili Tomingas
(University of Tartu)

Keywords: Livonian, Finnic languages, discourse particles, pragmatics, language contact

Courland Livonian is a critically endangered language with approximately 40 L2 speakers left (Ernštreits 2013: 15). Over the years, Courland Livonian has been strongly influenced by neighbouring Latvian. Although there has been parallel research on Latvian and Livonian as close contact languages (see, e.g., Wälchli 2001, Halling 2006, Ernštreits & Kļava 2014), there has been little research with the special focus on discourse particles so far (see Grünthal 2015 and Tomingas 2022). Some Latvian-origin discourse particles in Livonian have been mentioned or briefly described in the article by Tomingas (2022), but the topic would still need a wider coverage to investigate the Latvian-influenced particles more precisely and to also understand their origin and development.

This paper focuses on three discourse particles of Latvian-origin: *riktig* 'actually, exactly', *nekā* 'like', and *nu* 'well'. Due to the influence of the Latvian particle *nu* 'well', it is more frequently occurring in Livonian than the similarly used *no* 'well', which is spotted in many other Finnic languages, e.g., Estonian, Finnish, Votic, Karelian. See example (1) on the use of *nu* in Courland Livonian:

(1) PK: ***nu*** *kui* *si'z* *`Valtõn ni* *lā̆-b.* (AEDKL: F1035-05)
PTCL how PTCL Valt-DAT now go-3SG
'**well**, how is Valt doing then now?'

The particle *riktig* originates from German *richtig* 'right, correct', but Livonian has borrowed it through Latvian *riktīgs* 'right, proper, actual' or *riktīgi* 'properly, truly, correctly'. It is used as a tonal particle in Livonian, giving a certain stress or tonality to a clause. Example (2) illustrates the use of Courland Livonian *riktig*:

(2) GK: *`tallõ* ***riktig*** *täs* *ä'b* *ūo* *k- kus* *`lā̆'-dõ*
in_winter PTCL here NEG.3SG be.CNG.SG where go-INF
*ju*. (AEDKL: DS0127-05)
PTCL
'in winter there is **actually** nowhere to go here'

The particle *nekā* has previously been compared to Latvian conjunction *ka* 'that' (Kettunen 1938: 244), but I also propose the influence of Latvian *kā* 'like' or *nekā* 'than' to be possible due to both similar forms and meanings. It can be used either as a connector or as a softener in the discourse, see example (3):

(3) ***nekā*** *`minā* *ju* *`kaŗŗõ* *lā̆'-nd* *u'm* *`pā̆giņ.*
PTCL1SG PTCL herding.ILL go-APP.SG be.1SG a_lot
(AEDKL: F1089-05)
'**like**, I have gone herding a lot'

The presentation focuses firstly on explaining the possible development of Livonian particles *riktig*, *nekā* and *nu* based on the dictionaries and earlier text resources, comparing them also to their Latvian lexical equivalents. Secondly, the particle examples in spoken Courland Livonian are looked at to give a brief overview of their syntactic and pragmatic use during the past decades. The research material comes from the audio recordings in the University of Tartu Archives of Estonian Dialects and Kindred Languages (AEDKL), 17 recordings with six different native speakers are used with total length of 7 hours and 13 minutes. The material is analysed qualitatively, also considering intonation, stress and self-repairs that help to interpret the functions of the aforementioned discourse particles in the recordings.

Acknowledgements: This work was supported by Estonian Research Council grant (PRG1290).

**References**

AEDKL = Archives of Estonian Dialects and Kindred Languages. [Tartu Ülikooli eesti murrete ja sugulaskeelte arhiiv]. Available at http://www.murdearhiiv.ut.ee. (accessed 15.11.2023)

Ernštreits, Valts. 2013. *Liivi kirjakeel*. Tartu: Tartu Ülikooli kirjastus

Ernštreits, Valts and Gunta Kļava. 2014. "Grammatical changes caused by contact between Livonian and Latvian". *Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics*, 5(1), 77–90.

Grünthal, Riho. 2015. "Livonian at the crossroads of language contacts". In Santeri Junttila (ed.), *Contacts between the Baltic and Finnic languages*. (Uralica Helsingiensia 7), 97– 150. Helsinki: University of Helsinki.

Halling, Tiina. 2006. "Liivi keel läti keele mõjusfääris". *Mitmõkeelisüs ja keelevaihtus Õdagumeresoomõ maiõ pääl = Mitmekeelsus ja keelevahetus Läänemeresoome piirkonnas*. (Võro Instituudi toimõtisõq 18), 55–70. Võro: Võro Instituut.

Kettunen, Lauri. 1938. *Livisches Wörterbuch mit grammatischer Einleitung*. *Lexica Societatis Fenno-Ugricae* V. Helsinki: Suomalais-Ugrilainen Seura.

Tomingas, Marili. 2022. "Sagedasemad diskursusepartiklid suulise liivi keele salvestistes". *Keel ja Kirjandus* 65(1–2), 95–109.

Wälchli, Bernhard. 2001. "Lexical evidence for the parallel development of the Latvian and Livonian verb particles". In Östen Dahl, Maria Koptjevskaja-Tamm (eds.), *Circum-Baltic Languages: Typology and Contact*, 2: *Grammar and Typology*. (Studies in Language Companion Series 55), 413–442. Amsterdam–Philadelphia: John Benjamins Publishing Company.

# WS13 Passive and passivization across languages in dynamic and typological perspectives: Conceptual and methodological challenges

# Complex predicate as analytical voice marking of antipassivization

Katarzyna Janic, Krzysztof Stroński & Mohammad Tavakoli
(Adam Mickiewicz University)

Keywords: antipassive, analytical voice marking, complex predicates, analytical antipassive marking, multivariate typology

Aim. Although recent studies extended the voice definition from an inflectional category to synthetic (affixation) and analytic marking (analytic verb forms, i.e., complex predicates CPs) (Zúñiga & Kittilä 2019; Creissels 2024), a noticeable disparity remains in how analytical voice marking is treated in voice domain. While CPs received much attention in valency-increasing alternations like causatives (Creissels 2024), there is little research on their detransitivization function. And when compared to passives, antipassives remain under-investigated (Zúñiga & Kittilä 2019: 114). This study aims to bridge this gap in the existing literature, seeking to demonstrate that analytical antipassive marking is more prevalent than previously assumed.

1. Mopan (Mayan; Hofling 2011, cited in Zúñiga & Kittilä 2019: 113-114)

a. *Walak-u-loch-ik-ech*.
INCOMPL-3.I-bend-INCOMPL.TR-2SG.II
'S/he bends you (sg).'

b. *Uch-i* *u-loch*.
**happen**-3SG.I[COMPL.ITR] 3.I-bend
'S/he bent (something).'

Approach, data, method, and research question: We conducted our research in a functional-typological approach based on a sample of 56 unrelated languages from six macroareas. We collected data through grammar mining and analyzed it in multivariate typology (Bickel 2011). Each antipassive construction was thus decomposed into formal variables, with attention to voice marking and its synthetic vs. analytical opposition. The investigation was guided by two research questions: What is the grammatical nature and evolution of the analytic voice form?

Results: Our analysis showed that CPs cover light verbs in the auxiliary function. In Mopan, Hindi, and Worrorra, light verbs serve a functional purpose, coding tense, aspect and/or person. In addition to the aspectual input (aktionsart), some (Mopan 'happen' [1], Hindi 'fall' [2], Worrorra 'be' [3]) act as analytic voice markers, (semi-)detransitivizing transitive construction. They yield antipassives and antipassive-lookalikes, where the ergative (A) argument becomes absolutive (Hindi), or the P argument loses a verb index (Mopan, Worrorra).

2. Hindi (Indo-European; Butt 1995: 107)

a. *us-ne* *gānā-∅ gāyā*
s/he-ERG song.M.SG sing.PST.M.SG.
'S/he sang a song.'

b. *vo* *gānā-∅* *gā* *paṛ-ā*
he.NOM song.M.NOM sing **fall**-PST.M.SG
'He fell to singing (burst out into song).'

3. Worrorra (Worrorran; Clendon 2014: 331)

a. *minjarl* *nyiN-Ø-mnya=bwu-na*
eat 3F-3-DD=hit-PST
'he ate her.'

b. *mangarri* *minjarl-minjarl* *kaarr=nu-na*
veg.food eat-eat 3PL=**be**-PST
'They used to eat vegetable food.'

Certain light verbs deriving antipassives manifest a well-established evolution. In French, *faire* 'do/make', which is not distinct from tense auxiliaries (Abeillé et al. 1998), can form causative (*faire acheter les chaussures* 'make [someone] buy the shoes') and antipassive periphrases (*faire des achats* 'do [some] shopping'), the latter taking a deverbal event noun in an object role (Sansò 2017). In Makalero, the light verb *kini* 'do' grammaticalized into a synthetic voice marker *-ini* when deriving antipassives. Finally, in Soninke, the antipassive *-ndì* and causative *-ndí* result from the grammaticalization of **tin* 'do' in causative and antipassive periphrases (Creissels 2012). The same holds for *-aɢan* in Mocoví (Juárez & González 2021).

By analogy with synthetic/analytic causative marking and its correlation with the direct vs. indirect causation meaning (Haspelmath 2018), we will check whether similar correlations also exist between the formal properties of antipassive marking and specific aspects of meaning.

Acknowledgments
This research is part of the project No. 2021/43/P/HS2/01395 cofunded 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, headed by Katarzyna Janic as Principle Investigator.

**References**

Abeillé, Anne, Danièle Godard, and Ivan A. Sag. 1998. 'Two Kinds of Composition in French'. In *Complex Predicates in Nonderivational Syntax*, edited by Erhard Hinrichs, Andreas Kathol, and Tsuneko Nakazawa, 1–41. San Diego: Academic Press.

Bickel, Balthasar. 2011. 'Multivariate Typology and Field Linguistics: A Case Study on Detransitivization in Kiranti (Sino-Tibetan)'. In *Proceedings of Conference on Language Documentation and Linguistic Theory 3*, edited by Peter Austin K., Oliver Bond, David Nathan, and Lutz Marten, 3:3–11. University of London: SOAS.

Butt, Miriam. 1995. *The Structure of Complex Predicates in Urdu*. Stanford (CA): CSLI Publications.

Clendon, Mark. 2014. *Worrorra: A Language of the North-West Kimberley Coast*. Adelaide: University of Adelaide Press.

Creissels, Denis. 2012. 'The Origin of Antipassive Markers in West Mande Languages'. Presented at the 45th Annual Meeting of the Societas Linguistica Europaea, Stockholm, September 29.

———. 2024. *Transitivity, Valency, and Voice*. Oxford (UK): Oxford University Press.

Haspelmath, Martin. 2018. 'Universals of Causative and Anticausative Verb Formation and the Spontaneity Scale'. Edited by Katarzyna Janic and Nicole Nau. *Lingua Posnaniensis* 58 (2): 33–64.

Hofling, Charles A. 2011. *Mopan Maya-Spanish-English Dictionary*. Salt Lake City: The University of Utah Press.

Juárez, Cristian, and Albert Álvarez González. 2021. 'Explaining the Antipassive-Causative Syncretism in Mocoví (Guaycuruan)'. In *Antipassive: Typology, Diachrony, and Related Constructions*,

edited by Katarzyna Janic and Alena Witzlack-Makarevich, 315–47. Amsterdam: John Benjamins.

Sansò, Andrea. 2017. 'Where Do Antipassive Constructions Come from? A Study in Diachronic Typology'. *Diachronica* 34: 175–218.

Zúñiga, Fernando, and Seppo Kittilä. 2019. *Grammatical Voice*. Cambridge: Cambridge University Press.

**Abbreviations**

| | | | |
|---|---|---|---|
| COMPL.ITR | completive intransitive | M | masculine |
| DD | discourse diectic | NOM | nominative |
| ERG | ergative | PL | plural |
| F | feminin | PST | past |
| INCOMPL | incompletive | SG | singular |
| INCOMPL.TR | incompletive transitive | | |

# ‘Bashfulnesse is banished’: passive constructions as a gateway to state-incorporation in Old French and Middle English contact

Tara Struik
(University of Mannheim)

Keywords: Early English, passives, resultativity, language contact, accommodation bias

This talk presents the results of a diachronic corpus study on the integration of French verbal copies into the semantic and syntactic structure of Middle and Early Modern English. We test the hypothesis that the non-finite accommodation bias for French verbs reported by Shaw & De Smet (2022) is due to an overreliance on passivation of French verbs.

We argue that the passive construction facilitates the integration of French verbal copies because of semantic and syntactic similarity. Struik, Kaltenbach & Trips (2023) show that (transitive) French verbs mostly contribute verbs which encode an end-state and add a novel way to express state in English, which was originally typically expressed by secondary predication, or one-argument predicates, such as unaccusatives or passives. Passives in Old English are ambiguous between adjectival or eventive readings, and are frequently stative in interpretation, as in (1) (cf. Toyota 2008).

(1) Þær wæron **gehælede** [...] fela adlige menn
there were healed many ill men
‘There many ailing men were healed.’
(ÆLS20.113, adapted from Jones & Macleod, 2016)

The use of a passive structure with French copies as in (2) is thus a good match with the native structure for two reasons: 1) it expresses a state, 2) it does so by means of a native structure, realizing an originally two-argument verb with only one argument.

(2) Al Troy is þus **destroyed**
‘All Troy is thus destroyed’

(a1400(?a1350) Siege Troy(1) (Eg 2862)233, MED)

To support our hypothesis we rely on data from the Penn Parsed Corpus of Middle English 2 (Kroch & Taylor 2000), Parsed Corpus of Middle English Poetry (Zimmermann 2018) the Parsed Linguistic Atlas of early Middle English for Middle English (Truswell et al. 2019), and the Penn Parsed Corpus of Early Modern English (Kroch, Santorini, & Delfs 2004) for Early Modern English. We use the lemmatized versions (Percillier & Trips, 2020) and the lemma-based Manner/Result annotation by Struik, Kaltenbach & Trips (2023) to determine whether French Result verbs more frequently integrate in the passive compared to other verb types and tenses, and whether they do so at different rates compared to native verbs.

Our findings show two important things. First, French Result verbs occur significantly more frequently in the passive compared to French Manner verbs. Second, French Result verbs occur in the passive significantly more frequently than native Result verbs, whereas French Manner verbs occur in the passive at comparable frequencies to native Manner verbs. These findings suggest that the non-finite accommodation bias for loan verbs reported in Shaw & De Smet (2022) is not merely a strategy

to avoid the morphological complexities of finiteness marking when using copied vocabulary, but that speakers are aware of the internal semantic and argument structural properties of both the source and donor language and that they aim to find the best structural match. Our findings furthermore suggest a diachronic consequence: the use of passives with native Result verbs, but crucially not Manner verbs, starts to increase after the Middle English period, suggesting a lasting impact of French on the native system.

**References**


dĕstroien v. (n.d.). In *Online edition in Middle English Compendium*. Retrieved November 13, 2023, from https://quod.lib.umich.edu/m/middle-english-dictionary/dictionary/MED11399/.

Jones, H., & Macleod, M. (2018). The status of passive constructions in Old English. *Transactions of the Philological Society*, *116*(1), 59–90.

Kroch, A., Taylor, A., & Santorini, B. (2000). The Penn-Helsinki Parsed Corpus of Middle English, second edition (PPCME2). University of Pennsylvania. http://www.ling.upenn.edu/hist-corpora/PPCME2-RELEASE-4/index.html.

Kroch, A., Santorini, B., & Delfs, L. (2004). The Penn-Helsinki Parsed Corpus of Early Modern English, first edition (PPCEME), release 3. University of Pennsylvania. http://www.ling.upenn.edu/ppche/ppche-release-2016/PPCEME-RELEASE-3.

Shaw, M., & De Smet, H. (2022). Loan word accommodation biases: Markedness and finiteness. *Transactions of the Philological Society, 120*(2), 201–217.

Struik, T., Kaltenbach, L. & Trips, C. (2023). The Results of Contact Tracing changes in resultativity encoding in Middle and early Modern English in contact with Old French. Paper presented at *Diachronic Generative Syntax (DiGS 24).* Université Paris Cité, Paris.

Toyota, J. (2008). *Diachronic change in the English passive*. Basingstoke: Palgrave Macmillan.

Trips, C., & Percillier, M. (2020). Lemmatising verbs in Middle English corpora: The benefit of enriching the Penn-Helsinki Parsed Corpus of Middle English 2 (PPCME2), the Parsed Corpus of Middle English Poetry (PCMEP), and A Parsed Linguistic Atlas of Early Middle English (PLAEME). In Nicoletta Calzolari et al. (eds.), *Proceedings of the 12th Language Resources and Evaluation Conference* (pp. 7170–7178). Paris: ELRA – European Language Resources Association.

Truswell, R., Alcorn, R., Donaldson, J., and Wallenberg, J. (2019). A Parsed Linguistic Atlas of Early Middle English. In Rhona Alcorn et al. (eds), *Historical Dialectology in the Digital A*ge, (pp. 19–38). Edinburgh: Edinburgh University Press.

Zimmermann, R. (2018). The Parsed Corpus of Middle English Poetry (PCMEP). https://pcmep.net/index.php

# The role of passive constructions in the evolution of argumenthood: Evidence from Indo-European

Eystein Dahl
(Universytet im. Adama Mickiewicza w Poznaniu)

Keywords: passive constructions, argumenthood, diachronic typology, grammaticalization, Indo-European

This paper explores the role played by passive constructions in the development of the morphosyntax of verbal arguments, drawing on evidence from the Indo-European linguistic family. There is broad agreement that no dedicated passive construction existed in the reconstructible prehistoric stages of this family. In the course of time, however, most of its daughter branches developed specialized passive forms, typically originating from passivation strategies that have left reflexes in several branches (cf. e.g., Luraghi et al 2021 for discussion). Apart from some minor and not fully productive strategies, such as lexical suppletion and lability, Luraghi et al (2021) identify three important source constructions:

- Inflectional valency-decreasing categories, such as middle/mediopassive constructions (e.g. Greek, Italic, Anatolian)
- Derivational categories, notably stative constructions (e.g. Indo-Iranian)
- Periphrastic constructions involving p-oriented verbal adjectives/participles (e.g. Italic, Celtic)

An important question arising from these observations is whether or to what extent the various source constructions develop into functionally equivalent passive constructions or not. To explore this set of problems, we draw on Kiparsky's (2013: 8) analytical framework for passive constructions which enables a precise investigation of the variation among passive constructions. The parameters of variation include whether there are any restrictions upon what types of verbs may passivize, whether subjectless/impersonal passives are allowed, whether an agent phrase can occur freely, whether non-canonical object case marking is maintained in passivization, whether the second or third object argument of ditransitives is passivized, and whether passive constructions show stacking (cf. Kiparsky 2013: 8). Some of these parameters clearly seem to be of relevance for the diachrony of passives as well. For example, as noted by Luraghi et al (2021), agent phrases frequently serve to disambiguate the passive reading of middle/mediopassive forms, suggesting that, at least in such cases, the agent is defocused rather than demoted and therefore less likely to be omitted. Under this analysis, one might hypothesise a development from agent defocusing to agent demotion with passives from certain source constructions.

In view of Siewierska's (2013) finding, that just 162 out of the 373 languages in her survey have a passive construction, the fact that most of the Indo-European branches develop specialized passive constructions is remarkable. In recent work, Cotticelli and Dahl (2022) observe that there is a tendency across the Indo-European languages to develop more consistently accusatively oriented morphosyntax over time, which they label 'accusativization', resulting in an increasing number of subjecthood properties. This is, amongst other things, indicated by the development of a more complex infinitive syntax in some of the Indo-European languages. There is a growing body of evidence that grammaticalisation is conditioned and facilitated by more general typological features of a language (cf. e.g., Narrog 2017, Narrog and Heine eds. 2018). It is therefore tempting to regard the rise of passive

constructions in the Indo-European languages as a key feature in the development of accusative morphosyntax.

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**

Cotticelli, Paola and Eystein Dahl. 2022. Split alignment, mixed alignment 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.

Kiparsky, Paul. 2013. Towards a null theory of the passive. In *Lingua* 125, 7-33.

Luraghi, Silvia, Guglielmo Inglese and Daniel Kölligan. 2021. The passive voice in ancient Indo-European languages: inflection, derivation, periphrastic verb forms. In *Folia Linguistica Historica* 42(2), 339-391.

Narrog, Heiko. 2017. Typology and Grammaticalization. In Alexandra Y. Aikhenvald and R. M. W. Dixon (eds.) *The Cambridge Handbook of Typology*. Cambridge, CUP, 151-177.

Narrog, Heiko and Bernd Heine (eds.). 2018. *Grammaticalization from a Typological Perspective*. Oxford: OUP.

Siewierska, Anna. 2013. Passive Constructions. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) WALS Online (v2020.3) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.7385533 (Available online at http://wals.info/chapter/107, Accessed on 2024-01-15.)

# Oceanic passives in synchrony and diachrony

Lidia Federica Mazzitelli
(University of Cologne)

Oceanic languages form a subgroup within the larger Austronesian family, of which they are the easternmost representatives. Morphological passives are found in all languages belonging to the Central Pacific branch of Oceanic (Fijian, Rotuman, and Polynesian languages), and are already relatively well studied. Outside of this branch, conversely, morphological passives seem to be extremely rare (Lynch et al 2002: 45) and are, so far, still under-investigated.

In this talk, I provide a first review of the passive constructions of 15 languages from three non-Central Pacific branches of Oceanic (Table 1.), analysing the following criteria: the formal nature of passive morphology; the promotional status of the passive; the possibility of taking an agent phrase; and their diachronic path of grammaticalization. I also show that, in some languages, the passive is a still unstable construction, with some attested ongoing variation. The sources of my data are published descriptive materials as well as my own fieldwork on Kara, Tigak, and Lakurumau; in Kara, two different passive constructions are found.

| **Branch** | **Languages** |
|---|---|
| Micronesian | Mokilese, Marshallese, Woleian |
| Meso-Melanesian | Kara, Tigak, Lakurumau, Roviana, Hoava, Ughele, Vitu, Bola |
| Southern Oceanic | Raga, Abma |
| Reefs-Santa Cruz | Natügu, Nalögo |

Table 1. The sample

In most sample languages, the passive marker is an affix: in two languages, Marshallese and Kara, passives are formed through stem-altering processes or zero coding. Most sample passives (9/16) are promotional: the non-agent argument is coded as a subject. In two languages (Abma and Raga), the formal passive subject is a dummy 3SG pronoun and the non-agent argument is coded as a non-subject; in Vitu and Tigak, the status of the non-agent argument is unclear. In Kara, a promotional and a non-promotional variant of the same passive construction co-exist; in Lakurumau, a promotional passive is on its way to develop from an underlying non-promotional impersonal construction.

Most passives in the sample (8/16) are agentless. In Kara and Tigak, there is an ongoing tendency towards the acceptance of animals as quasi-agents, not shared by all speakers. The Roviana passive is described by Corston (1996) as being agentless; more recently, Schuelke (2020) describes it as allowing agent phrases: possibly, this testifies to an ongoing development towards a fully-fledged passive.

The passive-marking strategies in my sample derive from three different diachronic sources: anti-causative (3/16), nominalisations (9/16), and deobjective constructions (4/16). The latter two strategies are particularly interesting, as, cross-linguistically, they seem to be relatively rarer sources of passive grams (Haspelmath 1990; Sansò 2016). There seem to be no relation between the diachronic source of the sample passives and their (non-)promotional status; conversely, all passives derived from a nominalisation strategy are agentless (ony some de-anticausative and de-deobjective passives accept agents). I suggest this may be due to the fact that in these languages (except Natügu and Nalögo), nominalisations can only occur with one argument, either the patient or the agent: constructions such as English *the destruction of the city by the enemies*, with both a patient and an agent phrase, are not allowed.

**References**

Corston, Simon. 1996. Ergativity in Roviana, Solomon Islands. Canberra: Pacific Linguistics.

Haspelmath, Martin. 1990. The grammaticization of passive morphology. *Study of Language* 14:1, 25–72.

Lynch, John, Ross, Malcolm, & Crowley, Terry (eds.) 2002. *The Oceanic languages*. Richmond: Curzon Press.

Sansò, Andrea. 2016. Agent-defocusing constructions from nominalized VPs: A cross-linguistic type?, *Studies in Language* 40(4), 894-954.

Schuelke, Peter. 2020. *Grammatical Relations and Syntactic Ergativity in Roviana: A little-described language of Solomon Islands.* University of Hawai'I doctoral dissertation.

# Emerging passives in Permic languages

Nikolett F. Gulyás
(Eötvös Loránd University)

Keywords: passives, impersonals, grammaticalization, Komi-Permyak, Udmurt

Previous grammars of Komi-Permyak and Udmurt (Permic, Uralic) do not consider passives a distinct category in these languages (Batalova 2002, Winkler 2011), however, the passive usage of certain participles and reflexive forms has already been reported (Asztalos 2011, F. Gulyás – Speshilova 2014). The border between the two construction types is not clear-cut; furthermore, it remains debatable if distinct impersonal and passive domains can be distinguished in Permic.

The aim of this talk is to give an overview on possible paths of the ongoing grammaticalization from participles and middles to different types of R-impersonals and then to canonical passives. From a functional viewpoint, impersonal constructions depict events where the agent is demoted (Malchukov – Siewierska 2011), which is also a prominent feature of passives but the main function of passivization is patient promotion (cf. Siewierska 2013).

I assume that passive has become a distinct but syncretic category in both languages. I claim that the original resultative function of the predicative participle (*-emyn* in Udmurt and *-öma* in Komi-Permyak) has first evolved into an impersonal with intransitive verbs (1) and then with transitive ones (2). The latter type has recently been reanalysed as a canonical passive by some speakers (3) given the demoted agent may be overt and marked with the instrumental case. Both languages show the same process, cf. the following elicited examples from Udmurt:

(1) Tatyn ekt-emyn.
here dance-PTCP.PRED
'There was dancing here.' (Lit. (It) was danced here.)

(2) Šyd-ez śi-emyn.
soup-ACC eat-PTCP.PRED
'The soup has been eaten.'
'Somebody has eaten the soup.'

(3) Šyd śi-emyn (anaj-en.)
soup eat-PTCP.PRED mother-INS
'The soup has been eaten (by mother).'

Results also show the grammaticalization of reflexives and middles to impersonals and passives in both languages which is common cross-linguistically (cf. Givón 2009). The reflexive suffix *-ś/-ć* shows a wide range of polyfunctionality in Komi-Permyak. It can mark middle voice (4) and antipassives (5). I assume that constructions describing non-volitional events (6) have evolved from middles. These can be interpreted as A-impersonals (cf. Malchukov – Siewierska 2011) on the basis of agent demotion as the genitive marking of the demoted argument indicates it. Additionally, some speakers also accept canonical, transitive passive constructions where the demoted, non-canonical agent (i. e. a natural force as in example (7)) takes the instrumental case:

(4) Ybös oś-ś-ö.
door open-REFL-3SG
'The door opens.'

(5) Pon-ys pur-ś-ö.
dog-3SG bite-REFL-3SG
'The dog bites.'

(6) Menam onmöśśi-ś-öma.
I.GEN fall_asleep-REFL-PST2.3SG
'I fell asleep (unintentionally).'

(7) Ybös oś-ś-i-s töl-ön.
door open-REFL-PST-3SG wind-INS
'The door has been opened by the wind.'

Given that the reflexive/middle > impersonal > passive change is currently going on, some speakers tend to reject canonical passives of either the participle or the reflexive type. Additionally, differential object marking and the usage of syncretic verb forms also often make the interpretation of certain examples opaque, I will discuss these issues in details. Nevertheless, data show strong support for grammaticalization patterns mentioned above.

**References**


Asztalos, Erika 2011. Transitive and intransitive passivization in Udmurt. In Csúcs, Sándor (ed.), *Congressus XI Internationalis Fenno-Ugristarum*. Piliscsaba: Reguly Társaság, 53–62.

Batalova, Raisa M. [Баталова, Раиса М.] 2002. *Кудымкарско-иньвенский диалект коми-пермяцкого языка*. Mitteilungen der Societas Uralo-Altaica. Heft 23. Moskva – Groningen.

F. Gulyás, Nikolett – Speshilova, Yulia 2014. Impersonals and passives in contemporary Udmurt. *Finnisch-Ugrische Mitteilungen* 38, 59–91.

Givón, Talmy 2009. *The genesis of syntactic complexity. Diachrony, ontogeny, neuro-cognition, evolution*. Amsterdam: John Benjamins.

Malchukov, Andrej – Siewierska, Anna 2011. Introduction. In Malchukov, Andrej – Anna Siewierska (eds.), *Impersonal constructions. A cross-linguistic perspective*. Amsterdam: John Benjamins, 1–15.

Siewierska, Anna 2013. Passive Constructions. In Dryer, Matthew S. – Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. [Data set]. Zenodo. https://doi.org/10.5281/zenodo.7385533 (Available online at http://wals.info/chapter/107, accessed on 2.11.2023.)

Winkler, Eberhard 2011. *Udmurtische Grammatik*. Wiesbaden: Harrassowitz.

# Emergence of Passive Voice in Central Borneo

Alexander Smith & Michael Yoshitaka Erlewine
(National University of Singapore & National University of Singapore)

We investigate the grammaticalization pathways for the passive voice in several understudied and endangered Austronesian languages of Central Borneo, which developed after the loss of the Proto-Austronesian “Philippine-type" voice system. Such systems remain understudied and provide unique opportunities to study passive voice innovations. In Central Borneo, many languages exhibit an innovative analytic passive with a dedicated preverbal marker (1b).

(1) a. Active: A V P

b. Passive: P PASS (A) V

Voice systems that resemble the paradigm in 1 are reported in numerous languages of the region, including Sa’ban (Clayre 1996), Lebo’ Vo’ Kenyah (Smith, Laing, and Tang 2022), Kayan (Clayre 2014), and Salako (Adelaar 2005). Many such languages share similarities in the shape of the passive marker, involving a nasal, n, and either a preceding or following vowel of varying quality.

We propose that such passives likely reflect the reanalysis of object focus constructions (specifically, null copula pseudoclefts; see e.g. Potsdam (2009)), a pathway towards passive voice rarely attested in studies on diachronic voice (Haspelmath 1990; Wiemer 2011).

(2) P [NP REL A V ] → P PASS A V

‘P is [the one that A V-ed ]’ ‘P is V-ed (by A)’

Primary evidence comes from the phonological form of passive markers, many of which may reflect common relative complementizer forms (such as the Proto-Austronesian “linker” *na, (Blust, Trussel, and Smith 2023)). For example, consider the Kelai (Bornean; Kayan group) object pseduocleft in (3), with relative complementizer ni. Such a structure produces the correct word order and passive marker form as in the Matéq (Bornean; Land Dayak group) passive in (4).

(3) Kelai pseudocleft (Smith 2018)

kuy [ni seP b@wp ] 1SG REL 3SG hit

‘I am the one that he hit.’

(4) Matéq passive (Connell 2013, p. 105)

balo rua karék ni koq mpulua. some seed rubber PASS 1sg gather

‘Some rubber seeds were gathered by me.’

This analysis is complicated by the existence of more divergent passive marker forms, as well as the near homophony between relativizers and genitive case markers from Proto-Austronesian *ni, *na, and

*nu, which suggest multiple passivization pathways. We discuss the range of attested patterns and their analysis at the talk.

**References**

Adelaar, K. Alexander (2005). Salako or Badameà, sketch grammar, texts and lexicon of a Kanayatn dialect in West Borneo. Berlin: Harrasowitz Verlag.

Blust, Robert, Stephen Trussel, and Alexander D. Smith (2023). CLDF dataset derived from Blust's "Austronesian Comparative Dictionary" (v1.2). URL: https:// doi. org/ 10 . 5281 / zenodo. 7741197.

Clayre, Beatrice (1996). "The changing face of focus in the languages of Borneo". In: Papers in Austronesian Linguistics, 3. Ed. by Hein Steinhauer. Canberra: Pacific Linguistics, pp. 51–88.

# Exploring the combinatorics of voice markers: A case study on passivization in Turkish

Niklas Wiskandt
(HHU Düsseldorf)

Within the field of research on grammatical voice marking, most studies have concentrated on describing the voice system of a language, analyzing the properties of a particular voice, or establishing a contrast between two voices, e.g. between active and passive voice.

Meanwhile, phenomena of iteration of voice marking, i.e. marking of two or more voice operations on the same predicate, have received little attention (e.g. Lyutikova & Tatevosov, 2015; Letuchiy, 2015). However, this behavioural dimension of voice marking can potentially shed light on much-debated questions in the voice domain, e.g. whether voice is a grammatical category with different voices as its values, or where voice should be placed on the continuum between inflection and derivation.

This talk presents a case study on combinatorics of voice markers in Turkish (Turkic; nucl1301), with a focus on the behaviour of passive markers. In agglutinating languages such as Turkish, iteration of voice marking occurs by means of agglutination of two or more voice suffixes. For instance, Turkish has causative, passive, reflexive and reciprocal suffixes, and it has been described in the literature (e.g. Kornfilt, 1997; Göksel & Kerslake, 2005) that two of these suffixes can be attached to a verb stem; attested combinations include -caus-caus, -pass-pass (e.g. (1a)), -caus-pass (e.g. (1b)), -refl-pass, -recp-pass and -recp-caus (e.g. (1c)).

(1) a. *Harp-te vur-ul-un-ur.*
war-loc shoot-pass-pass-hab
'One is shot in war.'

b. *Bütün öğrenci-ler-e resim-ler öğretmen tarafından yap-tır-ıl-dı.*
all student-pl-dat picture-pl teacher by paint-caus-pass-pret
'All the students were made to paint pictures by the teacher.'

c. *Baba-ları kardeş-ler-I öp-üş-tür-dü*
father-3pl.poss sibling-pl-acc kiss-recp-caus-pret
'Their father made the siblings kiss each other.' (Turkish)

In contrast, some combinations of voice suffixes are infelicitous. While -caus-pass (1b) works perfectly, -pass-caus (2) does not:

(2) **Bu toplantı birisi tarafindan düzenle-n-dir-di.*
this meeting somebody by organize-pass-caus-pret
intended: 'This meeting was caused to be organized by somebody.' (Turkish)

For several other logically possible combinations of voice suffixes, it remains unclear whether they are just under-described or actually ungrammatical, and why.

After a review of the literature, our case study starts off from the entire set of theoretically possible Turkish voice suffix combinations, in order to provide a comprehensive picture of the (un-)grammaticality of the different combinations. The focus lies on describing and testing all possible combinations involving passive markers. The analysis is based on elicitation data and acceptability judgements: While a systematic elicitation study provides the main database, acceptability judgements

are needed to provide negative evidence for ungrammatical suffix com- binations, and for testing combinations that are infrequent and difficult to target in elicitation, but nevertheless possible and thus essential for an exhaustive analysis of the phenomenon. We develop an explanation for the asymmetries in grammaticality, and discuss implications about the functional range of Turkish passive suffixes and their grammatical status. As a typolog- ical contextualization, the patterns identified in Turkish are compared to other agglutinating languages with similar sets of voice markers. The comparison gives rise to evidence for cross- linguistic parallels: We carve out a set of potentially language-independent constraints operating on combinations of passive markers with other voice markers, which can be tested in subsequent studies. Thereby the talk contributes to the fundamental discussion about the nature of passive markers and passivization.

Abbreviations

3 = third person, acc = accusative, caus = causative, dat = dative, hab = habitual, loc = loc- ative, pass = passive, pl = plural, poss = possessive, pret = preterite, recp = reciprocal, refl = reflexive.

**References**

Göksel, Aslı & Celia Kerslake. 2005. *Turkish: A Comprehensive Grammar* Routledge Compre- hensive Grammars. London: Routledge.

Kornfilt, Jaklin. 1997. *Turkish* Descriptive Grammars. London/New York: Routledge. Letuchiy, Alexander. 2015. Scope Versus Ordering of Operations: Causativization and Or- dering of Valency-Changing Operations in Adyghe. In Stela Manova (ed.), *Affix Ordering Across Languages and Frameworks*, 53–81. Oxford University Press. doi:10.1093/acprof:oso/9780190210434.003.0003. https://academic.oup.com/book/36257/chapter/316291072.

Lyutikova, Ekaterina & Sergei Tatevosov. 2015. Recursive Passivization: A Causative Co- ercion Account. In Stela Manova (ed.), *Affix Ordering Across Languages and Frame- works*, 17–52. Oxford University Press. doi:10.1093/acprof:oso/9780190210434.003.0002. https://academic.oup.com/book/36257/chapter/316290250.

# Passive voice constructions and uncoded passive-like constructions in creoles and worldwide

Martin Haspelmath & Susanne Maria Michaelis
(MPI-EVA Leipzig & Leipzig University)

Keywords: passive voice, valency alternation, uncoded alternation, creole languages, bare passive

Passive voice constructions have a verbal marker that indicates an unusual distribution of semantic roles over argument positions, and they are widely found across the world's languages, including in creole languages (Siewierska 2013; APiCS 2013). But since Holm (1997), creolists have been aware that quite a few creoles have a passive-like construction that lacks verbal coding, as in (1).

(1) Jamaican Creole (Kouwenberg 2023: 246)
*Lai a tel.*
lie IMPF tell
'Lies are being told.'

Such constructions had earlier been claimed to be inexistent (Haspelmath 1990), but more recently, they have been discussed for a number of languages, especially of West Africa (Cobbinah & Lüpke 2012).

In this presentation, we advance the discussion in two ways: First, we situate the discussion of creole "bare passives" in the context of recent research on valency and voice constructions (Zúñiga & Kittilä 2019; Haspelmath 2022; Creissels 2024), where a consensus has emerged that voice constructions have verbal coding by definition, so that the pattern in (1) is at most "passive-like", but not a passive construction. Given that traditional grammatical relations like "subject" are not easily applicable across languages around the world, we cannot simply say that (1) is "passive-like" in that "the object becomes the subject". Indeed, without verbal coding, it is unclear which of the alternating constructions should be regarded as basic, and in an alternation like (2), one might suggest that the second alternant (2b) is not a "bare passive" or a "passive-like ambitransitive construction", but an ergative construction. Thus, a deeper understanding of the cross-linguistic picture requires a more sophisticated approach to syntactic comparison.

(2) Bambara (Creissels 2024: §15.3.2)
a. *Wùlû má sògô dún.*
dog.D NEG meat.D eat
'The dog did not eat the meat.'

b. *Sògô má dún wùlú fɛ̀.*
meat.D NEG eat dog.D by/ERG
'The meat was not eaten by the dog.'

In a second step, we will re-examine the worldwide distribution of various passive and passive-like constructions in order to assess the likelihood of an explanation in terms of African substrate influences. This will be based on the APiCS survey, on the rich observations in Holm & Patrick (2007),

and on selected parallel-text studies of creole languages. While historical scenarios will always remain conjectural, we suggest that the earlier literature has been biased by the idea of "creole simplicity" and has not sufficiently appreciated the cross-linguistic rarity of uncoded passive-like constructions in the world's languages.

**References**


APiCS: Michaelis, Susanne Maria, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.). (2013). *Atlas of pidgin and creole language structures*. Oxford: Oxford University Press.

Cobbinah, Alexander & Friederike Lüpke (2012. Not cut to fit: Zero coded passives in African languages. In Matthias Brenzinger & Anne-Maria Fehn (eds.), *Proceedings of the 6th World Congress of African Linguistics*, 153–165. Cologne: Köppe.

Creissels, Denis (2024). *Transitivity, valency and voice*. Oxford: Oxford University Press (to appear).

Haspelmath, Martin (1990). The grammaticization of passive morphology. *Studies in Language* 14(1). 25–72.

Haspelmath, Martin (2022). Valency and voice constructions. Leipzig, MPI-EVA, ms. https://ling.auf.net/lingbuzz/005941.

Holm, John (1997). Passive-like constructions in English-based and other creoles. In Edgar Werner Schneider (ed.), *Englishes around the world (Volume 1: General studies, British Isles, North America)*. Amsterdam: Benjamins.

Holm, John & Peter Patrick (eds.) (2007). *Comparative creole syntax: Parallel outlines of 18 creole grammars* (Westminster Creolistics Series 7). London: Battlebridge.

Kouwenberg, Silvia (2023). The creole passive. In Patricia Cabredo Hofherr, Herby Glaude & Elena Soare (eds.), *Grammar is a moveable feast: A webschrift for Anne Zribi-Hertz*, 239–249. Paris: Zenodo. https://doi.org/10.5281/zenodo.8055461.

Siewierska, Anna (2013). Passive constructions. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: MPI-EVA. https://wals.info/chapter/107.

Zúñiga, Fernando & Seppo Kittilä (2019). *Grammatical voice*. Cambridge: Cambridge University Press.

# Active and inactive uses of the Middle Low German verb *blîven* as a trigger to renew the debate on diathesis

Marta Woźnicka, Alicja Piotrowska & Maria Klejnowska
(Adam Mickiewicz University, Adam Mickiewicz University & Adam Mickiewicz University)

Keywords: active, passive, inactive, diathesis, Middle Low German

From a typological point of view, the passive is generally associated with verbal forms having certain passive morphology. According to Haspelmath (1990: 170), there are no passive constructions without passive morphology, and he qualifies as markers of the passive voice, among others additional stem affix, auxiliary verb, particle etc. However, both contemporary (see Arka/Kosmas 2005) and historical linguistic data (see Luraghi et al. 2021: 371-378) indicate that passive morphology is not the rule. Further, synchronically, the sense of the diathesis *active* vs. *passive* is usually grasped as the opposition *agent* = *subject* vs. *patient* = *subject*. Adopting this opposition, however, significantly limits the number of verbs that can form it to the so-called transitive. Consequently, treating such asymmetrical opposition as grammaticalized seems incorrect - hence renewed attempts to deny diathesis, e.g. in German (Kotin 2003: 239).

An analysis of the Middle Low German verb *blîven* reveals a number of shortcomings of these synchronic oppositions. *Blîven* has an unusually rich inventory of meanings, i.e., among others, 'remain', 'become' and 'die' in combination with the adjective 'dead', as well as 'leave' and 'be left'. Although the verb, unlike in Scandinavian languages (into which it is borrowed in the late 1200s) (cf. Skrzypek 2020), never attains the status of an auxiliary verb in the passive voice, its uses in the sense of 'leave' (1) and in the sense of 'be left' (2) may constitute, on the basis of the unambiguous relation between the complement of the first sentence (*lxxxix punt*) and the subject of the second sentence (*dhat*), an *active* vs. *passive* opposition, even if the passive morphology is missing.

1. jck **blef** ein lxxxix pund
   I **left** them 89 pounds
   'I left them 89 punds'
   (Hild._Veck._Briefe, 292r,03)

2. dhat gvt to valeberghe scal eme **bliuen**
   The manor to Valeberghe shall him **be left**
   'The manor at Valeberghe shall be given/left to him'
   (Brs._Ält._DegB_Altst._II, 27rb,26-27)

The aim of the paper is to present the ambiguous cases in which the verb *blîven* is used in active and inactive constructions without any morphological changes. The basis for the analysis is an annotated corpus of 73 Middle Low German texts from the ReN and DiaPass corpus from the period 1200–1500. As *blîven* is used with active agents and without them, we argue that it can receive active or passive interpretation. In view of the above, and the limitations of assuming an *agent* vs. *patient* opposition in understanding diathesis (cf. Kotin 2003: 239), we assume that the sense of diathesis is better captured by the opposition of *active* vs. *inactive* proposed by Weisgerber (1963), which assumes that an active sentence expresses what X does (to the Y), while an inactive sentence expresses what happens (to Y), regardless of how it is codified. The optionality of Y justifies the rejection of the opposition *agent* vs. *patient* in the understanding of diathesis as well as significantly expands the stock

of verbs that can form it. The recognition of other, not only morphological ways of codifying the inactive makes it possible to consider sentences with *blîven* in the sense of 'be left' as inactive.

**References**


Arka, I Wayan, Kosmas, Jeladu. (2005). Passive without passive morphology? Evidence from Manggarai. In: I W. Arka and M.D. Ross (eds.), *The many faces of Austronesian voice systems: Some new empirical studies* (p. 87–117). Canberra: Pacific Linguistics.

Haspelmath, Martin. (1990). The grammaticization of passive morphology. *Studies in Language* 14 (1), 25–72.

Kotin, Michail L. (2003). *Die werden-Perspektive und die werden-Periphrasen im Deutschen. Historische Entwicklung und Funktionen in der Gegenwartssprache*. Frankfurt am Main: Peter Lang Verlag.

Luraghi, Silvia, Inglese, Guglielmo, Kölligan, Daniel. (2021). The passive voice in ancient Indo-European languages: inflection, derivation, periphrastic verb forms. *Folia Linguistica Historica* 42(2), 339–391.

Skrzypek, Dominika. (2020). The Swedish bli-passive in a diachronic perspective. *LEGE ARTIS Language yesterday, today, tomorrow* V(2), 301–334.

Weisgerber, Leo. (1963). *Von den vier Stufen in der Erforschung der Sprache*. Stuttgart: Schwann.

ReN: Referenzkorpus Mittelniederdeutsch/Niederrheinisch (1200–1650): https://www.slm.uni-hamburg.de/ren.html

DiaPass: an in-progress corpus of Danish, Swedish and Middle Low German annotated texts created for the project *Constructionalization and construction change. A diachronic perspective on the periphrastic passive in Danish and Swedish*: DiaPass: Home (amu.edu.pl)

# Passives and Related Constructions in Unangam Tunuu: Origins in Language Contact

Anne Berge
(University of Alaska Fairbanks)

Unangam Tunuu (UT) is an Eskaleut language; however, it differs substantially from the Yupik/Inuit (Y/I) branch of the language family. One way in which the branches diverge is in their passive constructions. In this paper, I show how such constructions in UT more closely resemble those in Dene languages than those in Y/I.

Y/I and UT passives are derived similarly, from combinations of nominalizers + a verbalizing suffix (1) or from a perfective derivational suffix (2):

| | | | |
|---|---|---|---|
| (1) | CAY *-gaq,* WG *-saq*<br>'passive participle' | + CAY *-(ng)u-* WG *-u-*<br>'copula' | CAY *-gau-,* WG *-saa-*<br>'passive' |
| (2) | UT *-lig- 'provided with'<br>Y/I *–(u)ma-* 'perfective'<br>UT *-x̂ ta-* 'active perfective' | + UT *-a-* 'copula' | *-lga-* 'passive'<br>*-(u)-ma-* with passive sense<br>*-ĝa-* 'passive perfective' |

Syntactically, they behave differently. Y/I passives take transitive verbs as input and detransitivize the verbs; promote direct objects to subject position; and are dispreferred (3).

| (3) | Alaskan Yup'ik (Miyaoka 2012) | West Greenlandic (Fortescue 1984) |
|---|---|---|
| | *tuqu-t-a-u-guq* | *tuqu-sa-a-voq* |
| | die-CAUS-PAS.PRT-COP-3SG.IND | die-(CAUS).PAS.PRT-COP-3SG.IND |
| | 'it is one that is killed' = 'it is killed' | it is one that is killed' = 'it is killed' |

UT passive constructions are ubiquitous, with very different characteristics:

- Passives are non-promotional, demoting the agent without necessarily promoting the object (4) (Bergsland 1997:172)

| (4) | *Ula-m ilan ting su-ku-x̂ txin* | *Ula-m ilan ting su-lga-ku-x̂* |
|---|---|---|
| | house-REL.SG in.LOC 1SG take-IND-2SG | house-REL.SG in.LOC 1SG take-PAS-IND-3SG |
| | 'You took me into the house.' | 'Someone took me into the house.' |

- Intransitive verbs can be morphologically passivized (5) (Bergsland 1997:167)

| (5) | *Angaĝi-qali-na-x̂* | *angaĝi-lga-qali-qa-x̂* |
|---|---|---|
| | live-begin-PART-3SG | live-PAS-begin-AN.PART-3SG |
| | 'He began to live.' | 'Someone began to live.' |

- A noun can be passivized with specific meanings, e.g. 'take…' (6) (Bergsland 1997:170)

(6) *Piitra-x̂ ukina-x̂ su-na-x̂* ʔ *Ukina-x̂ Piitra-lga-qa-x̂*
Peter-ABS.SG knife-ABS.SG take-PART-3SG knife-ABS.SG Peter-PAS-AN.PART-3SG
'Peter took the knife.' 'The knife was [taken] by Peter.'

- A passive construction can express a benefactive reflexive (7) (Bergsland 1997:169)

(7) *Nung iqa-x̂ si-lga-qa-qing*
1SG.DAT baidarka-make-PAS-AN.PART-1SG 'I had a baidarka made for me.'

- The passive is a distancing strategy: 3SG passives express 1PL (8) (Bergsland 1997:91)

(8) *Ayuxta-lga-ax̂ ta-x̂*
go.in.boat-PAS-OPT-3SG 'Let's go out in a boat.'

These characteristics are found in Dene languages; although the morphology is different, passivization suppresses verbal arguments with the same effects. The Dene D-'classifier' is associated with argument suppression and creates nonpromotional passives (1); impersonal passives (2); benefactive reflexives (3); detransitivization with indeterminate object (4) and 1PL (also in Eyak) (Krauss 2012, 2017):

| | | [no D-element] | D-element |
|---|---|---|---|
| 1. | Eyak | *k'u-x-kus*<br>IND-1-wash<br>'I am washing something' | *xu-da-kus*<br>1OBJ-CL-wash<br>'I am being washed' |
| 2. | Tlingit | *kóox woo-too-si-.ée* rice<br>PRF-1P-CL-cook 'we cooked rice' | *kóox woo-doo-dzi-.ée*<br>rice PRF-someone-CL-cook<br>'someone cooked rice' (Thompson 1996) |
| 3. | Koyukon | *le-tɬ-baats*<br>PRF-1:CL-boil<br>'I boiled it' | *daa-l-ge-baats*<br>TH-PRF-1CL-boil<br>'I boiled it for myself' |
| 4. | Eyak *kihsh* *OBJ-kihsh*<br>'dipnet' 'catch OBJ in dipnet' | | *'i-dA-ki:sh-k'*<br>3OBJ-CL-dipnet-customary<br>'he dipnets' |
| 5. | Mattole | *ʔi-ʃ-aŋ*<br>IND-1SG-eat<br>'I eat' | *ʔi-di-d-aŋ*<br>IND-1P-CL-eat<br>'we eat' (Thompson 1996) |

I discuss how this supports linguistic, archaeological, and genetic evidence for prehistoric contact between UT and Dene (Berge 2012, Fortescue 1998)

**References**

Berge, Anna. 2012. Object Reduction in Aleut. *Asian and African Languages and Linguistics* 7. ILCAA, Tokyo University of Foreign Studies 7. 5-23 Bergsland, Knut.

1997. *Aleut Grammar*. Fairbanks: ANLC.

Fortescue, Michael. 1998. *Language Relations Across Bering Strait: Reappraising the Archaeological and Linguistic Evidence.* London: Cassell.

Krauss, Michael. 2012. *Eyak Dictionary*, Draft. Fairbanks: Alaska Native Language Archive. Krauss, Michael. 2017. *Eyak Grammar*, Draft. Uncatalogued copy held at the Alaska Native Language Archive, Fairbanks, AK.

Thompson, Chad. 1996. The Na-Dene Middle Voice: An Impersonal Source of the D-Element. *International Journal of American Linguistics* 62:4:351-378.

# Agent marking in passive constructions in Modern Hebrew: A corpus-based diachronic study

Yael Reshef
(Hebrew University of Jerusalem)

Keywords: passive, language change, agent, Modern Hebrew, standardization, grammaticalization

This paper traces the formation process of fixed means for agent marking in passive constructions in Modern Hebrew. Whereas the morphological means to form passive verbs existed already in ancient Hebrew, like in other Semitic languages passive verbs were intrinsically impersonal (Rabin 1999). Therefore, they were rarely accompanied by specified agents (Notarius 2013). Modern Hebrew, by contrast, employs the inherited verbal patterns as a productive structural mechanism for passivization. The agent may be either left unspecified or explicitly mentioned by the adverbial phrase *ʕal yede* (or less often: *bi-yede*) 'by' (literally: 'on/in the hands of'), originally an instrumental (Halevy 2020), e.g.:

(1) *gever* *hutqaf* *(ʕal-yede* *neʕar-im)*
man attack.PASS.PST.3MSG (in-hands.CS youngster-PL)
'a man has been attacked (by youngsters)'

A wide distribution of passive constructions, with or without a specified agent, is a conspicuous feature of contemporary journalistic style (Taube 2013). Hence, the examination of early journalistic texts is particularly suitable from the methodological viewpoint for tracing the factors involved in the change process from impersonal usage to full-fledge passive. Our analysis is based on data retrieved from a vast morphologically-parsed textual corpus of journalistic texts from the seminal period of modernization (1870s-1930s), divided into 7 particularly significant historical time points throughout the period of transformation of Hebrew from a traditional language in a state of diglossia into a modern national tongue.

The textual data indicated an initial vacillation between various means of expression in passive constructions, as writers searched for suitable linguistic means to explicitly specify the agent, a feature extant in their native European contact languages. These included *mi-* 'from', *meʔet* 'from', *mi-taʕam* 'on behalf', *mi-tsad* 'from side', *be-yad* 'in hand', *bi-yede* 'in hands', *ʕal yede* 'on hands'. In light of this initial variation, we will claim that modernization occurred in this domain of the linguistic system by a process of selection rather than by the introduction of structural innovations. This finding integrates into a growing body of literature indicating that selection between inherited linguistic features, either in their original functions or by the allocation of new functions, played a major role in the modernization and standardization of Hebrew (see e.g. various papers in Doron 2016 and Doron et al. 2019, Bar-Ziv Levy 2021). Consequently, the processes of modernization often involved a reduction, rather than an increase, in the number of options potentially available to language users in achieving their communicative goals.

The transformation of *ʕal yede* from instrumental into an agent marker in passive constructions occurred despite repeated prescriptivist objections, reflecting the tendency of ordinary speakers to adopt linguistic means that were felt as suitable to fill the functional needs triggered by the contact languages. The consolidation of *ʕal yede* its this new function contributed not only for establishing a clearer distinction between the impersonal and the passive voice, but also for the formation of new formal registers, essential for the ability of Hebrew to develop into a modern

language, capable of meeting all the communication needs of the newly-created speech community (Rosén 1992).

This research was supported by the ISRAEL SCIENCE FOUNDATION (grant no. 474/20).

**References**

Bar-Ziv Levy, Miri (2021), Regularization in the Crystallization of Modern Hebrew: The case of counterfactual conditionals, *Hebrew Studies* 62, 381–402.

Doron, Edit (ed) (2016), *Language Contact and the Development of Modern Hebrew*, Leiden: Brill.

Doron, Edit, et al. (eds) (2019), *Language Contact, Continuity and Change in the Genesis of Modern Hebrew*, Amsterdam: John Benjamins.

Halevy, Rivka (2020), Impersonal and pseudo-impersonal constructions, in Ruth A. Berman (ed.), (2020), *Usage-Based Studies in Modern Hebrew: Background, Morpho-lexicon, and Syntax*, Amsterdam / Philadelphia: John Benjamins, 539–582.

Notarius, Tania (2013), Argument, in Geoffrey Khan (ed), (2013), Encyclopedia *of Hebrew Language and Linguistics*, Leiden: Brill (consulted online).

Rabin, Chaim (1999), givun signoni baguf hastami bamikra (Stylistic variation in the use of impersonals in Biblical Hebrew), in Moshe Bar-Asher & Baraq Dan (eds), (1999), *Linguistic Studies: Collected Papers in Hebrew and Semitic Languages*, Jerusalem: The Academy of Hebrew Language & The Bialik Institute, 244–253.

Rosén, Haiim B. (1992), zutot mehitgabšuta šel haʕivrit hayisreʔelit [Trifles from the emergence of Israeli Hebrew], *Haħug hayisreʔeli shel ħavre haħevra haʔeropit levalšanut, divrei ha-mifgaš hašmini* [*Societatis Linguisticae Europaeae Sodalicium Israelense: Studia*] V, 33–39.

Taube, Dana (2013), Passive: Modern Hebrew, in Geoffrey Khan (ed), (2013), Encyclopedia *of Hebrew Language and Linguistics*, Leiden: Brill (consulted online).

# The passive in South Slavic: Factors determining the distribution of passive constructions in Slovene and Macedonian

Mladen Uhlik, Liljana Mitkovska & Eleni Buzarovska
(University of Ljubljana; Fran Ramovš Institute of Slovenian Language; Research Centre of the Slovenian Academy of Sciences and Arts Ljubljana, FON; Faculty of Applied Foreign Language & Ss. Cyril; Methodius University; Faculty of Philology "Blaze Koneski")

Keywords: passive, voice, South Slavic languages, Slovene, Macedonian

This paper explores the complexities of the passive voice in the South Slavic languages, focusing on two distinct passive constructions: the reflexive passive and the periphrastic passive formed using 'be' and a past participle of the verb (Siewierska 1984, Sussex and Cubberley 2006). The primary objective is to investigate the potential applications of these passive constructions in Slovene and Macedonian, the two South Slavic languages positioned respectively at the northern and southern ends of the linguistic continuum.

The two constructions have various functions, but when used with a passive sense, both serve as agent-defocusing strategies, yet diverge in event structure. This results in unique usage constraints which may differ in the two languages. For example, when the predicate denotes a single (non-generic) event, Slovene precludes the use of the reflexive form in (1), while Macedonian allows both constructions, as do Bosnian, Croatian, Montenegrin and Serbian (BCMS) (cf. Uhlik & Žele 2022, Mitkovska 1998).

(1)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Sn | Vse | blag-o | je | bil-o | prod-an-o | do | poldn-eva. |
| | all | goods-NOM.SG.N | be.PRS.3SG | be.PST-3SG.N | sell.PFV-PTCP.PASS-SG.N | by | noon-GEN.SG.N |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Mk | Seta | stok-a | se | prodad-e/ | be-še | prodad-en-a | |
| | all | goods-SG.F | REFL | sell.PFV-AOR.3SG | be-IMPF.2/3SG | sell.PFV-PTCP.PASS-SG.F | |
| | do | pladn-e. | | | | | |
| | by | noon-SG.N | | | | | |

'All goods were sold by noon.'

On the other hand, Slovene periphrastic passive constructions encoding past events utilize past auxiliaries (cf. Uhlik & Žele 2022) in contrast to Macedonian, which allows present forms (2). Thus, in Slovene, unlike Macedonian, a formal distinction between the past passive (utilizing the past form of the auxiliary) and the present resultative (using the present form of the auxiliary) is introduced by the auxiliary verb's form (cf. Uhlik & Žele 2022).

(2)

| | | | | | | |
|---|---|---|---|---|---|---|
| Sn | Kdaj | je | bil | ubi-t | Franc | Ferdinand? |
| | when | be.PRS.3SG | be.PST-3SG.M | kill.PFV-PTCP.PASS-SG.M | Franz-NOM.SG.M | Ferdinand-NOM.SG.M |

| | | | | | |
|---|---|---|---|---|---|
| Mk | Koga | e | ubie-n | Franc | Ferdinand? |
| | when | be.PRS.3SG | kill.PFV-PTCP.PASS-SG.M | Franz-SG.M | Ferdinand-SG.M |

'When was Franz Ferdinand killed?'

The paper aims to identify the similarities and differences in the distribution of periphrastic and reflexive passives in Slovene and Macedonian. The focus will be on investigating the potential influence of the following parameters on the distribution of the constructions: the actional class of the predicate, the presence of temporal reference, and the semantics of the promoted argument. Additionally, the realization strategies of the demoted argument and its referential status will be explored.

The paper draws on examples from parallel corpora, such as *KonText concordancer*, and translation equivalents from contemporary fiction works, including translations in both directions. Original examples and their translations are annotated using the aforementioned parameters, enabling both a quantitative and qualitative analysis. To the best of our knowledge, the distribution of passive constructions in South Slavic languages has not been systematically analyzed using contrastive and empirical methods.

Corpora

Křen, M. Cvrček, V. Henyš, J. Hnátková, M. Jelínek, T. Kocek, J. Kováříková, D. Křivan, J. Milička, J. Petkevič, V. Procházka, P. Skoumalová, H. Šindlerová & J. Škrabal, M. (2020), *The KonText concordancer*. *SYN2020: reprezentativní korpus psané češtiny*. Praha: Ústav Českého národního korpusu FF UK. Dostupný z WWW: http://www.korpus.cz

**References**

Mitkovska, Liljana (1998), The SE Passive in Macedonian Compared to the Neighbouring Slavic Languages. *Studia Linguistica Polono-Meridianoslavica*, Tom 9, Torun. 47–60.

Siewierska, Anna (1984), *Passive: a comparative linguistic analy*sis. London: Croom Helm.

Sussex, Roland and Paul Cubberley (2006), *The Slavic Languages*, Cambridge: Cambridge University Press, New York.

Uhlik, Mladen, Žele, Andreja (2022), Expressing Voice in Slovene, Russian and Štokavian: Similarities and Differences, *Slavistična revija*, 70(3), 269–287.

# WS14 Pathways to insubordination

# The diachrony of "past imperatives" in Dutch

Daniel Van Olmen
(Lancaster University)

Keywords: conditional, Dutch, imperative, insubordination, optative

The Dutch constructions in (1) are often called "past imperatives" (Wolf 2003, Boogaart and Jansen 2010). They share the lack of an explicit subject and a certain directivity with the imperative (e.g. reprimanding addressees and urging them to own up to their behavior in 1a). However, as argued in Van Olmen (2018), regarding such constructions –in Dutch *and* other languages– as proper imperative paradigm members requires too broad a characterization of the imperative to capture its (im)possibilities and also leaves their cross-linguistic scarcity unexplained.

(1) a. *Was* *gekomen!*
be.PST.SG come.PST.PTCP
'You should have come!'
b. *Was* *maar* *eens* *kwaad!*
be.PST.SG MP MP angry
'You should be angry for once!'

We therefore propose to analyze them as separate constructions and, following Duinhoven (1995, 1997), as the idiosyncratic products of a variety of interacting source constructions and diachronic processes. A central role here is reserved for the insubordination (Evans 2007) of inverted conditionals. Unlike the previous work, though, we conduct a systematic corpus study of data from the 1500s to the 2000s. Our investigation mainly relies on the Nederlab interface (Brugman et al. 2016) for a multitude of diachronic corpora. They are complemented by resources like the Corpus Spoken Dutch (Dutch Language Union 2004) and the nlTenTen20 corpus of online Dutch (Jakubícek et al. 2013) for recent data.

We first confirm prior assumptions that counterfactual (1a) predates less insistently directive (1b), with first attestations from the 19th and the late 20th century respectively, and that (1b) remains a marginal phenomenon. It is also shown that, contrary to Duinhoven (1995) but consistent with our own earlier hypothesis, (1a) does not directly originate from a conditional subclause: the historically primary part played by insubordination in its development is turning inverted conditionals into optative constructions. The constructed example in (2a) illustrates this evolution, in the 1600s in our data, and (2b) the roughly simultaneous one with simple past verbs.

(2) a. *Was jij maar gekomen, dan …* 'If you had just come, then …'
a'. → *Was jij maar gekomen!* 'If only you had come!'
b. *Kwam jij maar, dan …* 'If you just came, then …'
b'. → *Kwam jij maar!* 'If only you came!'

In the second person, these optatives –*and* their full-fledged conditional counterparts before and after the changes in (2)– are found to be used pragmatically as addressee-oriented reproaches and soft directives. In our view, it is this functional contiguity with the imperative that motivated the analogy

that, if 'you' can be left out in the imperative (e.g. *zwijg (jij) eens!* '(you) be quiet!'), it can also be omitted in (2a'). This entire process, finally, has given rise to the construction in (1a), whose distinctness/autonomy is evidenced by its 20th-century ability to occur without modal particles (which could be seen as cues for an implicit apodosis). For (1b), however, it is incomplete: its tentative emergence is likely attributable to analogy in some speakers of (2b) with (2a) and (1a) and it also does not seem able to appear without modal particles.

**References**


Boogaart, Ronny, and Theo Janssen (2010), Mood in Dutch, in Björn Rothstein, and Rolf Thieroff (eds), *Mood in the Languages of Europe*, Amsterdam: John Benjamins, 117–132.

Brugman, Hennie, Martin Reynaert, Nicoline van der Sijs, René van Stipriaan, Erik Tjong Kim Sang, and Antal van den Bosch (2016), Nederlab: Towards a single portal and research environment for diachronic Dutch text corpora, *Proceedings of the International Conference on Language Resources and Evaluation* 10, 1277–1281.

Duinhoven, A.M. (1995), *Had gebeld!* De irreële imperatief, *Tijdschrift voor Nederlandse Taal- en Letterkunde* 111, 346–364.

Duinhoven, A.M. (1997), *Stopte dan ook!* De imperatief in ontwikkeling, in Els H.C. Elffers-van Ketel, Joop M. van der Horst, and Wim G. Klooster (eds), *Grammaticaal Spektakel: Artikelen Aangeboden aan Ina Schermer-Vermeer*, Amsterdam: Vakgroep Nederlandse Taalkunde, 47–58.

Dutch Language Union (2004), *Corpus Spoken Dutch: Release 1.0*, The Hague: Dutch Language Union.

Evans, Nick (2007), Insubordination and its uses. In Irina Nikolaeva (ed), *Finiteness: Theoretical and Empirical Foundations*, Oxford: Oxford University Press, 366–431.

Jakubícek, Miloš, Adam Kilgarriff, Vojtěch Kovár, Pavel Rychlý, and Vít Suchomel (2013), The TenTen corpus family, *Proceedings of the International Corpus Linguistics Conference* 7, 125–127.

Van Olmen, Daniel (2018), Reproachatives and imperatives, *Linguistics* 56(1), 115–162.

Wolf, Henk (2003), Imperatieven in de verleden tijd, *Taal en Tongval* 55, 168–187.

# Insubordinate hypothetical manner clauses from Latin to Spanish: From *quasi* to *como si*

An Van linden, María Sol Sansiñena & Ezra la Roi
(University of Liège, KU Leuven & Ghent University; FWO)

Key words: clause-combining, hypothetical manner, insubordination, diachronic corpus study, pathways of change

While most work on Spanish insubordinate constructions has adopted a synchronic perspective (see Schwenter 2016; papers in Beijering et al. 2019; Gras & Sansiñena 2021; Lastres-López 2021), this paper focuses on the diachrony of underdescribed Spanish insubordinate clauses introduced by *como si* 'as if', e.g. (2) and (3). Whereas syntactically dependent clauses introduced by *como si* generally express hypothetical manner, as in (1), marked for past subjunctive, insubordinate clauses like (2) and (3) have different semantic-pragmatic values.

(1) *Te abrí las puertas de mi vida* ***como si*** *fueras a quedarte.* (Twitter)
'I opened my life's doors for you as if you were to stay.'

(2) *"Sexualidad, chicos… conoce sus puntos débiles."*
*¡****Como si*** *no lo supiéramos, de verdad!* (Val.Es.Co)
'"Sexuality, boys… know their weaknesses." As if we didn't know that, to be honest!'

(3) A: *¿Cuándo queréis hacer el rodaje?*
B: *Si por mí fuese, comenzaríamos mañana mismo.*
C: *Lo mismo digo. Por mí* ***como si*** *empezamos esta tarde.* (esTenTen18)
A: 'When do you want to start shooting?'
B: 'If I had my way, we would start tomorrow already.'
C: 'Same here, we can (even) start this afternoon for all I care.'

In (2), the (past subjunctive) *como si*-clause also gets a counterfactual interpretation, but it is used turn-initially and functions as an interactional challenge. Specifically, the speaker uses it to deny the assumption that they would *not* know boys' weaknesses and to express their indignation at the speaker of the preceding utterance entertaining that very assumption. In (3), by contrast, the speaker uses an independent (present indicative) *como si*-clause to express their indifference with respect to when exactly they would want to start filming by setting up a scale with a hypothetical case at its extreme (cf. Royo Viñuales & Van linden ms). The scale at issue is a timeline, in which 'this afternoon' is extreme relative to speaker B's baseline of 'tomorrow'.

This paper aims to investigate the source(s) and developmental pathway(s) of *como si*-constructions in Peninsular Spanish. Interestingly, precursors of the type in (2) have been observed in Archaic and Classical Latin, featuring the conjunction *quasi* 'as if', as in (4), marked for present subjunctive (la Roi 2022: 40-41).

(4) Davus *modo introii.*
Simo ***quasi*** *ego quam dudum rogem*. (Ter. *And.* 850 – Archaic Latin)
Davus 'I went in just now'
Simo 'As if I asked you how long ago!'

Precursors of the scalar evaluation construction in (3), by contrast, have not been recorded in Latin. Moreover, absence of this type in South American varieties may suggest that it was not entrenched in Peninsular Spanish before the late 15th century.

Our diachronic study relies on data extracted from the *Corpus del Diccionario histórico de la lengua Española* (CDH nuclear; 12th–21st centuries), more specifically random 200-hit samples per century, restricted to Peninsular Spanish. For Latin and earlier Spanish data, we will rely on the corpora of Archaic and Classical Latin used in la Roi (2022), as well as the *Corpus Diacrónico del Español* for the 8th–11th centuries.

Acknowledgments
The research reported on in this talk was made possible by the research grant "How do grammar and discourse interact? Answers from subordination, coordination and insubordination" (PDR T.0065.20), awarded to Liesbeth Degand and An Van linden by the Fund for Scientific Research (F.R.S.-FNRS).

**References**

Beijering, Karin, Gunther Kaltenböck, María Sol Sansiñena (eds) (2019), *Insubordination: theoretical and empirical issues*, Berlin: De Gruyter.

Gras, Pedro and María Sol Sansiñena (2021), Discourse structure, constructions and regional variation: Non-embedded indicative *que*-clauses in three regional varieties of Spanish, in Miriam Bouzouita, Renata Enghels and Clara Vanderschueren (eds), *Convergence and divergence in Ibero-Romance across contact situations and beyond*, Berlin / Boston: De Gruyter, 245-280.

Lastres-López, Cristina (2021), *From subordination to insubordination: A functional-pragmatic approach to* if/si-*constructions in English, French and Spanish spoken discourse*, Bern: Peter Lang.

la Roi, Ezra (2022), Insubordination in Archaic and Classical Latin: Commands, Requests, Wishes and Assertives, *Journal of Latin Linguistics* 21(1), 23–45.

Royo Viñuales, Víctor & An Van linden (Submitted Manuscript), Beyond hypothetical manner: A functional typology of insubordinate *como si*-clauses.

Schwenter, Scott (2016), Independent *si*-clauses in Spanish: Functions and consequences for insubordination, in Nicholas Evans and Honoré Watanabe (eds), *Insubordination*, Amsterdam: John Benjamins, 89-112.

# Insubordination pathways in three millennia of Greek: origins, morphosyntactic change and pragmatic enrichment

Klaas Bentein & Ezra la Roi
(Ghent University & Ghent University)

Keywords: insubordination, morphosyntactic change, pragmatic enrichment, history of Greek, mood and modality

Research on insubordination, the conventionalized independent use of a formally subordinate clause (Evans 2007, 377), has thrived in recent years (e.g. Evans and Watanabe 2016; Beijering et al. 2019). Scholars have pointed out the variety of their interpersonal functions and developed diachronic hypotheses about the mechanisms behind insubordination (e.g. Evans 2007, ellipsis of the main clause; Mithun 2008 a dependency extension; Heine et al. 2016 cooptation, Cristofaro 2016:395 clausal disengagement). Yet, studies have typically based their findings on synchronic data, also to infer the diachronic processes involved. This might explain why most 'diachronic' research has focused on the *origins* of insubordinate constructions, since they "can emerge from a variety of sources via a variety of mechanisms" (Mithun 2019, 31), but disregarded the question which continuity and change insubordinate constructions themselves show across different periods of a language (exceptions are Narrog 2016; la Roi 2021; and 2022 on different stages of Japanese, Ancient Greek, and Latin).

Therefore, we use the rich evidence from the history of Greek (VIII BCE – now) to fill in several gaps in our understanding of insubordination pathways, from their origins to their present-day outcomes. We focus on three insubordinate schemas, each with a different source: conditional insubordinate *eíthe* 'if only' (la Roi 2021), insubordinate infinitives, and complement insubordinate (*í)na* 'that' (see la Roi forthc for an overview). The corpus data which is to supplement and revise the limited descriptions in grammars has been collected using (mainly lemmatized) searches in the *Thesaurus Linguae Graecae* (for literary texts), *Trismegistos words* (for the papyri), and *Sketch Engine* (ElTenTen 2019). After describing the various origins of their insubordination, we describe (i) the different morphosyntactic changes which these insubordinate constructions undergo (e.g. mood changes, insubordinator fusion and doubling), and (ii) the pragmatic spread to different speech acts in later stages of the Greek language (e.g. from directive to assertive speech act). Whereas some insubordinate strategies seem eventually to be abandoned (e.g. insubordinate infinitives due to the loss of the infinitive or *eíthe* wishes after specialization to high register in Medieval Greek.), others become an integral part of the modal system (e.g. the insubordinate *na* as directive is viewed by grammarians as part of the imperative paradigm in Modern Greek, e.g. Mackridge 1987: 283). Summarizing, the evidence from the history of Greek offers further insights not only into the pathways *into* insubordination, but also of the *long-term histories* of insubordinate constructions (la Roi forthc. b.).

**References**


Beijering, Karin, Gunther Kaltenböck, and Maria S. Sansiñena, eds. 2019. *Insubordination: Theoretical and Empirical Issues*. Berlin; Boston: De Gruyter.

Cristofaro, Sonia. Routes to insubordination: A cross-linguistic perspective, in *Insubordination,* N. Evans and H. Watanabe (eds.), Amsterdam ; Philadelphia: John Benjamins Publishing Company, 393-422.

Evans, Nick. 2007. Insubordination and Its Uses, in *Finiteness: Theoretical and Empirical Foundations*, I. Nikolaeva (ed.), New York: Oxford University Press, 366–431.

Evans, Nick, and Honoré Watanabe. 2016. *Insubordination*. Amsterdam ; Philadelphia: John Benjamins Publishing Company.

Heine, Bernd, Gunther Kaltenböck, and Tania Kuteva. 2016. On Insubordination and Cooptation, in *Insubordination,* N. Evans and H. Watanabe (eds.), Amsterdam: John Benjamins Publishing Company, 39–64.

Mackridge, Peter. 1987. *The Modern Greek language : a descriptive analysis of standard Modern Greek*. Oxford: Clarendon Press.

Mithun, Marian. 2008. The Extension of Dependency Beyond the Sentence, *Language* 84 (1), 69–119.

———. 2019. Insubordination, in *Theoretical and Empirical Issues*, K. Beijering, G. Kaltenböck, and M. Sansiñena (eds.), Berlin; Boston: De Gruyter, 29–54.

Narrog, Heiko. 2016. Insubordination in Japanese Diachronically, in *Insubordination*, N. Evans and H. Watanabe (eds.), Amsterdam ; Philadelphia: John Benjamins Publishing Company, 247–82.

la Roi, Ezra. 2021. The Insubordination of If- and That-Clauses from Archaic to Post-Classical Greek: A Diachronic Constructional Typology, *Symbolae Osloenses* 95 (1), 2–64.

———. 2022. Insubordination in Archaic and Classical Latin: Commands, Requests, Wishes and Assertives, *Journal of Latin Linguistics* 21 (1), 23–45.

———. forthc.a. Insubordination, in *Encyclopedia of Greek Language and Linguistics*, G. Giannakis, E. Crespo, P. Filos, B. D. Joseph, and T. Markopoulos (eds.). Leiden: Brill.

———. forthc.b. The impact of insubordinate (ἵ)να on the history of Greek: cross-linguistic comparability, morphosyntactic innovation and pragmatic spread in the mood system, *Journal of Greek linguistics.*

# Insubordination and swearing in Akkadian (and beyond)

Rodrigo Hernáiz
(The Open University)

Keywords: insubordination, complementation, conditional clauses, Semitic, Akkadian

Oath formulas preserved in Akkadian written documents present special morphosyntactic properties (Von Soden 1995: 292) which set them apart from 'ordinary statements' (Huehnergard 2005: 436). Two frequently attested illocutionary expressions of oaths —and other 'oath-like assurances' (Kouwenberg 2017: 734)— exhibit an unequivocal syntax whereby seemingly *independent* clauses

a) appear as protases without their apodoses:
   (1) Akkadian (Afro-Asiatic, Semitic; Huehnergard 2005: 438).
   ***šumma*** *ṣibût-ki* *lā* *ē-te-puš*
   **if** wish-your.F NEG 1SG-PERF-do
   lit. "if I have not carried out your wish" (translated as: "I will carry out your wish." [content of an oath, with no apodosis]).

b) bear subordinative marking (SUBR) otherwise restricted to subordinate clauses:
   (2) Akkadian (Afro-Asiatic, Semitic; Huehnergard 2005: 437).

| *kiam* | *l-izkur-ū* | *dā'ik* | *PN* | *lā* | *ī-d-**û*** |
|---|---|---|---|---|---|
| thus | PREC-swear-M.PL | murderer | PN | NEG | 1SG-know-**SUBR** |

   "So they must swear: I do not know the murderer of (PN)".

Although the morphosyntactic characteristics of oath/swearing formulations in Akkadian have attracted the attention of scholars for more than a century (Sjörs 2021), they remain to be fully understood. This paper will assess their characteristics from the perspective of insubordination and extend this scope to other Semitic languages.

Akkadian *conditional* oaths (type *a,* above) were recently mentioned in passing in a footnote by Sjörs (2021) as 'presumably' the result of insubordination. Previous literature on Akkadian grammar —predating the inception of the concept of insubordination (Evans 2007) — assume indeed the ellipsis of matrix clauses (apodoses) in these formulations (cf. Von Soden 1995: 293). In contrast to more widely known instances of insubordinate *if*-clauses —as in contemporary English (Lastres-López 2020) — this syntactic reduction has been associated in Akkadian scholarship with taboo conventions, by which an apodosis with negative connotations is left unexpressed: "if I do not X [*I will be cursed*]".

The second type of independent clause in Akkadian oath formulations (type *b,* above) has not been previously described as insubordination. However, it will be shown that recent observations on Akkadian clause combining (Hernáiz 2024), and diachronic Akkadian textual evidence suggest that these independent clauses with subordinate marking might also be seen as a type of insubordination, resulting from the conventionalised ellipsis of matrix clauses as well as complementisers.

Comparative analyses of similar constructions in Semitic languages, particularly the oath formulas in Biblical Hebrew studied by Conklin's (2011), indicate that insubordination can explain the syntactic characteristics of swearing constructions in these ancient languages, extending beyond Akkadian.

The presentation will interrogate the potential role of taboo and language contact in the inception of independent clauses with subordinate marking. It will also provide an insight into the methodological challenges for the diachronic analysis of these constructions (Cristofaro 2016), and the potential contribution of ancient Semitic oath formulations to crosslinguistic patterns of insubordination.

**References**


Conklin, Blane (2011), *Oath formulas in Biblical Hebrew*, Einsebrauns.

Cristofaro, Sonia (2016), Routes to insubordination, in N. Evans, and H. Watanabe (eds), (2016) *Insubordination*, John Benjamins, 393-422.

Evans, Nicholas (2007) Insubordination and its uses, in I. Nikolaeva (ed) *Finiteness: Theoretical and empirical foundations.* Oxford University Press.

Hernáiz, Rodrigo (2024), The grammaticalization of manner expressions into complementizers: Insights from Semitic languages, *Linguistics: An Interdisciplinary Journal of the Language Sciences.* De Gruyter. https://doi.org/10.1515/ling-2021-0197

Huehnergard, John (2005), *A grammar of Akkadian*, Eisenbrauns.

Kouwenberg, N.J.C. (2017), *A Grammar of Old Assyrian*. Brill. https://doi.org/10.1163/9789004472846

Lastres–López Cristina (2020) Subordination and insubordination in contemporary spoken English: If-clauses as a case in point. *English Today*, (2020) 36(2): 48-52. doi:10.1017/S026607841900021X

Sjörs, Ambjörn (2021), "Notes on the Use of the Subordinative in Oaths in Akkadian", in R. Hasselbach-Andee and N. Pat-El (eds), *Bēl Lišāni: Current Research in Akkadian Linguistics*, Penn State University Press, 201-218. https://doi.org/10.1515/9781646021598-009

Soden, Wolfram von (1995), *Grundriss der akkadischen Grammatik*, Editrice pontificio istituto biblico.

# Pity that semi-insubordination (SIS) is so hard to retrieve from corpora: A diachronic study of SIS in the recent history of British English

Xulia Sánchez-Rodríguez
(Universidade de Vigo)

Keywords: insubordination, semi-insubordination, diachronic variation, BNC1994, BNC2014

The linguistic strategy of insubordination illustrated in (1), which involves "the conventionalized main clause use of what, on prima facie grounds, appear to be formally subordinate clauses" (Evans 2007:367), has attracted a great deal of scholarly attention over the past years (see Mithun 2008, Verstraete et al. 2012, Van Linden and Van de Velde 2014, Sansiñena et al. 2015, Gras 2016, Beijering et al. 2019, Lastres-López 2020, Kaltenböck and Keizer 2022). However, the study of semi-insubordinate (SIS) constructions, that is, "complex sentences with 'incomplete' matrix clauses" (Kaltenböck 2021:126) such as (2), remains virtually non-existent.

(1) That I should live to see such ingratitude! (Evans 2007:403)
(2) Well, *funny you should ask, Florence*. (Kaltenböck 2021:127)

The term 'semi-insubordination' was coined in 2014 by Van Linden and Van de Velde, the only study on SIS in (American) English being Kaltenböck's (2021). According to Kaltenböck, a SIS construction in English consists of a matrix element – adjectival (in (3)), nominal (4), verbal (5) or prepositional (6) – and a (non-)finite subordinate clause.

(3) *Strange* how the Lebanon should be cover with them. (Kaltenböck 2021:133)
(4) *Pity* that wasn't filmed. (Kaltenböck 2021:133)
(5) *Seems* he chomped down on a sandwich wrap he says contained a, quote, "dangerous" substance", specifically an olive pit. (Kaltenböck 2021:129)
(6) Yes, *about time* we got out of this show. (Kaltenböck 2021:129)

This study thus aims to widen the literature on SIS in British English in recent diachrony, with the purpose of providing a comprehensive understanding of its characteristics and usage patterns. To this end, I use data from the *British National Corpus*, more specifically from BNC1994 (BNC Consortium 2007) and BNC2014 (Love et al. 2017, and Brezina et al. 2021). The data obtained after manual pruning were coded for (i) structural factors such as type of matrix element (with a focus on adjectival SIS), lexical head in the matrix element and type of subordinate clause; (ii) discourse functions, using Kaltenböck's (2021) distinction between commenting SIS (in (7)), which express the speaker's subjective evaluation of the proposition in the subordinate clause, and presentative SIS (in (8)), which comment and introduce a proposition into the discourse; and (iii) time period.

(7) Boone: Poor Ann-Margret sang so beautifully when she was on several years later, but she was beat that night by a guy playing a classical melody on a leash in his hands.
Phillips: *Funny that you should mention that Pat Boone*. (Kaltenböck 2021:147)
(8) Well, I guess he doesn't have the magic crystal. A crystal shop, full of customers too. *Funny, I never noticed this place before*. (Kaltenböck 2021:148)

The results point towards an ongoing process of functional and lexical productivity affecting SIS constructions in the recent history of British English. Functionally, the data show a

tendency towards a lower distributional differentiation among complement-clause types in SIS and an increase of complement-clause types fulfilling the prototypical commenting function. Lexically, this study reports a noticeable diachronic rise of adjective types occurring in SIS constructions.

Acknowledgements
This research has been funded by a predoctoral grant from the University of Vigo, as well as the Spanish State Research Agency and the European Regional Development Fund (grant no. PID2020-117541GB-I00).

**References**

Beijering, Karin, Gunther Kaltenböck, and María Sol Sansiñena (2019), Insubordination: Central issues and open questions, in K. Beijering, G. Kaltenböck, and M. S. Sansiñena (eds.), (2019), *Insubordination. Theoretical and empirical issues*, Berlin / Boston: De Gruyter Mouton, 1-28.

Brezina, Vaclav, Abi Hawtin, and Tony McEnery (2021), The Written British National Corpus 2014 – design and comparability, *Text & Talk* 41(5-6), 595-615.

British National Corpus Consortium (2007), *British National Corpus: XML edition*, Oxford: Oxford Text Archive.

Evans, Nicholas (2007), Insubordination and its uses, in Irina Nicolaeva (ed.), (2007), *Finiteness: Theoretical and empirical foundations*, Oxford: Oxford University Press, 366-431.

Gras, Pedro (2016), Revisiting the functional typology of insubordination: *que*-initial sentences in Spanish, in N. Evans, and H. Watanabe (eds.), (2016), *Insubordination*, Amsterdam / Philadelphia: John Benjamins, 113-144.

Kaltenböck, Gunther (2021), *Funny you should say that*. On the use of semi-insubordination in English, *Constructions and Frames* 13(1), 126-159.

Kaltenböck, Gunther, and Evelien Keizer (2022), Insubordinate *if*-clauses in FDG: Degrees of independence, *Open Linguistics* 1, 675-698.

Lastres-López, Cristina (2020), Subordination and insubordination in contemporary spoken English. *If*-clauses as a case in point, *English Today* 36(2), 48-52.

Love, Robbie, Claire Dembry, Andrew Hardie, Vaclav Brezina, and Tony McEnery (2017), The Spoken BNC2014: Designing and building a spoken corpus of everyday conversations, *International Journal of Corpus Linguistics* 22(3), 319–344.

Mithun, Marianne (2008), The extension of dependency beyond the sentence, *Language* 84(1), 69-119.

Sansiñena, María Sol, Hendrik De Smet, and Bert Cornillie (2015), Between subordinate and insubordinate. Paths toward complementizer-initial main clauses, *Journal of Pragmatics* 77, 3-19.

Van Linden, An, and Freek Van de Velde (2014), (Semi-)autonomous subordination in Dutch: Structures and semantic-pragmatic values, *Journal of Pragmatics* 60, 226-250.

Verstraete, Jean-Christophe, Sarah D'Hertefelt, and An Van Linden (2012), A typology of complement insubordination in Dutch, *Studies in Language* 36(1), 123-153.

# Insubordination from relative clause structures and resultant alignment patterns

Rebecca Paterson
(Princeton University)

Previous studies have investigated the diachronic origins of competing argument alignment systems (Cristofaro & Inglese 2021; Mithun 2005; Zúñiga 2018 inter alia), but very little is known about the origins of typologically-rare marked nominative systems (König 2008). Northwest Kainji languages (Nigeria; Benue-Congo) demonstrate how a marked nominative pattern can arise from the reanalysis of relative clause structures and accompanying obligatory noun class agreement (Paterson, submitted). The result is also a case of insubordination (Evans 2007: 367), such that a clause that bears the formal markers of a subordinate relative clause now occurs as a conventionalized main clause under specific conditions. In the NWK data explored here, the conditions of occurrence are based on the configuration of the NP subject form, rather than the verbal or clause structure.

Based on an internal reconstruction of Ụt-Ma'in [gel], I argue that speakers have reanalyzed a relative clause structure as a main clause and reanalyzed the relativizer, which agrees with the class of the head noun, as a unique subject (S=A) marker. In Ụt-Ma'in (1), the class 7 agreement marked relativizer *jə̀* occurs immediately after the head noun *tʃāmpá* 'man'. The bracketed noun phrase can be understood as the object argument of the matrix verb *hʲáŋ* 'see.PST' as in translation (a). Alternately, this could be understood as two separate finite clauses as in (b). Reanalysis involves a realignment of phrasal boundaries from an NP [*tʃāmpá* [***jə̀*** *hɛ̄ :g*]RELCL]NP to an NP VP sequence [[*tʃāmpá* ***jə̀***]NP *hɛ̄ :g*]]VP.

(1) *ə̄ m* *hʲáŋ* [*tʃāmpá* ***j-ə̀*** *hɛ̄ :g* ]
1SG.SUBJ see.PST man AG7-REL fall.PST
(a)'I saw a man who fell.'/ (b) I saw; a man fell.'

The fully-reanalyzed structure occurs in (2) where the class 7 *jə̀* occurs immediately after the head noun *kɔ́ːt* 'guinea fowl', taken from a narrative where there is no option to understand this sequence as an NP containing an embedded relative clause.

(2) *sɛ̄* [*kɔ́ːt-**jə̀***] *rwə̄ n* *ə̄ rvástɛ̀*
then guinea.fowl-C7.NOM exit last
'Then guinea fowl exited last.' (*then guinea fowl that exited last…)

The Ụt-Ma'in citation/accusative form of *kɔ́ːt* has the class 7 prefix *ū-*.

(3) *á=b* *hján* [***ū**-kɔ́ːt*]
COND=2SG see C7.ACC-guinea.fowl
'If you see a guinea fowl, …'

Three other NWK languages (ụt-Hun [uth], ụs-Saare [uss], and C'Lela [dri]) display this pattern of post nominal noun class marking on unexpanded preverbal subjects, parallel to (2), which is also the noun class flagging form following a noun modified by a relative clause (parallel 1a, with noted variation in other components of the relative clause). Elsewhere in these languages, nouns are flagged with noun class prefixes, parallel to (3).

Proposed mechanisms for the emergence of insubordinate structures include ellipsis (Evans 2007), extension (Mithun 2008), and clausal disengagement (Cristofaro 2016). Regardless of the mechanism that gave rise to this NWK pattern, the synchronic structure is obligatory. It is the only way for speakers to use an unexpanded (indefinite) noun phrase in preverbal subject position of basic clause. In this study, I explore how each of the proposed mechanisms may have led to the conventionalized synchronic structure.

**References**


Paterson, Rebecca. Submitted.

Cristofaro, Sonia. 2016. Routes to insubordination: A cross-linguistic perspective. In Nicholas Evans & Honoré Watanabe (eds.), *Insubordination* (Typological Studies in Language 115), 393–422. Amsterdam: John Benjamins Publishing Company. doi:10.1075/tsl.115.15cri.

Cristofaro, Sonia & Guglielmo Inglese. 2021. IntroducIng DEmA: the Pavia diachronic emergence of alignment database. In Erica Biagetti, Chiara Zanchi & Silvia Luraghi (eds.), *Building New Resources for Historical Linguistics*, 183–200. Pavia: Pavia University Press.

Evans, Nicholas. 2007. Insubordination and its uses. In Irina Nikolaeva (ed.), *Finiteness: theoretical and empirical foundations* (Oxford Linguistics), 366–431. Oxford; New York: Oxford University Press.

König, Christa. 2008. *Case in Africa* (Oxford Linguistics). Oxford, England; New York: Oxford University Press.

Mithun, Marianne. 2005. Beyond the Core: Typological Variation in the Identification of Participants. *International Journal of American Linguistics* 71(4). 445–472. doi:10.1086/501247.

Mithun, Marianne. 2008. The Extension of Dependency Beyond the Sentence. *Language*. Linguistic Society of America 84(1). 69–119. doi:10.1353/lan.2008.0054.

Zúñiga, Fernando. 2018. The diachrony of morphosyntactic alignment: Diachrony of Morphosyntactic Alignment. *Language and Linguistics Compass* 12(9). e12300. doi:10.1111/lnc3.12300.

# Toward a grid of domains, settings, processes, and purposes for comparing cases of (alleged or real) insubordination

Björn Wiemer
(Johannes-Gutenberg-Universität, Mainz)

Keywords: (semi-)insubordination, dependency shift, Slavic languages, reported discourse, metadiscursive justification

While Evans' (2007) original understanding of insubordination implies a diachronic perspective (with 'main clause ellipsis' as a core prerequisite), subsequent approaches also have captured insubordination as a synchronic phenomenon arising in dyadic turn-taking (e.g., Gras & Sansiñena 2015). However, the synchronic and the diachronic perspective share two basic assumptions: first, an insubordinate structure is a formerly subordinate clause as a trunc of an originally biclausal construction. Second, the insubordinate structure highlights a function that either derives from the former matrix clause, or conventionalizes functions that were only contextually implied by the biclausal construction. This new function is, as a rule, accompanied by reduced commitment (Wiemer 2017: 271-273, 2022), which operates either in the domain of cognition (e.g., weakening of epistemic support, or mirativity), of volition (e.g., directive speech acts), or of both at once (e.g., apprehension). However, assumptions about the subordinate nexus relation within the biclausal construction, used as 'reference point' for the purported insubordinate structure, happen to be based on shaky empirical and/or conceptual premises (e.g., because the subordinating function of clause-initial connectives is debatable, or no stages preceding alleged insubordination can be traced).

On this backdrop, we suggest a comprehensive approach toward insubordination *sensu stricto*, in contrast to dependency shift (according to the CfP) and semi-insubordination (Van linden & Van de Velde 2014, D'Hertefelt 2018: 179-182). For this purpose, we will follow D'Hertefelt (2018), Mithun (2019), la Roi (2021), but complement their insights on the basis of a systematic account of pertinent cases across Slavic languages. We will elaborate on a grid that correlates (i) notional domains and (ii) structural "settings" of clause combining with (iii) micro-processes which yield new constructions (iv) characterized by particular ranges of (pragmatic) functions. This grid may (a) further serve as a heuristic for systematic checks on pathways based on insubordination or related, but different mechanisms (see the CfP), (b) provide guidelines for language-specific in-depth studies of assumed cases of insubordination (which "inputs" have led to insubordination and which ones have not?, do the involved micro-processes differ?, to which extent do the resulting function ranges overlap?), and (c) allow for more adequate cross-linguistic comparisons and the detection of areal or genealogical patterns (in particular, Slavic vs Germanic).

The survey of cases from Slavic languages primarily encompasses (i) the employment of would-be complementizers for the purposes of quotation (Guz 2019, Rojszczak-Robińska 2021, Letučij 2023) or information packaging in reported discourse, (ii) apprehensional and related optative clause types (Wiemer, forthcoming), (iii) metadiscursive justification (Kawka 1988, Mendoza 2003), (iv) nonverbal predicates ('predicatives') and their clausal arguments (Wiemer 2019). The survey reveals that, for the Slavic group, insubordination *sensu stricto* seldomly applies, whereas dependency shifts are easy to identify. We will propose an explanation for this finding.

Furthermore, we will also try to answer whether the rise of insubordination and of subordination (in complex sentences) can be seen as results of converse processes.

The talk will thus contribute to the first two aims of the workshop as formulated in the CfP.

**References**


D'Hertefelt, Sarah (2018), *Insubordination in Germanic (A Typology of Complement and Conditional Constructions)*, Berlin / Boston: De Gruyter Mouton.

Evans, Nicholas (2007), Insubordination and its uses, in I. Nikolaeva (ed.), (2007), *Finiteness: Theoretical and Empirical Foundations*, Oxford etc.: Oxford University Press, 366–431.

Gras, Pedro and María Sol Sansiñena (2015), An interactional account of discourse connective *que*-constructions in Spanish, *Text & Talk* 35(4), 505–529.

Guz, Wojciech (2019), *Quotative uses of Polish* że, Lublin: Katolicki Uniwersytet Lubelski.

Kawka, Maciej (1988), *Bo* metatekstowe [Metatextual *bo*], *Polonica* 13, 65–83.

la Roi, Ezra (2021), The Insubordination of *If*- and *That*-Clauses from Archaic to Post-Classical Greek: A Diachronic Constructional Typology, S*ymbolae Osloenses* 95, 2–64.

Letučij, Aleksandr B. (2023), Osobennosti razgovornogo upotreblenija sojuza *čto*: citativnost' i drugie nestandartnye svojstva [Peculiarities of the colloquial use of the conjunction *čto*: quotativity and other non-standard properties], *Voprosy jazykoznanija* 2023-3, 27–59.

Mendoza, Imke (2003), „Pragmatische Begründungen" im Polnischen, in L. Geist and R. Hammel (eds), (2003), *Linguistische Beiträge zur Slavistik* 10, Munich: Sagner, 174–191.

Mithun, Marianne (2019), Sources and mechanisms, in K. Beijering, G. Kaltenböck and M. S. Sansiñena (eds), (2019), *Insubordination. Theoretical and empirical issues*, Amsterdam / Philadelphia: John Benjamins, 29–54.

Rojszczak-Robińska, Dorota (2021), Mowa niezależna wprowadzana przez *iże* w staropolskich apokryfach. Perspektywa źródłowa [Direct speech introduced by *iże* in Old Polish apocryphs. A look at the sources], *Biuletyn Polskiego Towarzystwa Językoznawczego* 77, 227–243.

Van linden, An and Freek Van de Velde (2014), (Semi-)autonomous subordination in Dutch: Structures and semantic-pragmatic values, *Journal of Pragmatics* 60, 226–250.

Wiemer, Björn (2017), Main clause infinitival predicates and their equivalents in Slavic: Why they are not instances of insubordination, in Ł. Jędrzejowski and U. Demske (eds), (2017), *Infinitives at the Syntax-Semantics Interface: A Diachronic Perspective*. Berlin / Boston: De Gruyter Mouton, 265–338.

Wiemer, Björn (2019), On illusory insubordination and semi-insubordination in Slavic: Independent infinitives, clause-initial particles and predicatives put to the test, in K. Beijering, G. Kaltenböck and M. S. Sansiñena (eds), (2019), *Insubordination. Theoretical and empirical issues*, Amsterdam / Philadelphia: John Benjamins, 107–166.

Wiemer, Björn (2022), Insubordination, in M. L. Greenberg (ed.), *Encyclopedia of Slavic Languages and Linguistics*, Leiden: Brill. http://dx.doi.org/10.1163/2589-6229_ESLO_COM_034742

Wiemer, Björn (forthcoming), Two major apprehensional strategies in Slavic: a survey of their areal and grammatical distribution, in M. Faller, E. Schultze-Berndt and M. Vuillermet (eds), *Apprehensional constructions (in a cross-linguistic perspective)*, Berlin: Language Science Press.

# Causal relations in Ancient Greek: Problematizing insubordination in historical corpus data

Giuseppina di Bartolo
(University of Cologne; University of Pavia)

Keywords: historical discourse analysis, clause combining, causal clauses, insubordination, dialogic corpus data

Recent linguistic research has shown that historical data can contribute to the study of insubordination and to a better understanding of the diachrony of this phenomenon (among others, Ruiz Yamuza 2020, and la Roi 2021 for Ancient Greek).

This paper aims to contribute to the typological research on insubordination by addressing data from Ancient Greek (AG). In this respect, it focuses on finite causal clauses introduced by the conjunction *epeí* (meaning both 'when', 'since' and 'because') when expressing causal relations. In addition to bi-clausal constructions (i.e., matrix and subordinate *epeí*-clause), AG reference grammars (among others, van Emde Boas *et al*. 2019: 549) account for instances of 'independent' *epeí*-clauses. In other words, *epeí*-clauses that do not show syntactic integration. They usually provide the motivation for the preceding utterance and are not interpretable without the preceding context (e.g., 1 and 2).

(1) *mētròs* *tád'* *hēmîn* *ekphéreis* *zētḗmata?*
mother:GEN.SG DEM.ACC.PL.N 2PL.DAT bring.out:2SG.IND.PRS mean:ACC.PL
*epeí* *g'* *ho* *daímōn* *boúletai*
because PTCL ART god:NOM want:3SG.IND.PRS (PTCL= particle)
Ion: 'Are you laying out the means to find my mother here?'
Old servant: 'Yes! Because the god wants it.' (Eur. *Ion* 1352–3)

(2) *eísō* *komízou* *kaì* *sú* *Kassándran* *légō*
inside get:2SG.IMP.PRS and you:NOM.SG Cassandra:ACC.SG say:1SG.IND.PRS
*epeí* *s'* *éthēke* *Zeùs* *amēnítōs*
because you:ACC.SG set:3SG.IND.AOR Zeus:NOM not.angry:NOM.M
*dómois* *koinōnòn* *eînai* *kherníbōn*
house:DAT.M.PL common:ACC.SG be:INF.PRS vessel.for.water:GEN.N.PL
'Get inside! I'm talking to you, Cassandra! Because Zeus, not angry, placed you in this household to share the holy water.' (Aesch. *Ag*. 1035–1039)

The paper applies the broader categorization of adverbial clauses taken from language typology (among other, Thompson *et al*. 2007) and addresses issues of syntactic and pragmatic dependency by considering whether the instances of 'independent' *epeí*-clauses should be considered a case of insubordination. My corpus consists of dialogic texts from Classical Greek (5th–4th cent. BCE), including data from Aeschylus, Sophocles and Euripides. The data are extracted through the *Thesaurus Linguae Graecae* (https://stephanus.tlg.uci.edu) and manually annotated. This corpus provides the possibility to address the occurrence of 'cross-speaker dependency' (Evans and Watanabe 2016: 5, 27–28), and to verify whether the mechanisms of 'extension of dependency' and 'clausal disengagement' respectively underlie the formation of these types of clauses (Mithun 2008, 2019; Cristofaro 2016). Moreover, the paper addresses the methodological challenges of analyzing insubordination in historical corpus data (specifically in dialogic contexts in the absence of spoken data), and additionally considers the data in light of the recent problematization of insubordination by Couper-Kuhlen and Thompson (2022).

Furthermore, the paper provides a diachronic analysis of the use of *epeí* in AG, taking into account, in addition to the corpus under investigation, the data previously collected and analyzed by Muchnova (2011), which includes the *epeí* occurrences from Homer and Xenophon.

Finally, the paper draws cross-linguistic comparisons with analogous cases of causal clauses found in other modern Indo-European languages, such as English *because*-clauses (among others, Verstraete 2007; Kaltenböck 2019) and German *weil*-clauses (among others, Günthner 1996).

Acknowledgments: This research has been conducted with the support of the Alexander von Humboldt Foundation.

**References**

Couper-Kuhlen, Elisabeth and Sandra Thompson (2022), Can temporal clauses be insubordinate?, *Interactional Linguistics* 2(2), 165–189.

Cristofaro, Sonia (2016), Routes to insubordination, in N. Evans, and H. Watanabe (eds) (2016), *Insubordination,* Amsterdam / Philadelphia: John Benjamins, 393–422.

van Emde Boas, Evert, Albert Rijksbaron, Luuk Huitink and Mathieu de Bakker (2019), *The Cambridge Grammar of Classical Greek,* Cambridge: Cambridge University Press.

Evans, Nicholas and Honoré Watanabe (2016), 'The dynamics of insubordination', in N. Evans and H. Watanabe (eds) (2016), *Insubordination,* Amsterdam / Philadelphia: John Benjamins, 1–37.

Günthner, Susanne (1996), From subordination to coordination? Verb-second position in German causal and concessive constructions, *Pragmatics* 6(3), 323–356.

Kaltenböck, Gunther (2019), Delimiting the class: A typology of English insubordination, in K. Beijering, G. Kaltenböck and M. S. Sansiñena (eds) (2019), *Insubordination*, Berlin / Boston: De Gruyter, 167–198.

la Roi, Ezra (2021), The insubordination of if- and that-clauses from archaic to post-classical Greek: a diachronic constructional typology, *Symbolae Osloenses* 95(1), 1–63.

Mithun, Marianne (2008), The extension of dependency beyond the sentence, *Language* 84(1), 264– 280.

Mithun, Marianne (2019), Sources and mechanisms, in K. Beijering, G. Kaltenböck and M. S. Sansiñena (eds) (2019), *Insubordination*, Berlin / Boston: De Gruyter, 29–54.

Muchnova**,** Dagmar (2011), *Entre conjonction, connecteur et particule: Le cas de ἐπεί en grec ancien,* Prague: Karolinum.

Ruiz Yamuza, Emilia (2020), Insubordination in Ancient Greek? The case of ὥστε sentences, in M. Leiwo, S. Dahlgren and M. Vierros (eds) (2020), *Papers on Ancient Greek Linguistics: Proceedings of the Ninth International Colloquium on Ancient Greek Linguistics (ICAGL 9)*, Helsinki: Societas Scientiarum Fennica, 383–402.

Thompson, Sandra A., Rovbert E. Longacre and Shin J. Hwang (2007), Adverbial clauses, in T. Shopen (ed) (2007), *Language Typology and Syntactic Description* (Vol. II: Complex Constructions), Cambridge: Cambridge University Press, 237–300.

Verstraete, Jean-Christophe (2007), *Rethinking the coordinate-subordinate dichotomy*, Berlin / Boston: De Gruyter.

# Chasing trajectories of insubordination

Marianne Mithun
(University of California, Santa Barbara)

Keywords: co-optation, dependency extension, insubordination, matrix ellipsis, mechanisms, sources

Recent work has greatly enriched our appreciation of the role of insubordination, defined broadly, in shaping language. But especially since Evans (2007), it has been clear that insubordination is not a unified phenomenon (Malčukov 2013, Evans and Watanabe 2016, Cristofaro 2016, Gras 2016, Beijering, Kaltenböck, and Sansina 2019, Mithun 2019, etc.). It can develop from a variety of sources, via a variety of mechanisms, to serve a variety of functions. Sources include several types of dependent clauses, which vary in their bonds to main clauses and their formal signals of dependency. Mechanisms include the matrix clause ellipsis of Evans, dependency extension (Mithun 2008), co-optation (Heine et al. 2016), hypoanaysis (Van linden and Van de Velde 2014), dependency shift (D'Hertefelt and Verstraete 2014), and Cristofaro's (2016) clausal disengagement (Beijering et al 2019: 17). Functions include interpersonal control, modality, text structuring, speaker-hearer interaction management, and more.

An ongoing issue is how the sources and mechanisms shape functions. Links have been observed between matrix ellipsis and interpersonal control and modality, and between dependency extension and text structuring. But correspondences are not one-to-one, and not always unilinear. One source can result in multiple functions, and one function can develop from multiple sources. Additional factors may also affect developments. Such scenarios are exemplified here with conversational data from Mohawk, Central Alaskan Yup'ik, and Central Pomo.

A trajectory similar to that described by Evans, originating in complement constructions with various matrix clauses, apparently occurred in Mohawk. Insubordinate constructions deemed ungrammatical in isolation by speakers have developed from irrealis complement clauses via matrix ellipsis. They now serve a variety of functions, signaling respect, dynamic, deontic or epistemic modality, and more.

Mohawk courteous request: Watshenní:ne' Sawyer, speaker

| | |
|---|---|
| ***A**ontahskhawíhten'* | *iehnekarihsià:tha'?* |
| **aa**-onta-hsk-hawi-ht-en-' | ie-hnek-a-rihsi-a't-ha' |
| **IRR**-CSL-2SG>1SG-carry-CAUS-BEN.APPL-PFV | INDEF.AGT-liquid-LK-cut-INS-HAB |

'**Would** you bring me a lemon?'

Mohawk deontic modality: Margaret Hill, speaker (to a child)

*T**a**hsateweièn:ton'.*
t-**aa**-hs-ate-weien'ton-'
DV-**IRR**-2SG.AGT-MID-know.how-PFV
'You **should** do your best.'

But respect constructions can also develop from other sources via other mechanisms. In Yup'ik they originated in adverbial clauses marked by subordinative mood inflection descended from nominalizers, and evolved via dependency extension.

Central Alaskan Yup'ik, Elena Charles and George Charles, speakers

GC *Tua-llu-qaa* *taukut* *aqva**lu**ci?*
tua=llu=qaa tauku-t aqva-**lu**-ci
then=too=Q that.RESTR-PL fetch-**SUB**-R>2PL
'Then com**ing** to get you too?' = 'So then did they come to get you?'

EC *Aqva**lu**ta-llu.*
aqva-**lu**-ta=llu
fetch-**SUB**-R>1Pl=too
'Com**ing** to get us too' = 'They did come to get us.'

Factors beyond sources and mechanisms can also shape the development of insubordinate constructions. The interpretations of both the Mohawk and the Yup'ik insubordinate constructions depend on context.

Other constructions available in a language might hinder development. Central Pomo contains a set of adverbial clause constructions distinguishing realis from irrealis, and simultaneous from sequential situations. On occasion, these formally dependent clauses are separated across turns: co-construction.

Central Pomo: Frances Jack, Florence Paoli, speakers

FJ ['There to the east, they say, they had a lot of acorns growing there. Early crop of clover growing too. Many different things growing. The Hopland people would camp around there.']
*Dú:du* *ma:* *hlátč'îc'iw.*
*dú:du* *ma:* *hlá-t-č'îc'iw.*
different place go.PL-MULTIPLE.AGENCY-FREQ
'They would go around to different places.'

FP *Só* *ʔá'č'i**n***
só ʔaʔ-č'i-**n**
clover gather-IPFV.PL-**SAME.SIM**
'Gather**ing** clover,'
*p'dú* *ʔáʔč'i**n**.*
p'dú ʔáʔ-č'i-**n**
acorn gather-IPFV.PL-**SAME.SIMULTANEOUS**
'Gather**ing** acorns.'

Yet insubordination is strikingly rare in Central Pomo. A contributing factor could be the pervasiveness of another construction. Ideas may be expressed in separate sentences, the first comparable to an adverbial

clause, the second linked to it semantically by a particle formed from the verb *ʔi-* plus any of the adverbial dependency markers, much like the semi-insubordination of Van linden and Van de Velde.

Central Pomo: Frances Jack, speaker

['You spread a canvas sack on top so the dirt won't fall inside there.']

*Mé:n* ***ʔíba***

me:n ʔi-**ba**

so be-**SAME.SEQUENTIAL**

'(So being then)

*qʰabé ʔel wná: mčá-m*

rock ART on.top set.multiple on

you put the rocks on top . . .'

We still have much to learn about the ranges of origins, trajectories, mechanisms, and external factors shaping insubordination constructions. More information about more languages is showing us that all are less categorical and more nuanced than we first imagined.

**References**


Beijering, Karin, Gunther Kaltenböck, and María Sol Sansiñena (2019). Insubordination: Central issues and open questions, in *Insubordination*: *Theoretical and empirical issues*, K. Beijering, G. Kaltenböck, and M. Sol Sansiñena (eds), Berlin: De Gruyter Mouton, 1-28.

Cristofaro, Sonia (2016). Routes to insubordination: A cross-linguistic perspective. in N. Evans and H. Watanabe (2016), *Insubordination*. Amsterdam: John Benjamins, 393-422.

D'Hertefelt, Sarah and Jean-Christophe Verstraete (2014). Independent complement constructions in Swedish and Daish: Insubordination or dependency shift? *Journal of Pragmatics* 60.89-102.

Evans, Nicholas (2007), Insubordination and its uses, in I. Nikolaeva (ed), (2007), *Finiteness: Theoretical and Empirical Foundations*, Oxford: Oxford University Press, 366-431.

Evans, Nicholas and Honoré Watanabe (2016). The dynamics of insubordination: An overview, in N. Evans and H. Watanabe (2016), *Insubordination*. Amsterdam: John Benjamins, 1-37.

Gras, Pedro (2016). Revisiting the functional typology of insubordination: Insubordinate *que*-constructions in Spanish, in N. Evans and H. Watanabe (2016), *Insubordination*. Amsterdam: John Benjamins, 113-143.

Heine, Bernd, Gunther Kaltenböck, and Bert Cornillie (2016). On insubordination and cooptation, in N. Evans and H. Watanabe (2016), John Benjamins, 65-87.

Malčukov, Andrej (2013), Verbalization and insubordination in Siberian languages, in M. Robbeets and H. Cuyckens (eds.), *Shared Grammaticalization: With special focus on the Transeurasian languages*, Masterdam: John Benjamins, 177-208.

Mithun, Marianne (2008), The extension of dependency beyond the sentence. *Language* 81(1), 69-119.

Mithun, Marianne (2019), Sources and mechanisms, in *Insubordination*: *Theoretical and empirical issues*, K. Beijering, G. Kaltenböck, and M. Sol Sansina (eds), Berlin: De Gruyter Mouton, 29-54.

Van linden, An and Freek Van de Velde (2014). (Semi-)autonomous subordination in Dutch: Structures and semantic-pragmatic values. *Journal of Pragmatics* 60. 226-250.

# WS15 Poetic language

# Measuring formularity through corpus linguistics methods: A co-varying collexeme analysis of Rigvedic similes

Erica Biagetti
(University of Pavia)

Keywords: Oral-formulaic language, Vedic Sanskrit, Construction Grammar, Corpus Linguistics, Collexeme Analysis

It has long been acknowledged in the literature that the building blocks of oral poetics consist of both fixed formulas and more schematic expressions (formulaic systems in Parry 1971[1928] and Lord 1960). Adopting a Construction Grammar (Fillmore and Kay 1993, Goldberg 1995) approach that equates formulas with constructions, i.e. learned pairings of form and function, Bozzone (2014) was able to account for different levels of flexibility that formulas can display, ranging from fixed to schematic expressions (on similar proposals, cf. Frog 2014, Antović & Pagán Cánovas 2016).

The constructionist approach to the study of formulas seems suited to the formulaic style of the *Rigveda*, a collection of religious hymns composed in Vedic Sanskrit that constitutes the oldest layer of Indic literature. While the form of the hymns depends on tradition, Vedic poets emphasize the novelty of their works, and this tension between tradition and innovation brings about continuous variations in the expression of conventional themes (Elizarenkova 1995: 22-24).

This paper tests the constructionist approach to the study of Rigvedic oral-formulaic language by presenting a co-varying collexeme analysis (CA; Stefanowitsch & Gries 2003, 2005) of Rigvedic similes. Co-varying CA measures the association between elements occurring in two slots of a macro-construction; such an association constitutes a clue for the existence of a more substantive meso-construction involving the elements in question. Rigvedic similes introduced by the comparative markers *ná*, *iva*, and *yáthā* 'like' can be regarded as a macro-construction instantiated by more fixed meso-constructions, as several types of similes can be identified based on the semantics of their parameter and standard of comparison. For instance, the comparison 'swell like a river' is encoded by different verbs for 'swelling, spreading, growing' (*prath-*, *pinv-*, *pyā-*, *vṛdh-*, *uruṣay-*) and different lexemes for 'river, stream, water' (*áp-*, *avatá-*, *nadī́-*, *síndhu-*, *vā́r-*, etc.).

Taking formulas as variably fixed constructions, CA allows isolating formulaic similes that fall at different points along the continuum between substantive and schematic.

The analysis is carried out by querying morpho-syntactic information from the Rigvedic portion of the Vedic Treebank (Hellwig et al. 2020), in which all similes have been manually annotated, and semantic information from the Sanskrit WordNet (Biagetti et al. 2021). The results point to the existence of a structured inventory of constructions in which the more substantive similes inherit features from the more schematic ones. We can postulate that such an inventory formed the poet's repertoire and allowed him to produce novel comparisons while at the same time remaining within the limits imposed by tradition.

The method employed in this paper is promising both from an empirical perspective, as it allows for the automatic extraction of variably substantive formulas, and from a cognitive perspective, as it enables the formulation of hypotheses about which entities were stored in the poet's mind and, therefore, about the mechanisms that regulated oral poetic composition in ancient India. The method can be applied to other Rigvedic constructions and to texts belonging to different poetic traditions, thereby providing a tool for comparative studies.

Acknowledgments

Research for this paper and for the creation of the Sanskrit WordNet 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)” project 2022YAPFNJ “Linked WordNets for Ancient Indo-European Languages”; CUP F53D2300490 000.

**References**

Biagetti, Erica, Chiara Zanchi & William Michael Short (2021), Toward the creation of WordNets for ancient Indo-European languages. In *Proceedings of the 11th Global Wordnet Conference*, 258-266.

Bozzone, Chiara (2014), *Constructions: A New Approach to Formularity, Discourse, and Syntax in Homer*. PhD Thesis, UCLA.

Elizarenkova, Tatyana Y. (1995), *Language and Style of the Vedic Ṛṣis*. Albany: State University of New York Press.

Fillmore, Charles J. and Paul Kay (1993), *Construction grammar coursebook*. Berkeley: University of California.

Frog (2014), Oral Poetry as Language Practice: A Perspective on Old Norse *dróttkvætt* Composition. In: Pekka Huttu-Hiltunen et al. (eds.), *Song and Emergent Poetics – Laulu ja runo – Песня и видоизменяющаяся поэтика*, 279-307. Kuhmo: Juminkeko.

Goldberg, Adele E. (1995), *Constructions: A construction grammar approach to argument structure*. University of Chicago Press.

Hellwig, Oliver, Salvatore Scarlata, Elia Ackermann, and Paul Widmer (2020), The Treebank of Vedic Sanskrit. In Nicoletta Calzolari, Frederic Bechet, Philippe Blache, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi et al. (eds.), *Proceedings of The 12th Language Resources and Evaluation Conference* (LREC 2020), 5137–5146.

Lord, Albert B. (1960), *The Singer of Tales*. Cambrigde, MA, Harvard University Press.

Pagán Cánovas, Cristóbal and Mihalo Antović (2016), Construction grammar and oral formulaic theory. In Cristóbal Pagán Cánovas and Mihalo Antović (eds.), Oral Poetics and Cognitive Science, Mouton de Gruyter. Berlin: De Gruyter, 79–98.

Parry, Milman (1971[1928]), The Traditional Epithet in Homer. In Adam Parry (ed.), *The Making of Homeric Verse: The Collected Papers of Milman Parry*. Oxford, Oxford University Press, pp. 1–190.

Pinault, Georges. 1997. Distribution des particules comparatives dans la Rik-Samhitâ, *Bulletin d’Études Indiennes* 13-14, 307–367.

Stefanowitsch, Anatol & Stefan T. Gries (2003), Collostructions: Investigating the interaction of words and constructions. *International journal of corpus linguistics* 8.2: 209–243.

Stefanowitsch, Anatol & Stefan T. Gries (2005), Covarying collexemes. *Corpus Linguistics and Linguistic Theory*, 1(1), 1–43.

# Fluency-misattribution in verse comprehension: Evidence from Dutch rhyming

Stefan Blohm, Jeroen Dera and Roel Willems
(Leibniz-Institute for the German Language, Radboud Universiteit & Netherlands Institute for Social Research)

Rhyme is a widespread euphonic ornament of verbal art and song whose cognitive and stylistic effects remain insufficiently understood. Perhaps counter-intuitively, rhyme's sound correspondence may have quasi-semantic effects during verse comprehension, e.g., making statements appear more accurate and convincing. Such rhyme-as-reason effects have been related to processing events during incremental comprehension. Specifically, it has been argued that "rhyme […] affords statements an enhancement in processing fluency that can be misattributed to heightened conviction about their truthfulness" (McGlone & Tofighbakhsh, 2000); similar claims have been made in fluency-based aesthetics (e.g., Graf & Landwehr, 2015). In this paper, we report evidence from a study of verse comprehension in Dutch that combined eye-tracking during reading with intuitive judgments to test hypotheses derived from fluency-misattribution accounts.

Hypotheses:
H1) Rhyme-as-reason effect: Rhyme enhances the perceived meaningfulness / aesthetic appeal of verse .
H2) Fluency effect: Rhyme facilitates word processing (e.g., Obermeier et al., 2016)
H3) Fluency-misattribution: Rhyme-induced fluency modulations predict rhyme's stylistic effects.

Method:
We selected 48 rhymed couplets of Dutch verse and created non-rhyming versions by replacing the first rhyme word with a synonym as in (1):

(1) wat niemand kan **weten/kennen** || kan ik niet **meten**
what nobody can know can I not measure
'I cannot measure what nobody can know'

Participants (n=54) read each couplet in either the original or the modified version while their eye movements were recorded and rated the aesthetic appeal, comprehensibility, and perceived meaningfulness of each text on a quasi- continuous scale (0-100). We used linear mixed-effects regression to test how rhyme affects ratings as well as several gaze-time measures (duration of first fixation, first-pass reading time, total reading time) for the second rhyme word; item-level linear regression allowed us to test whether rhyme's fluency effects predict its quasi-semantic and aesthetic effects.

Results and discussion:
Participants judged rhyming couplets as more meaningful, comprehensible and appealing than non-rhyming versions (H1), in line with numerous earlier results. However, the analyzed gaze-time measures failed to reveal a general fluency- enhancing effect of rhyme (H2), challenging a standard assumption about rhyme cognition. In fact, we observed detrimental effects of rhyme on early word processing (=first-fixation durations) at short distances between rhyme words—a pattern that is familiar from sentence comprehension (Frisson, Koole, Hughes, Olson, & Wheeldon, 2014), and that identifies rhyme distance as a relevant variable in verse comprehension. Crucially, we observed that rhyme-induced differences in total reading times of critical rhyme words indeed partly accounted for the quasi-semantic and aesthetic effects of rhyme. Seemingly supporting a key claim of fluency-

misattribution accounts (H3), this finding questions the notion that rhyme-as-reason effects reflect perceptual fluency during early word processing. Taken together, our results shed new light on rhyme's effects during reading and on the cognitive processes that underlie its quasi-semantic and aesthetic effects.

**References:**


Frisson, S., Koole, H., Hughes, L., Olson, A., & Wheeldon, L. (2014). Competition between orthographically and phonologically similar words during sentence reading: Evidence from eye movements. Journal of Memory and Language, 73, 148-173. doi: 10.1016/j.jml.2014.03.004

Graf, L. K., & Landwehr, J. R. (2015). A Dual-Process Perspective on Fluency-Based Aesthetics: The Pleasure-Interest Model of Aesthetic Liking. Personality and Social Psychology Review, 19(4), 395-410. doi:10.1177/1088868315574978

McGlone, M. S., & Tofighbakhsh, J. (2000). Birds of a feather flock conjointly (?): Rhyme as reason in aphorisms. Psychological Science, 11(5), 424–428. doi:10.1111/1467-9280.00282

Obermeier, C., Kotz, S. A., Jessen, S., Raettig, T., von Koppenfels, M., & Menninghaus, W. (2016). Aesthetic appreciation of poetry correlates with ease of processing in event-related potentials. Cognitive, Affective, & Behavioral Neuroscience, 16(2), 362-373. doi:10.3758/s13415-015-0396-x

# What do the ritual Afro-Baloch songs tell us about their origin?

Maryam Nourzaei
(Uppsala University)

Key words: African descendant, Balochi, orality, women, identity.

The aim of the presentation is to study the contents, forms and prosodic features of the ritual *Mamaby* songs among the Afro-Baloch (AfB) communities along the coast in Sistan and Balochistan province of Islamic Republic of Iran. The term AfB refers to the people of African descendants who brought them from Eastern Africa to Balochitan. They lost their original language in favour of Balochi. Nourzaei (2023) reports that it is only their practice of unique traditions that distinguishes them from other regional groups The term *Mamabies* (own coinage) refers to songs sung by women specifically for a pregnant woman during pregnancy, delivery and postpartum.

The data for the present talk stems from 75 songs recorded in natural setting in (Nourzaei's digital corpus accepted). The ritual songs are short with a lot of repetitions, and are sung only by women, as a group consists of one main singer and seven ordinary women.

The data demonstrates that the themes and motifs of these ritual songs have completely merged with Balochi counterparts. So far, in only one recorded song, performed by an old singer, is the name Zanzibar mentioned and including presumably two African words. Regarding their prosodic features, the significant commonality is their meter, this means that all the songs have a meter as in ex.1

Singer (S): *o kapūt wāšš nālagẽ morg mohammada rāzī kano byā*
*Chorus (C)*: *mohammade rāzī hamešẽ o mādare hakkā šenāset*

S: Oh dove, sweet voiced bird, satisfy [Prophet] Mohammad and come back.
*C*: To satisfy [Prophet] Mohammad (lit. Mohammad's satisfaction) is for (lit. and) you to recognize the mother's moral obligation.

The data further suggests that 15% of the songs contains a Refrain 'radif' as in *atay* in Ex.2. The attested Refrain forms are the same verbs, nouns, and a clause.

S: *hazrat belāl hazrat belāl to may maǰede bāngū* ***atay***
C: *may maǰede bāngū atay mesle čerāga rōk* ***atay***

S: Prophet Bilal, Prophet Bilal you were our mosque's muezzin.
C: You were our mosque's muezzin; you were as glorious as the light.

and 5% possess a rhyme as ***ār*** in Ex.3. The attested *rhyme* are '*ār*', *a, īyā, yā, ān*.

*S: may donyāe bāgo bah**ār**en mohammad*
*C:watī omate gamgos**ār**en mohammad*

S: [Prophet] Mohammad is our world's garden and spring.
C: [Prophet] Mohammad is the comfort of his own nation.

In the corpus, just one song contains both Refrain '***nemāz'*** and rhyme '*en*' as in Ex.4

S: *nemāz bewān ke dozahe dēmpān**en nemāz***
C: *dardo dōrānī marhamo darmān**en nemāz***

S: Pray! Because prayer is the protector from Hell
C: Prayer is the healing and cure of pains and illnesses.

The paper concludes that the content and themes of these songs have completely been merged with the Balochi. However, they have preserved their African melodies. Refrain and rhyme are new developments in their songs. In addition, the practice of performing these ritual songs along with their healing traditions has been preserved as an important part of the AfB identity, although their original language has been lost.

Acknowledgements: I am grateful to the Swedish Riksbankens Jubileumsfond (RJ) for their generous funding of this research (grant number: P20–0076). I wish to thank my main singers and Afro-Baloch and Baloch communities for supporting me during my field studies.

**References**

Nourzaei. Maryam (2024) "Afro-Baloch communities in modern Iran and their healing traditions" in McNeer, Kevin and Andreeva, Elena (eds.), Slavery in the Middle East, North Africa, published by Bloomsbury as part of its series: Sex, Marriage, and the Family in the Middle East.

Nourzaei. Maryam, (accepted) *A digital corpus of ritual Mamabies songs among the Afro-Baloch communities in Iran.*

Nourzaei. Maryam, (in prep) *Mamabies tradition among the Afro-Baloch communities and beyond.*

# Rhythmic reading of conventional poetry: Investigating a cognitive phenomenon

Judith Beck & Lars Konieczny
(University of Freiburg & University of Freiburg)

Keywords: poetry, meter, top-down, bottom-up, musicality

When investigating poetry, aspects of the production, reception and recitation level need to be considered (Blohm et al. 2021, Peer & Chesnokova 2023). One is the cognitive phenomenon of meter. Readers seem to realize a poetic rhythm from a 'silent' text by inferring a beat from the "rhythmic gestalt" and projecting a meter, a process elicited by the rhythm patterns derived from the phonemic quality of the composed sequence of syllables. Therefore, the main interest in our experiment (Beck & Konieczny 2023) was how top-down and bottom-up processes would interact when reading conventional poetry, i.e. applying a leading metrical grid while processing syllabic material.

We collected audio recordings of subjects' (n = 13) reading poetry aloud. These were analysed using PRAAT and a mixed effects regression model approach for a data set comprising on average of 981 syllables and 120 lines per participant. Stress and musical activity were coded as predictor variables. Poems were manipulated by replacing regular syllables at random positions with the syllable "tack". Syllable onset interval (SOI) and mean intensity were used as operationalization of how strongly a syllable was stressed.

Results show that articulation duration (SOIs) for strong syllables was on average longer (41.08 ms; $p < 0.001$) than for weak ones, however, the effect disappeared for "tacks". Musically inactive readers' SOIs for 'weak' tacks were longer (19.07 ms, $|d| = 0.51$). Syllable intensities captured metrical stress of "tacks" (strong: 1.32 dB louder; $p < 0.001$), but only for musically active participants. Additional normalized pairwise variability index (nPVI)-results indicated that for SOIs, lines were read less altering. This effect was proportional to the number of tacks per line, i.e. for each tack more the nPVI decreased by 2.98 ($p = 0.001$).

Our findings reveal that readers' rhythmic reading is not only shaped bottom-up such as by processing the poems' syllabic composition but also top-down, modulated by one's musical activity. Since this suggests that the perception of meter is not only stimulus dependent, we will discuss more generally how future experiments investigating conventional poetry reading can account for individual differences in musicality (for music compare Dalla Bella et al. 2024) as well as the influence of diachronic and synchronic aspects of speech rhythm and musical cultures (Patel 2023) in research on poetry.

**References**

Beck, J., & Konieczny, L. (2023). What a difference a syllable makes—Rhythmic reading of poetry. Frontiers in Psychology, 14, 1043651. https://doi.org/10.3389/fpsyg.2023.1043651

Blohm, S., Kraxenberger, M., Knoop, C. A., & Scharinger, M. (2021). Sound Shape and Sound Effects of Literary Texts. In D. Kuiken & A. M. Jacobs (Eds.), Handbook of Empirical Literary Studies (pp. 7–38). De Gruyter. https://doi.org/10.1515/9783110645958-002

Dalla Bella, S., Janaqi, S., Benoit, C.-E., Farrugia, N., Bégel, V., Verga, L., Harding, E. E., & Kotz, S. A. (2024). Unravelling individual rhythmic abilities using machine learning. Scientific Reports, 14(1), 1135. https://doi.org/10.1038/s41598-024-51257-7

Peer, W. van, & Chesnokova, A. (2023). Experiencing poetry: A guidebook to psychopoetics (First edition). Bloomsbury Academic.

Patel, A. D. (2003). Rhythm in language and music: Parallels and differences. Annals of the New York Academy of Sciences, 999, 140–143. https://doi.org/10.1196/annals.1284.015

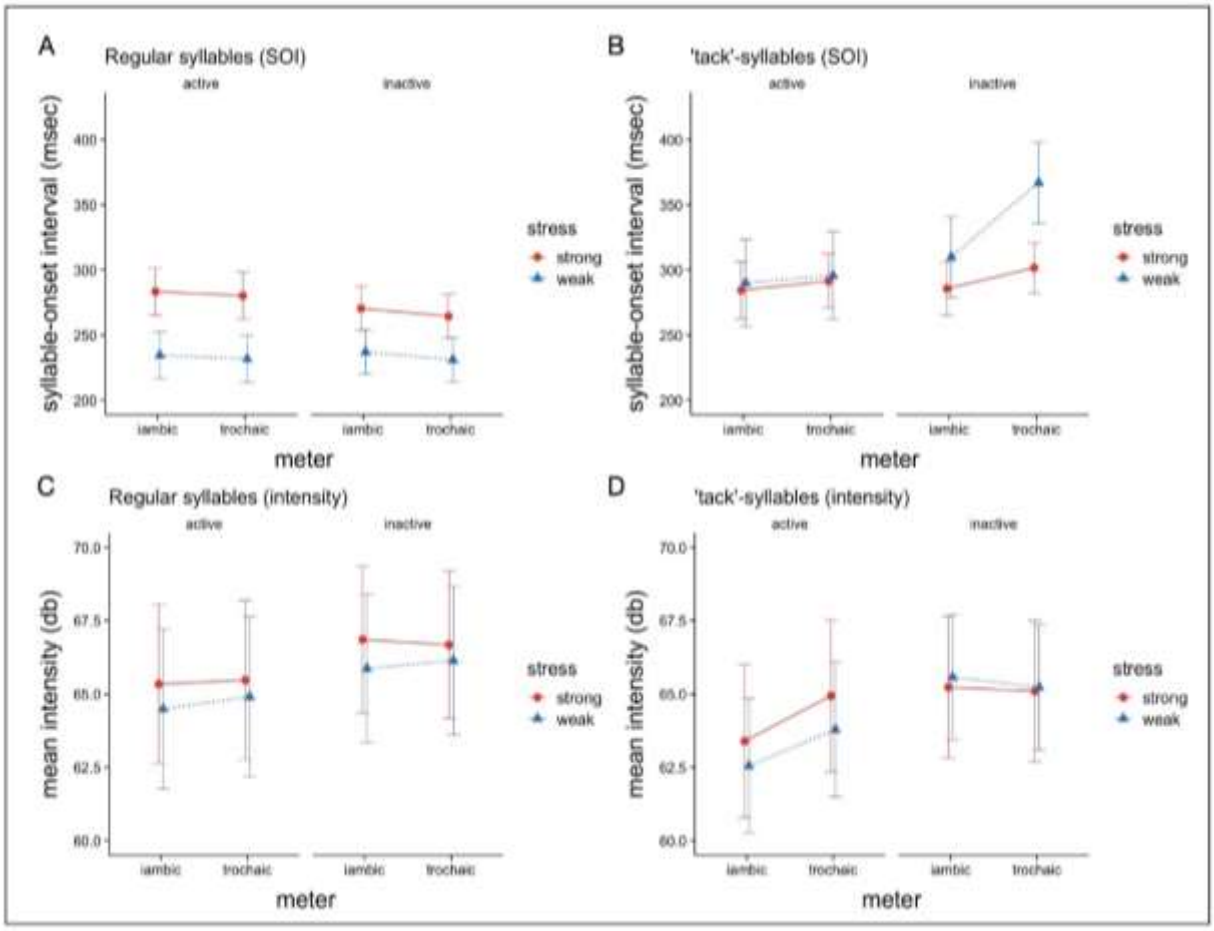



**FIGURE 1** Syllable-onset intervals **(A,B)** and syllable intensities **(C,D)** for *regular syllables* **(A,C)** and *'tack'-syllables* **(B,D)**, as a function of stress (strong vs. weak), meter (iambic vs. trochaic), and musical activity (active vs. not active). The whiskers depict 95% confidence intervals.

# Musical surrogate languages:
# How poetics and music shape their communicative niche

Laura McPherson & Alexis Han Bit Kang
(Dartmouth College & Dartmouth College)

Keywords: music, poetics, speech surrogates, typology, discourse

At the intersection between language and music, we find musical surrogate languages, systems of communication in which linguistic structure is transposed onto a musical instrument. A slowly growing body of literature explores the structural relationships between language and music in a wide range of surrogate traditions, including different Yorùbá drumming systems (Villepastour 2010, Akinbo 2019, Durojaye et al. 2021), Bora slit log drumming (Seifart et al. 2018), the Sambla balafon (McPherson 2019), the Northern Toussian balafon (Struthers-Young 2021), and the Igbo òjà aerophone (Carter-Énỳì et al. 2021), among others, though much work remains to be done. In contrast, the purpose or communicative niche of musical surrogate languages cross-linguistically has received far less attention. Why would a community develop a musical surrogate language, and why is it used only in particular contexts? This talk lays out a preliminary typology of contexts of use for musical surrogates and explores the question of why these communicative practices may be well adapted to these uses.

A key finding is that regardless of purpose (long distance communication, courtship, spiritual communication, etc.), most musical surrogates involve some degree of poetic language or rhetorical devices in how the message is encoded. We argue that this poetic language, coupled with musical form, allows musical surrogate languages to take advantage of the communicative properties of music and poetry, such as emotional expressivity, transcendence, deindividuation, and promotion of group cohesion. Thus, the apparent limitations of speech surrogates in terms of ambiguity and comprehensibility are compensated for by other cognitive benefits that accompany their modality. To support these claims, we draw on four case studies that span a range of communicative purposes. First, poetic verses on the Hmong jaw harp are used in courtship (Catlin 1982), capturing both emotional expressivity and deindividuation, as the identity of the lover is masked by the use of the instrument. Second, the Sambla balafon's ability to communicate socially inappropriate messages (demands for money, criticism; McPherson 2019) demonstrates a different angle on deindividuation, as it is viewed as the instrument rather than the performer who is communicating the message; the ability to give criticism also plays a role in social cohesion. Third, the Hmong *qeej* (mouth organ) demonstrates a surrogate's power of transcendence over the quotidian; in funeral settings, its messages are meant for the spirits of the dead to direct them to the afterlife (Falk 2003). Fourth, the Yorùbá *dùndún* (tension drum) draws on a rich vocabulary of praises and proverbs to address festival-goers and celebrate their lineage (Euba 1990). This shared repertoire and public display of praise promotes group cohesion, with emotional expressivity amplified by the musical medium. In short, these musico-linguistic traditions shed light on the division of labor between language and music and how the two modalities can combine to fulfill a unique position in a community's communicative ecology.

Acknowledgements: This material is based upon work supported by the National Science Foundation under grant BCS-1943896.

**References**

Akinbo, Samuel (2019), Representation of Yorùbá tones by a talking drum: An acoustic analysis. *Linguistique et Langues Africaines* 1: 11-23.

Carter-Ényì, Aaron, Nnaemeka C. Amadi, Quintina Carter-Ényì, Charles Chukwudozie, Jude Nwankwo, and Ebruphiyor Omodoro (2021), Igbo speech surrogacy: Preliminary findings based on the oja flute. *Frontiers in Psychology* 12: 653068.

Catlin, Amy (1982), Speech Surrogate Systems of the Hmong: from singing voices to talking reeds. *The Hmong in the West: observations and reports*: 170-197.

Durojaye, Cecilia, Kristina L. Knowles, K. Jakob Patten, Mordecai J. Garcia, and Michael K. McBeath (2021), When music speaks: An acoustic study of the speech surrogacy o the Nigerian dùndún talking drum. *Frontiers in Communication* 6:652690.

Euba, Akin (1990), *Yorùbá drumming: The* dùndún *tradition.* Bayreuth: Bayreuth African Studies.

Falk, Catherine (2003), "If You Have Good Knowledge, Close It Well Tight": Concealed and Framed Meaning in the Funeral Music of the Hmong qeej. *British Journal of Ethnomusicology*: 1-33.

McPherson, Laura (2019), The talking balafon of the Sambla: Grammatical principles and documentary implications. *Anthropological Linguistics* 60.3: 225-294.

Seifart, Frank, Julien Meyer, Sven Grawunder, and Laure Dentel (2018), Reducing language to rhythm: Amazonian Bora drummed language exploits speech rhythm for long-distance communication. *Royal Society open science* 5.4: 170354.

Struthers-Young, Anthony (2022), A preliminary account of the Northern Toussian balafon surrogate language. *Frontiers in Communication* 6: 652740.

Villepastour, Amanda (2010), *Ancient text messages of the Yorùbá bàtá drum: Cracking the code.* Farnham: Ashgate Publishing, Ltd.

# Tonal textsetting in ancient Greek vocal music: Revisiting the grave

Dieter Gunkel & Kevin M. Ryan
(University of Richmond & Harvard University)

Keywords: metrics, music, pitch accent, phonology, textsetting, tone, tone-melody matching

This paper examines the relationship between linguistic tone and musical melody in ancient Greek vocal music. The goal is to reconstruct the phonological representation and phonetic realization of the grave accent, which has been a subject of debate since the early twentieth century. There are two defensible views on the grave: it was either a downstepped high tone or a deleted one.[1] The analysis offered here has several advantages vis-à-vis previous ones, including the sophisticated analysis of Devine & Stephens (1994: 181–4 *et passim*). It is based on cleaner data, thanks to ongoing philological work on the musical fragments (e.g. Hagel 2000, Pöhlmann & West 2001), and it builds on significant recent advances in tonal textsetting (cf. Ladd & Kirby 2020) and the corpus linguistic analyses thereof (e.g. McPherson & Ryan 2017). We find better evidence for the grave as a deleted high tone.

The music studied here provides evidence for linguistic pitch movements because the poets actively match the ups and downs of the melody with the direction of such movements in the corresponding text.[2] Consider, for instance, the setting of the phrase αἰόλοις μέλεσιν [ai.ió̩ .loiṣ̩ . mé.le.sin] 'shimmering:DAT.PL tunes:DAT'. The melody rises to peak at the high-toned syllables [ió] and [mé], then falls to word-end. That closely matches the pitch movements of the text, given what we know from non-musical and therefore independent sources such as the descriptions of the ancient grammarians (cf. Probert 2006: 1–124 *et passim*).

(1) Setting of αἰόλοις μέλεσιν (DAGM 20.14f.)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | D | | | | | |
| | Db | | Db | | Db | | |
| C | | | | C | | C | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | Ab |
| aiṣ̩ | | ió̩ | loiṣ̩ | | mé | le | si |

[1] For references to the literature, see Gunkel 2023: 236.

[2] Important foundational work on tone-melody matching in ancient Greek includes Crusius 1894: 113–123, Wackernagel 1896: 305, and Winnington-Ingram 1955. For recent overviews with further references, see Pöhlmann 1970: 140, West 1992: 199, and Probert 2006: 48. For statistical evidence that the matches should not be attributed to chance, see Cosgrove & Meyer 2006.

While the matching is not perfect, of course, it is quite strict. A pilot study (Gunkel 2023: 239–244) finds that within words, tone and melody move in parallel ca. 75%, obliquely 25% (tone rises/falls and melody stays level, or vice versa), and never in opposition (the one rises and the other falls). A further study (Sandell & Gunkel) estimates opposing settings at less than 5% overall, i.e. within or across words. In the aggregate, then, the pitch movement of the melody reflects the pitch movement of the text. With the help of careful statistics, this allows us to ask the same sorts of questions of the setting data that we can ask of pitch tracks.

In pitch tracks, on the hypothesis that the grave represents a deleted high tone, we would expect to find interpolation between the surrounding syllables. In other words, the grave-accented syllable should be set near the average of the surrounding syllables. On the hypothesis that the grave represents a downstepped high tone, we would expect to find evidence of an active high target, i.e. a local peak or elbow in the pitch track. In a series of tests (e.g. comparing grave to the interpolation line, comparing grave to unaccented syllables), we find no support for an active high target.

Abbreviations

DAGM = Pöhlmann, Egert, and Martin L. West. 2001. *Documents of Ancient Greek Music*. Oxford: Clarendon.

**References**


Cosgrove, Charles H., and Mary C. Meyer. 2006. "Melody and word accent relationships in ancient Greek musical documents: The pitch height rule." *Journal of Hellenic Studies* 126: 66–81.

Crusius, Otto. 1894. *Die Delphischen Hymnen: Untersuchungen über Texte und Melodien*. (Ergänzungsheft zum *Philologus*, 53). Göttingen: Dieterich.

Devine, A. M., and Laurence D. Stephens. 1994. *The Prosody of Greek Speech*. Oxford: Oxford Univ. Press.

Gunkel, Dieter. 2023. "Musical evidence for low boundary tones in ancient Greek." *Greek and Roman Musical Studies* 11.2: 231–284.

Hagel, Stefan. 2000. *Modulation in altgriechischer Musik. Antike Melodien im Licht antiker Musiktheorie*. Frankfurt am Main: Lang.

Kirby, James, and D. Robert Ladd. 2021. "Tone-melody matching in tone-language singing." In *The Oxford Handbook of Language Prosody*, ed. Carlos Gussenhoven and Aoju Chen, 676–688. Oxford: Oxford University Press.

McPherson, Laura, and Kevin M. Ryan. 2018. "Tone-Tune Association in Tommo So (Dogon) Folk Songs." *Language* 94.1: 119–156.

Pöhlmann, Egert. 1970. *Denkmäler altgriechischer Musik*. Nürnberg: H. Carl.

Pöhlmann, Egert, and Martin L. West. 2001. *Documents of Ancient Greek Music*. Oxford: Clarendon.

Probert, Philomen. 2006. *Ancient Greek Accentuation: Synchronic Patterns, Frequency Effects, and Prehistory*. Oxford: Oxford University Press.

Sandell, Ryan P., and Dieter Gunkel. Submitted. "On the representation and realization of the Ancient Greek acute: Evidence from tone-tune mappings in Ancient Greek music."

Wackernagel, Jakob. 1896. "Das Zeugniss der delphischen Hymnen über den griechischen Accent." *Rheinisches Museum für Philologie*, N. S. 51: 304–305. West,

Martin L. 1992. *Ancient Greek Music*. Oxford: Clarendon.

Winnington-Ingram, R. P. 1955. "Fragments of unknown Greek tragic texts with musical notation: II The music." *Symbolae Osloenses* 31: 29–87.

# Mapping Oral Art across Time: A new data visualization tool for poetic structures

Uta Reinöhl & T. Mark Ellison
(University of Freiburg & University of Cologne)

In this talk, we present and test recurrence maps as powerful visualization tools for poetic structures. These maps color-code degrees of similarity between sequenced units, such as intonation units or regular temporal intervals. They reveal stability and variation across time, visualizing, for instance, alliterative chains or clusters of morpho-syntactic constructions. They thus function as potent tools in the analysis of poetic language, helping us map the diversity of poetic forms present in the languages of the world. Recurrence maps can be generated for any level of structure - e.g. full spectrograms, f0 (as in Figure 1), glosses or translations. The potential of this data visualization tool extends beyond poetic constructions, and can reveal recurring structures in any type of connected language data, such as prose works or spontaneous conversation (Tannen 1987). While repetition and variation are the underlying structuring principles of any type of language use, repetition is particularly prominent, even defining, for many kinds of poetry (Jakobson 1966, Fox 1988).

We illustrate the versatility of recurrence maps with Vedic (VedaWeb 2.0, https://vedaweb.uni-koeln.de/) and Igu. The poetic structures of Vedic hymns have been intensively studied (e.g. Staal 1979, Watkins 1995, Klein 2002, 2012). They can therefore serve as a standard of comparison for how well recurrence maps highlight known similarity and dissimilarity structures, or reveal novel, less obvious or well-studied (dis)similarity relations.

Igu is the ritual language of the Kera'a (isocode: *clk*; Reinöhl 2022), a Tibeto-Burman society on the border of northeastern India and southern Tibet (see Choudhuri 2018, Dele 2018, 2021, Mene & Miso 2022 for literary-anthropological descriptions). Of 8 hours of recorded ritual, approximately 90 mins have been glossed and translated by xxx [anonymized reference]. Whereas Vedic poetic forms have been intensively studied in the past, recurrence maps of Igu data illustrate the explorative potential for undescribed or underdescribed poetic structures.

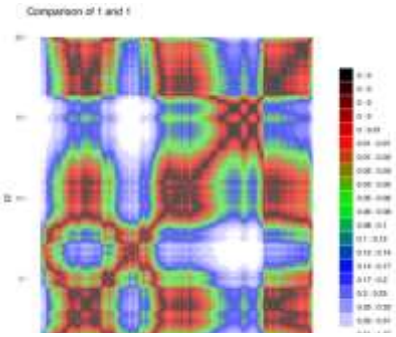


*Figure 1. Partial recurrence map of Kaliwu ritual.*

We illustrate application options on two levels, for both languages: f0 and transcriptions, which reveal melodic contours and segmental patterns respectively. In addition, we show how different maps of the same Igu ritual - based on distinct performances by the same or different shamans - highlight distinct degrees of transmission faithfulness (see Bravi & Proto 2019, Oras 2019). Such information, in turn, provides cues for assessing the time depth of oral transmission, and so feeds into the historical and genealogical analysis of Igu and Kera'a within their Tibeto-Burman context.

**References**


Bravi, Paolo & Teresa Proto. Intra-line and inter-individual variation in Sardinian arrepentina, *Open Linguistics* 5 (1), 496-510. https://doi.org/10.1515/opli-2019-0027

Choudhuri, Sarit K. 2008. Plight of the igus: Notes on Shamanism Among the Idu Mishmis of Arunachal Pradesh, India. *Revue European Bulletin of Himalayan Research* 32, 84-108.

Dele, Razzeko. 2018. *Idu Mishmi shamanic funeral ritual Ya: a research publication on shamanic oral chants and rituals related to mortuary behaviour among the Idu Mishmis of Arunachal Pradesh.* Delhi: Bookwell.

Delley [Dele], Razzeko. 2021. *Igu. Study on four Idu Mishmi shamanic rituals: Kanliwu, Machiwu, Anongko, A-Taye (Birth ritual).* Delhi: Bookwell.

Fox, James. 1988. Introduction. In James Fox (ed.). T*o speak in pairs. Essays on the ritual languages of Eastern Indonesia*. Cambridge: Cambridge University Press, 1-28.

Jakobson, Roman. 1966. Grammatical parallelism and its Russian facet. *Language* 42, 398-429.

Klein, Jared S. 2002. Responsion in the Rigveda. *Journal of the American Oriental Society* 122 (2), 311–317. doi.org/10.2307/3087625

Klein, Jared S. 2012. Categories and Types of Interstanzaic Repetition in the Rigveda. *Historische Sprachforschung 1*25, 191–210.

Mene, Tarun & Rajiv Miso. 2022. Origin Myth and the Making of Shaman: A Case Study of the Idu Mishmis of Arunachal Pradesh. *The Researchers* 8 (1), 22-36. doi.org/10.21276/tr.2022.8.1.AN3

Oras, Janika. 2019. Individual Rhythmic Variation in Oral Poetry: The Runosong Performances of Seto Singers, *Open Linguistics* 5 (1), 570-582. doi.org/10.1515/opli-2019-0031

Reinöhl, Uta. 2022. Locating Kera'a (Idu Mishmi) in Its Linguistic Neighbourhood: Evidence from Dialectology. In: Mark W Post, Stephen Morey & Toni Huber (eds.), *Ethnolinguistic Prehistory of the Eastern Himalaya*, 232–263. Leiden: Brill. doi.org/10.1163/9789004518049_010

Staal, Frits. 1979. Ritual Syntax. In: M. Nagatomi, B.K. Matilal, J.M. Masson, E.C. Dimock (eds.). *Sanskrit and Indian Studies. Studies of Classical India, vol 2.* Springer, Dordrecht. doi.org/10.1007/978-94-009-8941-2_7

Tannen, Deborah. 1987. Repetition in Conversation: Toward a Poetics of Talk. *Language* 63 (3), 574-605.

Watkins, Calvert. 1995. How to kill a dragon: Aspects of Indo-European poetics. New York/Oxford: Oxford University Press.

# Subjective experience of poetry in translation

Filip Nenadić & Marija Nenadić Žurka
(University of Potsdam & "G. Călinescu" Institute for Literary History and Theory)

Keywords: subjective experience, translation, poetry, adjective rating scales, non-expert reader

The variety present in poetic language can be described through a poem's fairly objective characteristics, but also through its reception by the reader. One important technique utilized in empirical investigation of subjective and aesthetic experiences of art forms in general are adjective rating scales. Such scales enable researchers to describe the basic dimensions of subjective experience of artwork and to inspect how various artwork, perceiver, and context characteristics influence reception. The goal of the present work is to argue in favor of developing subjective rating scales for poetry in different languages and to showcase how a carefully developed rating scale that measures the basic dimensions of subjective experience of poetry can be used to inform us about an important aspect of poetry reading – reading poetry in translation.

The majority of the wealth of world literature is available to the reader solely through translations. Previous empirical studies indicated that there are differences between subjective ratings of the same poem across different cultures and languages. Although they provide valuable information, these studies often use a single poem each and collect ratings on adjectives previously elicited for those specific poems. More importantly, these studies cannot tease apart the effects induced by the translation procedure from the effects of reader culture or language, as the participants read either the original poem version in a foreign language or a single translation of the original poem. In the present study, we investigate differences in the subjective experience of poem translation variants in order to assess the impact translator choices have on reading experience.

We presented 219 layman readers (90% female, 9% male, 1% other or do not wish to disclose; age range 18 to 55, $M$ = 21.62, $SD$ = 5.97) with seven poems translated into Serbian language. Baudelaire, Lorca, and Blok were authors of two poems each, and one poem was written by Mayakovsky. All seven poems were available in two published translations by different translators. Each participant only read one translation version randomly assigned to them and the poem order was randomized for each participant. The task for the participants was to rate these poems on 33 adjectives on a Likert-type five-point scale. A preliminary analysis indicates that although different poems receive different ratings, translation variants of the same poem do not significantly differ across three major latent dimensions of layman subjective experience of poetry (Pleasantness, Opacity, and Poignancy).

Based on these results, it would seem that valid, published translations of poems, despite their differences, affect a non-expert reader similarly. Future studies should test whether presenting the same layman reader with both translation variants is sufficient for some differences in subjective experience of those variants to emerge, or whether expert readers are more sensitive to differences in translation variants. In order to fully answer the question how poems are experienced by native versus non-native readers in their original version versus translation variants we would need to develop subjective experience scales for additional languages.

# WS16 Psycholinguistic approaches to the study of heritage, Indigenous, and minoritized languages

# Navigating structural changes: The impact of new speakers on minority language revitalization (The case study of Kalmyk)

Vlada Baranova
(University of Helsinki)

There have been numerous successful examples of revitalizing indigenous 'sleeping' languages in recent decades (Hinton 2001). Grassroots revitalization initiatives aim to transform language practices within specific communities, increasing the number of minority language speakers and expanding the spheres of language use. Despite its goal of re-establishing the language within the speech community, language revitalization paradoxically induces partial changes in the language's structure.

For a minority language to survive, it is crucial that its usage among younger age groups increases, as they are the ones who will speak it in the future. However, involving more young and socially active participants in the revitalization process leads to a greater number of 'new speakers' whose competence may differ from that of older traditional speakers. The revitalization process manifests various starting points in the language shift continuum, resulting in a diverse range of bilingual speakers with incomplete competence. The language production of individuals who have not previously used a minority language differs from traditional competence.

While recent studies on ‘new speakers’ primarily focus on attitudes and sociolinguistic issues, linguistic analysis remains understudied (Smith-Christmas et al. 2018 (eds); McLeod, O’Rourke 2015). Changes in revitalized languages can take different directions, often involving overgeneralization influenced by dominant languages. For instance, young Palenquero speakers may use the Palenquero prenominal plural marker in singular contexts (Lipski 2021). It also can lead to innovation, for example, the revived Cornish has a new form to express possession with abstract nouns that differs from traditional Cornish (Arbes 2019). At the same time, new speakers may adopt purism, avoiding syntactic patterns associated with the influence of dominant language, as seen with young Maya speakers in Guatemala (Barrett 2008).

This paper explores the structural consequences of the emergence of new speakers within minority language revitalization efforts, focusing on the Kalmyk language. Utilizing data from stand-up comedy performances on Instagram and YouTube channels, along with other online sources and interviews with speakers, the research analyzes the language of young Kalmyk speakers (17-35 years old). Despite non-family transmission, new speakers acquire fluency and communicate in Kalmyk. The study deals with word order, coding of the subordinate clause (coverbal forms vs ‘conjunction-like’ strategy for adverbial clauses) and some other morphosyntactic features. Some features appear to be distinctive of young speakers, such as non-standard word order in Kalmyk, but their consolidation in the future remains uncertain.

The findings illuminate the dynamic interplay between language revitalization efforts, the emergence of new speakers, and resulting structural shifts within minority languages. With the increasing number of "new speakers" among minority language communities, linguistics should consider their influence on language in the future.

**References**

Arbes D. (2019). Predicative possession in revived Cornish. In.: L. Johanson, L. Mazzitelli, I. Nevskaya. *Possession in Languages of Europe and North and Central Asia.* Amsterdam: John Benjamins.

Barrett R. (2008). Linguistic differentiation and Mayan language revitalization in Guatemala *Journal of Sociolinguistics* 12/3, 2008: 275–305.

Lipsky J.M. Language revitalization as L2 shadow boxing: The case of Palenquero plural- marking. *Studies in Language Acquisition* 43: 220–235.

McLeod W., O'Rourke B. (2015) "New Speakers" of Gaelic: Perceptions of linguistic authenticity and appropriateness. *Applied Linguistics Review*6(2), 151-172.

Smith-Christmas C., Ó Murchadha N., Hornsby M. &Moriarty M. (eds).2018. *New Speakers of Minority Languages*. *Linguistic Ideologies and Practices*. London: Palgrave Macmillan.

# Being the Other:
# Bilingual effects arising in the shadow of big, standard languages

Evelina Leivada
(Universitat Autònoma de Barcelona; Institució Catalana de Recerca i Estudis Avançats)

Keywords: bilingualism, dialect, diversity, minority languages

Scientists have recently been taking critical steps towards achieving a better level of transparency and inclusivity. In terms of transparency, significant progress has been made. The effort to pre-register hypotheses, properly archive and make available raw materials, elicitation protocols, code used for the analyses, and make preprints available in an open format is a laudable initiative that already produces useful results (e.g., Simonsohn 2013). However, inclusivity is a more slippery terrain (Blasi et al. 2022). In fields that study some aspects of human identity, the absence of inclusivity may be linked to the prevalence of various types of stereotypes. Taking the field of psycholinguistics as an example, concerted efforts have been directed towards the identification of racist, nationalist, ableist, and gender stereotypes in psycholinguistic research. For instance, recent analyses of the mainstream term 'native speaker' have critically discussed the uses of this notion, showing how it may exclude certain populations from scientific representation (Cheng et al. 2021). Precisely because characterizations like 'native language' and 'first/second language' often embody powerful ideologies, they may be used as the basis for language assessments such as 'broken grammar', 'bad English', and 'incomplete acquisition' that may function as racialized signs of deviance, foreignness, and otherness (Rosa 2016). Who exactly is the Other?

In this talk, I will focus on one of the most debated aspects of bilingualism: the benefits and costs that bilingualism confers to cognition. Being bilingual has been linked to an improvement in executive functions and other domains of cognition (Bialystok et al. 2022), but this claim has been contested by scholars who fail to replicate the effect (Paap et al. 2021). While many explanations for the variation in the results have been offered (Leivada et al. 2021), we lack an understanding of what populations are being represented and to what degree they are linked to bilingual adaptations. In this talk, I will present the results of a population re-analysis of the biggest systematic review of bilingual effects to this date (Masullo et al. 2023). The results suggest that from the original sample of 368 studies, 35% of the represented language pairs correspond to 'Others'. Otherness can be defined as a bilingual or multilingual mega-group that consists of people who use one big, standard language (usually English or Spanish) and another language. Quite often the individual identities behind otherness are not explicit. This talk will discuss the implications of this sampling bias, which amounts to using big bilingual samples that involve a mix of different ages, languages, and backgrounds, arguing that it (i) obscures our understanding of bilingual effects that arise in the shadow of big, standard languages, and (ii) hampers inclusivity and visibility of certain linguistic communities.

**References**

Bialystok, Ellen and Fergus I. M. Craik (2022), How does bilingualism modify cognitive function? Attention to the mechanism, *Psychonomic Bulletin & Review* 29, 1246–1269.

Blasi, Damián, E., Joseph Henrich, Evangelia Adamou, David Kemmerer and Asifa Majid (2022), Over-reliance on English hinders cognitive science, Trends in Cognitive Sciences 26, 1153–1170.

Cheng, Lauretta S.P., Danielle Burgess, Natasha Vernooij, Cecilia Solís-Barroso, Ashley McDermott and Savithry Namboodiripad (2021), The problematic concept of native speaker in psycholinguistics: Replacing vague and harmful terminology with inclusive and accurate measures, *Frontiers in Psychology* 12, 715843.

Leivada, Evelina, Marit Westergaard, Jon Andoni Duñabeitia and Jason Rothman (2021), On the phantom-like appearance of bilingualism effects on neurocognition: (How) should we proceed?, *Bilingualism: Language and Cognition* 24, 197–210.

Masullo, Camilla, Vittoria Dentella and Evelina Leivada (2023), 73% of the observed bilingual (dis)advantageous effects on cognition stem from sociolinguistic factors: A systematic review. *Bilingualism: Language and Cognition*, 1–15.

Paap, Kenneth R., Lauren Mason and Regina Anders-Jefferson (2021), Predictions about the cognitive consequences of language switching on executive functioning inspired by the Adaptive Control Hypothesis fail more often than not, *Brain Sciences* 11, 1217.

Rosa, Jonathan D. (2016), Standardization, racialization, languagelessness: Raciolinguistic ideologies across communicative contexts, *Journal of Linguistic Anthropology* 26, 162–183.

Simonsohn, Uri (2013), Just post it: The lesson from two cases of fabricated data detected by statistics alone, *Psychological Science* 24, 1875–1888.

# Examining Lexical Variation in the Galician Version of MultiPic: A Psycholinguistic Inquiry

María Álvarez de la Granja, Montserrat Comesaña, Isabel Fraga Carou & M. Carmen Parafita Couto
(University of Santiago de Compostela, University of Minho, University of Santiago de Compostela & Leiden University)

Keywords: Galician, MultiPic, lexical variation, psycholiguistics

In the realm of psycholinguistic research, the meticulous selection and control of stimuli constitute imperative prerequisites for reaching insights into cognitive processes. Within psycholinguistics, images often serve as stimuli, thereby necessitating recourse to image databases for the examination of linguistic, mnemonic, and visual perceptual phenomena. The Multilingual Picture (MultiPic) database (Duñabeitia et al. 2018, 2022), represents a collection of 500 meticulously curated color depictions extracted from a larger pool of 750 images. Each of these images serves as a visual representation of specific concepts and general knowledge. Duñabeitia and colleagues (2022) detailed the protocol they employed, which involved engaging speakers of 32 languages or language varieties in the assignment of linguistic labels to the images contained within the MultiPic database, while concurrently assessing their degree of familiarity with the depicted concepts. In an expansion of this research, we included Galician as one of the languages for which MultiPic is accessible. Galician is a Western Ibero-Romance language primarily concentrated in Galicia, an autonomous community situated in northwestern Spain, where it holds official status alongside Spanish. Eighty-eight Galician speakers participated in the naming study. They also responded to a series of initial questions covering sociolinguistic variables like age, gender, place of residence, educational level, and social status.

This presentation will delve into the considerable lexical variation that we have encountered, a variation stemming from intricate challenges related to the unequivocal identification of specific concepts, the choice between hypernyms and hyponyms, as well as the presence of synonyms or quasi-synonyms, frequently influenced by the Spanish language. We will establish connections between the lexical variation and the previously mentioned sociolinguistic variables. Importantly, we will reflect upon the challenges posed by this variation in the development of standardized materials for utilization in forthcoming psycholinguistic investigations conducted within the Galician linguistic milieu.

**References**

Duñabeitia J. A., et al. 2018. MultiPic: A standardized set of 750 drawings with norms for six European languages. *Q J Exp Psychol* (Hove). 2018 Apr;71(4), :808-816. doi: 10.1080/17470218.2017.1310261.

Duñabeitia, J. A. et al. 2022. The Multilingual Picture Database. *Sci Data* 9, 431. doi:10.1038/s41597-022-01552-7

# Measuring and evaluating proficiency in the context of language shift

Lenore Grenoble
(The University of Chicago)

Widescale language shift provides opportunities to understand language change in process across typologically diverse languages. But studying change in such communities is challenging: in order to evaluate change in process, we need to understand what linguistic structures we would expect without shift, and what the innovations are. Speakers in shifting language ecologies are not homogenous, and thus we need to study speakers with different backgrounds. This requires some evaluation of their proficiency. In this talk I first present research methods that are intended to capture the grammars of shifting speakers and then move to an assessment of proficiency based on a number of factors that derive from the data. We find tremendous variation in their grammars in both controlled elicitation and freer, more spontaneous speech, but this variation is systematic as can be determined by examining the source of variation, and how it correlates with proficiency.

Our research shows that proficiency can be determined by a nexus of correlated phenomena, linguistic and non-linguistic. Our conclusions are illustrated by data from fieldwork in the Berezovka Even (Tungusic) community, with a sample of 40 speakers, using a mixed methods approach of targeted elicitation tasks, elicited narratives, and sociolinguistic interviews conducted in the Even language. Even and Russian are both used in the community; the more proficient speakers have L1 Even and L2 Russian, while the younger heritage speakers have a Russian L1 and varying degrees of proficiency in Even. We provide evidence of shift across generations of individual families, showing systematic patterns across the younger speakers who are Russian-dominant.

A commonly used measure in heritage language research is speech rate (Polinsky 2020, 2008; Daller et al. 2011, Anstatt 2017), but we find that this is not an accurate measure in many Indigenous communities, where older, more proficient L1 speakers often speak more slowly than younger L2 speakers; this is shown to be the case for Laz-Turkish speakers (Eren 2023) and Nigua-Spanish speakers (Adamou 2021). Measuring errors in targeted elicitation is also not consistently reliable, as highly proficient speakers often fail to perform as expected because the task itself is artificial, not because they lack the linguistic skills to do so, but rather because the task is unnatural. Moreover, Adamou (2016) has shown that differences in turn-taking, overlap and have more to do with individual speaker style than proficiency. Increasingly, research shows the relevance of both social factors and linguistic variants in proficiency (Meakins 2023).

In our study, certain phenomena routinely co-occur among the speakers we expect to have the lowest relative proficiency (the youngest generation, dominant in Russian): recurring grammatical errors, in both nominal and verbal morphology; discourse-level phenomena, such as hesitations, self-corrections, addressing the researcher in a more dominant language for clarification or assistance in performing a task and heavy code-mixing. An analysis of their performance in Russian for the narrative tasks shows that these changes cannot be explained by individual speaker style but rather proficiency as linked to language shift. In addition, speakers' self-assessment of proficiency corresponds to their performance.

**References**


Adamou, Evangelia. 2021. The adaptive bilingual mind: Insights from endangered languages. Cambridge: Cambridge University Press.

Adamou, Evangelia. 2016. *A corpus-driven approach to language contact. Endangered languages in a comparative perspective*. Berlin: De Gruyter Mouton.

Anstatt, Tanja. 2017 Language attitudes and linguistic skills in young heritage speakers of Russian in Germany. In Ludmila Isurin and Claudia Maria Riehl (Eds.), *Integration, identity and language maintenance in young immigrants: Russian Germans or German Russians*, 197–224). John Benjamins Publishing Company.

Daller, Michael H., Cemal Yıldız, Nivja H. de Jong, Seda Kan and Ragıp Başbaği. 2011. Turkish-German bilinguals: methodological aspects of measurements in structurally different languages. *International Journal of Bilingualism* 15(2). 215–236. https://doi.org/10.1177/1367006910381197

Eren, Omer. 2023. *Morpho-syntactic variation in (Heritage) Laz*. PhD Dissertation, University of Chicago. (ProQuest Dissertations Publishing No. 30528803)

Meakins, F. 2023. Qualifying the language dynamics of bilingual communities. *Linguistic Approaches to Bilingualism* 13(1). 76-82. https://doi.org/10.1075/lab.22059.mea

Polinsky, Maria. 2020. *Heritage languages and their speakers*. Cambridge: Cambridge University Press.

Polinsky, Maria. 2008. Gender under incomplete acquisition: Heritage speakers' knowledge of noun categorization. *Heritage Language Journal* 6(1). 40–71. https://doi.org/10.46538/hlj.6.1.3

# Semantic change in Abui: A panel study

George Saad
(Palacky University Olomouc)

This paper combines psycholinguistic and sociolinguistic methods to study variation and change in the lexical semantics in Abui, an indigenous language spoken in eastern Indonesia. Using a panel study, I report on findings comparing production and comprehension data in 2015 with data in 2023. The main question I address is: What happens to the lexical semantic innovations observed in 2015 eight years later?

Abui is a minority language which is in contact with Alor Malay, the regional lingua franca spoken in the Alor-Pantar Archipelago. Rather than this being a straightforward shift scenario, contact with Alor Malay manifests itself into a novel type of bilingual pattern recently termed Late/Delayed/Adult Vernacular Production (Saad, Arnold, and Peddie in prep; Peddie 2021; Anderbeck 2015)children are raised in Alor Malay, but acquire passive competence in Abui through overheard speech (Kratochvíl 2007; Saad 2020). However, they only become active speakers of Abui in early adulthood (17+). This widespread, yet under-reported pattern of acquisition has exciting implications for the fields of bilingual language acquisition, language contact and change, and language endangerment (Saad, Arnold, and Peddie in prep).

Testing both production (using a video elicitation task) and comprehension (using a forced choice task), I investigate lexical semantic innovations taking place among (pre)adolescents (< 16 years) and young adults (<26 years) in wave 1 (2015) and compare them to the same speakers' language use in wave 2 (2022). I therefore address the question of how the apparent-time construct as well as age-grading apply to such a population. In so doing, I also tackle the validity of the methods used to investigate production and comprehension. The paper thus may prove useful for scholars of language contact and change, bilingualism, heritage language studies, and endangered language studies.

**References**

Anderbeck, Karl. 2015. "Portraits of Language Vitality in the Languages of Indonesia." In *Language Documentation and Cultural Practices in the Austronesian World: Papers from ICAL 12 (A-PL 019)*, edited by I Wayan Arka, L. N. Seri Malini, and I.A.M Puspani, 19–47. Canberra: Pacific Linguistics.

Kratochvíl, František. 2007. "A Grammar of Abui: A Papuan Language of Alor." PhD dissertation, Leiden: Leiden University.

Peddie, Emma. 2021. "Delayed Vernacular Production in Melanesia: A Case Study-Based Analysis." MA (Hons.) thesis, Edinburgh: University of Edinburgh.

Saad, George. 2020. "Variation and Change in Abui: The Impact of Alor Malay on an Indigenous Language of Indonesia." PhD dissertation, Leiden: Leiden University.

Saad, George, Laura Arnold, and Emma Peddie. in prep. "Late Vernacular Production in Island South East Asia and the Pacific." Manuscript.

# Changes in the degree of word order flexibility: An experimental investigation of Sakha

Jessica Kantarovich
(The Ohio State University)

Keywords: word order, Sakha (Turkic), bilingualism, language contact, language change

Sakha (iso sah) is an under-studied language of Siberia, whose speakers are increasingly shifting to Russian, especially in urban areas. Like other Turkic languages, Sakha is agglutinating and verb-final, with SOV being the most unmarked word order; however, descriptions of Sakha note its relatively free word order (Baker & Vinokurova 2010). In this study, we investigate two questions: (1) how flexible is word order among conservative speakers (and what factors condition the choice of non-SOV order), and (2) how is word order changing due to increasing Russian bilingualism?

We targeted speakers living in Yakutsk, the largest growing city in northeastern Siberia, and asked them to participate in two tasks examining word order in basic finite clauses in Sakha: an acceptability task in which they rated 68 sentences on a 1-5 scale (with 5 being the most acceptable), and a production task in which they produced 27 sentences based on picture stimuli with the target lexical items provided. In both tasks, the stimuli included intransitive, transitive, and ditransitive verbs; in the acceptability task, speakers rated all theoretically possible orders for each valency (except verb-initial orders in ditransitives). A total of 31 bilingual Sakha-Russian speakers participated: 5 were born and raised in Yakutsk; the others hailed from all over the Sakha Republic.

Verb-final orders emerged as the overall preference in unmarked contexts: all verb-final orders (including OSV, which is dispreferred cross-linguistically, see Dryer 2013) were significantly preferred to the other orders ($p < 2x10^{-16}$, with V-final sentences rated 4.1 on average compared to 2.8 for non-V-final ones). There were significant differences between speakers who spent their formative years in Yakutsk and those who had moved to Yakutsk only recently: Yakutsk speakers were generally more permissive of non-final orders ($p < 0.004$). In some cases, they showed a greater preference for SVO compared to SOV, a pattern that is also borne out in their production data. A preference for verb-medial orders was also captured in their ditransitive ratings: Yakutsk speakers rated SVOO sentences significantly higher than non-Yakutsk speakers, and rated SOOV sentences significantly lower. The fact that this preference was so robust among longtime urban dwellers is likely due to the widespread use of Russian in the city, and overall higher rates of Russian bilingualism among the Sakha in Yakutsk.

Another telling pattern emerged from the collective ratings of both urban and non-urban speakers. Although the overall trend was to rate verb-final sentences better than verb-medial sentence (on average, 4.1 for SOOV and 2.8 for SVOO and SOVO), sentences where the verb intervened between the two object arguments were significantly preferred to SVOO across the board ($p < 2x10^{-12}$). This pattern suggests that Sakha is indeed not rigidly verb-final, even among conservative Sakha-dominant speakers, who also produce non-verb-final sentences in their production. The availability of these orders complicates explanations for word order change among the Yakutsk speakers: although the pattern is likely induced by Russian interference, it is also bolstered by verb-medial constructions that can be heard even in conservative Sakha speech.

**References**

Baker, Mark C. and Nadya Vinokurova. 2010. Two modalities of Case assignment: Case in Sakha. *Natural Language and Linguistic Theory* 28(3). 593-642.

Dryer, Matthew S. 2013. Order of subject, object and verb. In Dryer, Matthew S. & Martin Haspelmath (eds.), *WALS Online* (v2020.3). http://wals.info/chapter/81

# Loosing evidentiality: A reading experiment and a judgment task with speakers of a minoritized Slavic language, Pomak

Seckin Arslan & Evangelia Adamou
(CNRS & CNRS)

Keywords: evidentiality; processing; reading experiment; judgment task; Pomak

Background: Evidentiality refers to the obligatory encoding of the source of information through grammar. Acquiring evidentiality occurs later in development (Aksu-Koç 1988) and is susceptible to breakdown in individuals with aphasia (Arslan et al. 2014). In addition, studies with heritage speakers show that evidentiality is more susceptible to changes than tense (in Turkish see Arslan et al. 2015, 2017). Bilingual experience is also relevant, e.g., low daily exposure to the heritage language predicts non-target uses of evidentiality (Arslan and Bastiaanse 2020).

Goals: We investigate loss of evidentiality in a minoritized Slavic variety, Pomak, spoken in Greece (Adamou, 2010). Based on published folktales and spoken data, Adamou (2013) reports that older Pomak speakers use a specialized verbal form for evidentiality in tales, while younger speakers have replaced it with the perfect form.

Method: Using a self-paced reading experiment, we tested 25 Pomak speakers residing in Greece (age M=39, SD=17). They were equally proficient in Pomak and Greek, and spoke Turkish to a lesser extent (self-reported-proficiency Pomak: M=4.4/5, Greek M=4/5, Turkish M=2.96/5). Participants read a total of 40 Pomak sentences in two conditions—perfect (1) or evidential (2)—and provided an end-of-sentence acceptability rating. Sentences were constructed around a traditional character in local folk tales, Nasradin Hodza, and were interspersed with 80 fillers.

| | | | | | |
|---|---|---|---|---|---|
| 1. Nasradin, | ye | klal, | şiker, | faf, | çayen. |
| ‘Nasradin | has | added | sugar | in | his tea.’ |

| | | | | |
|---|---|---|---|---|
| 2. Nasradin, | klal, | şiker, | faf, | çayen. |
| ‘Nasradin | added(EVD) | sugar | in | his tea.’ |

Results: *Judgments.* Results reveal significant effects of Condition and an integration between Age and Condition, suggesting that overall, participants found perfect forms more acceptable than evidential forms, and that this effect was modulated by age, with younger participants considering evidential forms less acceptable; see Figure 1.

*Reading times.* Critical differences emerged in the reading times data, especially in region 2 where the verb appeared. We found significant effects of Age (ß=0.38, SE=0.15, $t$=2.47, $p$=0.02) and Condition (ß=0.16, SE=0.05, $t$=2.87, $p$=0.004), indicating that older participants read more slowly, and that both groups read evidential verbs with longer RTs; see Figure 1.

Discussion: Our results suggest that Pomak speakers have reduced sensitivity to grammatical evidentiality, comparable to Turkish heritage speakers in Europe (Arslan et al. 2015, 2017). We

interpret these results as evidence that Pomak evidential forms are undergoing reduction, possibly due to contact with Greek, a language with no grammatical evidentiality (Adamou 2013).

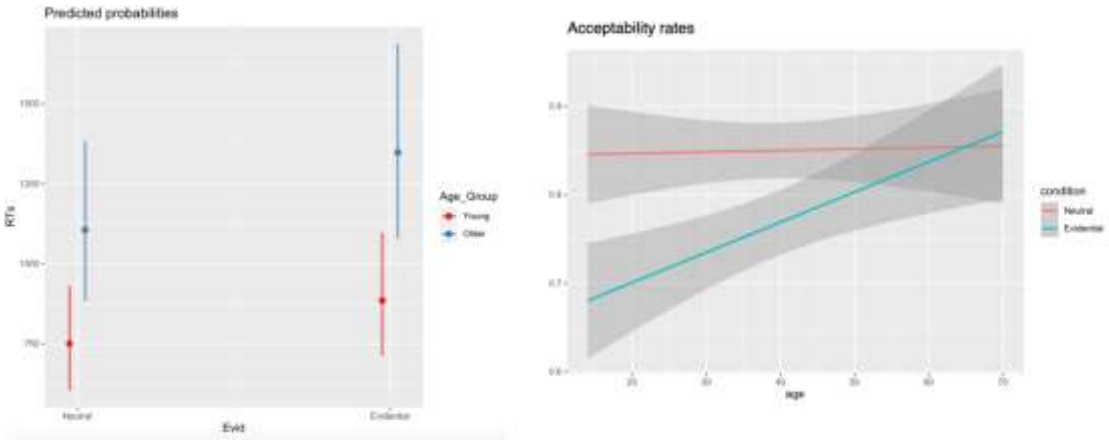



Figure 1. Pomak participants' reading times at region 2 (left, perfect/evidential verb), and end-of-sentence acceptability rates (right).

Acknowledgement: We are grateful to our anonymous research assistant who helped during data collection. Seckin Arslan acknowledges funding support by the National Research Agency (ANR) and Université Côte d'Azur through Académie d'Excellence 5, project BALKANS (grant no. ANR-15-IDEX-01). Evangelia Adamou acknowledges funding support by the National Research Agency, project Atlas of the Balkan Linguistic Area – ABLA (grant no. ANR-21-CE27-0020).

**References**

Adamou, Evangelia (2010), Bilingual speech and language ecology in Greek Thrace: Romani and Pomak in contact with Turkish*. Language in Society*, 39(2), 147–171.

Adamou, Evangelia (2013), Change and variation in a trilingual setting: Evidentiality in Pomak (Slavic, Greece). In I.Léglise and C. Chamoreau (eds), *The Interplay of Variation and Change in Contact Settings*. John Benjamins, 229–252.

Aksu-Koç, Ayhan (1988). *The Acquisition of Aspect and Modality: The Case of Past Reference in Turkish* Cambridge: Cambridge University Press.

Arslan, Seckin, Aksu-Koç, Ayhan, Maviş, İlknur,and Bastiaanse, Roelien (2014), Finite verb inflections for evidential categories and source identification in Turkish agrammatic Broca's aphasia. *Journal of Pragmatics,* 70, 165–181.

Arslan, S., & Bastiaanse, R. (2020). First language exposure predicts attrition patterns in Turkish heritage speakers' use of grammatical evidentiality. In F. Bayram (Ed.), *Studies in Turkish as a heritage language,* John Benjamins Publishing Company, 105–126.

Arslan, Seckin, Bastiaanse, Roelien, and Felser, Cloudie (2015), Looking at the evidence in visual world: Eye-movements reveal how bilingual and monolingual Turkish speakers process grammatical evidentiality. *Frontiers in Psychology*, 6, 1-12.

Arslan, Seckin, De Kok, Doerte, and Bastiaanse, Roelien (2017), Processing grammatical evidentiality and time reference in Turkish heritage and monolingual speakers. *Bilingualism: Language and Cognition*, 20(3), 457–472.

# WS17 Spatial (a)symmetries across languages

# Goal bias and deixis in Dutch: An elicitation study

Benjamin Fagard, Machteld Meulleman & Thomas Hoelbeek
(Laboratoire Lattice, CNRS, ENS|PSL and Paris Sorbonne Nouvelle; Université de Reims Champagne-Ardenne; & Vrije Universiteit Brussel)

Keywords: Dutch, Source-oriented events, Goal-oriented events, Source-Goal asymmetry, deixis

We investigate Goal bias (Ikegami 1987) in Dutch, a satellite-framed language (Talmy 1985, 2000) which as such mainly relies on satellites (Talmy 2000: 222, e.g. *in* "in" in example (1) below), in order to encode Source and Goal. Analyzing new data obtained for Netherlandic Dutch, we explore the extent and expression of the expected Goal bias, and focus on differences in frequency, morphological complexity and semantic granularity. Our data consist of descriptions of self-agentive motion, collected with a set of visual stimuli (Ishibashi et al. 2006). The corpus includes a total of 14 descriptions of 55 short video clips representing Goal-oriented and Source-oriented events.

In this presentation, we focus on a subset of 224 descriptions of Source- and Goal-oriented scenes, in which we identified and analyzed Source and Goal expressions (lemma, word class, construction). Our analysis reveals that the main differences lie in the expression of Goal in Source-oriented scenes (43%) and non-expression of Source in Goal-oriented scenes (7%), with the former scenes frequently yielding descriptions including both Source and Goal expression, as in (1).

(1) *Een vrouw loopt een grot uit het bos in.*
a woman walk.PRS.3SG a cave out the wood in
'a woman walks out of a cave into the woods'

This comes with another bias, the (unexpected) overexpression of the ventive *komen* 'to come', as in (2), and underexpression of the itive *gaan* 'to go'.

(2) *Een jongen komt weer uit de zee gerend.*
a boy come.PRS.3SG again out the sea run.PSTPART
'a boy comes running out of the sea'

In order to determine whether the overexpression of *komen* and the underexpression of *gaan* are linked, we explore the hypothesis that the cognitive load of Source expression hinders the conflation of manner and path in the verb stem, favoring instead the division of information across main verb (*come* indicating path) and infinitive (*run* indicating manner).

The combination of *komen* with a manner verb (appearing as an infinitive or a past participle) and a directional phrase (as a satellite) is known to designate an unfolding deictic motion event, and the variant with the past participle, as in (2), "highlights the *end* of a process" (Beliën 2016: 30, italics in the original). In our data, *komen* always combines with the preposition *(van)uit* 'out (of)', and thus the motion events only involve a short path, where the agent completes the path of motion from the inside to the outside. In this type of motion event, the past participle variant is typically found to be

much more frequent than the infinitive variant (Beliën 2016: 29). In our data, however, the situation is more balanced.

This ventive-itive asymmetry presents an interesting overlap with the source-goal asymmetry: all occurrences of *komen* are found in source-oriented scenes, i.e. scenes in which someone can be seen exiting or leaving from a given place, and is not strictly dependent on the observer's perspective. This is in line with the findings of Matsumoto et al. (2017), pointing to the importance of the functional component.

**References**


Beliën, Maaike (2016), Exploring semantic differences in syntactic variation: Dutch *komen* 'come' with a past participle or an infinitive, in A. Bannink, and W. Honselaar (eds), (2016), *From Variation to Iconicity: Festschrift for Olga Fischer on the Occasion of her 65th Birthday*, Amsterdam: Pegasus, 17–32.

Ikegami, Yoshihiko (1987), 'Source' vs. 'goal': A case of linguistic dissymmetry, in R. Dirven, and G. Radden (eds), (1987), *Concepts of case*, Tübingen: Narr, 122-146.

Ishibashi, Miyuki, Anetta Kopecka, and Marine Vuillermet (2006), *Trajectoire : Matériel visuel pour élicitation des données linguistiques*, Laboratoire Dynamique du Langage (CNRS / Université Lyon 2) – Fédération de Recherche en Typologie et Universaux Linguistiques: CNRS, France.

Matsumoto, Yo, Kimi Akita, and Kiyoko Takahashi (2017), The functional nature of deictic verbs and the coding patterns of Deixis: An experimental study in English, Japanese, and Thai, in I. Ibarretxe-Antuñano (ed), (2017), *Motion and space across languages*, Amsterdam / Philadelphia: John Benjamins, 95–122.

Talmy, Leonard (1985), Lexicalization patterns: Semantic structure in lexical forms, in T. Shopen (ed), (1985), *Language Typology and Syntactic Description: Grammatical Categories and the Lexicon*, Cambridge: Cambridge University Press, 225-282.

Talmy, Leonard (2000), *Towards a Cognitive Semantics II: Typology and Process in Concept Structuring*, Cambridge: MIT Press.

# Asymmetries in the use of deictic motion verbs in actual and non-actual motion descriptions: Evidence from Finnish and Estonian

Emilia Tuuri & Maija Belliard
(Tampere University & Tampere University)

Keywords: actual motion, non-actual motion, deictic motion verbs, Finnish, Estonian

In our presentation, we discuss asymmetries in the use of deictic verbs in Finnish and Estonian descriptions of actual and non-actual (fictive) motion. We analyze data collected from 50 native Finnish and 25 native Estonian speakers with two elicitation tools: the *Trajectoire* videos (Ishibashi et al. 2006) for actual motion and a picture set (Blomberg 2014) for non-actual motion. The study design allows questions regarding variation between and within the two closely related languages and between actual and non-actual motion. The latter has rarely been discussed from the point of view of deixis.

Our analysis on the Finnish verbs *mennä* 'go' and *tulla* 'come' and the Estonian verbs *minema* 'go, leave' and *tulema* 'come' revealed asymmetrical patterning (Tuuri & Belliard 2021; accepted). In general, ventive verbs tend to focus on the expression of Source, while itive verbs have generic use alongside the expression of Goal (e.g., Wilkins & Hill 1995). In our data, this tendency was stronger in the non-actual motion descriptions relying on a more limited verb selection than the actual motion descriptions (see also Taremaa 2013). Finnish and Estonian also displayed this asymmetry to a different extent, which relates to differences in verb semantics and to variation in lexicalization strategies leading to more frequent and flexible use of deictic verbs in the Estonian data. For example, Estonian participants typically used the itive verb *minema* for videos depicting a rather long walk before arrival at the Goal, as in (1) describing a video of a woman walking on the lawn and stepping into a forest in the end (the verb form *läheb* is a part of the suppletive paradigm of *minema*). In contrast, Finnish participants clearly opted for Manner verbs (e.g., *kävellä* 'walk') in such contexts.

(1) *Naine* ***lähe-b*** *metsa*. (Estonian #19, video 057)
woman go-PRS.3SG forest.ILL
'A woman goes to the forest.'

Furthermore, in both actual and non-actual motion descriptions, asymmetry arises in the expression of different types of Sources and Goals. Specifically, deictic verbs tend to appear in the expression of boundary-crossing situations (entering and exiting) (see (2)), while other verbs are preferred in connection to Sources and Goals without boundary-crossing. Similar connection to boundary-crossing has been reported for ventive verbs in *Trajectoire* data collected from Swedish, German, and Thai participants (Fagard et al. 2016).

(2) *Tie* ***tule-e*** ***ulos*** *tunneli-**sta***. (Finnish #6, picture 024)
road come-PRS.3SG out tunnel-ELA
'A road comes out of the tunnel.'

Deictic verbs appear in boundary-crossing contexts even when the direction of motion is neutral with respect to the speaker's viewpoint (e.g., a video portraying horizontal motion pictured from the side) (e.g., Fagard et al. 2016; Matsumoto et al. 2017). This patterning can possibly be interpreted in terms of functional deixis that associates the use of deictic verbs to the speaker's visual or interactive circle of attention alongside spatial direction (see Matsumoto et al. 2017). Our analysis suggests that these functional properties extend to the use of these verbs in non-actual motion expressions.

Acknowledgments

This work was supported by the Kone Foundation under Grant 201609105 and by the Finnish Cultural Foundation under Grant 00190986.

**References**

Blomberg, Johan (2014), *Motion in language and experience: Actual and non-actual motion in Swedish, French and Thai*. PhD diss., Lund University.

Fagard, Benjamin, Jordan Zlatev, Anetta Kopecka, Massimo Cerruti, and Johan Blomberg (2016), The expression of motion events: A quantitative study of six typologically varied languages, in M. Faytak, M. Goss, N. Baier, J. Merrill, K. Neely, E. Donnelly, and J. Heath (eds), (2016), *Proceedings of the Thirty-Ninth Annual Meeting of the Berkeley Linguistics Society*, Berkeley: Berkeley Linguistics Society, 364–379.

Ishibashi, Miyuki, Anetta Kopecka, and Marine Vuillermet (2006), *Trajectoire: Matériel visuel pour élicitation des données linguistiques*, Laboratoire Dynamique du Langage (CRNS/Université Lyon 2) – Fédération de Recherche en Typologie et Universaux Linguistiques, CRNS, France.

Matsumoto, Yo, Kimi Akita, and Kiyoko Takahashi (2017), The functional nature of deictic verbs and the coding patterns of deixis: An experimental study in English, Japanese, and Thai, in I. Ibarretxe-Antuñano (ed), (2017) *Motion and space across languages*, Amsterdam: John Benjamins, 95–122.

Taremaa, Piia (2013), Fictive and actual motion in Estonian: Encoding Space. *SKY Journal of Linguistics* 26, 151–183.

Tuuri, Emilia, and Maija Belliard (2021), Liikkeen kielentämisen variaatio lähisukukielissä: Väylän alun ja lopun ilmaisu suomessa ja virossa [Variation in encoding motion in closely related languages: expressing beginning and end of Path in Finnish and Estonian], *Finnish Journal of Linguistics* 34, 257–299.

Tuuri, Emilia, and Maija Belliard (accepted for publication in *Acta Linguistica Hafniensia*), Boundary-crossing situations and the use of deictic verbs in Finnish and Estonian expressions of non-actual motion.

Wilkins, David P., and Deborah Hill (1995), When 'go' means 'come': Questions on the basicness of basic motion verbs. *Cognitive Linguistics* 6(2–3), 209–259.

# The goal bias in Classical Armenian

Petr Kocharov
(Julius-Maximilians-Universität Würzburg)

Keywords: Classical Armenian, case marking, goal bias

The paper offers an overview of different manifestations of the Goal bias phenomenon in Classical Armenian (CA). Cross-linguistic studies have shown that the Goal of motion tends to be more salient in discourse than the Source, which often entails asymmetry in the morphosyntactic coding of these participants (Ikegami 1987, Bourdin 1997, Wälchli & Zúñiga 2009, and Kopecka & Vuillermet 2021 with further references). Languages differ in the extent to which the asymmetry is formally expressed. CA represents an intermediate type, which combines clear manifestations of the Goal bias with symmetrical expressions of Goal, Place and Source. The paper addresses three key aspects of this phenomenon: 1) case-marking, 2) prepositional constructions; 3) postpositional constructions.

In CA, the allative (Goal), essive (Place) and ablative (Source) meanings are predominant functions of prepositions *i*, *aṙ* and *ənd* with nouns in the accusative (for allative), locative (for essive), and ablative (for ablative) cases (Meillet 1913, Jensen 1959, and Aṙakʻelyan 2010). Therefore, asymmetries in the morphological complexity and distribution of these expressions can be associated with their prototypical uses to code the Goal, Place, and Source. Tables 1-3 demonstrate manifestations of such asymmetries.

| Decl. | *-o-* | *-a-* | *-i-* | *-u-* | -C- | *-o-* | *-a-* | *-i-* | *-u-* | -C- |
|---|---|---|---|---|---|---|---|---|---|---|
| | Singular | | | | | Plural | | | | |
| Acc | -Ø | -Ø | -Ø | -Ø | -$S_1$-Ø | *-s* | *-s* | *-s* | *-s* | -$S_1$-*s* |
| Loc | -Ø / *-oǰ* / *-um* | -Ø | *-i* | *-u* | -$S_2$-Ø | *-s* | *-s* | *-s* | *-s* | -$S_1$-*s* |
| Abl | *-oy* / *-oǰ-ē* / *-m-ē* | *-ay* | *-ē* | *-ē* / *-u-ē* | -$S_3$-*ē* | *-oc῾* | *-ac῾* | *-ic῾* | *-uc῾* | -$S_2$-*c῾* |

Table 1. The accusative, locative, and ablative cases in Classical Armenian

| | | Prepositions | | |
|---|---|---|---|---|
| | | *i* | *aṙ* | *ənd* |
| Cases | Acc | 812x | 218x | 109x |
| | Loc | 539x | 30x | 159x |
| | Abl | 495x | – | 8x |

Table 2. Frequency of prepositional constructions in the Gospels

| | | Prepositons | | |
|---|---|---|---|---|
| | | Proximal | Medial | Distal |
| Cases | Acc | – | – | 10x (*andr*) |
| | Loc | 5x (*ast*) | – | – |
| | Abl | 11x (*asti*) | 5x (*ayti*) | 77x (*anti*) |

Table 3. Constructions with local postpositions in the Gospels

The accusative of Goal is typically either unmarked or has a comparable morphological complexity as the locative of Place, as opposed to a more complex ablative of Source (Table 1). Tables 2 and 3 contain frequency data, extracted from the UD treebank of Classical Armenian (ca. 82K tokens; https://universaldependencies.org/treebanks/xcl_caval/index.html). Table 2 shows that the relative frequency of prepositional constructions which typically express the Goal and to a lesser extent the Place are common with three different prepositions, whereas only one construction is common for expressing the Source. Thus, the Goal and Place are coded on average by a wider range of prepositions. Table 3 shows the relative frequency of adverbial postpositions, which can code spatial adjuncts cumulatively with the constructions in Table 2. The postposition most often occurs in the constructions typical for the Source, which therefore most often has a morphologically more complex expression. Each of the aforementioned aspects will be discussed in a cross-linguistic perspective.

Acknowledgments
The presented study is part of the "CAVaL: Classical Armenian Valency Lexicon" project funded by the Deutsche Forschungsgemeinschaft (DFG, 518003859).

**References**

Ařak‘elyan, Varag (2010), *Grabari k‘erakanut‘iwn* [A grammar of Old Armenian]. Yerevan: Viamir.

Bourdin, Philippe (1997), On Goal-bias across languages: Modal, configurational and orientation parameters, in B. Palek, O. Fujimura, and J. Neustupný (eds), (2009), *Proceedings of LP’96. Typology, item orderings and universals.* Prague: Karolimun, 185–218.

Ikegami, Yoshihiko (1987), ‘Source’ and ‘Goal’: A case of linguistic dissymmetry, in R. Dirven, and G. Radden (eds), (1987), *Concept of Case*. Tübingen: Narr, 122–146.

Jensen, Hans (1959), *Altarmenische Grammatik.* Heidelberg: Winter.

Kopecka, Anetta, and Vuillermet, Marine (2021), Source-Goal (a)symmetries across languages, *Studies in Language* 45(1), 2–35.

Meillet, Antoine (1913), *Altarmenisches Elementarbuch.* Heidelberg: Winter.

Wälchli, Bernhard, and Zúñiga, Fernando (2009), Source-Goal (in)difference and the typology of motion events in the clause, *STUF – Language Typology and Universals* 59(3), 284–303.

# Ventive vs Itive asymmetry in vertical motion: Chinese Directionals encoding upward motion

Christine Lamarre
(Inalco; CRLAO / Paris)

Keywords: Chinese, ventive, directionals, asymmetry, upward motion

This study investigates an asymmetry pertaining to the Ventive and Itive components of Chinese bimorphemic directionals encoding upward motion. Standard Chinese directionals typically occur in bimorphemic templates, combining a component expressing motion 'in, out, up, down, back, over/across', and a component expressing either Ventive (i.e., towards the Deictic Center (DC)) or Itive (away from the DC) orientation, as in *pǎo-chu-lai* (run-OUT-VEN) 'run out towards the DC'. In Standard Chinese, upward motion is peculiar in that it may be encoded by two distinct directionals: *-shang* encodes Goal-oriented upward motion ($UP_G$) and *-qi* encodes Source-oriented upward motion ($UP_S$, see Liu 1998: 21, Lamarre 2008). Compare for instance:

(1) *pá-shang-lai*
climb-$UP_G$-VEN
'climb up (a slope, a tree) towards the DC'
(2) *pá-qi-lai*
climb-$UP_S$-VEN
'crawl up, get to one's feet (from the floor)'.

Like other Chinese directionals, the Goal-oriented directional *-shang* ($UP_G$) may combine with both Ventive and Itive directionals, as in:

(3) *pǎo-shang-qu*
run-$UP_G$-ITV
'run up away from the DC'

(4) *pǎo-shang-lai*
run-$UP_G$-VEN
'run up towards the DC'.

The Source-oriented *-qi* ($UP_S$) stands out among directionals in that it may only take Ventive *-lai*, as in *zhàn-qi-lai* (stand-$UP_S$-VEN) 'stand up'. Based on the distribution of *-qi* 'up' and *-shang* 'up' in a written literary corpus (4 modern novels), we propose several explanations to account for this asymmetry. Note that verbs taking *-qi* ($UP_S$) often "incorporate UP in their most unmarked use" (Viberg 2021 about some Swedish verbs taking *upp*). 'Walk up' takes-*shang*, while 'raise one's hand' takes *-qi.*

(1) Unlike other directionals, *-qi* 'up' does not encode any translational motion, where an entity's "basic location shifts from one point to another" (Talmy 2000:35-36), and is therefore not a **path** directional. Aurnague's classification of French motion verbs (2011) also excludes change of posture verbs from "motion verbs" in a narrow sense. We argue that Deictic orientation may be irrelevant for *-qi* ($UP_S$) because it specializes in upward changes of posture or in body part motion: the most frequent verbs cooccurring with *-qi* are *zhàn* 'stand' for autonomous motion, and 'raise' (head, arm) for caused motion,

as in *tái-qi-lai* (raise-UP$_S$-VEN) (Liu 1998: 336; 374). The Ventive directional *-lai* could here be a mere "slot-filler", added under the influence of the pervasive bimorphemic template of Chinese directionals.

(2) In the case of caused motion, the choice of the Ventive directional may nevertheless be motivated: with inanimate objects, e.g., *jiǎn-qi-lai* (pick.up-UP$_S$-VEN) 'pick up' (something on the ground), or *ná-qi-lai* (take-UP$_S$-VEN) 'lift' (a suitcase), the verbs indicate that the object enters the personal domain of the agent (contact with a body part). This is consistent with a motion towards the DC. Note that *-qi* expresses motion up within the range of body movements. A comparison with the (single) directional encoding downward motion, *-xia*, is provided to support this analysis.

To conclude, we consider this phenomenon from the angle of Goal-Source asymmetry: both Itive and Ventive orientation are possible for Goal-oriented upward motion. This choice is not available for Source-oriented upward motion.

Acknowledgements: French ANR Project SALTA (SPATIAL ASYMMETRIES ACROSS LANGUAGES: A TYPOLOGICAL APPROACH, 2021-2025)

**References**

Aurnague, Michel (2011), How motion verbs are spatial: The spatial foundations of intransitive motion verbs in French. *Lingvisticae Investigationes*, Philadelphia; Amsterdam: John Benjamins, 34(1), 1-34.

Lamarre, Christine (2008), The Linguistic Categorization of Deictic Direction in Chinese: With Reference to Japanese. In D. Xu (ed.) *Space in Languages of China: Cross-linguistic, synchronic and diachronic perspectives*. Dordrecht: Springer, 69-97.

Liu, Yuehua (1998), *Qūxiàng bǔyǔ tōnshì* [A compendium of Directional Complements], Beijing: Beijing Language University Press.

Talmy, Leonard (2000), *Towards a cognitive semantics*. *Vol. II – Typology and Process in Concept Structuring*. Cambridge, MA: MIT Press.

Viberg, Åke (2021), Moving up and down in real space**.** The verbal particles *upp* 'up' and *ner* 'down' and the typological profile of Swedish. In Anna Čermáková, Signe Oksefjell Ebeling, Magnus Levin and Jenny Ström Herold (eds) *Crossing the Borders: Analysing Complex Constrastive Data*. *BeLLS* 11(1), 69-90.

# Spatial asymmetries and variation: On the role of the Ground in Place, Goal, and Source constructions

Julia Nintemann
(University of Bremen)

Keywords: spatial relations, types of Ground, typology, asymmetries, variation

It has already been recognized in linguistic research that there are asymmetries in the coding of spatial relations, especially between the two dynamic relations Goal and Source (cf., e.g., Ikegami 1987, Bourdin 1997, or, more recently, Kopecka & Vuillermet 2021). However, it still has to be investigated to what extent the type of Ground (cf. Talmy 1978) influences these (a)symmetries. This talk aims to fill this gap by analyzing different types of Grounds in Place, Goal, and Source constructions. It will be shown that asymmetries and also variation in the coding of spatial relations do indeed often occur with different types of Ground. The Ground position may be filled by common nouns, proper names, spatial adverbs, but also personal and interrogative pronouns. In many languages, the type of Ground determines which kind of construction is used to express Place, Goal, or Source. As demonstrated by Stolz et al. (2014), for example, toponyms are more often zero-marked in Ground position than common nouns cross-linguistically. Stolz & Nintemann (2024) take up this issue and show that name categories often obey rules different from those of common nouns – not only but especially – when they are involved in constructions with spatial meanings. However, it is not only names that follow their own rules. Incorporating other types of Ground into this comparison, it will be shown that many languages have specific rules for different categories in constructions of spatial relations.

In Gurindji [Pama-Nyungan], for example, there is both asymmetry and variation in the coding of Goal, which is generally expressed with an allative suffix (e.g., *kaja-**ngkurra*** [land-**ALL**] '**to** the land'; Meakins & McConvell [2021: 391]). However, depending on the type of Ground, different rules apply. Toponyms are the only category that allows for zero-marking (e.g., *Jinparrak-**Ø*** [Jinparrak-**Ø**] '**to** Jinparrak'; Meakins & McConvell [2021: 493]). Animate goals, on the other hand, usually take the dative suffix (e.g., *nyununy-ku jaju-**wu*** [your-DAT grandmother-**DAT**] '**to** your maternal grandmother'; Meakins & McConvell [2021: 76]). With spatial interrogatives, Goal is more often expressed with the locative marker (*wanyji-**ka*** [which-**LOC**] 'where') than with the allative marker (*wanyji-**kurra*** [which-**ALL**] 'where **to**') (cf. Meakins & McConvell [2021: 309]), although both forms exist and the locative marker cannot be used to mark a Goal with other types of Ground. There is thus asymmetry in Gurindji Goal constructions in that there is one type of Ground that allows for zero-marking as opposed to the obligatory overt marking of other Grounds. Furthermore, there is variation as different case markers are used for the same purpose with different types of Grounds.

The talk will show that asymmetries and variation in the coding of Place, Goal, and Source relations depending on the type of Ground are common phenomena in languages all over the world. For this purpose, a sample of 50 languages from different areas and of different affiliations will be examined with regard to spatial relations from a functional-typological perspective. The data will primarily be taken from descriptive grammars, complemented by primary sources and input from language experts.

**References**


Bourdin, Philippe (1997), On Goal-bias across languages: Modal, configurational and orientation parameters, in B. Palek (ed.), *Proceedings of LP'96: Typology, item orderings and universals,* (Proceedings of the Conference held in Prague, August 20-22, 1996), Prague: Karolinum - Charles University Press, 186–218.

Ikegami, Yoshiko (1987), 'Source' and 'Goal': A case of linguistic dissymmetry, in R. Dirven, and G. Radden (eds.), *Concept of case*, Tübingen: Gunter Narr Verlag, 122–146.

Kopecka, Anetta and Marine Vuillermet (2021), Source-Goal (a)symmetries across languages, *Studies in Language* 45(1), 2–35.

Meakins, Felicity and Patrick McConvell (2021), *A grammar of Gurindji: As spoken by Violet Wadrill, Ronnie Wavehill, Dandy Danbayarri, Biddy Wavehill, Topsy Dodd Ngarnjal, Long Johnny Kijngayarri, Banjo Ryan, Pincher Nyurrmiari and Blanche Bulngari* (Mouton Grammar Library 91), Berlin, Boston: De Gruyter.

Stolz, Thomas and 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, and Sander Lestrade (2014), *The Crosslinguistics of Zero-Marking of Spatial Relations* (Studia Typologica 15), Berlin: De Gruyter.

Talmy, Leonard (1978), Figure and ground in complex sentences, in J. Greenberg, C. A. Ferguson, and E. A. Moravcsik (eds.), *Universals of human language: Volume 4: Syntax*, Palo Alto/Ca.: Stanford University Press, 625–649.

# Spatial asymmetries in Path and Manner encoding: A comparative study of narratives in English and French

Adélie Buclier, Anetta Kopecka & Efstathia Soroli
(Université de Lille, Université Lumière Lyon II & Université de Lille)

Key words: spatial language – spatial asymmetries – lexicalization patterns – English – French

Languages differ importantly in the means they offer to express Path and Manner of motion (cf. Verb-framed vs. Satellite-framed languages) (Talmy 2000, Hickmann et al. 2017). Some recent studies focus on specific sub-categories of these components and have shown asymmetries revealing both across- and within-language variation (Pourcel and Kopecka, 2005) such as: (i) Path asymmetries in Source/Goal expression (i.e., high variation and frequency of devices that express the Goal, as in French *de* 'from' vs. *à* 'to'*, vers* 'towards'*, jusqu'à* 'all the way to') (Ikegami 1987, Kopecka and Ishibashi 2011, Lakusta and Landau 2012, Vuillermet 2020, Johanson et al. 2019, Kopecka and Vuillermet 2021); (ii) Manner asymmetries in Fast/Slow Motion expression (i.e., use of a diverse lexicon to express speed/Fast Motion in a more frequent way than Slow Motion, as in Estonian or Russian) (Taremaa and Kopecka 2023, Plungian and Rakhilina 2013); and (iii) presence of some privileged combinations (e.g., in Estonian, Fast Motion verbs that tend to combine more often with Goal) (Taremaa and Kopecka 2023). Despite a growing interest in the expression of Path and Manner components, the role of their specific sub-categories in spatial asymmetries is yet to be finely explored.

The present corpus study focuses on Path, Manner, and their sub-categories (Source/Goal; Fast/Slow motion respectively) in English and French narratives, and aims to examine: (i) the relative weight of these components as expressed in verbs, satellites (e.g., particles, prefixes), and other peripheral devices, (e.g., prepositions, adverbials); (ii) their asymmetries in terms of frequency, morphosyntactic complexity, and semantic granularity; and (iii) their potential combinations.

The analysis based on the Frog Story narratives (Mayer, 1969) of 24 adult speakers (12 per language) shows that, in this dataset:

(i) Lexicalization patterns: Path is more often expressed in verbs in French and in satellites in English. On the other hand, Manner is expressed in both verbs and peripheral devices in French but is exclusively lexicalized in English. These statements confirm previous findings about the expression of these two components in Romance and Germanic languages (Talmy 2000, Hickmann et al. 2017);

(ii) Asymmetries: In both languages, the Source is the predominant component of Path, which is expressed in a more frequent and more diverse way than the Goal. More specifically, English speakers primarily use peripheral devices encoding the Source (e.g., *away, off, out*), and occasionally use Goal markers (e.g., *in, into*). Similarly, English speakers show a preference for Source verbs over Goal verbs. French speakers follow a similar pattern, while also conflating Manner and Path (e.g., *s'évader* 'to escape'*, s'enfuir* 'to run away'*, s'échapper* 'to escape'). As regards the expression of Speed, French speakers favor Fast Motion in terms of both frequency and diversity of linguistic devices. English speakers show a great diversity of usage of Slow Motion verbs – a finding that differs from previous observations with respect to the salience of Goal and Fast Motion (Ikegami 1987, Stefanowitsch and Rohde 2004, Stefanowitsch 2018);

(iii) Combinations: In both languages, the preferred combination is the expression of Fast Motion with Source markers (e.g., *to run away, 's'enfuir'*), as opposed to what has been previously observed in other languages (e.g., for Estonian in Taremaa and Kopecka 2023).

The present study discusses current research on Source/Goal and Fast/Slow Motion asymmetries and investigates their language-specific characteristics and manifestations in a fine-grained fashion.

**References**


Hickmann, M., Engemann, H., Soroli, E., Hendriks, H. & Vincent, C. (2017). Expressing and categorizing motion in French and English: Verbal and non-verbalcognition across languages. In I. Ibbartexe- Antunano (ed.), *Motion and space across languages. Theory and applications*, 61-94. Amsterdam: John Benjamins.

Ikegami, Y. (1987). 'Source' vs 'Goal': A case of linguistic dissymmetry. In Dirven R. & Radden G. (eds.) *Concepts of case*, 122–46. Tübingen: Günter Narr Verlag.

Johanson, M., Selimis, S., & Papafragou, A. (2019). The Source-Goal asymmetry in spatial language: language-general vs. language-specific aspects. *Language, Cognition and Neuroscience*, *34*(7), 826-840.

Kopecka, & Ishibashi, M. (2011). L'(a-)symétrie dans l'expression de la Source et du But : perspective translinguistique. *Faits de Langues*, 38(2), 131– 149.

Kopecka, A., & Vuillermet, M. (2021). Source-Goal (a)symmetries across languages. *Studies in Language*, *45*(1), 2-35.

Lakusta, L. & Landau, B. (2012). Language and Memory for Motion Events: Origins of the Asymmetry Between Source and Goal Paths. *Cognitive Science 36*, 517- 544.

Mayer, M. (1969). *Frog, where are you?* New York: Dial Press.

Plungian, V. & Rakhilina, E. (2013). Time and speed: Where do speed adjectives come from? *Russian Linguistics*, 37(3), 347–359.

Pourcel, S., & Kopecka, A. (2005). Motion expression in French: typological diversity. *Durham & Newcastle working papers in linguistics*, *11*, 139-153.

Stefanowitsch, A., & Rohde, A. (2004). The goal bias in the encoding of motion events. In Radden G. & Panther K-U. (eds.), *Studies in linguistic motivation*, 249–267. Berlin/New York: Mouton de Gruyter.

Stefanowitsch, A. (2018). The goal bias revised: A collostructional approach. *Yearbook of the German Cognitive Linguistics Association* 6(1), 143-166.

Talmy, L. (2000). *Toward a Cognitive Semantics*. Cambridge MA: MIT Press.

Taremaa, P., & Kopecka, A. (2023). Manner of motion in Estonian: A descriptive account of Speed. *Studies in Language*, 47:1, pp. 32-78.

Vuillermet, M. (2020). Source-Goal asymmetries in Ese Ejja. *Studies in Language,*45(1), 235–275.

# Spatial asymmetries in contemporary dance instructions

Chiara Minoccheri
(CLLE Laboratory (UMR 5263); University Toulouse Jean Jaurès; CNRS)

Keywords: contemporary dance oral instructions, manner of motion, spatial semantics, speed, source *vs* goal

For this workshop I will look at some aspects of spatial asymmetries in a corpus of 1200 French motion instructions (10348 words) coming from four advanced-level contemporary dance classes.

The first aspect I will explore concerns the phenomenon known as "Goal-bias" (Ikegami, 1987) in motion events that imply a change of placement of the moving entity (*sortir* 'to get out', *aller à* 'to go to', *courir* 'to run', *marcher* 'to walk') (Aurnague, 2011). As expected, dance data show some features related to the Goal-bias, namely a more frequent expression of the final landmark ("Goal") (*on revient au centre* 'we return to the centre') than of the initial landmark (*les mains se sépareraient du corps* 'the hands would separate from the body'). This result not only confirms that dance data reflect a general tendency observed in languages (Bourdin, 1997; Lakusta & Landau, 2005; Regier & Zheng, 2007), but it can also be interpreted as supporting the pragmatic principle highlighted by Aurnague (2019): dancers (like people in general) need to be told where to go (final landmark) throughout the motion process, while they are perfectly aware of their location at the beginning of the movement (initial landmark).

The second issue I will focus on is the fast *vs.* slow asymmetry in the expression of speed, whether it is encoded in the verb semantics or conveyed by manner adverbials. Speed is one of the parameters that trigger Manner (Stosic, 2019), and it is the fourth most frequent (over 16) in the analysed corpus. In the verbal domain, fast motion is more frequent than slow motion, but, in terms of items, only two of the five verbs conveying speed express fast motion (*s'élancer* 'set off', *lancer* 'throw'). Nevertheless, the three verbs denoting slow motion all derive from the same root (*déposer*), and the coding of slowness is limited to specific usages found within the choreographers' utterances as documented in this corpus (*le front se dépose* 'the forehead is laid down', *redéposez l'omoplate* 'lay down the shoulder blade'). Concerning speed adverbials, only two tokens were found and they express slow motion, although a predominance and greater variety of fast motion expressions was expected (Taremaa & Kopecka, 2022).

Dance data offer a huge amount of varied and creative motion expressions in which different types of asymmetries can be investigated. In addition to the two mentioned above, which are more common outside the dance domain, the expression of forward (*on va vers l'avant* 'we move forward') *vs.* backwards (*vous marchez vers l'arrière* 'you walk backwards') direction and centripetal (*revenez au centre* 'come back to the centre') *vs.* centrifugal (*on va aller sur le côté droit ici* 'we are going to the right side') motion can be investigated. By comparing the utterances of the four choreographers with regard to these asymmetries, this presentation will provide some insights into their different conceptualisations of movement through language across a corpus of attested, spontaneous and spoken dance instructions that is nowadays still unexplored in the field of semantics.

**References**


Aurnague, M. (2011). How motion verbs are spatial: The spatial foundations of intransitive motion verbs in French. *Lingvisticae Investigationes*, *34*(1), 1-34. https://doi.org/10.1075/li.34.1.01aur

Aurnague, M. (2019). About asymmetry of motion in French: Some properties and a principle. In M. Aurnague & D. Stosic (Éds.), *The Semantics of Dynamic Space in French* (p. 31-66). John Benjamins. https://doi.org/10.1075/hcp.66.01aur

Bourdin, P. (1997). On goal-bias across languages: Modal, configurational and orientational parameters. In *Proceedings of LP '96: Typology, prototypes, item orderings and universals, conference of Prague, August 20-22, 1996* (p. 185-218). na.

Ikegami, Y. (1987). Source'vs.'goal': A case of linguistic dissymmetry. In R. Dirven & G. Radden (Éds.), *Concepts of case* (p. 122-146). Gunter Narr Verlag.

Lakusta, L., & Landau, B. (2005). Starting at the end: The importance of goals in spatial language. *Cognition*, *96*(1), 1-33.

Regier, T., & Zheng, M. (2007). Attention to Endpoints: A Cross-Linguistic Constraint on Spatial Meaning. *Cognitive Science*, *31*(4), 705-719. https://doi.org/10.1080/15326900701399954

Stosic, D. (2019). Manner as a cluster concept: What does lexical coding of manner of motion tell us about manner? In M. Aurnague & D. Stosic (Éds.), *The semantics of dynamic space in French* (Vol. 66, p. 142-177). John Benjamins Publishing Company. https://doi.org/10.1075/hcp.66.04sto

Taremaa, P., & Kopecka, A. (2022). Speed and space: Semantic asymmetries in motion descriptions in Estonian. *Cognitive Linguistics*.

# Uncommon ground: ground conflation in Naduhup motion verbs

Karolin Obert
(Lund University)

Keywords: Motion; Amazonian languages; Ground; GIS; Typology

Cross-linguistic research has extensively explored how languages encode motion events, revealing diverse patterns across different linguistic systems (cf., Talmy 1985; Slobin 2004; Bohnemeyer et al. 2006; Goschler & Stefanowitsch 2013). While much attention has been devoted to the encoding of Manner and Path components, Ground— the object or surface in relation to which the motion takes place—has received comparatively limited attention. In fact, the possibility to be a structuring part of motion verb inventories has been previously excluded in typological studies, treating verbs like 'to box' as outliers in semantic theory (Talmy 2000:60-62). However, recent advancements in data collection methods and insights from underrepresented languages challenge this perspective (e.g., Wnuk 2016; Burenhult & Purves 2020). For example, data collected with the Dâw and Nadëb people (Naduhup, Amazonia) while walking could reveal how properties of the terrain can systematically structure the basic motion verb inventory as displayed in (1). This raises questions with respect to the prior predicted impossibility of Ground conflation in motion verbs in motion event typology, and, allows for insights into less conventional ways of how languages can structure and represent the domain of motion. Specifically, I ask: What type of grounds are conflated in Naduhup motion verbs and (how) do they map onto identifiable characteristics of the terrain? And, what is the grammatical design space of these ground-encoding motion verbs?

| | | | | |
|---|---|---|---|---|
| (1) | *tén* | *id* | ***bəʃ*** | *xəd* |
| | now | 1PL | cross.waterway | PROG |
| | 'Now, we are crossing (the creek)' | | | Dâw |

Here, I present the results of an investigation into Naduhup representations of motion in context, i.e., while speakers are walking. Specifically, speakers of Dâw and Nadëb were videorecorded during foraging trips with action cameras in tandem with a time-aligned GIS record, allowing for posterior analysis of motion verb instantiations in the landscape. Additional data comes from speaker descriptions and labelling of these motion events, as well as from our corpus of narratives and conversations.

The results indicate that basic motion verbs systematically combine the general Motion component with specific Ground information represented by landscape elements. These features often reflect hydrological aspects, such as movement with respect to bodies of water, or other topographic features like clearings and slopes. Additionally, all motion verbs encode Path, indicating directionality with respect to these topographic features, including horizontal (up and down), vertical (across, upriver/downriver), and goal/source-orientation. Further evidence for the deep integration of the Ground component in basic motion verbs comes from the systematic absence of Ground-denoting locative adjuncts with these motion verbs (see (1)).

Finally, the findings underscore how novel ways of *in-situ* data collection can provide us with typological surprises revealing intangible Indigenous categories as present in language and grammars, which account for the interactive, cultural, and historical contexts in which these languages are spoken.

This also opens windows into what is conceptually important to speakers as visible in grammatical and lexical systems (see Talmy 1983; Heine 1997; Enfield 2002; Palmer et al. 2017), and shows that linguistic structure, cultural and environmental aspects can play a role in lexical elaboration.

**References**


Beavers, John; Levin, Beth; Wei Tham, Shiao (2010), The typology of motion expressions revisited, *Journal of Linguistics, 46*(2), 331—377.

Bohnemeyer, Jürgen; Enfield, Nicholas J.; Essegbey, James; Ibarretxe-Antuñano, Iraide; Kita, Sotaro; Lüpke, Frederike; Ameka, Felix (2007), Principles of event segmentation in language: The case of motion events, *Language*, 83(3), 495—532.

Burenhult, Niclas; Purves, Ross (2020), The spatial properties of forager motion categories: Evidence from Jahai, *Hunter Gatherer Research*, 4(3), 411—426.

Enfield, Nicholas J. (2002), Ethnosyntax: Introduction. In N. J. Enfield (ed.), (2002), *Ethnosyntax: Explorations in culture and grammar,* Oxford: Oxford University Press, 1–30.

Goschler, Juliana; Stefanowitsch, Anatol (2013), *Variation and Change in the Encoding of Motion Events*, John Benjamins.

Heine, Bernd (1997), *Cognitive foundations of grammar*, New York: Oxford University Press.

Palmer, Bill; Lum, Jonathon; Schlossberg Jonathan; Gaby, Alice (2017), How does the environment shape spatial language? Evidence for sociotopography, *Linguistic Typology,* 21(3), 457—491.

Talmy, Leonard (1983), How language structures space. In H. L. Pick, and L. P. Acredolo (eds), *Spatial orientation: Theory, research, and application*, New York: Plenum, 225–282.

Talmy, Leonard (1985), Lexicalization patterns: semantic structure in lexical forms. In T. Shopen (ed), *Language typology and syntactic description*, Cambridge: Cambridge University Press, 36—149.

Talmy, Leonard (2000), *Toward a cognitive semantics*. Vol 2, Cambridge: MIT Press.

Slobin, Dan (2004), The many ways to search for a frog: Linguistic typology and the expression of motion events. In L. Verhoeven, and S. Stromqvist (eds), *Relating events in narrative: Vol. 2. Typological and Contextual Perspectives*, Mahwah: Lawrence Erlbaum Associates, 219—257.

Wnuk, Ewelina (2016), *Semantic specificity of perception verbs in Maniq.* PhD Thesis, Radboud University Nijmegen, Nijmegen.

Zlatev, Jordan; Blomberg, Johan; Devylder, Simon; Naidu, Viswanatha; van de Weijer, Joost (2021), Motion event descriptions in Swedish, French, Thai and Telugu: a study in post-Talmian motion event typology, *Acta Linguistica Hafniensia*, 53(1), 58—90.

# With or without :
# Asymmetries in meaning components of manner of motion verbs

Sarda Laure
(Lattice CNRS; ENS-PSL; Université Sorbonne Nouvelle)

Keywords : Motion event; Manner/Path; asymmetries in directionality; Force/ speed/ effort/ control; Aspect

I explore the manner-of-motion verbs in French, exemplified by *courir* (run), and their possible boundary-crossing interpretation as in *Jean a couru à la plage*. Various authors, including (Kopecka 2009, Stosic 2009, Sarda 2019 and Laporte & al. fc), have explored this shift in meaning in verb-framed languages, resulting in a potential boundary-crossing reading of atelic motion verbs. The question arises: if neither the verb (*courir*) nor the static preposition (*à*) encodes Path, where does the boundary-crossing interpretation come from?

Aurnague (2011, 2019, fc) extensively addresses this issue, identifying three key meaning components — speed, direction, and force — grouped under the concept of *tendentiality*. These components influence the telic interpretation of manner-of-motion verbs, explaining the shift in meaning. Furthermore, Stosic (2009, 2019) offers a detailed analysis of the manner encoding in motion verb lexical meanings, defining fourteen parameters, the most frequent being *motor pattern*, *speed*, *Figure (re)configuration*, and *force power*. Most of these parameters can be encoded in both telic and atelic verbs.

The theoretical contributions of Aurnague on the one hand, and Stosic on the other hand, lay the groundwork for motion description, offering a conceptual toolbox with potential meaning components that can be associated with the motion component. They provide a lexical decomposition of the notion of Manner and Path related to a gradient of dynamicity. Departing from the prevailing notion of manner/result complementarity (Levin & Rappaport 2013), I argue in this framework that a single verb root can lexicalize both manner and path components.

To scrutinize this conceptual shift, a corpus study is conducted on a series of French motion verbs. The focus is on [Verb + PP] constructions with a particular emphasis on two sets of motion verbs: *marcher* (walk) and *courir* (run) on the horizontal axis, and *grimper* (climb up) and *dévaler* (downhill) on the vertical axis. The investigation entails a quantitative overview of the data (extracted from the Intercorp Korpus (Čermák and Rosen 2012, Nádvorníková et al. 2012), and Frantext database), and a detailed analysis of examples and counter-examples, considering the semantics components of the verbs. Additionally, attention is paid to co-textual elements, such as adverbials, which may activate tendentiality features contributing to the boundary-crossing reading.

The study investigates the impact of asymmetries associated with directionality in the interpretation of manner verbs of motion. In exploring the co-textual elements and tendentiality features, the study uncovers asymmetrical patterns that influence the boundary-crossing interpretation. For example, on the horizontal axis, asymmetry may be evident in the contrast between fast and non-fast motions. Similarly, on the vertical axis, *grimper* (climb) and *dévaler* (downhill) present distinct asymmetrical tendencies with regard to effort and control. By scrutinizing these directional asymmetries, the study contributes to a more comprehensive understanding of how the semantics components and co-textual elements interact in shaping the boundary-crossing interpretation of manner-of-motion verbs.

This works is supported by the ANR ANR-20-CE27-0015-01 SALTA : *Spatial Asymmetries Across Languages: A Typological Approach*.

**References**

Aurnague, Michel (2011). How motion verbs are spatial: the spatial foundations of intransitive motion verbs in French. Lingvisticae Investigationes, 34(1), 1–34.
https://doi.org/10.1075/li.34.1.01aur

Aurnague, Michel (2019). About asymmetry of motion in French: Some properties and a principle. In M. Aurnague & D. Stosic (Eds.), The semantics of dynamic space in French: Descriptive, experimental and formal studies on motion expression (pp. 31–65). Amsterdam & Philadelphia: John Benjamins.

Aurnague, Michel (forthcoming) Telic motion constructions in French and the notion of tendentiality.

Čermák, František and Rosen, Alexandr (2012): The case of InterCorp, a multilingual parallel corpus. In *International Journal of Corpus Linguistics*, 17(3), 411–427.

Kopecka, Anetta (2009). L'expression du déplacement en français: l'interaction des facteurs sémantiques, aspectuels et pragmatiques dans la construction du sens spatial. Langages, 173, 54–75. https://doi.org/10.3917/lang.173.0054

Laporte Eric, Leclère Christian, Marque-Pucheu Chrsitiane and Nakamura Takuya (forthcoming). J'ai couru au bar 'I ran to the bar': A purpose-oriented verb of manner of motion. In E. Corre & L. Sarda (Eds.), *Neglected aspects of motion-event description: deconstructing motion events*. Amsterdam & Philadelphia: John Benjamins.

Levin, Beth and Rappaport Hovav, Malka (2013). Lexicalized Meaning and Manner/Result Complementarity'. In B. Arsenijević, B. Gehrke, & R. Marín, (Eds.), Subatomic Semantics of Event Predicates (pp. 49–70), Springer, Dordrecht.

Nádvorníková, Olga, Vavřín, Martin and Zasina, Adrian Jan: *Korpus InterCorp – francouzština, verze 15 z 11. 11. 2022*. Ústav Českého národního korpusu FF UK, Praha 2022. Dostupný z WWW http://www.korpus.cz

Sarda, Laure (2019). French motion verbs: insights into the status of locative PPs. In M. Aurnague M. & D. Stosic (Eds.), The semantics of dynamic space inFrench: descriptive, experimental and formal studies on motion expression (pp. 67–107). Amsterdam & Philadelphia: John Benjamins.

Stosic, Dejan (2009). La notion de « manière » dans la sémantique de l'espace. Langages, 175, 103–121. https://doi.org/10.3917/lang.175.0103

Stosic, Dejan (2019). Manner as a cluster concept: What does lexical coding of manner of motion tells us about manner? In M. Aurnague & D. Stosic (Eds.), The semantics of dynamic space in French: Descriptive, experimental and formal studies on motion expression (pp. 141–177). Amsterdam & Philadelphia: John Benjamins.

Corpus :

Frantext : https://www.frantext.fr/

Intercorp : https://www.korpus.cz

# Source, Trajectory, and Vehicle syncretism in Ryukyuan

Aleksandra Jarosz
(Nicolaus Copernicus University)

Keywords: Ryukyuan, ablative, Source, Trajectory, Vehicle

The goal of this communication is an analysis of the polysemic use of the Proto-Japonic marker *-kara in modern Ryukyuan languages. The essential function of descendants of *-kara shared in virtually all known Ryukyuan varieties, North and South alike, is Source, cf.:

| | | | | | | |
|---|---|---|---|---|---|---|
| (1) | Ffima-kara | basu-n | nuːr-i | Psara-nu | Izzatu-n | uri-z |
| | Kurima-ABL | bus-DAT | get.on-CVB | Hirara-GEN | Nishizato-DAT | descend-CNCL.NPST |

‘[You] get on the bus from Kurima and get off in Nishizato in Hirara.’
(Kurima-Miyako; author’s fieldwork, 2019)

Source in individual Japonic varieties, including Ryukyuan lects, has been observed to be encoded maximally by two markers: all varieties use ablative originating from *-kara, and some also employ accusative specifically to indicate interelative (e.g. Nakachi *bass-u uri-z* bus-ACC descend.NPST-CNCL ‘to get off of a bus’, Tomihama 2013: 83).

The marking of Source and Goal in Ryukyuan shows asymmetry referred to in the literature as the “goal bias” (Bourdin) or “goal-over-source principle” (Ikegami; cf. Stefanowitsch 2018: 143, Luraghi et al. 2017). All examined Ryukyuan varieties have more Goal than Source markers. Source seems to be encoded maximally by two markers: all varieties use ablative originating from *-kara, and some also employ accusative specifically to indicate interelative (e.g. Nakachi *bass-u uri-z* bus-ACC descend.NPST-CNCL ‘to get off of a bus’, Tomihama 2013: 83). In contrast, individual lects often have multiple Goal markers. To exemplify this disproportion, North Ryukyuan Yoron has allative *-kati*, allative/locative *-nan* and mostly essive *-nonti* with their allomorphs, directive *-nnai*, limitative (“goal extension” in Bourdin’s 1997 terms) *-tana*/*-ntana*, and, sporadically, an allative/essive use of dative *-n*, contrasted with the single marker of Source, ablative *-kara* (cf. Nohara 1998, Kiku 2014).

This single marker displays a specific spatial polisemy pattern in a significant number of Ryukyuan topolects. Namely, descendants of *-kara can be used as perlative, signifying Trajectory (02), and as motion instrumental, signifying Vehicle (03), which refers to the means of transport employed in a motion event. Crucially, ablative is not used as a general instrumental, despite this being a cross-linguistically widespread pattern (Kuteva et al. 2019: 37–38) – this polisemy is therefore strictly tied to the spatial properties of the ablative.

| | | | | |
|---|---|---|---|---|
| (2) | tiːra-kara | ʔattɕ-iːneː | tɕimburu | jami-n-roː |
| | sun-ABL | walk-CND | head | hurt-NPST-EMP |

‘If you walk under the sun, you are going to get a headache.’
(Maja-Okinawan; Nohara 1998: 270)

(3) ʔUtɕina:-katɕi-ja funi-kara-ru ʔits-uru
Okinawa-ALL-TOP ship-ABL-FOC go-ATR.NPST
'I'm going to Okinawa by ship.' (Maja-Okinawan; Nohara 1998: 269)

Like reflexes of Source-expressing *-kara, the pattern is attested in the entire Ryukyuan-speaking area, although less pervasively. Topolects which have one of these semantic extensions of the ablative often also have the other, which implies a diachronic link between the two.

There may be a common centrifugal denominator between these Source, Trajectory and Vehicle roles encoded by the Ryukyuan ablative. Trajectory and Vehicle are perhaps conceptualized as signifying an outward movement, focusing respectively on the "going *from* through", "going *from* with/by the means of" aspects of the motion.

Such metaphoric productivity of Source rather than Goal in Trajectory and Vehicle conceptualizations are surprising in the light of the "goal bias". Perhaps it represents what Heine and Stolz (2008: 343) call "experiential metaphor", that is, a description of "abstract or conceptually complex phenomena [moving through a space, or moving with the assistance of a means – AJ] in terms of concrete or less complex phenomena [out-movement – AJ]". This implies Source to be a cognitively robust device of semantic extensions. The strikingly limited Ryukyuan variation in case-marking means of expressing Source as opposed to Goal, which apparently results from a human's universal behavioral and/or perceptual emphasis on Goal over Source (Lakusta and Landau 2005: 30–31), does not hamper the metaphoric flexibility of Source.

Abbreviations

| | | | |
|---|---|---|---|
| ABL | ablative | DAT | dative |
| ACC | accusative | EMP | emphatic |
| ALL | allative | FOC | focus |
| ATR | attributive | GEN | genitive |
| CNCL | conclusive | NPST | non-past |
| CND | conditional | TOP | topic |
| CVB | converb | | |

Acknowledgments: This research was supported by the funding "Excellence Initiative – Debuts IV", grant no. [4101*66, 03.01*2356] as part of the project "Case marking and valence patterns in Ryukyuan: a typological and comparative study".

**References**

Bourdin, Philippe (1997), On Goal-bias across languages. Modal, configurational and orientation parameters, in Bohumil Palek, Osamu Fujimura, and Jiři Václav Neustupný (eds.), (1997), *Proceedings of LP'96. Typology, item orderings and universals*, Prague: Karolinum, 185–218.

Heine, Bernd, and Thomas Stolz (2008), Grammaticalization as a creative process, *Language Typology and Universals (STUF)* 61, 326–357.

Kiku Hidenori 菊秀史 (2014), *Yunnu futuba-de hanasō 4. Keiyōshi, joshi, hyōgen ito* 与論の言葉で話そう 4 形容詞、助詞、表現意図 [Let's talk in the Yoron speech 4. Adjectives, particles, expressions of intention], Yoron: Yoron Minzokumura.

Kuteva, Tania, Bernd Heine, Bo Hong, Haiping Long, Heiko Narrog, Seongha Ree (2019), *World lexicon of grammaticalization. Second, extensively revised and updated edition*, Cambridge: Cambridge University Press.

Lakusta Laura, and Barbara Landau (2005), Starting at the end: the importance of goals in spatial language, *Cognition* 96, 1–33.

Luraghi, Silvia, Tatiana Nikitina, Chiara Zanchi (2017), Space in diachrony: an introduction, in: Silvia Luraghi, Tatiana Nikitina, Chiara Zanchi (eds.), (2017), *Space in diachrony*, Amsterdam/Philadelphia: John Benjamins Publishing Company, XI–XVII.

Nohara Mitsuyoshi 野原三義 (1998), *Shimpen Ryūkyū hōgen joshi-no kenkyū* 新編　琉球方言助詞の研究 [a study of particles in Ryukyuan dialects, new edition], Tokyo 東京都: Kumaya Insatsu 熊谷印刷.

Stefanowitsch, Anatol (2018), The goal bias revisited: A collostructional approach, *Yearbook of the German Cognitive Linguistics Association* 6, 143–166.

Tomihama Sadayoshi 富浜定吉 (2013), *Miyako Irabu hōgen jiten* 宮古伊良部方言辞典 [a dictionary of the Irabu-Miyako dialect], Naha 那覇市: Okinawa Taimusu-sha 沖縄タイムス社.

# Structural (a)symmetries in the spatial case systems of Permic languages

Riku Erkkilä
(Ludwig-Maximilians-Universität München; University of Helsinki)

Keywords: source-goal asymmetry, case system, spatial cases, Permic languages, semantics

In this paper, I discuss the spatial case systems of the three Permic languages (Udmurt, Zyrian, and Permyak; a branch of Uralic languages). These languages possess notably large spatial case inventories. Literary Udmurt has seven spatial cases (Перевощиков et al., 1962), and literary Zyrian has eight (Бубрих, 1949), and the number can be even higher in dialects. For example, the Kudymkar-In'va dialect of Permyak is said to have at least 12 spatial cases (Batalova, 2002). The languages show both symmetries and asymmetries in their spatial case systems, both on synchronic and diachronic levels. I analyze how the different (a)symmetries in coding of SOURCE and GOAL in Permic varieties seems to reflect the interplay between general cognitive properties and language specific characteristics (cf. Kopecka & Vuillermet, 2021: 16–17).

Asymmetries in the spatial case systems include, e.g., the larger number of GOAL cases than SOURCE cases in different Permic varieties (e.g., Некрасова, 2015), the conflation of GOAL and LOCATION cases in the possessive inflection of nouns (e.g., Edygarova, 2022: 510; Kuznetsov, 2022: 491), and diachronic developments, where SOURCE cases are morphologically more complex than GOAL cases (cf. Csúcs, 2005: 184–193). Furthermore, some varieties are developing new spatial case series based on a GOAL case that can take additional spatial case inflection. Similar phenomena have been mentioned as exponents of SOURCE-GOAL asymmetry (cf. e.g., Kopecka & Vuillermet, 2021: 7–11).

On the other hand, different Permic varieties do exhibit symmetries in coding of GOAL and SOURCE. For example, they have a case for coding both what has been called by Bourdin (1997: 190–196) GOAL-ATTAINMENT and GOAL-EXTENT, but also their SOURCE counterparts, i.e., 'from' and 'all the way from'. Some Udmurt dialect also possess a spatial case system with a case for GOAL-APPROXIMATION and its SOURCE counterpart 'approximately from the direction' (cf. Bourdin, 1997 ibid.).

This paper gives an insight on the effects of SOURCE-GOAL (a)symmetry on the level of language structure in a group of languages that exhibit notably large spatial case inventories and thus use case inflection in expressing spatial relation. SOURCE-GOAL (a)symmetry has not been studied before in Permic languages, so this study is rather exploratory. The study shows, however, that source-goal asymmetry is clearer diachronically than synchronically, i.e., GOAL cases develop earlier and have morphologically simpler, but synchronic spatial case systems can equally well be symmetric or asymmetric, in languages with large spatial case inventories, presumably because a large case inventory can incorporate more distinctions over time.

I restrict the analysis to previous grammatical and historical descriptions of Permic varieties, as there are no suitable corpora available at the moment. I look at the descriptions of spatial case systems and their development and look if the varieties have more cases which code GOAL than SOURCE, do the varieties exhibit syncretism in GOAL or SOURCE cases, and do the case forms can be regarded diachronically simple or complex.

**References**

Batalova, Raisa M. (2002), *Kudymkarsko-in'venskij dialekt komi-permjackogo jazyka*, Societas Uralo-Altaica: Moskva.

Bourdin, Philippe (1997), On Goal-bias across languages. Modal, configurational and orientation parameters, in P. Bohumil, O. Fujimura, & J. Václav Neustupný (eds), *Proceedings of LP'96. Typology, item orderings and universals*, Prague: Karolinum, 185–218.

Csúcs, Sándor (2005), *Die Rekonstruktion der permischen Grundsprache*, Budapest: Akadémiai Kiadó.

Edygarova, Svetlana (2022), Udmurt, in M. Bakró-Nagy, J. Laakso, & E. Skribnik (eds), (2022), *The Oxford guide to the Uralic languages*, Oxford: Oxford University Press, 507–522, https://doi.org/https://doi.org/10.1093/oso/9780198767664.003.0027

Kopecka, Anetta, & Vuillermet, Marine (2021), Introduction. Source-Goal (a)symmetries across languages, *Studies in Language*, *45*(1), 2–35, https://doi.org/https://doi.org/10.1075/sl.00018.kop

Kuznetsov, Nikolay (2022), Komi, in M. Bakró-Nagy, J. Laakso, & E. Skribnik (eds), *The Oxford guide to the Uralic languages*, Oxford: Oxford University Press, 487–506, https://doi.org/http://doi.org/10.1093/oso/9780198767664.003.0026

Бубрих, Д. В. (1949), *Грамматика литературного коми языка*, Ленинград: Издательство Ленинградского государственного ордена Ленина университета им. А. А. Жданова.

Некрасова, Галина (2015), Альтитерминатив в падежной системе коми-пермяцкого языка, *Lingvistika Uralica*, *LI*(3), 202–211, https://doi.org/http://dx.doi.org/10.3176/lu.2015.3.04

Перевощиков, П. Н., Вахрушев, В. М., Алатырев, В. И., Поздеева, А. А., & Тараканов, И. В. (eds) (1962), *Грамматика современного удмуртского языка. Фонетика и морфология*, Ижевск: Удмуртское книжное издательство.

# Path, Deixis or Both? Slovenian motion prefixes under German contact

Ruprecht von Waldenfels
(Friedrich Schiller University Jena)

Motion events as in English *come running out* (see below example 1) involve, at the same time, motion towards a deictic center (Deixis) together with motion in relation to a landmark (Path); see Matsumoto et al. (2017) for why these are distinct categories. A survey of such events in a parallel corpus shows an asymmetry of expression: even strongly deictic languages (Ricca 1993) in Europe, particularly the Romance languages, use path verbs here to mark Path but do not mark Deixis (see below French), while Germanic languages tend to mark both (see below the German and English translation).

In the Slavic languages, like in Romance, Path generally takes precedence if Path and Deixis compete, even in those languages such as Bulgarian and Serbian that are strongly deictic; cf. below the Russian, Serbian and Czech translation. However, the South Slavic language Slovene stands out by expressing both Path and Deixis:

(1) Next second, Quirrell **came hurrying out** of the classroom straightening his turban.
Im nächsten Moment **kam** Quirrell [...] **aus dem Klassenzimmer gestürzt**. [German]
in next Moment came Quirell out.of the classroom hurry-PTCPL
V sledujuščuju sekundu Belka toroplivo **vy-šel** **iz** **klassa** [...] [Russian]
In next second Quirell quickly out-went from classroom.
Sledećeg trenutka, Kvirel **iz-lete** **iz** **učionice** [Serbian]
Next second Quirell out-ran from classroom.
V příštím okamžiku Quirrell spěšně **vy-běhl** z učebny [...] [Czech]
In Next moment Quirell quickly out-ran from classroom.
Un instant plus tard, il **sortit** **en hâte** de la salle en redressant son turban. [French]
one instant later, he exit.PST in hurry from the room straighten.PTCPL his turban
V naslednjem trenutku je [...] **pri-hitel iz učilnice**. [Slovene]
in next moment AUX hither-hurried from classroom.
(Rowling, Harry Potter 1, from ParaSol Corpus)

In this paper, I claim that the inherited system of Slavic motion prefixes in Slovene evolved under the influence of German to give precedence to a **deictic** perspective, triggering a reorganization of the system of satellites that express path in motion events.

Most Slavic languages use inherited verbal prefixes to express Path, much like English or German. In the example above Russian, Czech and Serbian use the particles *vy-/iz-* 'out of'. Slovene, in contrast, departs from this pattern by using the innovated ventive verbal prefix *pri-* 'towards the deictic center', i.e., Deixis. Path, in contrast, is expressed solely by the PP introduced by the preposition *iz* 'out of'. Slovene is thus the only Slavic language to express both Deixis and Path, as does Germanic.

I show that in Slovene, the spatial prefixes are still there, but they have largely lost their function to show Path with motion events. Rather, the prefix slot is specialized to convey Deixis and express either motion towards (*pri-*) or from (*od-*) the deictic center. Path *strictu sensu* is now primarily expressed by prepositions, or by new spatial particles (Giger 2004), and consequently, many verbal prefixes are no longer used in spatial function.

This organization allows the Slovene system to represent deictic and relative path symmetrically, as in German. In the talk I explore the idea how contact pressure to express the notion of Deixis has led to a wholesale reorganization of the system of Slavic path prefixes in Slovene.

**References**


Di Meola C (2003) Non-deictic uses of the deictic motion verbs 'kommen' and 'gehen' in German. In: *Lenz F (ed) Deictic Conceptualisation of Space, Time and Person.*, Benjamins, Amsterdam/Philadelphia, pp 41–67

Filipović L (2007) Talking about motion. A crosslinguistic investigation of lexicalization patterns. John Benjamins

Giger, M. (2004). Präliminarien zu einem Studium von Spatialpartikeln im Slavischen. In: *Beiträge der Europäischen Slavistischen Linguistik (POLYSLAV 7)* (pp. 53-62). München: Sagner.

Matsumoto, Y, Akita, K, and Takahashi K (2017) The functional nature of deictic verbs and the coding patterns of Deixis. In *Ibarretxe-Antuñano I (ed) Motion and space across languages: Theory and applications*. John Benjamins, Amsterdam/Philadephia, pp 95-122.

Ricca, D (1993) I verbi deittici di movimento in Europa: una ricerca interlinguistica. La Nuova Italia

Vogel, P. M. (2005). Neue Überlegungen zu den Fügungen des Typs sie kamen gelaufen. Zeitschrift für germanistische Linguistik, 33(1), pp 57-77.

# On the spatial asymmetry of itive and ventive Associated Motion in Baoding (Sinitic)

Na Song
(INALCO-CRLAO)

Keywords: Spatial asymmetry, Itive and ventive, Associated Motion, Baoding, Sinitic

The Source-Goal asymmetry of motion events is garnering increasing attention cross-linguistically (Bourdin 1997, Kopecka & Vuillermet 2021, among others). While previous literature on spatial asymmetries among Sinitic languages has primarily focused on this topic (Lamarre 2008, Song 2017, Lamarre & Song 2021), this paper presents a preliminary investigation into the spatial asymmetry of itive and ventive associated motion (AM) in the Baoding dialect of Jilu Mandarin (Sinitic), a Mandarin dialect spoken in northern China.

The Associated Motion in Baoding, encoded by final enclitic particles, forms the 'motion – CUM – purpose' type. The dedicated paradigm is explored based on a binary opposition: 'VP+ *tɕʰi*' 'go and do something', or 'VP+ *lɛ*' 'come and do something'. In addition to AM, final enclitic particles in Baoding also encode TAME (Tense-Aspect-Modality-Evidentiality). As asserted by Wilkins (1991:211), aspect and AM are distinct categories. In Baoding, AM and TAME particles appear in different paradigms, with AM particles positioned closer to the VP than other final enclitic particles. Our focus lies on their interaction with TAME in Baoding through combinatorial possibilities. By examining these possibilities and the inventory of final particles that can co-occur with itive and ventive AM, we observe that itive AM is more likely to interact with TAME than ventive AM. For instance, itive AM can co-occur with a tense marker, such as the imminent future particle. This paper also aims to investigate whether and how the expression of itive and ventive AM relates to the deictic orientation of motion asymmetry (Wilkins and Hill 1995, Koga et al. 2008) by examining the correlation between the motion orientation and the situation type, as well as the semantic features of the predicates, as the motion in the AM event in Baoding consistently anchors deictically (towards or away from the deictic center). The data are taken from the author's personal fieldwork, which includes elicited sentences, in addition to approximately 40 hours of daily conversation.

**References**

Bourdin, Philippe (1997), On Goal-bias across languages: modal, configurational and orientation parameters. In P. Bohumil (ed.), Proceedings of LP'96, *Typology, item orderings and universals*, (Proceedings of the Conference held in Prague, August 20-22, 1996), 185-218.

Koga, Hiroaki, Yulia Koloskova, Makiko Mizuno and Yoko Aoki (2008), Expressions of spatial motion events in English, German, and Russian with special reference to Japanese. In C. Lamarre, T. Ohori, & T. Morita (eds.), *Typological studies of the linguistic expression of motion events*: Vol. 2, A contrastive study of Japanese, French, English, Russian, German, and Chinese: Norwegian Wood, Tokyo: 21st Century COE Program Center for Evolutionary Cognitive Sciences.13–44.

Kopecka, Anetta and Marine, Vuillermet (eds.), (2021), Source-Goal Asymmetries across Langages. Thematic issue of *Studies in Language* 45:1.

Lamarre, Christine (2008), The linguistic categorization of deictic direction in Chinese — With reference to Japanese. In D. Xu (ed.), *Space in languages of China: Cross-linguistic, synchronic and diachronic perspectives*, Berlin / Heidelberg / New York: Springer, 69-97.

Song, Jinke (2017), *Source-Goal (A)symmetry: A study of spontaneous and caused motion events in Mandarin Chinese*. NAMED, International workshop on 'Neglected Aspects of Motion-Event Description'. ENS. Paris (France). 19-20 may 2017.

Lamarre, Christine and Song, Na (2021), *Morphosyntactic manifestations of Goal-related asymmetry in Mandarin Chinese: Insights from non-standard Mandarin varieties*. NAMED-3, International workshop on 'Neglected Aspects of Motion-Event Description'. ENS. Paris (France). July 1-2, 2021.

Wilkins, David P. (1991), The semantics, pragmatics and diachronic development of "associated motion" in Mparntwe Arrernte. *Buffalo Papers in Linguistics* vol 1, 207–257.

Wilkins, David P. and Hill, Deborah (1995), When go means come: Questioning the basicness of basic motion verbs. *Cognitive Linguistics* vol 6(2/3), 209-259.

# The Goal-over-Source Asymmetry in Korean and Thai

Kultida Khammee & Seongha Rhee
(University of Phayao; Faculty of Liberal Arts, Mahidol University & Hankuk University of Foreign Studies)

Keywords: the Goal-over-Source asymmetry, allative, ablative, Korean, Thai

Korean and Thai have large inventories of adpositional particles, including the source and goal markers. As reported in many languages (Kopecka and Vuillermet 2021), Korean and Thai adpositions also prominently exhibit 'the goal-over-source asymmetry' or 'the goal bias' (Ikegami 1987; Bourdin 1997; Fagard and Kopecka 2021) at multiple levels. This paper supports this asymmetry hypothesis from two typologically and genealogically distinct languages, i.e., Korean and Thai.

The goal-over-source asymmetry has often been attributed to the cognitive and pragmatic salience of goal (e.g. Lakusta and Landau 2005, 2012, Regier and Zheng 2007, as cited in Fagard and Kopecka 2020). In the two languages, at the level of paradigm, the goal markers exceed the source markers (34 vs. 13 in Korean; 29 vs. 6 in Thai), and exhibit more fine-grained specialization, e.g., Directionals, Allatives, and Datives, whereas the source markers form a conceptually uniform category, i.e., Ablatives. The asymmetric salience of goal-marking in Korean is confirmed by the fact that the aggregate token frequency of goal-marking is 38,721 pmw in a 24-million-word contemporary corpus, whereas that of the source-marking is 12,907 pmw. Similarly, in Thai, the aggregate token frequency of goal markers is 19,144 pmw in a 33.4 million-word contemporary corpus, whereas that of the source markers stands at only 5,614 pmw. These states of affairs suggest that goal is conceptually more salient and pragmatically more significant.

In both languages, the multiplicity of forms in the two polar categories is largely due to the stacking of multiple markers of (nearly-)synonymous adpositions as a strategy to reinforce the meaning or to add further fine shades of meaning. Formal reinforcement often yields emphatic effect on the form. This type of doubling or tripling of forms is a strategy also used, though to a much lesser degree, for creation of source markers, hence confirming the goal-over-source hypothesis again. The multiplicity of forms is also due to frequent innovation of new forms, especially the goal markers, in an effort to enhance expressivity and to entertain the desire of creativity (Heine and Stolz 2008, Rhee 2023). This is evident in that the forms being innovated tend to carry more lexical content than the older, fully grammaticalized forms, and thus carry more expressive potential. Emphasis meaning results from stacking of the markers in both languages, but in Korean, unlike Thai, the forms thus created tend to carry the speaker's negative evaluation about the event being described. Furthermore, Korean has forms specializing in marking [+Honorific] nominals, i.e., ones higher in hierarchy than the speaker. Another peculiarity in Korean is that a number of forms show Allative-Ablative syncretism, a crosslinguistically rare phenomenon. Even among these syncretic forms, the proportion of use for goal-marking far exceeds that for source-marking, again upholding the hypothesis.

Drawing upon corpus data, this paper addresses the goal-over-source asymmetry in Korean and Thai from the pragmatic and grammaticalization perspectives, focusing on their conceptual and functional asymmetry.

**References**


Bourdin, Philippe (1997). On Goal-bias across languages: modal, configurational and orientation parameters, in B. Palek (ed), (1997), *Proceedings of LP '96. Typology, item orderings and universals, held in Prague, August 20-22, 1996*, Prague: Charles University, 185-218.

Fagard, Benjamin and Anetta Kopecka (2020). Source/Goal (A)symmetry. *Studies in Language* 45(1), 130-171.

Fagard, Benjamin and Anetta Kopecka (2021). Source/Goal (A)symmetry: A Comparative Study of German and Polish. *Studies in Language* 45(1), 130-171.

Heine, Bernd, and Thomas Stolz (2008). Grammaticalization as a Creative Process. *Language Typology and Universals (STUF)* 61, 326-357.

Ikegami, Yoshihiko (1987). 'Source' and 'Goal': A Case of Linguistic Dissymmetry, in R. Dirven and G. Radden (eds), *Concept of Case,* Tübingen: Gunter Narr Verlag, 122-146

Kopecka, Anetta and Marine Vuillermet (eds) (2021). Source-Goal (A)symmetries across Languages. Special issue of *Studies in Language* 45(1).

Lakusta, Laura and Barbara Landau (2005). Starting at the End: The Importance of Goals in Spatial Language. *Cognition* 96(1), 1-33.

Lakusta, Laura and Barbara Landau (2012). Language and Memory for Motion Events: Origins of the Asymmetry between Source and Goal Paths. *Cognitive Science* 36(3), 517-544.

Regier, Thierry and Mingyu Zheng (2007). Attention to Endpoints: A Cross-linguistic Constraint on Spatial Meaning. *Cognitive Science* 31, 705-719.

Rhee, Seongha (2023). (Im)politeness in SNS communication in Korea. Russian Journal of Linguistics 27(1), 39-66.

# Source-Goal Asymmetry in Chaozhou

Adeline TAN
(University of Tours; LLL)

Keywords: Asymmetry, deictic directionals, Source-Goal, zero marker, Chaozhou

This paper presents the Source-Goal (starting and the ending point) asymmetry in the expression of path in the Chaozhou dialect, a Southern Min variety mainly spoken in Guangdong, China. Previous studies have shown that speakers tend to express Goals more frequently than Sources [Ikegami 1979, Bourdin 1997: 190, Fagard & Kopecka 2021, and Kopecka & Vuillermet (eds.) 2021]. This phenomenon is also observed in Chaozhou.

The present study explores the morphosyntactic resources available in Chaozhou to encode motion expressions and investigates the nature of the Source-Goal asymmetry in Chaozhou. Our analysis is based upon data collected through television series and sentences elicited using the "Put & Take" and "Trajectoire" (Ishibashi, Kopecka, and Vuillermet 2006) videos with native language consultants in China. Our preliminary results show that Source information is encoded more often in a more complex strategy (Song 2022) compared to Goal information, such as a prepositional phrase combined with non-deictic and deictic directionals as shown in (1). As for Goal, it tends to be expressed either by directionals (2a) or by a preposition (2b). Example (2a) illustrates that the deictic directional $k^hɯ^{213}$ in Chaozhou has a prepositional use. Like Beijing dialect of Mandarin (Chirkova & Lamarre 2005, Zhu Dexi 1982), the Goal Noun Phrase (NP) can directly follow the deictic directional (2a), whereas the Source NP generally requires the addition of a Source preposition "$to^{35}$" (1). This complexity for Source information applies to both caused and spontaneous motion (Talmy 2000).

(1) $i^{33}$ **$to^{35}$** $ts^hiu^{11}$=$teŋ^{42}$ $k^hio$ $ʔ^2$ $tiao^{55}$ $soʔ^2$ **$loʔ^5$** **$lai^{55}$**
3SG **from** tree=upside take CLF rope **descend** **come**
"She removes the rope from the tree branch." [Source] (Put_HSH_17)

(2a) $i^{33}$ $p^hua^2$ $tiao^{55}$ $soʔ^2$ **$k^hi^{42}$** **$k^hɯ^{213}$** $ts^hiu^{11}$=$ko^{213}$
3SG hang CLF rope **ascend** **go** tree=place
"She hangs a rope on the branch of a tree." [Goal] (Put_SQS_3)

(2b) $i^{33}$ $k^hioʔ^5$ $tiao^{55}$ $soʔ^2$ $p^hua^2$ **$to^{35}$** $xɯ^{42}$ $ts^hiu^{11}$=$teŋ^{42}$
3SG OM CLF rope hang **at** that tree=upside
"She hangs a rope on the branch of a tree." [Goal] (Put_HSH_3)

Moreover, Chaozhou exhibits a possible omission of the Goal marker for even greater simplicity in Goal expression (3a-3b). This omission is not possible in Source expression.

(3a) $i^{33}$ $k^hioʔ^5$ $ki^{33}$ $pek^2$ $paŋ^{213}$ **$to^{35}$** $ts^hiu^{11}$ $k^huk^2$=$toĩ^{42}$
3SG OM CLF pen put **at** tree hole=inside
"She places the pen in the hollow of the tree." [with preposition $to^{35}$] (Put_57_HSH)

(3b) i[33] kʰioʔ[5] ki[33] pek[2] paŋ[213] tsʰiu[11] kʰuk[2]=toĩ[42]
3SG OM CLF pen put tree hole=inside
"She places the pen in the hollow of the tree." [omission of the preposition to[35]] (Tan 2020)

**References**


Bourdin, Philippe (1997), On goal-bias across languages: modal, configurational and orientational parameters, Proceedings of LP 96: Typology: prototypes, item orderings and universals, proceedings of the conference held in Prague, August 20-22, 1996, 185-216.

Chirkova, Katia and Lamarre, Christine (2005), The paradox of the construction [V zai NPloc] and its meanings in the Beijing dialect of Mandarin, *Cahiers de linguistique - Asie orientale*, 34(2), 169-220.

Fagard, Benjamin and Kopecka, Anetta (2021), Source/goal (a)symmetry: A comparative study of German and Polish, *Studies in Language* 45(1), 130-171.

Ikegami, Yoshihiko (1979), 'Goal' over 'Source': A Case of Dissymmetry in Linguistic Orientation, *Angol Filológiai Tanulmányok / Hungarian Studies in English*, 12, 139-157.

Ishibashi, Miyuki, Kopecka, Anetta, and Vuillermet, Marine (2006), Trajectoire : matériel visuel pour élicitation des données linguistiques, Laboratoire Dynamique du Langage, CNRS / Université Lyon 2. Projet de Fédération de recherche en Typologie et Universaux Linguistiques. http://tulquest.huma-num.fr/fr/node/132.

Kopecka, Anetta and Vuillermet, Marine (eds), (2021), Source-Goal (a)symmetries across languages, Special issue of Studies in Language 45(1).

Song, Jinke, (2022), Source-Goal asymmetry in Chinese Standard: A comparative study in spontaneous and caused motion events. In Sarda, L. and Fagard, B. (eds.), *Neglected Aspects of Motion Events Description: Deixis, Asymmetries, Constructions*, [Human Cognitive Processing, 72] John Benjamins, 149-172.

Talmy, Leonard, (2000), *Toward a Cognitive Semantics vol. II: Typology and process in Concept Structuring*, Cambridge, MA: MIT Press.

Zhu, Dexi, (1982), *Lectures on Grammar*, Beijing: Shangwu Yinshuguan.

# The Source/Goal asymmetry in Negidal motion verbs

Laurène Barbier
(Laboratoire Dynamique Du Language CNRS UMR 5596; Université Lumière Lyon 2)

Keywords: Source/Goal asymmetry, motion verbs, spatial events, Path, Tungusic

In the expression of Path of motion (Talmy 1895, 2000), a tendency among speakers of many languages to favour the expression of Goal-Grounds (an entity of reference that is the final part) to the detriment of Source-Grounds (the initial part) has been described in the literature (Ikegami 1987, Bourdin 1997). This phenomenon creates a spatial asymmetry, that has been observed at different levels of the linguistic structure, both semantic (types of semantic distinctions) and morphosyntactic (types and diversity of devices). It has also been shown that the degree and the manifestation of such asymmetry might vary across languages depending on the types of linguistic devices available in a given language and the semantic distinctions languages are sensitive to (Kopecka & Vuillermet 2021).

Therefore, research on languages less often described as part of studies on the Source/Goal asymmetry can shed new light on our understanding of cross-linguistic variability in this domain. The aim of the present study is to investigate the existence of such asymmetry in a language that has not been the subject of study in this domain of expression: Negidal (Tungusic). If attested, how does the asymmetry manifest itself in the morphosyntactic and semantic structure in Negidal? What is the degree of asymmetry? Is there variation in the expression of Source and Goal information depending on the types of motion event expressed?

In this study, the focus is on the Source/Goal asymmetry in Negidal motion verbs that convey Path meaning, through their frequency and co-occurrence with Source vs. Goal Grounds. Among verbs lexicalizing Path, some are inherently Goal-oriented (*i*ː- ‘enter’, *es*- ‘reach’) while others are Source-oriented (*ju*ː- ‘exit’). Therefore, a congruence between the orientation of the verb and the type of Grounds with which it appears is expected.

To conduct this study, I examined a corpus of narratives produced by 9 Negidal speakers (Pakendorf & Aralova 2017), which includes 3746 utterances expressing motion events. The data was coded following the Berkeley coding manual elaborated by Slobin (2005). Results reveal an asymmetry between Source-oriented vs. Goal-oriented verbs in motion descriptions in Negidal. Source-oriented verbs occur more frequently without Grounds, as in (1), while Goal-oriented verbs are used more frequently with an overtly expressed Ground, as in (2). Regarding combinability, Source-oriented verbs can co-occur with Source- and Goal-Grounds, whereas Goal-oriented verbs tend to co-occur with Goal-Grounds only. Finally, when Grounds are overtly expressed, Goal-Grounds are strikingly more frequent (67%) than Source-Grounds (11%).

| | | | |
|---|---|---|---|
| 1) | *uže* | *haŋka-ʧi-l-li-du-j* | ***ju***ː*-m* |
| | already.R | choke[INTR]-TAM2-INCH-PRS.PTCP-DAT.ESS-PRFL.SG | **exit**-NP.1SG |

‘When I was already starting to suffocate, I came out.’ (Pakndorf & Aralova 2017: DIN_tonuli: 21)

2) *ɟoː-la-n* ***es****-saːn=da* *iː-je-n*
house-LOC-3SG **reach**-SS.ANT=ADD enter-NFUT-3SG
'It reached her house and entered.' (Pakendorf & Aralova 2017: APK_1chindakan: 131)

In the corpus, the encoding of Goal-Grounds predominates over that of Source-Grounds. Constructions with Source- and Goal-oriented verbs not only show that the expression of Goal is less constrained, but also that Goal is often distributed in different morphosyntactic elements of the utterance, leading to finer semantic granularity in Goal-oriented constructions. The study of Negidal, a language that has not been investigated in this domain, will contribute to a better understanding of the diversity of strategies used to express spatial relations and language-specific characteristics that play a role in spatial asymmetries.

Acknowledgements
This research was funded, in whole or in part, by l'Agence Nationale de la Recherche (ANR), project ANR-20-CE27-0015.

**References**

Bourdin, Philippe. 1997. On goal-bias across languages: modal, configurational and orientational parameters, proceedings of *LP '96: Typology: prototypes, item orderings and universals*, Prague, Cezch Republic, August 20-22-1996, 185–216.

Ikegami, Yoshihiko. 1987. 'Source' vs. 'Goal': A case of linguistic dissymmetry. *In* R. Dirven and G. Radden *Concept of case*, Tübingen: Gunter Narr Verlag, 122–146.

Kopecka, Anetta & Marine Vuillermet. 2021. Source-Goal (a)symmetries across languages, in *Studies in Language* 45(1), John Benjamins Publishing Company, 2–35.

Pakendorf, Brigitte & Natalia Aralova. 2017. *Documentation of Negidal, a nearly extinct Northern Tungusic language of the Lower Amur*. http://hdl.handle.net/2196/b644db81-725c-4031-935c-f33c763df152. (Last access: 27/05/2024)

Slobin, Dan I. 2005. *Coding of Motion Events in Texts*.

Talmy, Leonard. 1985. Lexicalization patterns: Semantic structure in lexical forms. In Shopen, T *Language typology and semantic description, Grammatical categories and the lexicon*, vol. 3, Cambridge: CUP, 225–282.

Talmy, Leonard. 2000. *Toward a Cognitive Semantics*, MIT Press.

# The expression of vertical motion in light of the goal-over-source asymmetry

Piia Taremaa, Johanna Kiik & Ann Veismann
(University of Tartu, University of Tartu & University of Tartu)

Keywords: semantics, motion verbs, spatial language, corpus analysis, Estonian

The *goal-over-source* asymmetry is attested in many languages (Ikegami 1987; Kopecka & Vuillermet 2021). This asymmetry suggests that when expressing motion, its endpoint is mentioned more frequently and in more varied ways than its origin. Research also shows that the expression of Goal is associated with various factors, most relevantly with the semantics of the verb (Stefanowitsch & Rohde 2004). However, the majority of the studies focus mainly on horizontal motion and pay limited attention to vertical motion. Alternatively, the studies compare the expression of horizontal and vertical motion (Naigles & Terrazas 1998; Łozińska & Pietrewicz 2018). Consequently, little is known about how the semantics of the verb relates to the expression of spatial relations, including Goal, when describing vertical motion.

The current exploratory study focuses on vertical motion and aims to establish differences between verbs of upward vs. downward motion in their relation to the *goal-over-source* asymmetry in Estonian. Because upward and downward motion are significantly distinct – upward motion needs much energy but its endpoint may not be as concrete and salient as it is for downward motion – the expectation is that expressing spatial and manner information varies across the two types of vertical verbs.

The study reports on the use of 24 verbs of vertical motion in Estonian. According to a previous classification study, 10 of these express upward motion (e.g. *tõusma* 'rise', *hüppama* 'jump') and 14 express downward motion (e.g. *kukkuma* 'fall', *sukelduma* 'dive'). For each of the verbs, 200 clauses of actual motion were taken from the Estonian National Corpus 2021. The resulting 4800 clauses were coded for a range of semantic variables (e.g. mover animacy, Source, Goal, Manner).

Multiple correspondence analysis suggests that verbs of upward motion and verbs of downward motion somewhat overlap in their clausal structure for expressing space and manner, while also showing distinct patterns from each other. Conditional random forests and inference trees further show that the main distinctive factors are Goal, Result, Direction, and Source (whether they are expressed or not). Aligning with the *goal*-bias, verbs of downward motion combine frequently with Goal expressions (e.g. *Münt vajub vee põhja* 'The coin sinks to the bottom of the water'). Conversely, verbs of upward motion combine with Goal expressions significantly less frequently, and instead tend to co-occur with Direction and Source expressions (mainly directional verbal particles, such as *üles* 'up' and *välja* 'out'; e.g. *Maa seest [Source] purskab välja [Dir] kuum vesi* 'The hot water erupts from the ground' and *Maapinnalt [Source] kerkis tolmu* 'Dust rose from the ground'). This illustrates Goal-pattern variability with regard to verb semantics and highlights the importance of spatial interrelations expressed within a single clause.

Acknowledgments
The study has been supported by the Research fund of Kadri, Nikolai, and Gerda Rõuk and by the Estonian Research Council grant (PSG929) "Speed at the Crossroads of Language, Perception and Action"

**References**

Ikegami, Yoshihiko. 1987. 'Source' vs. 'Goal': A case of linguistic dissymmetry. In René Dirven & Günter Radden (eds.), *Concepts of case* (Studien Zur Englischen Grammatik 4), 122–146. Tübingen: Narr.

Kopecka, Anetta & Marine Vuillermet. 2021. Source-Goal (a)symmetries across languages. *Studies in Language*. John Benjamins 45(1). 2–35. https://doi.org/10.1075/sl.00018.kop.

Łozińska, Joanna & Barbara Pietrewicz. 2018. Lexicalisation of vertical motion: A study of three satellite-framed languages. *Cognitive Studies | Études cognitives*. Institute of Slavic Studies, Polish Academy of Sciences 18. https://doi.org/10.11649/cs.1601.

Naigles, Letitia R. & Paula Terrazas. 1998. Motion-verb generalization in English and Spanish: Influences of language and syntax. *Psychological Science* 9(5). 363–369. https://doi.org/10.1111/1467-9280.00069.

Stefanowitsch, Anatol & Ada Rohde. 2004. The goal bias in the encoding of motion events. In Günter Radden & Klaus-Uwe Panther (eds.), *Studies in linguistic motivation* (Cognitive Linguistics Research 28), 249–267. Berlin/New York: Mouton de Gruyter.

# WS18 Talking about truth, lies, and deception across languages and cultures

# From authenticity to sincerity and beyond: The Hand(s) on Chest gesture in Hebrew face-to-face interaction

Anna Inbar & Yuval Geva
(Levinsky-Wingate Academic Center & University of Haifa)

Keywords: sincerity, intensification, gesture, Hebrew, grammaticalization

The present study examines the pragmatic functions of the Hand(s) on Chest gesture (HOC) in Hebrew face-to-face interaction and suggests the grammaticalization path that the gesture could undergo. The HOC is performed by placing either one hand or two hands on the upper torso, with the two-handed variant including the crossing of one hand over another. In Hebrew culture, the gesture is strongly associated with the concepts of *honesty* and *sincerity*, as evident, for example, in the Hebrew idiom *'im yad 'al ha-lev* 'honestly' (lit. 'with a hand on the heart'), in which the chest metonymically represents a heart (cf., Baranyiné Kóczy and Sipőcz 2023). However, we will show that the analysis of the gesture as it is used in naturally occurring conversation, reveals other concepts that potentially may give rise to the concept of *sincerity* or be derived from it.

The HOC has sparked the interest of researchers in various fields, including art history (Lange 1887), embodied cognition (e.g., Cantarero et al. 2017; Parzuchowski et al. 2014), and nonverbal communication (Farley et al. 2021). In the field of multimodal interaction analysis, however, the gesture was mentioned only in passing (e.g., Quasinowski et al. 2022; Streeck 2009). The present study seeks to fill this lacuna.

Our data are drawn from the *Haifa Multimodal Corpus of Spoken Hebrew*, which consists of video recordings of naturally occurring casual conversations collected from 2016 through 2023, with approximately 18 hours in total. We identified 85 occurrences of the gesture in total. However, since the same gestural form can be used for indexing first person, we excluded from our analysis 47 ambiguous cases in which the gesture could be interpreted as representing such propositional information. Then, employing the methodologies of interactional linguistics (Couper-Kuhlen and Selting 2018) and multimodal conversation analysis (e.g., Goodwin and Goodwin 2000; Mondada, 2016), we analyzed the sequential environments in which the 38 remaining occurrences of the gesture were found. The analysis revealed three broad contexts: (a) 10 occurrences of the gesture were co-produced with expressions of feelings (mostly with verbal expressions of fear or surprise); (b) 11 occurrences were deployed when speakers conveyed content that (was assumed to) violate co-participants' expectations; in these contexts, the gesture was often co-produced with verbal expressions of sincerity, such as *be-'emet* (lit. 'in true'); and, finally, (c) 17 occurrences were associated with intensification. In the present talk, we focus on what this gestural resource accomplishes in interaction in the moment in which it is produced. Moreover, based on our findings, on the principles of grammaticalization (e.g., Hopper and Traugott 1993), and on the principles of the 'emancipation' of gestures as semiotic resources used in interaction (Müller 2014; Streeck 2009, 2021), we will suggest and elaborate the following path of grammaticalization of the HOC.

**Table 1:** The Hand(s) on Chest Gesture:
The Origin and the Path of Grammaticalization

| | | | | | | |
|---|---|---|---|---|---|---|
| **Physiological response** to palpitations caused by excitement | ⇨ | Enacting the physiological response as a means to convey **authenticity** | ⇨ | Pragmatic marker of **sincerity** | ⇨ | Pragmatic marker of **intensification** |

**References**


Baranyiné Kóczy, J. and Sipőcz, K. (eds.). (2023), *Embodiment in cross-linguistic studies: The 'Heart',* Brill.

Cantarero, K., Parzuchowski, M., and Dukala, K. (2017), White lies in hand: Are other-oriented lies modified by hand gestures? Possibly not, *Frontiers in Psychology*, 8:814.

Couper-Kuhlen, E. and Selting, M. (2018), *Interactional linguistics*, Cambridge University Press.

Farley, S. D., Akin, K., and Hedgecoth, N. (2021), Exploring the meanings of the "heartfelt" gesture: A nonverbal signal of heartfelt emotion and empathy, *Journal of Nonverbal Behavior* 45, 567–585.

Goodwin, C. and Goodwin, M. H. (2000), Emotion within situated activity, in A. Duranti (ed.), *Linguistic Anthropology: A Reader*, Basil Blackwell, 239–257.

Hopper, P. J. and Traugott, E. C. (1993), *Grammaticalization*, Cambridge University Press.

Lange, J. (1887/2021), *The hand on the breast tr. Karl Johns, Journal of Art Historiography*, 25.

Mondada, L. (2016), Challenges of multimodality: Language and body in social interaction, *Journal of Sociolinguistics* 20(3), 336–366.

Müller, C. (2014), Gestural modes of representation as techniques of depiction, in C. Müller, A. Cienki, E. Fricke, S. H. Ladewig, D. McNeill, and J. Bressem (eds), *Body—language—communication: An international handbook on multimodality in human interaction*, De Gruyter Mouton, 1687–1702

Parzuchowski, M., Szymkow, A., Baryla, W., and Wojciszke, B. (2014), From the heart: Hand over heart as an embodiment of honesty, *Cognitive Processing* 15, 237–244.

Quasinowski, B., Solmaz, A., Bachmann, C., Chen, W., Elcin, M., Kamisli, C. Liu, T., Maass, A. H., Merse, S., Morbach, C., Neumann, A., Neumann, T., Sommer, I., Stoerk, S., Weingartz, S., Weiss, A., and Wietasch, G. (2023), Hearts in their hands – physicians' gestures embodying shared professional knowledge around the world, *Sociology of Health and Illness* 45, 1101–1122.

Streeck, J. (2009), *Gesturecraft: The manu-facture of meaning*, John Benjamins.

Streeck, J. (2021), The emancipation of gestures, *Interactional Linguistics* 1(1), 90–122.

# Slavic epistemic pragmatic markers of truth-roots origin: Their etymology, development and functions

Piotr Sobotka
(Institute of Slavic Studies of the Polish Academy of Sciences)

Keywords: truth, etymology, epistemic words, grammaticalization, abductive reconstruction

The Proto-Slavic roots **jьst-* or **jìst-*, **pravьd-*, **rěsn-* and **věr(n)-* are common ancestors of most words related to the concept of truth in Slavic languages, cf. Ukr. *istynnist*' 'truthfulness', Pol. *prawdziwy* 'true', Sln. *resníca* 'truth', LSorb. *wěrnosć* 'truth', Polab. *vernə* 'true'. The distribution of the reflexes of these roots in the Slavic languages is basically territorially specialised, cf. the map below.


These stems are also the source of so-called pragmatic markers, such as assertion operators (cf. Cz. *jistě* 'certainly'), gradation markers (cf. Ukr. *spravžnie* 'true X'), sentence adverbs (cf. Cro. *doista* 'indeed'), concessives (cf. Pol. *wprawdzie _, ale _* 'although _, but _'), etc., cf.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| (1) | Rus. | *Bumažnik,* | ***pravda,*** | *ostavili,* | *no* | *v* | *nem* |
| | | wallet.NOM.SG | CONC | leave.PST.3PL | but.CONJ | in.PREP | 3SG.LOC |
| | *uže* | *ne* | *bylo* | *ničego.* | | | |
| | already.ADV | NEG | be.PST.3SG | nothing.GEN.SG | | | |

'Although they had left the wallet, there was nothing inside.'

Pragmatic markers are usually used as nonrepresentative adjuncts to comment on the actual speech (cf. Fraser 1996, Aijmer & Simon-Vandenbergen 2011). Prominent linguistically orientated logical analyses of the subject show that the meaning structure of these words must include components of speaking, such as 'the speaker notes…', 'the speaker underlines…', 'the speaker doubts…', etc. (cf. Reichenbach 1947). As Brinton (1996: 38) asserts, the interpersonal use of such markers includes expressions "of the speaker's attitudes, evaluations, judgments, expectations, and demands".

Despite the importance of the topic, the motivation for the transformation of representative truth--items such as nouns, adjectives, verbs into pragmatic markers is not well understood. Moreover, previous etymologies of the roots in question (cf. Trubačev, Žuravlev & Varbot 1974–; Žuravlev 2006) do not elucidate the nature and mechanisms of these shifts, as they do not include components of speaking as identified above. It seems that Slavic diachronic studies are somewhat inconsistent in the interpretation of the semantic and syntactic changes of the truth-roots descendants, which are sometimes very surprising, such as the shift from the domain of assertion (2a) to the domain of epistemic hesitation (2b), cf.

| | | | | | | |
|---|---|---|---|---|---|---|
| (2) | a. | MPol. | *odpowiedział* | ***wierę,*** | *niewiem* | *złali,* |
| | | | answer.PST.3SG.M | tell(ing)_the_truth | NEG-know.PRS.1SG | bad.SG.F-IM |
| | | | *dobrali* | *byłá* | [*woda*] | |
| | | | good.SG.F-IM | be.PST.3G.AUX | water.ACC.SG | |

'He replied, I'm telling the truth: I don't know if the water was dirty or good.'

| | | | | | | |
|---|---|---|---|---|---|---|
| (2) | b. | Pol. | *temat,* | ***wierzę,*** | *jest* | *opisany* |
| | | | subject.NOM.SG | believe.PRS.1SG | be.PRS.3SG.AUX | describe.PPP |

*dokładnie*

thoroughly.ADV

'the subject, I believe, is described in details.'

The talk outlines a map of the categorial and conceptual shifts of the studied lexis based on a detailed analysis of dictionary and corpus-based Slavic data, which covers those languages with extensive documented histories. The study focuses on the emergence of non-representative items and the mutual relations between source and target categories. Furthermore, the paper both provides arguments for a reconstruction of the etymological structure of the truth-etyma and aims to validate the theory of internal reconstruction as a device of etymological investigations.

The abductive (cf. Douven 2022) reconstruction of the etyma in this case requires the retrospective study to consider the directions and types of semantic and functional shifts that led to the emergence of the actual speech comments. The approach shows that the speech--related components of the non-representative items in question are etymologically motivated, e.g. contrastively, structurally and functionally, cf. *$V_{DIC}$ + *wierę*$_{APRT}$' to speak, speaking the truth '> _ *wierę* _ 'I speak: this is: _'.

**References**

Aijmer, Karin & Anne-Marie Simon-Vandenbergen (2011), Pragmatic markers. In J. Zienkowski, J.-O. Östman, and J. Verschueren (eds), *Discursive Pragmatics* (Handbook of Pragmatics Highlights 8), 223–247. Amsterdam / Philadelphia: John Benjamins Publishing Company.

Brinton, Laurel J. (1996), *Pragmatic markers in English : grammaticalization and discourse functions* (Topics in English Linguistics 19). Berlin/New York: Mouton de Gruyter.

Douven, Igor (2022), *The Art of Abduction*. Cambridge, MA: Massachusetts Institute of Technology.

Fraser, Bruce (1996), Pragmatic markers. *Pragmatics* 6(2). 167–190.

Reichenbach, Hans (1947), *Elements of symbolic logic*. New York: Macmillan Co.

Trubačev, Oleg N., and Anatolij F. Žuravlev, and Žanna Ž. Varbot (eds) (1974–), *Ètimologičeskij slovar' slavjanskix jazykov: praslavjanskij leksičeskij fond*, vol. 1–. Moskva: Izdatel'stvo Nauka.

Žuravlev, Anatolij F. (2006), Po povodu*istiny*. In I. Janyšková, and H. Karlíková (eds), *Studia Etymologica Brunensia*, vol. 3, 443–455. Praha: Nakladatelství Lidové noviny.

# Hiding yourself as source of false information: A lie or not a lie?

Francesca Giardini, Stanka Fitneva & Anne Tamm
(University or Groningen, Queens University & Károli Gáspár University of the Hungarian Reformed Church)

Keywords: deception, epistemic markers, grammar, lexicon, intercultural comparison of hiding deception by linguistic means

Introduction: Epistemic markers, e.g., grammaticalized evidentials, are employed to obfuscate the source or provenance of information. Such mechanisms are pivotal in communication, serving to mask the identity of the message's originator. Giardini et al (2019) found via a cross-linguistic online vignette experiment that epistemic markers are used as instruments of deception. Their findings suggest cultural variances not the type of the structure of epistemic devices (i.e., grammaticalized evidentials vs. lexical epistemic means) as potential underpinnings for the observed discrepancies in the application of such linguistic tools.

Research Question: The hypothesis is that the differential employment of epistemic markers, like evidentials, in perpetuating a self-beneficial falsehood while simultaneously safeguarding one's reputation as a credible informant, is intrinsically influenced by cultural perceptions of deceit specifically related to the epistemic device.

Materials: The dataset for this analysis comprises unexamined free-text responses from the dataset of Giardini et al (2019), documenting participant choices in information request scenarios across Turkish and Estonian contexts. These responses were coupled with subsequent data from the Hungarian cohorts. The Hungarian and Estonian free texts form the corpus for analysis.

Method: The analysis classifies the textual data by the reference to lies in the choice to employ or eschew epistemic markers when fabricating self-serving falsehoods amid a competitive milieu. The discourse pertaining to the act of lying, specifically when utilizing an epistemic marker to conceal direct evidence was coded and compared.

Results: While Estonian languages features evidential verb morphemes, facilitating the concealment of one's role as a purveyor of misinformation, Hungarian relies on lexical epistemic markers to insulate the speaker from the repercussions of being branded a fabricator. The preliminary results show that the Estonian corpus exhibit homogeneity with scant references to deceit, whereas the Hungarian responses display a pronounced divergence, reflecting on the strategic use of epistemic markers to either gain a competitive edge or to sidestep the issue of lying altogether.

Discussion: The linguistic choices surrounding epistemic markers and deception are not solely contingent on language structure. If that were the case, Turkish and Estonian (cf Johanson 2018, Erelt et al 2006) would exhibit a uniform pattern, as opposed to English and Italian, which was not found to be the case (Giardini et al 2019). The preliminary textual analysis reveals considerable intra-cultural variation in the perception and categorization of epistemic markers used as a shield in competitive advantage scenarios. These findings tie with the insights of Aikhenvald (2004:359) and Trudgill

(2011:182), highlighting the sociocultural aspects of evidentials. Our preliminary results suggest that the conceptualization of the role of epistemic markers in concealing deception varies across and within cultures.

**References**

Aikhenvald, A (2004), Evidentiality. OUP.

Giardini, F.; Fitneva, S.A.; Tamm, A. (2019), "Someone told me": Preemptive reputation protection in communication. *PLoS ONE 14(4):* e0200883. https://doi.org/10.1371/journal.pone.0200883

Johanson, L (2018), *Turkic indirectivity*. In Aikhenvald, A. ed. *The Oxford Handbook of Evidentiality.* OUP. 510–524

Erelt, M, H Metslang, K Pajusalu (2006), Tense and evidentiality in Estonian. In: B Cornillie, N Delbecque (Ed.). Topics in Subjectification and Modalization (125 −136). Amsterdam: Benjamins.

Trudgill, Peter (2011), Sociolinguistic Typology: Social Determinants of Linguistic Complexity. OUP.

# On Truth and Trickery:
# A compara5ve study of Bislama *giaman* 'trick, fib' and Anglo English *lying*

Carsten Levisen
(Roskilde University)

In Anglowestern discourses of truth, core dis4nc4ons revolve around "*truth* v. *lies*" and "*true* v. *false*". These opposites reflect the moral and scien4fic tradi4ons of Anglo English, and as such they can be viewed as discourses and dis4nc4ons that are the products of a par4cular conceptual universe, rather than of a universally shared discourse of truth (cf. Wierzbicka 2002, LaIas 2010).

In this paper, my main aim is to compare the Anglo linguis4c tradi4ons with the conceptual universe of Bislama (Vanuatu). As a "postcolonial language variety" (cf. Perez and Sippola 2021), Bislama is based on a lexical surface consis4ng mainly of words of English origin, but the conceptual universe associated with Bislama words, including the meaning of words and the discourse prac4ces that they habitually are embedded in, are not reflec4ve of Anglo English, but of urban South Pacific life, moral and values.

In the first part of the paper, I will survey the Bislama domain that corresponds roughly to Anglo English "truth talk", in order to iden4fy some of the most important overall differences, but also point to some shared aspect of meaning-making. The paper then turns to a compara4ve analysis of the Bislama concept of *giaman* 'trick, fib' (etymon: English *gammon*) as a discursive opposite of *talem tru '*saying (what is) true'.

The Natural Seman4c Metalanguage (NSM) approach (Goddard 2018) is employed to provide a high-resolu4on compara4ve seman4c analysis of Bislama *giaman* v. Anglo English *lying.* The study of *giaman* is based on ongoing seman4c fieldwork and discourse studies of *giaman* in interac4on (Levisen 2016, in press/2024). In a final discussion, the paper brings the lessons from Bislama and the compara4ve case study into a more principled discussion about postcolonial seman4cs (Levisen, in press/2024), contras4ve metapragma4cs (Goddard and Levisen 2023), and the diversity of truth-like constructs, truth-opposites, and truth- related dis4nc4ons in the world's languages.

**References**

Lattas, Andrew. 2010. *Dreams, Madness and Fairy Tales in New Britain*. Durham: Carolina Academic Press. Goddard, Cliff. 2018. *Ten Lectures on the Natural Seman;c Metalanguage: Exploring Language, Thought and Culture*. Amsterdam: Brill.

Goddard, Cliff and Levisen, Carsten. 2023. Contrastive metapragmatics and the shifting semantics of "sarcasm" in English and Danish. *Contrastive Pragmatics*.

Levisen, Carsten. 2016. The Ethnopragmatics of Speech Acts in Postcolonial Discourse: "Truth" and "Trickery" in a Transculturated South Pacific Tale. In C. Schubert, & L. Volkmann (Eds.), *Pragmatic Perspectives on Postcolonial Discourse: Linguistics and Literature* (pp. 41- 64). Cambridge Scholars Press.

Levisen, Carsten. In press/2024. *Postcolonial Semantics: Meaning and Metalanguage in a Multipolar World*. De Gruyter Mouton.

Perez, Danae & Sippola, Eeva. 2021. *Postcolonial Varieties in the Americas*. Berlin: de Gruyter Mouton.

Wierzbicka, A. 2002. Russian cultural scripts: The theory of cultural scripts and its applications. *Ethos* 30(4), 401–432.

# A corpus analysis of the Finnish quasi-construction

Daniel Ebner
(Humboldt-Universität zu Berlin)

Keywords: *corpus analysis*, *deception*, *Finnish language*, *quasi-construction*, *verb union*

This study discusses the Finnish *kvasirakenne* 'quasi-construction', based on data drawn from the Finnish Web 2014 (fiTenTen14) corpus at Sketch Engine (Kilgarriff et al. 2014).
The Finnish quasi-construction is one of the many verbal constructions in Finnish called verb unions (Karlsson 2017). It consists of an auxiliary, *olla* 'to be', and a non-finite form of the main verb containing the suffix *-vinA-* followed by a possessive suffix (VISK §435). Although the suffix *-vinA-* can etymologically be analysed as a plural form of the present participle in essive case (with clear-cut segmentation due to the agglutinative nature of Finnish), modern day speakers may – according to VISK (§ 57) – also perceive it as a morphological unit that is not segmented further; cf. example (1), where the verb *soittaa* 'to play' is used in the quasi-construction.

(1) *oli soittavinaan*
ol-i soitta-v-i-na-an [*or* soitta-vina-an]
be-PST.3SG play-PRS.PTCP-PL-ESS-POSS.3SG [*or* play-NMZL-POSS.3SG]
'seemed to play'

The quasi-construction modifies the meaning of the main verb to convey a range of different (but related) semantic concepts, among them pretence/inauthenticity, imagined/fictitious acts and uncertainty. The meaning intended by the speaker/writer of an utterance may be inferred from the co(n)text (but may also remain ambiguous); cf. example (2) and its different English translations, based on the Finnish periphrases provided by Salminen (1997: 40):

(2) *Matti **oli soittavinaan** kitaraa*. (Salminen 1997: 40)
a. Matti pretended to play the guitar.
b. Matti imagined himself playing the guitar.
c. It appears to me that Matti is playing the guitar.
d. It seems that Matti is playing the guitar.

According to Salminen (1997), the range of semantic concepts can on the one hand be surmised under the term 'polyphony' (*moniäänisyys*), i.e. conveying different simultaneous perspectives on an action; on the other hand, they may also be systematized regarding the implication of truthfulness, i.e. nonfactuality (the utterance does not specify whether the presented proposition of e.g. playing a guitar is truthful or not; such as translation (2c) above) and counterfactuality (where the utterance classifies an action as a deliberate deception (as, e.g. translation (2a) above). In the latter case, the quasi-construction can be described not as a lie told, but a lie or deceptive action acted out, called out and thematized in the discourse (for the relation between – and different conceptualisations of – the terms 'lie' and 'deception', cf. e.g. Meibauer 2018).
Using a randomised sample of 1000 hits of the quasi-construction drawn from the fiTenTen14 corpus at Sketchengine (Kilgarriff et al. 2014), I address the following questions in this study:

- Taking the co(n)text of the hits into account, how large is the proportion of counterfactual hits in the sample?
- Which verbs or semantic verb groups are attracted to the construction (and how is their distribution regarding the analysis of a sentence as counterfactual or not)? Previous research has shown that verbs of perception frequently occur in the quasi-construction (Kumpulainen 2019); however, for verbs of perception the modification of meaning can be surmised to indicate uncertainty/nonfactuality rather than counterfactuality. Therefore, it is expected that perception verbs will be less frequent in counterfactual contexts.
- For the counterfactual hits, is the deceptive nature of the quasi-construction reinforced by the left/right context and if so, by what means (e.g. scare quotes, adverbs, verbal repetition)?

**References**


Karlsson, Fred (2017), *Finnish. A Comprehensive Grammar*, London: Routledge.

Kilgarriff, Adam, Vít Baisa, Jan Bušta, Miloš Jakubíček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý and Vít Suchomel (2014), The Sketch Engine: ten years on, *Lexicography* 1, 7-36.

Kumpulainen, Aaro (2019), "Olen aistivinani jotain sellaista"*: kvasirakenteelle tyypillisimmät verbit Suomi24-keskusteluissa*. BA thesis. Jyväskylä: Jyväskylän yliopisto.

Meibauer, Jörg (2018), Introduction: What is lying? Towards an integrative approach, in J. Meibauer (ed.), (2018), *The Oxford Handbook of Lying*, Oxford: Oxford University Press, 1-10.

Salminen, Taru (1997), Kvasirakenteen kieliopillistuminen, in T. Lehtinen, Tapani and L. Laitinen (eds.), (1997), *Kieliopillistuminen. Tapaustutkimuksia suomesta*. Kieli 12, Helsinki: Helsingin yliopiston suomen kielen laitos, 38-65.

VISK = Auli Hakulinen, Maria Vilkuna, Riitta Korhonen, Vesa Koivisto, Tarja Riitta Heinonen and Irja Alho (2004), *Iso suomen kielioppi*, Helsinki: Suomalaisen Kirjallisuuden Seura. Online version, accessed 15. January 2024, http://scripta.kotus.fi/visk URN:ISBN:978-952-5446-35-7

# The Patterns of Deception:
# The lying/misleading distinction from the perspective of truth evaluators

Shirly Orr
(Tel Aviv University)

Background: The two dominant definitions of Lying face a challenge when trying to distinguish lying from misleading claims. The traditional view asserts that lying involves false *explicit* content, while misleading claims involve false *implicated* content (1,5, but see 2).
Recent empirical studies show, however, that speakers can be seen as lying even when the false content is implicated (e.g., 3,7,4). A more recent view defines lying based on the speakers' perceived commitment to the false communicated content (4,6). While this view effectively accounts for the empirical evidence, it, nevertheless, obscures the boundary between lies and misleading claims, as even misleading claims involve some degree of commitment.

Objectives: The current study aims to draw a clearer distinction between lies and misleading claims by addressing the empirical methods used to gauge the distinction and focusing on the oft-overlooked role of truth evaluators in shaping this distinction.

The current study: I examined participants' disposition when they truth evaluated a literally true claim that indirectly conveyed a believed-false claim. Indirectness was manipulated using the underexplored in this domain, epistemic modals, *possible* and *possibly*, which express uncertainty (*N*=320, ~80 per condition). To establish that these expressions are *merely* misleading, I compared them to a literally true claim and a full-fledged lie (thus addressing Goal#1). Figure-1 outlines the experimental procedure.

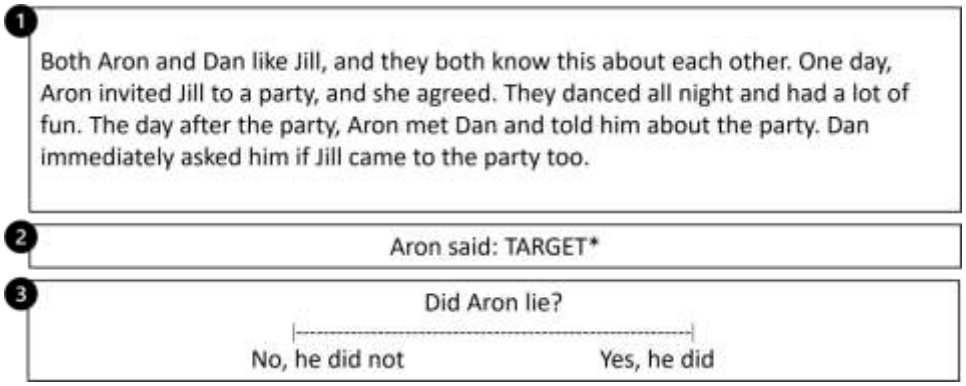



Figure-1 – Experimental Procedure

*See Table-1 for target sentences.

Table-1 – Targets' Conditions

| Condition | Target sentence | Description |
|---|---|---|
| Condition-1 | "It is possible that Jill came to the party" | literally true *p* + possible |
| Condition-2 | "Jill possibly came to the party" | literally true *p* + possibly |
| Condition-3 | "Jill did not come to the party" | literally false *p* |
| Condition-4 | "Jill came to the party" | literally true *p* |

Results: The means' lie ratings show that the literally false claim received a high lie rating (Condition-3; *M*=97.3, *SD*=11.4) and the literally truthful claim received a low lie rating (Condition-4; *M*=6, *SD*=12.2)—as expected. The hedged literally true claims received an intermediate lie rating (Condition-1: *M*=50.9, *SD*=32.6; Condition-2: *M*=56.6, *SD*=12.2). Data was analyzed using a Bayesian Zero-One-Inflated-Beta (ZOIB) analysis, which indicated that all conditions were significantly different from one another, except for Condition 1 with 2.

Discussion: These findings indicate that hedging a literally true claim using epistemic modals is a misleading act. As such, they show that not only is it possible to lie using implicated content, but it is also possible to merely mislead using false explicit content. These findings, nevertheless, do not directly explain the lying/misleading distinction (Goal#2). To do this, I adopted the truth evaluators' perspective. A closer look at the truth evaluators' behavioral patterns suggests that two distinct mindsets underlie the evaluation of different forms of deception (Figure-2). In misleading claims, participants are conflicted, likely due to the presence of two opposing truth-values.

They resolve this conflict by leaning towards one of the truth-values (evident through the bimodal distribution and its wide range). In full-fledged lies and truthful claims, truth evaluators experience no such conflict (evident through the skewed distribution with its narrow range).

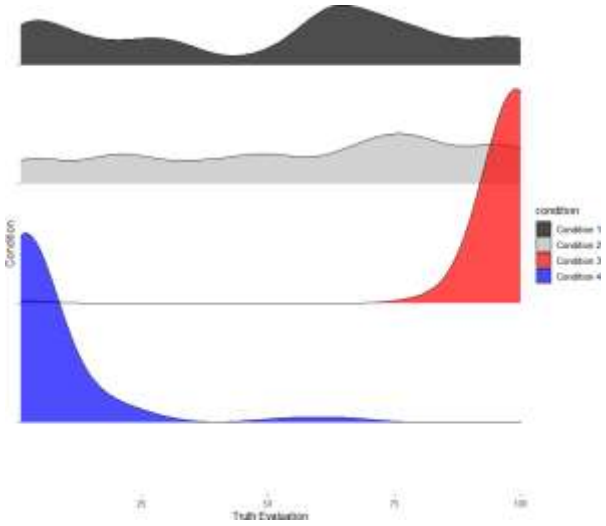



Figure-2 – The Patterns of Deception

**References**

Adler, J. E. (1997). Lying, Deceiving, or Falsely Implicating. *The Journal of Philosophy, 94(9)*, 435–452. https://doi.org/10.2307/2564617

Meibauer, J. (2005). Lying and falsely implicating. *Journal of Pragmatics, 37(9)*, 1373–1399. https://doi.org/10.1016/j.pragma.2004.12.007

Orr, S., Ariel, M., & Peleg, O. (2017). The case of literally true propositions with false implicatures. In I. Chiluwa (Ed.), *Deception and deceptive communication: Motivations, recognition techniques and behavioral control* (pp. 67–107). Nova Science Publishers, Inc.

Reins, L. M., & Wiegmann, A. (2021). Is Lying Bound to Commitment? Empirically Investigating Deceptive Presuppositions, Implicatures, and Actions. *Cognitive Science, 45(2)*, e12936. https://doi.org/10.1111/cogs.12936

Saul, J. M. (2012). *Lying, Misleading, and What is Said: An Exploration in Philosophy of Language and in Ethics*. Oxford University Press.

Viebahn, E. (2021). The Lying-Misleading Distinction: A Commitment-Based Approach. *The Journal of Philosophy, 118(6)*, 289–319. https://doi.org/10.5840/jphil2021118621

Weissman, B., & Terkourafi, M. (2019). Are false implicatures lies? An empirical investigation. *Mind & Language, 34(2)*, 221–246. https://doi.org/10.1111/mila.12212

# Truly speaking in Europe:
# A corpus-based study of truth-related words in thirteen European languages

Maria Koptjevskaja-Tamm & Volker Gast
(Stockholm University & Friedrich Schiller University)

Keywords: lexical semantics; subjectivity; intersubjectivity; discourse functions; concessives

Orr & Ariel (2021) demonstrate that *true* predications in spoken American English corpora have two major uses – "descriptive" vs. "performatory" (following Austin 1950 vs. Strawson 1950) – and argue that the notion of "stance triangle" (Du Bois 2007) provides a useful framework for understanding this distinction. Expressions of truth imply three linguistic actions: the speaker evaluates an object, subjectively positions themselves relative to that object, and aligns themselves with (an)other participant(s) in a convergent or divergent manner. Different readings of *true* predications arise when specific elements of the action are fore– or backgrounded. Descriptive vs. performatory uses of *true* predications in Orr & Ariel (2021) constitute a subset of the various descriptive and (inter)subjective meanings expressed by truth-related words, including the more grammaticalized ones (e.g., markers of contrast, emphasis, intensification, concessives and tag questions; cf. Wierzbicka 2002b, Apresyan 2010, Bardenstein & Ariel 2021). As shown in Beekhuizen et al. (2023), different parts of speech differ in their propensity to be used for descriptive and performatory ([inter]subjective) purposes, following the hierarchy 'Noun > Adjective > Adverb'.

In our study, based on comparative corpus data across thirteen European languages from the Europarl corpus (Koehn 2005) – English, German, Swedish; French, Spanish, Italian; Polish, Czech, Slovene; Latvian, Lithuanian; Estonian, Finnish – we focus on the discourse functions of *truth* expressions. In parliamentary speech (and, arguably, in everyday conversation), *truth* predications are typically used to signal either agreement (cf. 1) or disagreement (cf. 2), and if they signal agreement, they are often concessive or adversative (cf. 3). In accordance with their communicative functions, truth expressions moreover differ in their degrees of (inter)subjectivity.

(1) I should like to say to Frau Gebhardt and Mr De Clercq who mentioned this that we are certainly working hard at realising the possibility of having a Community patent [...] It is indeed *true* that the language problems are considerable. [EP]
(2) Mr Dary suggested that if we were to move towards a 'tariff only' WTO system we would somehow be surrendering ourselves to chaotic commercial forces. The *truth* is quite the reverse. [EP]
(3) Everything that has been said here about the relation to trade policy, access to our markets and the need to diversify agricultural production and the economy in these countries is *true*, but it is also *true* that there is a sort of de-linkage from all this affecting the very poor subsistence farmers. [EP]

Our data suggest correlations (i) between the hierarchy of discourse functions 'disagreement > concession > agreement' and the part-of-speech hierarchy 'Noun > Adjective > Verb'; and (ii) between the hierarchies 'objective > subjective' and 'Noun > Adjective > Adverb'. Moreover, certain collocates are symptomatic of specific types of truth expressions. For example, positive-emotive content words seem to pattern with attributive *true* (*democracy*, *solidarity*, *social*, *equality*), *but* and *not* tend to be found in the context of concessive uses of *true*, negatively connotated words (*crimes*, *commited*, *war*) are associated with the noun *truth*, and positive-emotive words (*hope*, *inclusive*, *open*, *thank*) with the adverb *truly*.

Considering the discourse contexts in which *truth* expressions occur will allow us to gain a better understanding of the network structure of this lexical domain. Our guiding hypothesis is that in dyadic discourse, *true* acquires an interpersonal function because the predication it relates to is under discussion,

and is re-evaluated (confirmed, strengthened, relativized, challenged). Strengthening typically accompanies positively evaluated propositions, relativizing and challenging are often found with negatively evaluated propositions. In this way *truth* expressions acquire (inter)subjective connotations which, due to their discourse function, are associated with different parts of speech. For example, in corrective uses the *truth* predicate is in the background and the predication is in focus, while in confirmatory uses the predication remains in the background and the *truth* predicate is in focus.

**References**


Apresjan, Valentina (2010), From 'Truth' to Concessives: Semantic Development, in Zybatow, G. et al. (eds), *Formal Studies in Slavic Languages*, 175–192, Frankfurt: Lang.

Austin, John L. (1950), Truth, *Aristotelian Society Supplementary Volume* 24, no. 1: 111–128, https://doi.org/10.1093/aristoteliansupp/24.1.111.

Bardenstein, Ruti and Mira Ariel (2021), The missing link between truth and intensification, *Studies in Language* 46(2): 285–322, https://api.semanticscholar.org/CorpusID:233583765.

Beekhuizen, Barend, Maya Blumenthal, Lee Jiang, Anna Pyrtchenkov, and Jana Savevska (2023), Truth be told: A corpus-based study of the cross-linguistic colexification of representational and (inter)subjective meanings, *Corpus Linguistics and Linguistic Theory* (ahead of print), https://doi.org/10.1515/cllt-2021-0058.

Du Bois, John W. (2007), The stance triangle, in: Englebretson, R. (ed.), *Stancetaking in Discourse: Subjectivity, Evaluation, Interaction*, 139–182, Amsterdam: John Benjamins.

Koehn, Philipp (2005), Europarl: A Parallel Corpus for Statistical Machine Translation, in *Proceedings of Machine Translation Summit X*: Papers, 79–86, Phuket, Thailand. https://aclanthology.org/2005.mtsummit-papers.11.

Orr, Shirly and Mira Ariel (2021), Predicating truth: An empirically based analysis, *Journal of Pragmatics* 185: 131–45, https://doi.org/10.1016/j.pragma.2021.09.005.

Strawson, Peter F. (1950), Truth, *Aristotelian Society Supplementary Volume* 24, no. 1: 129–156, https://doi.org/10.1093/aristoteliansupp/24.1.111.

Wierzbicka, Anna (2002), Philosophy and discourse: The rise of 'really' and the fall of 'truly', *Cahiers de Praxématique* 38(3): 85–112.

# Extended functions of ‘truth’-words

Anna Sjöberg, Alice Bondarenko & Anastasia Panova
(Stockholm University, Stockholm University & Stockholm University)

Keywords: grammaticalization, truth, lexical typology, semantics, pragmatics

As follows from the cooperative principle (Grice 1975), in conversation one is generally expected to speak the truth, and consequently people tend to assume that what is being said is true. In addition, speakers are expected to be relevant, which makes explicit mention of the truthfulness of a proposition redundant (cf. Orr & Ariel 2021: 131-132). Therefore, ‘truth’-words tend to carry other functions, including intensification (Kuteva et al 2019; Beekhuizen et al 2023: 20-21), tag questions (Svennevig 2007) and concessive marking (König 2006).

In this paper, we analyse the extensions of ‘truth’-words in a diachronic and areal-typological perspective, using a sample of 50 spoken and signed Eurasian languages. We identify types of extensions and their relations to each other and consider possible genealogical biases and contact effects.

Our preliminary results show that the development from ‘truth’ to intensification is the most frequent type of extension. This function is illustrated in (1), from Ukrainian.

| | | | | | | |
|---|---|---|---|---|---|---|
| (1) | ***Spravdi*** | *ljuba,* | *dyvnyj* | *son* | *tobi* | *prysnyvsja* |
| | **truly** | dear | amazing | dream | you.DAT | dream.PRS.3SG |

‘It was a curious dream, dear, certainly’ (RNC: Alice in Wonderland)

Upon a closer look, intensification is complex and at least two subfunctions related to ‘truth’-words can be distinguished: ‘really, indeed’ (see example 1 above) and ‘very’ (see example 2 from Baka, Atlantic-Congo).

| | | | | |
|---|---|---|---|---|
| (2) | *wósɛ́* | *ʔé* | ***ko*** | *jókò!* |
| | woman | 3:SG | **truly** | beauty |

‘The woman is **very** pretty!’ (Kuteva et al 2019: 443)

One aspect of the difference between the two functions is the scope of the intensifier — one of the parameters which we explore in connection with ‘truth’-words. In our pilot study, in contrast to the ‘really, indeed’ uses, the instances of ‘truth’ words that have developed into ‘very’ intensifiers (such as in English, Fr. *vrai* -> Eng. *very*) are infrequent.

Another frequent function of ‘truth’-words is as a tag-question, cf. Kalhori Kurdish (3):

| | | | | | |
|---|---|---|---|---|---|
| (3) | *dāya* | *bayāni* | *ǰoma* | *as,* | ***dorosa****?* |
| | Mom | tomorrow | Friday | is | **true** |

‘Mom, tomorrow is Friday, **isn’t it?**’ (elicited)

We argue that a possible secondary development to the use of 'truth'-words as tag-questions is their use as affirmation markers. This development might be facilitated by similar syntactic contexts of juxtaposition with a declarative clause, as illustrated in Swedish Sign Language in (4).

(4) *(...)* *MÅNGA* *HINNA.INTE* *PRATA* *MED* *ALLA* ***SANN***
[...] MANY NOT.HAVE.TIME SPEAK WITH EVERYONE **TRUE**
**'Yes**, there's too many to speak to everyone' (STS-korpus 2023, SSLC)

Finally, we identify the concessive function, best-described for Russian *pravda* 'truth' (Apresjan 2010). Moreover, we also find it, for example, in Adyghe (Northwest Caucasian), see (5), raising questions about possible contact-induced grammaticalization.

(5) *təʁe-r,* ***ŝəpqe,*** *qə-qʷe-ç'ə-ʁa-x,* *aw* *ǯ'ə-rjə* *čəʔe*
sun-ABS truth CSL-LOC-exit-PST-IAM but now-ADD cold
'Although the sun has already risen, it's still cold.' (Vodoždokov 1960: 672)

In our talk, we will generalise the results of our study as a semantic map, discuss our findings in light of grammaticalization theory and pay special attention to the interplay of grammar and discourse.

**References**


Apresjan, V. Y. (2010). From truth to concessives: Semantic development. *Formal Studies in Slavic Linguistics* (25), 175–192.

Beekhuizen, B., Blumenthal, M., Jiang, L., Pyrtchenkov, A. & Savevska, J. (2023) "Truth be told: a corpus-based study of the cross-linguistic colexification of representational and (inter)subjective meanings". *Corpus Linguistics and Linguistic Theory*. https://doi.org/10.1515/cllt-2021-0058

Grice, H. P. (1975). Logic and conversation. In Cole, P. & Morgan, J., (eds.): *Speech acts*, 41–58. New York: Academic Press.

Kuteva, T., Heine, B., Hong, B., Long, H., Narrog, H., & Rhee, S. (2019). *World Lexicon of Grammaticalization* (2nd ed.). Cambridge: Cambridge University Press.

König, E. (2006). Concessive clauses. In Asher, R. E. (ed.): *The encyclopedia of language and linguistics*, 679–681. Oxford: Pergamon.

Orr, S., & Ariel, M. (2021). Predicating Truth: An empirically based analysis. *Journal of Pragmatics* (*185*), 131–145.

RNC: Russian National Corpus. https://ruscorpora.ru

Svennevig, J. (2007) "Ikke sant" as a response token in Norwegian conversation. In Nilsen, R-A., N.A.A. Amfo & K. Borthen (eds.): *Interpreting Utterances. Pragmatics and Its Interfaces. Essays in Honour of Thorstein Fretheim,* 175-189. Oslo: Novus.

Tang, G. (2006). Questions and negation in Hong Kong Sign Language. In Zeshan, U. (ed.): *Interrogative and negative constructions in sign languages (Sign Language Typology Series No. 1)*. 200–224. Nijmegen: Ishara Press.

Vodoždokov, X. D. (1960). *Russko-adygejskij slovar'* [*Russian-Adyghe dictionary*]. Moscow: Gosudarstvennoe izd-vo innostrannyx i nacional'nyx slovarej.

# Falsehood exposed:
# On verbs telling us about getting rid of falsehood from discourse

Joanna Zaucha
(Cardinal Stefan Wyszynski University)

Keywords: semantics, verbs, falsehood, denying

As Sissela Bok (1999) noted, the amount of literature devoted to truth vastly exceeds the number of studies on falsehood. Even though recent works of Jörg Meibauer (2014, 2018), Thomas Carson (2010), Jennifer M. Saul (2012), Jennifer Lakcey (2013), Andreas Stokke (2018) and others bridge this gap to some extent, there are still issues undiscussed and unsolved.

In my presentation I would like to focus on the class of verbs, which I have called falsehood exposing verbs. They are used to describe struggling in getting rid of falsehood from discourse. This subclass includes verbs such as: *rectify* [something], *refute* [something], *true up* [something], *dispel* [something], *correct* [something], *deny* [something], *belie* [something]. What is common to these lexemes it that their meaning contains two presuppositions: 1) the content expressed in the subordinate sentence was an element of the discourse and 2) the complement proposition is false from the perspective of a speaker and an agent (i.e. a person who actually corrects something).

The insight into functional differences between English falsehood exposing verbs will reveal the ways of lexicalisation of processes they refer to. My analysis will be carried out taking into account so called negative material, i.e. syntactically well-built but logically defective sentences (Bogusławski 2009, 2023) as well as normal usage of the verbs mentioned. One of the most interesting differences within the considered subclass of verbs is the presence of an element referring to negative evaluation of the agent's action. We can see that the neutral verb *deny* can be used together with such an adjective, while *correct* not: + *Barbara **denied** the information that her company was in financial trouble, which was reprehensible*. vs. **Barbara **corrected** the information that her company was having financial problems, which was reprehensible.* Moreover, the negative material reveals that denial is semantically simpler than refutation. The latter implies the former, but not the other way around: **Miller **refuted** the information about his resignation, but did not **deny** it*. vs. *+ Miller **denied** the information about his resignation, but did not **refute** it.* A careful study of the semantic similarities and differences between verbs exposing falsehood is a good way to identify the logic of dealing with falsehood in discourse. Showing it as a whole is the main objective of my presentation.

**References**

Bogusławski, Andrzej (2009), *Myśli o gwiazdce i regule*, Warszawa: BelStudio.

Bogusławski, Andrzej (2023), *Metodologia lingwistyki. Zarys*, Warszawa: WUW.

Bok, Sissela (1999), *Lying. Moral choice in private and public life*, New York: Vintage Books/ Kindle Edition.

Carson, Thomas (2010), *Lying and Deception: Theory and Practice*, Oxford: UOP/ Kindle Edition.

Meibauer, Jörg (2014), *Lying at the Semantics-Pragmatics Interface*, Mouton Series in Pragmatics, De Gruyter Mouton/ Kindle Edition.

Meibauer, Jörg (2018)., Lying, implicating, presupposing, In Meibauer J. (ed.). *The Oxford Handbook of Lying*, Oxford: UOP, 193-202.

Saul, Jennifer Mather, 2012, *Lying, misleading and what is said. An exploration in Philosophy and Language*, Oxford: OUP.

Stokke, Andreas (2018), *Lying and Insincerity*, Oxford: UOP.

Lackey, Jennifer (2013), Lies and Deception: An Unhappy Divorce, *Analysis* 73(2), 236-248.

Mahon, James Edwin (2016), The Definition of Lying and Deception. In Edward N. Zalta (ed.), *The Stanford Encyclopaedia of Philosophy*, https://plato.stanford.edu/archives/win2016/entries/lying-definition, [online 15.07.2023].

Mahon, James Edwin (2018), Contemporary approaches to the philosophy of lying. In Meibauer J. (ed.), *The Oxford Handbook of Lying*, Oxford: UOP, 32-55.

# The Discourse of Quality in Mopan Maya

Eve Danziger
(University of Virginia)

Keywords: intentionality, Mayan, Grice, language ideology, lying

Grice's (1989) Maxim of Quality instructs co-operative speakers to "try to make your contribution one that is true; do not say what you believe to be false". But in Mopan Maya, violations of Quality are routinely taken to have occurred regardless of the belief and desire states of the utterer, and "reflects a culturally particular and not a culturally universal view of what constitutes a blameworthy violation" (Danziger 2010: 211). Danziger (2010, 2013) describes the traditional Mopan view that correct and 'respectful' human conduct is key to keeping the universe functioning, so that false speech entails cosmic danger regardless of the internal states or intentions of the speaker. The present paper extends our understanding of the cultural relativity of Gricean Quality by investigating the lexicon of 'trying' and 'believing' -- so central to Grice's formulation -- in Mopan Maya. It shows that, like Mopan *tus* 'lying', both Mopan *yaal* 'trying' and *tz'okes* 'believing' are centrally understood in ways that are independent of mental state attribution or calculation.

Corpus analysis of 170,000 words across 75 Mopan texts (Ventur 1976) shows that although a number of Mopan predicates exist to express mental states (for example *k'ati* 'wanting; desiring'), neither *yaal* 'trying' nor *tz'okes* 'believing' are among them. Mopan *yaal* 'trying' does not include the component of mental effort that is most centrally at issue in Grice's use of English *try* in his formulation of the Maxim of Quality. Instead, the Mopan term canonically refers to what in English could be re-formulated as 'trying out' or 'testing'. In various other languages, of course, this English polysemy is also not found, with two contrasting lexemes expressing the two distinct meanings (cf. Dutch *proefen* 'to test; to try out' in contrast to *probieren '*to move toward a goal by making a mental effort'). But in Mopan no second lexeme exists to encode the 'mental effort' aspects of English *try.* In parallel, Mopan *tz'okes* 'believing' is more centrally concerned with physical obedience than with mental commitment (Danziger 2013).

This paper concludes that the irrelevance of 'trying' and 'believing' to Mopan understandings of what constitutes a violation of the Maxim of Quality is not incidental. In fact, even if Mopan *yaal* 'trying' and *tz'okes* 'believing' were included in its formulation, the resulting Mopan Maxim would not centrally include reference to mental states.

**References**

Danziger, Eve, 2013. Conventional Wisdom: : Imagination, Obedience and Intersubjectivity. *Language and Communication* 33(3): 251-262.

Danziger, Eve. 2010. On Trying and Lying: Cultural Configurations of the Gricean Maxim of Quality. *Intercultural Pragmatics* 7(2):199-219..

Grice H. P. 1989. Logic and conversation. *Studies in the Way of Words*. Cambridge: Harvard University Press, pp. 22-40

Ventur, Pierre, 1976. Mopan Maya Folklore Texts. Smithsonian Institution (Accession #80-38)

# The unrealistic world of uso

Fumiko Nishimura
(University of Waikato)

Keywords: Japanese communication, deviation, proverbs, falsehood, cultural protocols

Telling or not telling the truth is one of the most fundamental issues of communication. When it comes to the study of this area in Japanese, quite a few studies focus on 'uso' as the word 'uso' is commonly regarded as the equivalent to 'lie' in English. However, 'uso' does not seem to be the same as 'lie'. For example, Yoshimura (1995) finds falsehood plays the most important role in defining 'uso.' In contrast, Coleman and Kay (1981) find that the speakers' belief on falsehood is more important than falsehood to define 'lie.' In this presentation, I examine the word 'uso' used in everyday conversation and the proverbs/synonyms of 'uso' aiming at Illustrating the characteristics of 'uso.'

It is well-known for Japanese people to use the word 'uso' like an interjection to give an instantaneous remark on the given information (Sakaba, 2020). This type of usage is not necessarily taken as offensive (i.e., it does not sound like 'that's a lie!'). An analysis of conversational data reveals such remarks are used to refer to deviations the speakers recognise. Here, deviations mean a difference between the information given to the speakers and the authenticity, ideal or realistic situations the speakers believe. Thus, 'uso' is not just about falsehood.

The collected 'uso'-related proverbs suggest the following significant elements on the practice of 'uso': 1) unethical aspect of lying, 2) useful side of lying and 3) desirable ways of lying. 'Uso mo hoben,' one of the most famous uso-related proverbs (*lit.* a lie could be a means), is an example of Point 2 above and suggests that it is good/necessary to use a lie to manage obstacles. This type of lie might be identified as a white lie which can be observed in any other country (Terkourafi, 2019). However, the lies this Japanese proverb refers to are not limited to white lies. Or we might have to say that Japanese people would define white lies differently.

The collected synonyms of 'uso' show three distinguish elements among them: 1) distance from the truth, 2) manners of distortion of the truth and 3) speakers' intention.

While 'uso' is commonly translated as 'lie,' the above data suggest that its meanings extend well beyond mere falsehood. The 'uso' or non-truth telling practice in Japanese encompasses not only the utterance of false statements but also the speakers' intention and manner by which distorted information is conveyed. These findings underscore the importance of considering cultural protocols in interpreting 'uso' comprehensively.

**References**

Coleman, L. and Kay, P. (1981), Prototype Semantics: The English word Lie. *Language* 57 (1), 26–44.

Sakaba, H. (2020). The Difference of Japanese Uso and English Lie from the Perspective of Speech Acts. *Osaka University Knowledge Archive : OUKA* (29), 35–55.

Terkourafi, M. (2019), Lying and politeness, in J. Meibauer (ed), *The Oxford handbook of lying*, Oxford: Oxford University Press, 382–396.

Yoshimura, K. (1995), *Ninchiimiron no hoho*. Tokyo: Jimbun Shoin.

# Talking the truth, lying and deceiving in Ancient Egyptian

Gaëlle Chantrain
(UCLouvain)

In this paper, I propose to give a general overview of lexical expressions pertaining to the semantic domains of truth/honesty and lie/deception in Ancient (pre-Coptic) Egyptian (AE), an Afro-asiatic language that was spoken between ca. 2700 BC and 400 AD. Attested on the Longue Durée (over 3000 years), AE offers a unique opportunity to look at the evolution of semantic fields in diachrony and permits to actually witness processes which most often can only be inferred in the frame of a synchronic approach. The first part of the paper will deal with the core lexicon of truth and lie. Truth is expressed in the first place in AE by the lexeme *mȝʿ.t* "truth", "rectitude" obtained as derived meaning from « straightness » (as concrete, physical quality), a semantic extension that is well attested cross-linguistically (Yu & al. 2016; Pagliaro & al. 2018; Zhu & al. 2023). In the domain of truth, the root *mtj/mtr* "precise correct" and, by extension, "honest, sincere" can also be added to the picture. They stand in partially asymmetrical antonymic relation with the lexemes *grg* "lie, deception", *jsf.t* "evil, lie" and *jwms* "untruth, lie", whose etymologies appear a bit more obscure. Their respective semantics and context of use will be unfolded in order to understand how they relate to each other in synchrony and diachrony.

The second part of the paper will give an overview of (compound) metaphorical expressions drawing from various domains that express the concepts of truth and lie/deception (e.g. *shȝj ḥȝty* « to make the heart go down » for « disappointing » or « cheating » someone).

The data presented in this study have been gathered systematically through targeted searches on several parameters in two main thesauri for ancient Egyptian texts: the *Thesaurus Linguae Aegyptiae* (https://thesaurus-linguae-aegyptiae.de/home) and the *Ramses* database (http://ramses.ulg.ac.be ). The parameters taken into consideration in primary data survey are: translation in context, semantic classifier (i.e. graphemic element in Egyptian that provides information on the semantics the lexeme), search on the two words for "heart" (*jb* and *HAty*) and their compounds, as well as on other body parts (e.g. head, face, belly, etc.), regularly found in emotion and cognition expression. Further lexical units pertaining to the domains of truth/honesty and lie/deception have been evidenced by the systematic analysis of the co(n)text of the primarily evidenced lexemes and expressions. The gathering of these data and their analysis lay the ground for further cross-linguistic comparison of the source conceptual domains of truth and lie (1) as regards their validation in chronological depth - modern vs ancient languages (Di Biase-Dyson & Egg 2020) and (2) in an areal perspective, between AE and Afro- asiatic languages for instance (see i.a. Vanhove & Ahmed 2022).

**References**

Di Biase-Dyson, C. & Egg, M. 2020. Drawing Attention to Metaphor. Case Studies across time periods, cultures and modalities, Figurative Thought and Language 5, John Benjamins.

Pagliaro, S. & al. 2018. Straight to heaven: Rectitude as spatial representation of morality. Eur. J. Soc. Psychol. 2018, 48, 567–724.

Vanhove, M. & Ahmed, M-T. 2022. Heart Metaphors in Beja, Language in Africa 3 (2), 368-388. https://doi.org/10.37892/2686-8946-2022-3-2-368-388

Yu, N., Wang, T., & He, Y. 2016. Spatial subsystem of moral metaphors: A cognitive semantic study. Metaphor and Symbol, 31, 195–211. https://doi.org/10.1080/10926488.2016.1223470

# Contrastive semantics of “truth” and “lying” in English and Russian (NSM approach)

Cliff Goddard, Anna Gladkova & Anna Wierzbicka
(Griffith University, Lomonosov Moscow State University & Australian National University)

The empirical focus of this paper is certain key English and Russian words related to truth and lying. The analytical framework is the NSM approach to meaning (Goddard and Wierzbicka 2014), wherein complex, language-specific meanings are paraphrased into simple, cross-translatable words (semantic primes). The paper has five parts.

(1) Foundations. We argue that TRUE is a semantic prime (i.e. impossible to satisfactorily paraphrase without circularity) and that it is expressible in all languages in two simple, cross-translatable sentences: ‘this is true’ and ‘this is not true’ (cf. Wierzbicka 1996, 2014a). Specifying these two canonical sentences helps resolve potential confusions involving polysemic meanings of the lexical items in question, e.g. English *true*, Russian *pravda*, Finnish *tosi~totta*, Malay *benar*, Mandarin *zhēn*. The paper illustrates how language-specific words and phrases can be revealingly analysed using TRUE along with other primes and cross-translatable expressions.

(2) Contrastive semantics of “lying” in English and Russian. We focus on the English verbs *lie* and the Russian verbs *lgat’* “lie” and *vrat’*, roughly, “distort truth” (cf. Gladkova 2007; Apresjan V. 2010). The analysis pinpoints similarities and differences in meaning between the three words. For example, Russian *lgat'* has a stronger negative evaluation (‘people know that it is very bad if someone does this’) than does English *lie*. Russian *vrat'* is based on ‘it is not like this’, rather than on ‘this is not true’. These meaning differences are shown to correlate with various distributional and phraseological properties of the respective verbs.

(3) Contrastive semantics of “truth”. We argue that the English expression *the truth* is not, as widely assumed, a mere noun equivalent of *true*, but has additional semantic components connected with something like importance (‘it is good if people want to know this’); for example, one can use the adjective *true* about something trivial, as in *It is true that I like/don’t like spinach,* but normally one would not use the noun *truth* like that (cf. Goddard 2020). As for Russian, we argue that predicative noun-like *pravda-1* corresponds to the semantic prime TRUE, whereas fully nominal *pravda-2* has additional semantic components (cf. Apresjan V. 2010; Wierzbicka 2014a).

(4) Cultural and ethnopragmatic considerations. We discuss differing cultural attitudes and traditions towards saying true and untrue things (cf. Šmelev 2005; Wierzbicka 2002a, 2002b). These considerations help explain different patterns of usage of key words in the Russian and English linguacultures (cf. Peeters 2018; Levisen 2016).

(5) Wider perspectives. Concluding, we argue that cross-cultural investigations of “lying” and “deception”, not only in linguistics but also in experimental philosophy, cross-cultural psychology, and other disciplines (Meibauer 2018), often go astray by relying too heavily on Anglo/Euro conceptions of “lying” and/or by using ‘true’ and ‘truth’ interchangably. Grounding our investigations in clear, cross-

translatable words allows for greater precision and reduced risk of Anglocentrism (Wierzbicka 2014b; Levisen in press/2024).

**References**


Apresjan, V. 2010. Slovarnye stat'i pravda, nepravda, istina, obmanut', obman, vrat', vran'e, lgat', lož', pravdivyj, lživyj. [Dictionary entries pravda, nepravda, istina, obmanut', obman, vrat', vran'e, lgat', lož', pravdivyj, lživyj]. In J. Apresjan (Ed.) *Prospekt aktivnogo slovarja russkogo jazyka.* [Prospectus of the Russian Active Dictionary], pp. 170-184. Moskva: Jazyki slavjanskix kul'tur.

Gladkova, A. 2007. *Russian Emotions, Attitudes and Values: Selected topics in cultural semantics.* PhD thesis, The Australian National University. [https://openresearch-repository.anu.edu.au/handle/1885/151599]

Goddard, C. and A. Wierzbicka 2014. *Words and Meanings: Lexical Semantics Across Domains, Languages and Cultures*. OUP.

Goddard, C. 2020. Overcoming the linguistic challenges for ethno-epistemology: NSM perspectives. In M. Mizumoto, J. Ganeri and C. Goddard (eds.), *Ethno-Epistemology: New Directions for Global Epistemology*, pp. 130-153. Routledge.

Levisen, C. 2016. The ethnopragmatics of speech acts in postcolonial discourse: "truths" and "trickery" in a transculturated South Pacific tale. In C. Schubert and L. Volkmann (eds.) *Pragmatic Perspectives on Post-colonial Discourse: Linguistics and Literature,* pp. 41–64. Newcastle upon Tyne: Cambridge Scholars.

Levisen, Carsten. In press/2024. *Postcolonial Semantics: Meaning and Metalanguage in a Multipolar World*. De Gruyter Mouton.

Meibauer, Jörg. (ed.) 2018. *The Oxford Handbook of Lying*. OUP.

Peeters, Bert, 2018. (White) lies and (pieux) mensonges: Ethnolinguistic elaborations on not telling the truth. *Etnolingwistyka,* 30, 169-188.

Šmelev, A. 2005. Kommentarii k stat'e Anny Vežbickoj [Comments on Anna Wierzbicka's article]. In Zalizniak, A., Levontina, I., and A. Šmelev. *Ključevye idei russkoj jazykovoj kartiny mira* [Key ideas of the Russian linguistic picture of the world], pp. 500-510. Moskva: Jazyki slavjanskoj kul'tury.

Wierzbicka, Anna. 1996. *Semantics: Primes and Universals*. New York: OUP.

Wierzbicka, Anna. 2002. Russian cultural scripts: The theory of cultural scripts and its applications. *Ethos*, 30(4), 401-432.

Wierzbicka, Anna. 2002b. Philosophy and discourse: The rise of "really" and the fall of "truly". *Cahiers de praxématique*, 38, 85-112.

Wierzbicka, Anna. 2014a. It is TRUE. Ms.

Wierzbicka, Anna. 2014b. *Imprisoned in English: The Hazards of English as the Default Language*. New York: OUP.

# “They say” makes good liars: A cross-linguistic investigation on evidentiality in language and deception

Selma Berfin Tanis, Cagla Aydin, Amanda Edmonds, Ayberk Kaan Gunes, Seyma Kalender, Junko Kanero, Fanny Meunier, Shona Whyte & Seckin Arslan
(Sabancı University, Sabancı University, Université Côte d'Azur, Sabancı University, Sabancı University, French National Center for Scientific Research (CNRS), Université Côte d'Azur/Nice & French National Center for Scientific Research (CNRS)

Keywords: Lie-telling, deception, evidentiality, Turkish, French

We know little about the grammatical composition of deception. Studies have found that, as compared to truthful ones, deceitful speakers produce relatively more negation words (Hauch et al. 2015), conditionals (Meibauer 2018), inconsistent person and tense marking (Porter and ten Brinke 2010). Little research has focused on other grammatical cues, however. One such grammatical category is evidentiality codifying through which sources of information a speaker knows about an event in his/her own statement (Aikhenvald 2004). In English, perception verbs are used to express reported information (e.g., *I hear that*; Whitt 2010). French employs a non-obligatory conditional form to denote unattested knowledge (Dendale 1993). Turkish inflectional marker *-dı* indicates direct firsthand information while *-mış* codifies inferred or reported information (Aksu-Koç and Slobin 1986). Japanese uses *-rashii* or *-souda* hearsay markers when referencing an indirect report as compared to unmarked direct reports of witness events to express evidentiality (Matsubara 2017). Turkish and Japanese evidentials are obligatory categories. The goal of the present study is to decipher whether speakers across two evidential languages (Turkish and Japanese) and two non-evidential languages (French and English) are using different grammatical “hints” and pragmatic strategies such as inconsistent uses of tense/evidentiality markers when producing deceptive statements.

We investigate written production of grammatical and/or non-grammatical evidential markers across four languages (English, French, Turkish, Japanese) in four conditions. We manipulate deceptive information (truth versus lying) and modality of access to information (direct access through video clips versus indirect access through linguistic reports of others). 160 participants are being recruited, with 40 native speakers of each four languages. A set of 12 mini stories were created, linguistically and culturally adapted. Each story comprises one human referent and four transitive verbs which the referent performs each of these four actions (e.g., *Mary is sitting at the kitchen table with a photograph in front of her. She pours juice into a glass and drinks it. She suddenly tears the photo in two and throws the pieces into the bin next to her.*). The stimuli have been normed for naturalness with 126 participants. In our written production experiment, participants watch half of the stories as silent animated videos, in which the events in the storylines are directly witnessed (see Figure 1 below); and hear the other half of the stories auditorily, in which the events are indirectly accessed. Then, they are instructed to immediately retell the story truthfully or by deliberately changing at least three of the actions. A 2-way crossed design was used, all 12 stories are seen in two conditions of modality across different blocks, fixed in either lie or truth condition.
Data collection is currently underway. A sneak-peek into the so-far collected pilot data indicated (i) a trend towards more frequent use of past tense forms in deception conditions over non-past forms in

all four languages, (ii) overuse of evidential forms in deception conditions in Japanese and Turkish. We further expect that deceitful retellings would present an elevated inconsistent and programmatically unmotivated evidential and tense forms.

Figure 1. Example fragments from an animated video used in the experiment.

Acknowledgements

This work was supported by the Cooperation Agreement between The Scientific and Technological Research Council of Türkiye (TÜBİTAK, Grant no: 122N747) and the French Ministry for Europe and Foreign Affairs and the Ministry for Higher Education and Research (MESR, No: 49074XK - PHC BOSPHORE 2023) granted to Çağla Aydın and Seçkin Arslan.

**References**

Aikhenvald, Alexandra Y. (2004), *Evidentiality*. Oxford: Oxford University Press.

Aksu-Koç, Ayhan, and Slobin, Dan I. (1986) A psychological account of the development and use of evidentials in Turkish. In W. Chafe and J. Nichols (eds), *Evidentiality. The linguistic coding of epistemiology,* New Jersey: Ablex Publishing Corporation, 159-167.

Dendale, Patrick (1993) *Le conditionnel de l'information incertaine: marqueur modal ou marqueur évidentiel?* Proceedings from the Congrès International de Linguistique et Philologie. 164-176.

Hauch, Valerie, Iris Blandón-Gitlin, Jaume Masip, and Siegfried L Sporer, (2015), Are computers effective lie detectors? A meta-analysis of linguistic cues to deception. *Personality and social psychology Review*, 19 (4), 307-342.

Matsubara, Julie (2017), *The semantics and pragmatics of the Japanese evidentials-rashii,-sooda, and-yooda: An experimental investigation.* Doctoral dissertation. Evanston, Illinois: Northwestern University.

Meibauer, Jörg (2018) The linguistics of lying. *Annual Review of Linguistics,* 4, 357-375.

Porter, Stephen, and ten Brinke, Leanne (2010), The truth about lies: What works in detecting high-stakes deception? *Legal and criminological Psychology* 15 (1): 57-75.

Whitt, Richard J. (2010), *Evidentiality and perception verbs in English and German*. Bern: Peter Lang