Emerging Topics in Typology 2025
Leipzig, Germany

Book of abstracts

Contents

Title: The typology of the numeral classifier system of Malieng (Vietic, Vietnam): a study on the syntax and semantics of body parts and fruits

Authors: Albert Badosa Roldós, Université Sorbonne Nouvelle / LACITO; Carla Ferrerós Pagès, Universitat de Girona / Universitat de Perpinyà Via Domitia

Keywords: syntax, semantics, typology, historical linguistics, numeral classifiers, body parts, fruits

Abstract:

Numeral classifiers are a type of nominal classification systems which help identifying noun referents for quantifying purposes (Aikhenvald 2000, Grinevald 2000). Malieng displays a numeral classifier system with a small-sized inventory (4 common classifiers) with similar semantic but different syntactic characteristics as the other languages in the Mainland Southeast Asian linguistic area (Vittrant & Allassonnière-Tang 2021; Bisang 1999).

Malieng language documentation began only recently, in 2022. While documenting its classifiers, we observed co-occurrences not present in other classifier languages—measure terms co-occurring with classifiers—or less common—classifiers co-occurring with class terms—, both of them clear differences with Vietnamese, the best described Vietic language.

Further observations in the Malieng corpus (Badosa 2023) triggered the design of an experiment testing the grammaticality of different combinations of 77 phrases containing classifiers in (mostly) quantified and non-quantified expressions, and with fruits and body parts as the lexical elements. The phrases with fruits and body parts were elaborated by the researchers on the basis of earlier elicitation sessions aimed at those semantic fields. The reasons of the selection of fruits and body parts as targets for classifier elicitation are that they show different semantic behaviour because of the inalienability of the second (Chappell and McGregor 1996; Clark 1996). Before designing the experiment we already knew that in Vietnamese classifiers are optional for body parts but not for fruits.

The experiment was carried out with 8 speakers from two different villages in August 2024. The aims of the experiment were (1) to corroborate the documented odd co-occurrences, (2) to check the optionality of the classifiers, (3) to evaluate the anaphoric use of classifiers, and we also obtained collateral data on the use of classifiers as individuators.

This paper offers, on one hand, the analysis of the Malieng classifier system on the basis of the aforementioned experiments. On the other hand, it discusses the position of the Malieng classifier system synchronically, typologically (Haspelmath 2025; Cinque 2022) and diachronically, within the Mainland Southeast Asian Linguistic Area Vittrant & Allassonnière-Tang 2021; Enfield and Comrie 2015). The discussions are supported by closer comparisons of the Malieng classifier system with Vietnamese (Austroasiatic, Nghiệu 2024; Phan 2019; Nguyễn 1957;) and other languages in the area: Black Hmong (Hmong-Mien, Mouton 2024),

Eastern Bru (Katuic, Miller 1964; Miller 2017), Western Bru (Katuic, Engelkemier 2010), Mày (Austroasiatic, Babaev and Samarina 2021) and Stieng (Austroasiatic, Bon 2014).

Finally, the paper argues that Malieng is in the final wave of classifier development for a Vietic/Austroasiatic, in a context where classifier systems in Austroasiatic languages are a more recent innovation than classifiers in Sinotibetan languages (LaPolla 2003:46). We further argue that the small classifier system displayed by Malieng is borrowed from other languages, which explains the co-ocurrences of borrowed classifiers with native elements typically involved in a classifier system, especially measure terms. These native elements have not yet been grammaticalised into the classifier system, an argument for considering Malieng a conservative Vietic/Austroasiatic language. This process is one of the outcomes of the tendency of these languages. The study of languages such as Malieng is key to the study the development of classifier systems as an areal feature (both semantically and syntactically), the historical linguistics and language contact of the area, and detailing how areal tendencies (towards monosyllabisation, the appearance of tonal systems and the disappearence of inflectional morphology, Michaud 2012) influence peripheral languages.

References:

Aikhenvald, Alexandra. 2000. *Classifiers. A typology of noun categorization devices*. Oxford University Press.

Babaev, Kirill and Samarina, Irina 2021. *A Grammar of May*. Leiden: Brill.

Badosa Roldós, Albert. 2023. *Documentation of the Malieng language, a Vietic language spoken in the Quảng Bình and Hà Tĩnh provinces of Vietnam*. Endangered Languages Archive. Handle: http://hdl.handle.net/2196/o54m9804-2479-4h8w-956s-y5n30245840f

Bisang, W. 1999. Classifiers in East and Southeast Asian languages: Counting and beyond. In: Gvozdanovic, Jadranka (ed.), *Numeral Types and Changes Worldwide*, 113–185.

Bon, Noëllie 2014. *Une grammaire de la langue stieng: langue en danger du Cambodge et du Vietnam*, PhD Thesis, Département de Sciences du Langage, Université Lumière Lyon 2, Lyon, 629 p.

Chappell, Hilary and McGregor, William. *The Grammar of Inalienability: A Typological Perspective on Body Part Terms and the Part-Whole Relation*, Berlin, Boston: De Gruyter Mouton, 1996. https://doi.org/10.1515/9783110822137

Cinque, Guglielmo 2022. Two types of sortal classifiers and their co-occurrence. Università Ca'Foscari: *Quaderni di lavoro* ASIt n. 24 (2022): 75-87.

Clark, Marybeth. "Where do you feel? – Stative verbs and body-part terms in Mainland Southeast Asia". *The Grammar of Inalienability: A Typological Perspective on Body Part*

Terms and the Part-Whole Relation, edited by Hilary Chappell and William McGregor, Berlin, Boston: De Gruyter Mouton, 1996, pp. 529-564. https://doi.org/10.1515/9783110822137.529

Enfield, N.J. and Comrie, Bernard 2015. Mainland Southeast Asian languages, in N.J. Enfield and Bernard Comrie, *Languages of Mainland Southeast Asia: The State of the Art*, Berlin, München, Boston: De Gruyter Mouton, 2015. https://doi.org/10.1515/9781501501685

Engelkemier, Jennifer M. 2010. *Aspects of Bru Khok Sa-at grammar based on narrative texts*. MA Thesis, Payap University.

Grinevald, Colette. 2000. A morphosyntactic typology of classifiers. In Senft, G. (ed.) Systems of Nominal Classification (p. 50-92). Cambridge University Press.

Haspelmath, Martin 2025. "Numeratives in the languages of Asia and beyond". Draft paper. lingbuzz/008754

LaPolla, Randy 2003. Overview of Sino-Tibetan morphosyntax, in Thurgood, G. i LaPolla, R. 2003. The Sino-Tibetan languages. 2007a ed. Language family series. New York: Routledge.

Michaud, Alexis 2012. "Monosyllabicization: patterns of evolution in Asian languages". Nicole Nau, Thomas Stolz & Cornelia Stroh. *Monosyllables: from phonology to typology*, Akademie Verlag, pp.115-130, 2012. ffhalshs-00436432v3f

Miller, J. 1964, "Word classes in Brôu", *The Mon-Khmer Studies Journal*, vol. 1, pp. 41-62.

Miller, Carolyn P. 2017. *Eastern Bru Grammar Sketch*. SIL International.

Mouton, Léa 2024. *Esquisse grammaticale du hmong noir : Etude comparative de l'expression de l'espace dans les langues hmong*. PhD Thesis, Université Lumière-Lyon 2.

Nghiệu, V. Đ. 2024. Ngữ pháp, ngữ nghĩa của những từ được gọi là 'loại từ' trong tiếng Việt. Vietnam Journal of Social Sciences & Humanities, 10(4).

Nguyễn Đình Hòa 1957. Classifiers in Vietnamese. Word, 13(1), 124-152. https://doi.org/10.1080/00437956.1957.11659631

Phan, Trang. 2019. "The absence of classifiers in numeral classifier constructions in Vietnamese". In M.Bağrıaçık, A.Breitbarth, and K.De Clercq, eds., *Mapping Linguistic Data*. Essays in honour of Liliane Haegeman. 208-213. Ghent University.

Vittrant, Alice & Allassonnière-Tang, Marc. 2021. "Classifiers in Southeast Asian Languages". In *The Languages and Linguistics of Mainland Southeast Asia* (p. 733-771). De Gruyter.

Morphological complexity of Catalan: a diachronic and sociolinguistic perspective

Alejandro García-Matarredona - University of Barcelona

One of the most hotly debated topics in linguistic typology currently is what is known as the Language Niche hypothesis (Lupyan & Dale, 2010, 2016) (henceforth LNH), which posits that languages adapt to their sociolinguistic niche. This hypothesis places the languages of the world on an *esotericity* continuum, esoteric languages being those which are spoken by generally small, tight-knit communities, and exoteric ones being those which are spoken by large populations and which have a high proportion of L2 speakers. Under the LNH, exoteric languages tend to simplify their morphology over time to accommodate L2 speakers, whereas esoteric languages tend to complexify it in order to aid in child acquisition (Wray & Grace, 2007).

In the last few years there have been papers published both for (Chen et al., 2024) and against (Shcherbakova et al., 2023) the LNH. However, no studies appear to have been done on what the evolution of a particular language's morphological complexity may look like tested against its population characteristics.

Catalan (ISO 639-3: cat) offers an interesting case study for how a language's morphological complexity might shift over time due to changes in its population, thanks to its multifaceted nature and history: once an international language under the Crown of Aragon, it has since maintained vitality despite variable state support and being considered a minoritized language across the territories where it is spoken (Baylac-Ferrer & Ferrerós-Pagès, in press). Taking the LNH as a guiding principle, I propose a longitudinal study of Catalan using texts from the Computerized Corpus of Old Catalan (Torruella et al., 2010). This corpus contains texts ranging from the 11th to the 18th century, and accounts for genre and dialect of the texts. I will use morphological complexity measures based both on language description and Information Theory (measures such as Shannon Entropy (Shannon, 1948) and Kolmogorov complexity (Kolmogorov, 1963), which are, in principle, theory-agnostic), in order to compare the resulting complexity scores to the evolution of the number of Catalan speakers and contact with Spanish and other neighboring languages. I will derive results with techniques such as Granger Causality (Granger, 1969), with which I will compare the time series of morphological complexity to others related to the sociolinguistic status of Catalan, such as population size or degree of contact with other languages, to be inferred from historical data. In this sense, I want to provide a historical test of the idea that languages change in complexity according to their sociolinguistic niche, in order to shed some light on what the specific mechanisms might be for these processes of simplification and complexification.

Keywords: historical linguistics, typology, morphosyntax, morphological complexity, Information Theory, Language Niche Hypothesis

References

Baylac-Ferrer, A., & Ferrerós-Pagès, C. (in press). Catalan across three states: A comparison of language minoritisation features. In M. Gandarillas (Ed.), *Fighting Stigma and Discrimination in Minoritised Languages*. Cambridge Scholars Publishing.

Chen, S., Gil, D., Gaponov, S., Reifegerste, J., Yuditha, T., Tatarinova, T., Progovac, L., & Benítez-Burraco, A. (2024). Linguistic correlates of societal variation: A quantitative analysis. *PLOS ONE*, *19*(4), e0300838. https://doi.org/10.1371/journal.pone.0300838

Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, *37*(3), 424–438. https://doi.org/10.2307/1912791

Kolmogorov, A. N. (1963). On Tables of Random Numbers. *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002)*, *25*(4), 369–376. https://www.jstor.org/stable/25049284

Lupyan, G., & Dale, R. (2010). Language Structure Is Partly Determined by Social Structure. *PLOS ONE*, *5*(1), e8559. https://doi.org/10.1371/journal.pone.0008559

Lupyan, G., & Dale, R. (2016). Why Are There Different Languages? The Role of Adaptation in Linguistic Diversity. *Trends in Cognitive Sciences*, *20*(9), 649–660. https://doi.org/10.1016/j.tics.2016.07.005

Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, *27*(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Shcherbakova, O., Michaelis, S. M., Haynie, H. J., Passmore, S., Gast, V., Gray, R. D., Greenhill, S. J., Blasi, D. E., & Skirgård, H. (2023). Societies of strangers do not speak less complex languages. *Science Advances*, *9*(33), eadf7704. https://doi.org/10.1126/sciadv.adf7704

Torruella, J., Pérez Saldanya, M., & Martines, J. (Eds.). (2010). *Corpus informatitzat del català antic*.

Wray, A., & Grace, G. W. (2007). The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua*, *117*(3), 543–578. https://doi.org/10.1016/j.lingua.2005.05.005

Content interrogative constructions compared to narrow focus declarative constructions: A cross-linguistic view

This talk investigates the relation between grammatical marking of narrow focus in declarative clauses and of content interrogatives against data from genetically and areally diverse languages. The goal is to provide empirical feedback to the theoretical discussion whether content interrogative constructions should be analysed as a type of focus construction.

The investigation assumes a non-derivational view of syntax and adopts Lambrecht's (1996: 213) definition of focus in declaratives as the non-presupposed semantic component of a proposition expressed by the clause.

Structural parallelisms between content interrogatives and narrow focus constructions observed in languages have led some linguists to claim that content interrogative constructions are in fact a type of focus constructions (cf. e.g. Horvath 1986: 118–122; Dik 1989: 278). On the other hand, other authors have since argued that content interrogative constructions and focus constructions should not be equated, basing their claim on detailed studies showing structural or pragmatic asymmetries between the two constructions in individual languages (cf. e.g. Aboh 2007: 299–307; Cable 2008).

The research questions addressed here are: 1) whether such languages are attested that allow grammatical marking of narrow focus but not of content interrogatives, 2) whether such languages are attested that feature obligatory grammatical marking of narrow declarative focus but no obligatory marking of content interrogatives, 3) how common it is for content interrogative and narrow focus declarative marking to coincide in structure. Under the assumption that content interrogatives are a type of focus construction, the hypothesis would be borne out that an interrogative construction in a language should be reducible to a focus construction, conceivably with additional interrogative marking. Prosodical marking is not considered.

The study builds on data from 55 languages from five continents and 21 families.

Three languages were found that allow marking of narrow focus declaratives but no content interrogative marking: Sebat Bet Gurage (Afro-Asiatic, Semitic), Kharia (Austroasiatic) and Marathi (Indo-European), demonstrated for Marathi in 0. Two studied languages from the Ethiopian Highlands obligatorily mark narrow focus in declaratives, but do not obligatorily mark content interrogatives: Zay (Afroasiatic, Semitic) and Zayse-Zergulla (Ta-Ne-Omotic). In 40 of the 45 languages featuring both relevant construction types, at least one interrogative construction is reducible to a focus construction, as exemplified by Basque syntactic positions in (2). On the other hand, nine languages featuring both construction types exhibit construction pairs not reducible to a common blueprint. The largest group of such construction pairs are comprised by syntactic positions available only to interrogative phrases and some other means of marking

narrow focus declaratives. For example, in Icelandic (Indo-European), interrogative phrases must be placed in a special left-peripheral slot, while narrowly focal constituents in declaratives may only be marked by clefting, as seen in (3). A similar pattern is observed in Central Khmer (Austroasiatic) and Colloquial French (Indo-European).

In conclusion, the data support the claim that content interrogatives should not be conceptualised as a type of focus construction, although the two constructions tend to be related.

Examples

Marathi is a canonically SOV language. The narrowly focal constituent g*haryach kamasathi* 'for the house work' in (1)a) is placed in a special syntactic slot, as evidenced by its position preceding the subject constituent *Ram*. On the other hand, content interrogatives feature no special grammatical marking beside the use of an interrogative word, as seen in (1)b):

(1) Marathi (Indo-European, South Asia) (Nayudu 2008: 29–30; Pandharipande 1997: 12)

a. *Gharyach* *kamasathi* *Ram* *kal* *Mumbaila*
house work.for Ram yesterday Mumbai.ACC/DAT
gela.
go.PST.3SG.M
'Ram yesterday went to Mumbai **for the house work**.'

b. *mohan* *klly* *khato?*
Mohan what eat.PRS.3SG.M
'What does Mohan eat?'

In Basque, interrogative phrases in content interrogatives and narrowly focal constituents in declaratives are placed in the same pre-verbal slot, as in (2)a) and (2)b):

(2) Basque (isolate, Europe) (Hualde & Ortiz de Urbina 2003: 459)

a. *Nori* *azaldu* *zion* *Jonek* *atzo* *bere* *erabakia?*
who.DAT explain AUX J.ERG yesterday his decision
'Whom did Jon explain his decision to yesterday?'

b. *Berari* *azaldu* *zion* *Jonek* *atzo* *bere* *erabakia*.
3SG.DAT explain AUX J.ERG yesterday his decision
'Jon explained his decision **to him** yesterday.'

(3) Icelandic (Indo-European, Europe) (Þráinsson 2007: 76, 360)

a. *Hvað* *sá* *María?*
what.ACC see.PST.3SG Mary.NOM
'What did Mary see?

b. *Það* *var* *lítinn* *hund* *sem*
EXPL be.PST.3SG little.M.SG.ACC dog.ACC.SG REL
María *sá*.
Mary.NOM see.PST.3SG
'It was a little dog that Mary saw.'

References

Aboh, Enoch O. 2007. Focused versus non-focused wh-phrases. In Enoch O. Aboh, Katharina Hartmann & Malte Zimmermann (eds.), *Focus strategies in African languages: the interaction of focus and grammar in Niger-Congo and Afro-Asiatic*, 287–314. Berlin: Mouton de Gruyter.

Cable, Seth. 2008. Wh-fronting (in Hungarian) is not focus-fronting. http://people.umass.edu/scable/papers/Focus-Wh-Hungarian.pdf. (3 February, 2025).

Dik, Simon C. 1989. *The theory of functional grammar. Part 1: The structure of the clause*. Dordrecht: Foris Publications.

Horvath, Julia. 1986. *FOCUS in the theory of grammar and the syntax of Hungarian*. Dordrecht: Foris Publications.

Hualde, José Ignacio & Ortiz de Urbina. 2003. *A Grammar of Basque*. Berlin: Mouton de Gruyter.

Lambrecht, Knud. 1996. *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge: Cambridge University Press.

Nayudu, Aarti. 2008. *Issues in the syntax of Marathi: a minimalist approach*. Durham University PhD Thesis. http://etheses.dur.ac.uk/1308/1/1308.pdf. (24 September, 2024).

Pandharipande, Rajeshwari. 1997. *Marathi*. London: Routledge.

Þráinsson, Höskuldur. 2007. *The syntax of Icelandic*. Cambridge: Cambridge University Press.

Writing does not Impact the Evolutionary Dynamics of Syntax

Writing and reading have been shown to influence how the brain processes language (Dehaene et al. 2010; Cilibrasi, Adani, and Tsimpli 2019). This raises the question of whether these cognitive effects have left an imprint on the evolution of languages as they transitioned to written forms and as literacy spread across various regions of the world. It has been hypothesized that the advent of writing facilitated expressive complexity, specifically by promoting hierarchical structures (subordination, hypotaxis) and reducing reliance on concatenation (coordination, parataxis) (Small 1924; Mitchell 1985; Dąbrowska 2015). Although the development of writing systems was gradual and their adoption among populations occurred over a prolonged time, it is possible that written expression influenced spoken language, particularly through socially prestigious innovators. Such changes could have initiated linguistic evolution toward more complex syntactic constructions (Karlsson 2009).

To test these hypotheses, we conducted two complementary analyses. The first examined the immediate effect of writing on language use, focusing on sentence-level syntactic patterns in written and spoken contexts. Using Universal Dependencies data (De Marneffe et al. 2021), we analyzed about 100,000 sentences from 30 languages across ten families and three genres: Spoken, Fiction, and Wikipedia. We applied Poisson and negative binomial hierarchical regression models to test whether writing influenced (i) the number of clauses per sentence and (ii) the depth of hierarchical clause embeddings. Our models adjusted for phylogenetic and areal relatedness and for dataset-specific idiosyncrasies.

The second analysis investigated the effect of written traditions on the evolution of grammatical constructions. We compiled a dataset of 765 clause-combining constructions from 59 languages in the Indo-European, Sino-Tibetan, and Tupi-Guaraní families, coding them for 18 syntactic features reflecting hierarchy in the form of structural asymmetries between clauses (Foley and Van Valin 1984; Cristofaro 2003; Bickel 2010; van Gijn, Galucio, and Nogueira 2015). To model grammatical evolution, we employed an Ornstein-Uhlenbeck process (Butler and King 2004) with two regimes ('writing' and 'non-writing') for specific branches and time spans, incorporating random effects for feature type and language to account for cross-linguistic variability.

Our findings reveal no significant influence of writing on syntax in either analysis. For language use, we found no evidence that written genres increased the number of clauses or embedding depth in sentences compared to spoken genres. Similarly, phylogenetic and areal effects showed no measurable impact. For grammatical evolution, the probability of syntactic asymmetry did not differ substantially between the 'writing' and 'non-writing' regimes. However, we also found marginal evidence for reduced variance and increased selective pressure under the 'writing' regime (particularly within Indo-European), tentatively suggesting normativizing effects of writing on grammar evolution.

While the scope of our analysis is limited by its coverage of global linguistic diversity, we tentatively conclude that the advent of writing had little impact on the structural evolution of syntax. However, the subtle normativizing effects observed in some lineages, particularly Indo-European, warrant further investigation into the role of writing traditions in shaping grammatical variation.

References

Bickel, Balthasar (2010). "Capturing particulars and universals in clause linkage". In: *Clause Linking and Clause Hierarchy: Syntax and Pragmantics*. Ed. by Isabelle Bril. Studies in Language Companion Series. Amsterdam: John Benjamins B.V., pp. 51–101. ISBN: 9789027287588.

Butler, Marguerite A and Aaron A King (2004). "Phylogenetic comparative analysis: a modeling approach for adaptive evolution". In: *The american naturalist* 164.6, pp. 683–695.

Cilibrasi, Luca, Flavia Adani, and Ianthi Tsimpli (2019). "Reading as a predictor of complex syntax. The case of relative clauses". In: *Frontiers in psychology* 10, p. 1450.

Cristofaro, Sonia (2003). *Subordination*. New York: Oxford University Press.

Dąbrowska, Ewa (2015). "Language in the mind and in the community". In: *Change of paradigms–New paradoxes*, pp. 221–235.

De Marneffe, Marie-Catherine et al. (2021). "Universal dependencies". In: *Computational linguistics* 47.2, pp. 255–308.

Dehaene, Stanislas et al. (2010). "How learning to read changes the cortical networks for vision and language". In: *science* 330.6009, pp. 1359–1364.

Foley, William A and Robert Jr. Van Valin (1984). *Functional Syntax and Universal Grammar*. Cambridge, UK: Cambridge University Press.

Karlsson, Fred (2009). "Origin and maintenance of clausal embedding complexity". In: *Language complexity as an evolving variable*. Ed. by Geoffrey Sampson, David Gil, and Peter Trudgill. Oxford: Oxford University Press, pp. 192–202.

Mitchell, Bruce (1985). *Old English Syntax: Concord, the parts of speech, and the sentence*. Oxford University Press.

Small, George W. (1924). *The comparison of inequality*. Baltimore: University Press.

van Gijn, Rik, Ana Vilacy Galucio, and Antonia Fernanda Nogueira (2015). "Subordination strategies in Tupian languages". In: *Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas* 10.2, pp. 297–324.

Classifying trans-phonologization: a typological study

This study proposes that trans-phonologization processes can be categorized as following one of four different patterns of behavior. Trans-phonologization is defined as an event in which "phonologization of one phonetic cue is often accompanied by the de-phonologization of another" (Yu 2013, p. viii). A frequent example of this phenomenon is tonogenesis where, for example, a voicing contrast between consonants in a syllable onset may be neutralized, but a pitch distinction becomes phonologized on the following vowel, thus maintaining the distinction in meaning between words. Though many studies have examined individual trans-phonologization phenomena in different languages, no study has yet examined the patterning of trans-phonologization across different process types to create a more holistic understanding of trans-phonological patterning. Processes that may include instances of trans-phonologization have typically been studied in their own typologies, including tonogenesis (Hyslop 2022), palatalization (Bateman 2011), nasalization (Michaud, Jacques & Rankin 2012) and labialization (Giampaolo 2020). This study, however, presents the first cross-linguistic comparative study of trans-phonologization processes that is not restricted to the type of phonetic cues transferred. This systematic study therefore addresses the following question: are the directions of change we see largely determined by the phonetic cues at play, or are there other properties of trans-phonologization that could be shared across different phonological process types, regardless of which cues are involved?

This study examines a sample of 20 languages representing 10 language families, shown in Table 1, where each language contains one trans-phonological process for comparison. As not every process is labeled as "trans-phonological" in its source material, an initial inspection was conducted to ensure that the processes in question meet the criteria necessary to be considered trans-phonologization (i.e., neutralization of a primary contrast and phonologization of a new contrast). The sampled processes were then coded for several variables of interest, such as type of phonological process (palatalization, nasalization, etc.), directionality of contrast migration (Right to Left or Left to Right), types of segments involved, segment adjacency, presence of morphological constraints, and the positioning of the segments within the syllable. Four major patterns of behavior were found in the sample, determined by whether the transfer of contrast moves from a consonant to a vowel (transfer from C to V), a vowel to a consonant (transfer from V to C), from a vowel to another vowel (from V to V), or in multiple directions. These four categories are shown in Table 2, and demonstrate the similarities present between trans-phonological phenomena across different conventional cue-based process categorizations. These findings also suggest that there are in fact broader patterns in how trans-phonologization may occur, independent of the specific cues involved. This can in turn supplement current understandings of phonological change by providing additional information and characterization of the phenomenon if it is trans-phonological in nature. This study is also the first attempt at a systematic discussion of trans-phonologization across different phonetic cues and therefore contributes to our understanding of these phenomena and phonological change patterns more broadly.

Works cited

Bateman, N. (2011). On the Typology of Palatalization. *Language and Linguistics Compass, 5*(8), 588–602.

Giampaolo, T. (2020). Labialization in Ægean and Nakh-Daghestanian Languages. *Language and Text, 7*(1), 109–120.

Hyslop. G. (2022). Toward a typology of tonogenesis: Revising the model. *Australian Journal of Linguistics*, *42*(3-4), 275-299.

Michaud, A., Jacques, G., & Rankin, R. L. (2012). Historical transfer of nasality between consonantal onset and vowel. *Diachronica, 29*(2), 201–230.

Yu, A. C. L. (2013). *Origins of Sound Change: Approaches to Phonologization*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199573745.001.0001

Table 1. Language Sample

Language	Reference Source	Family	Phonological Process
Chaha	Banksira 2013	Afro-Asiatic	labialization, palatalization
Panãra	Lapierre 2022	Jean	devoicing
Greek	Baltazani et al. 2016	Indo-European	palatalization
Laomian	Zhang, Jin & Liu 2023	Sino-Tibetan	palatalization
Arabic (Najdi)	Mahzari 2023	Afro-Asiatic	palatalization
Bamileke	Hyman 1976	Niger-Congo	aspiration
Kodi	Lovestrand, Balle & Edwards 2022	Austronesian	palatalization
Korean	Bang et al. 2018	Koreanic	tonogenesis
Chru	Brunelle et al. 2020	Austronesian	registrogenesis
Afrikaans	Coetzee et al. 2018	Indo-European	tonogenesis
Malagasy	Howe 2017	Austronesian	tonogenesis
Khmer (Phnom Penh)	Kirby 2014	Austro-Asiatic	tonogenesis
Kmhmu' (Northern, Eastern)	Kirby, Pittayaporn & Brunelle 2022	Austro-Asiatic	tonogenesis
Vietnamese	Thurgood 2002	Austro-Asiatic	tonogenesis
Tamang	Mazaudon & Michaud 2008	Sino-Tibetan	tonogenesis
Tai (Cao Bằng)	Pittayaporn 2017	Kra-Dai	registrogenesis
Chrau	Thanh, Brunelle & Nguyen 2022	Austro-Asiatic	registrogenesis
Raglai	Brunelle, Brown & Hà 2022	Austronesian	registrogenesis
Terena	Orphão de Carvalho 2021	Arawakan	accentual contrast
Goidelic	Jackson 1967	Indo-European	nasalization

Table 2. Trans-phonologization typological groupings

Type	Segment	Directionality	Adjacency	Morphology	Processes	Example
1	V to C	To preceding segment	Yes	No	Palatalization, Aspiration	**Aspiration** pu → p^{h}o po → po
2	C to V	To following segment	Yes	No	Nasalization, Tonogenesis, Registrogenesis	**Tonogenesis** pa → pá ba → pă
3	V to V	To following segment	Yes	Yes	Accentual contrast	**Accent Contrast** VCVCV → ØCV̂CV CVCV → CVCV
4	Mixed	To preceding segment, sometimes following	Variable	Variable	Palatalization, Labialization, Devoicing	**Palatalization** ajf → ef k'jt' → k'ʲit'

Emerging Topics in Typology - Leipzig, June 2025

A Semantic Map of Negation in Indo-European

Damiana Porcellato - Universität des Saarlandes

As Miestamo (2017) states, the semantic core of negation mirrors its definition in propositional logic, where negation is intended as an operator that changes the truth value of the proposition. This holds true despite the fact that in natural languages negation interacts with other functional domains, leading to semantic and pragmatic effects. This interaction results in the presence of different expressions of negation in human languages, depending on the subdomain being considered. In Table 1, I draw on Miestamo (2017) and van der Auwera & Krasnoukhova (2020: 91-92) to present an overview of the subdomains studied in previous works on negation. Table 1 illustrates the lexicalisation of negation across fifteen domains in six languages, chosen to represent different branches of the Indo-European language family. These data show both unity and variation: in all languages, there is one negator that is widely spread across contexts (*not, non, ne, nuk, na, na-*) and at least one less common negator (*no, no, net, mos, nei, nist*). The lexicalisation patterns of negators across subdomains are language-specific.

Table 1: Subdomains of negation

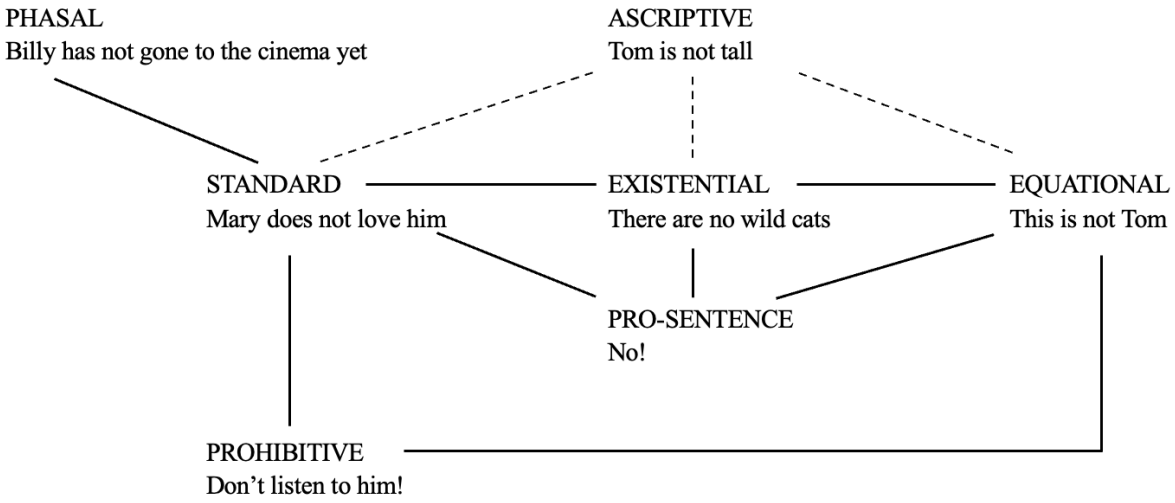
	ENGLISH	ITALIAN	RUSSIAN	ALBANIAN	BENGALI	SHUGHNI	…
STANDARD	Mary **does not** love him.	non	ne	nuk	na	na-	…
EMPHATIC	Mary **does not** love him **at all**.	non … mica	ne	nuk	na	ačaθ … na-	…
PHASAL	Mary **does not** live here **yet.**	non … ancora	ne	nuk	na	…	…
SUBORDINATE	I urge you **not** to talk to him.	non	ne	mos	na	na-	…
PROHIBITIVE	**Don't** listen to him!	non	ne	mos	na	mā-	…
INTERROGATIVE	**Doesn't** she love him?	non	ne	nuk	na	na-	…
EXISTENTIAL	There are **no** wild cats.	non	net	nuk	nei	nist	…
LOCATIVE	The cats are **not** in the garden.	non	net	nuk	nei	nist	…
POSSESSIVE	They have **no** money.	non	net	nuk	nei	nist	…
ASCRIPTIVE	Mary is **not** a teacher.	non	ne	nuk	nɔ-	nist	…
INDEFINITES	**Nobody** believes him.	nessuno	nikto ne	askush nuk	keu na	yičay(aθ)	…
PRO-SENTENCE	**No**!	no	net	jo	na	nay	…
PRIVATIVE	He was **without** money.	senza	bez	pa	chaṛa	be-	…
EXPLETIVE	Don't be surprised if it **doesn't** rain.	non	ne	nuk	na	na-	…
DERIVATIONAL	Mary **disagrees** with me.	essere in disaccordo	ne soglasen	nuk pajtohem	…	…	…
…	…	…	…	…	…	…	…

Building on previous works on diverse subdomains of negation with a focus on Indo-European (Bernini & Ramat 1996; Willis et al. 2013; Verkerk & Shirtz 2022), I aim to investigate the distribution of negative markers in the domains displayed in Table 1 across approximately 50 Indo-European languages. Data will be collected considering a mixture of grammars, questionnaire-based approach, and parallel corpora.

Interrelationships between the subdomains will be represented on a semantic map. Starting from Bond's (2009) pilot study, which identifies 7 functions of negation and represents them as shown in Figure 1, one goal is to investigate if Bond's map should be extended given cross-linguistic

data on the functions listed in Table 1, and to determine if Bond's proposal holds for the languages I intend to examine.

Figure 1: Conceptual space for negation, based on Bond (2009: 56)



Focusing on one language family allows for the integration of diachrony into the analysis. Considering that the overall aim of this study is to provide an overview of the subdomains of negation within the Indo-European language family, especially relevant are the areal patterns and the well-known cycles that lead to the renewal of negative markers. Some features found in the negation domain characterize in fact Standard Average European, such as the spread of Jespersen Cycle (Bernini & Ramat 1996), or the so-called 'compositional prohibitive' (Van Olmen & Van der Auwera 2016). From a diachronic perspective, Jespersen Cycle(s), the Negative Existential Cycle, and the Indefinite Cycle, which have been argued to be intertwined (van der Auwera et al. 2022), play a role in the distribution of negators.

In this talk, after presenting the semantic map, I will draw on the functional motivations provided in the literature (Miestamo 2005; van der Auwera 2005; Veselinova & Hamari 2022) to discuss a tentative analysis of why one negator is used widely across subdomains. Given Miestamo's (2017) notion that the semantic core of negation mirrors its definition in propositional logic, the existence of a highly multifunctional negator is to be expected. To examine the presence of different negators, I will point towards how diachrony can play a role in the analysis.

References

Bernini, Giuliano & Paolo Ramat. 1996. *Negative Sentences in the Languages of Europe: A Typological Approach*. Berlin: Mouton de Gruyter.

Bond, Oliver. 2009. Mapping negation in conceptual space. In Peter K. Austin, Oliver Bond, Monik Charette, David Nathan & Peter Sells (eds.), *Proceedings of Conference on Language Documentation and Linguistic Theory* 2. London: SOAS.

Georgakopoulos, Thanasis & Stéphane Polis. 2018. The semantic map model: State of the art and future avenues for linguistic research. *Language and Linguistics Compass* 12, e12270.

Miestamo, Matti. 2005. *Standard Negation. The negation of Declarative Verbal Main Clauses in a Typological Perspective*. Berlin: Mouton de Gruyter.

Miestamo, Matti. 2017. Negation. In Alexandra Y. Aikhenvald and R. M. W. Dixon (eds.), *Cambridge Handbook of Linguistic Typology*, 405 – 39. Cambridge: Cambridge University Press.

Shirtz, Shahar, Talamo, Luigi, & Verkerk, Annemarie. 2021. The Evolutionary Dynamics of Negative Existentials in Indo-European. *Frontiers in Communication*, 6, 661862.

van der Auwera, Johan, Olga Krasnoukhova & Frens Vossen. 2022. Intertwining the negative cycles. In Ljuba Veselinova & Arja Hamari (eds.), *The Negative Existential Cycle*, 611–650. Berlin: Language Science Press. DOI: 10.5281/zenodo.7353631.

van der Auwera, Johan. 2005. Prohibitives: why two thirds of the world's languages are unlike Dutch. *Proceedings of the Fifteenth Amsterdam Colloquium, December 19-21, 2005*. Paul Dekker and Michael Franke (eds.). Grafisch Centrum Amsterdam ISBN: 90-5776 1467.

van der Auwera, Johan & Olga Krasnoukhova. 2020. The typology of negation. In Veronica Déprez & María Teresa Espinal (eds.), *Oxford handbook of negation*, 91–116. Oxford: Oxford University Press.

Van Olmen, Daniël & van der Auwera, Johan. 2016. Modality and Mood in Standard Average European. in Nuyts, Jan & van der Auwera, Johan (eds), *The Oxford Handbook of Modality and Mood*, 362-383. Oxford: Oxford University Press.

Verkerk, Annemarie & Shahar Shirtz. 2022. Negative existentials in Indo-European: A typological and diachronic overview. In Ljuba Veselinova & Arja Hamari (eds.), *The Negative Existential Cycle*, 233–324. Berlin: Language Science. Press. DOI: 10.5281/zenodo.7353611.

Veselinova, Ljuba. 2013. Negative existentials: A cross-linguistic study. *Italian Journal of Linguistics* 25(1). 107–145.

Veselinova, Ljuba & Arja Hamari. 2022. Introducing the Negative Existential Cycle. In Ljuba Veselinova & Arja Hamari (eds.), *The Negative Existential Cycle*, 1–58. Berlin: Language Science. Press.

Willis, David, Christopher Lucas, and Anne Breitbarth (eds.). 2013. *The History of Negation in the Languages of Europe and the Mediterranean*, Vol. 1: *Case Studies*. Oxford: Oxford University Press.

The pervasive reality status distinction in the grammar of Yamalero (Guahiban, Colombia)

The crosslinguistic validity of a grammatical category called reality status (henceforth RS), for coding a reality status distinction independent from modality is still an open question (Bybee et al 1994; Mithun 1995; Bybee 1998; Elliot 2000; Palmer 2001; de Haan 2012; Cristofaro 2012). Michael (2014) has argued that Nanti (Arawak) has a RS system that may be considered a canonical system within the framework of canonical typology, and Rose (2014) and Danielsen & Terhart (2015) have contributed with similar data for South Arawak languages. In this paper, I will present the RS system of Yamalero, and I will show that it is typologically similar to that of Sikuani (the most extensively described Guahiban language) and of the mentioned Arawakan languages.

Yamalero is a Guahiban language spoken by some 300 people in the mid-Orinoco region of Eastern Colombia, with high rates of intergenerational transmission. The data presented here have been collected during 9 months of primary fieldwork between 2022 and 2024. They are part of a multi-genre corpus of 11 hours of naturalistic speech.

The RS system of Yamalero is expressed by nine pairs of obligatory verbal inflectional suffixes (slot +1 in the verbal morphological template), featuring a binary realis/irrealis distinction. The RS marking depends on several semantic parameters (first value indicates realis marking, second value indicates irrealis marking): temporality (non-future vs. future), polarity (positive vs. negative), prospectiveness (Ø vs. prospective complement), hypotheticality (actual vs. conditional and counterfactual), etc. Example (1) illustrates the polarity parameter.

(1) Semantic RS marking (polarity)
  (a) *wawai* *juna-**wa**-iba*
  white.people be.afraid.of-**REAL**-ITER
  'The white people were afraid of him [my grandpa Braulio]'
  (b) ***apa**-pa-juna-**wi**-n*
  NEG-PL-be.afraid.of-**IRR**-1SBJ
  'We were not afraid of them [of the guerrilla]'

In addition to these parameters, in some other contexts RS marking is also determined lexically, by verbal classes. This is the case of speaker-oriented modality (Ø vs. imperative) and aspect (Ø vs. semelfactive). The second values take irrealis marking with verbal class 9, while they take realis marking with verbal classes 1-8 (example (2) illustrates this in the case of aspect). In addition to the RS marking, the irrealis values are additionally marked with dedicated affixes indicating future, negation, etc. (see Table 1).

(2) Semantic and lexical RS marking (semelfactive aspect)
  (a) *bajapokoneje* *pe-ena* *kou* *x-**e-xaba***
  after.that 3POSS-mum EVD? eat-**IRR-SMLF**
  'Then they ate the mum' (yam0029: 88)
  (b) *na-boso* *ja-**na-xaba***

REFL?-tail break-**REAL-SMLF**
'He broke his tail' (yam0047: 69)

Table 1. Parameters determining RS marking.

Parameter	**Realis**	**Irrealis**	**Additional marking**
Tense	Non-future	Future	*-ena* 'FUT'
Polarity	Positive	Negative	*apa-* 'NEG'
Prospectiveness	Ø	Prospective complement	*-jetsa* 'PURP'
Hypotheticality	Actual	Conditional, counterfactual	*-tsipe* 'POT'
Speaker-oriented modality	Ø and imperative (verbal classes 1-8)	Imperative (verbal class 9)	*-de/-ma* 'IMP'
Aspect	Ø and semelfactive (verbal classes 1-8)	Semelfactive (verbal class 9)	*-xaba* 'SMLF'

Therefore, in this presentation I will show that the RS system of Yamalero supports Michael's (2014) claims to analyze this category separately from modality, not only because the RS marking does not involve the speaker attitude (future tense, negative polarity), but also because it is also determined by lexical grounds (verbal classes). In addition, these similarities between Guahiban and Arawakan languages are consistent with those presented by Meléndez (2014) on lexical borrowings from Arawakan languages into Guahiban, and might point to a situation of extensive language contact between these two language families in the Eastern Colombian Plains.

References

Bybee, Joan. 1998. "Irrealis" as a grammatical category. Anthropological Linguistics 40. 257–271.

Bybee, Joan, Revere Perkins & William Pagliuca. 1994. The evolution of grammar: Tense, aspect, and modality in the languages of the world. Chicago: University of Chicago Press.

Cristofaro, Sonia. 2012. Descriptive notions vs. grammatical categories: Unrealized states of affairs and 'irrealis'. Language Sciences 34. 131–146

Danielsen, Swintha y Lena Terhart. 2015. Realis/irrealis as a basic grammatical distinction in Southern Arawakan languages. En Revue de Semantique et Pragmatique, 38: 97-120.

de Haan, Ferdinand. 2012. Irrealis: Fact or fiction? Language Sciences 32. 107–130.

Elliott, Jennifer. 2000. Realis and irrealis: Forms of grammaticalization of reality. Linguistic Typology 5. 55–90.

Meléndez Lozano, Miguel Ángel. 2014. Préstamos Arawak (Achagua, Piapoco y Piapoco-Achagua) a la lengua sikuani. LIAMES: Línguas Indígenas Americanas 14(1), 173-218.

Michael, Lev. 2014. The Nanti reality status system: Implications for the typological validity of the realis/irrealis contrast. En Linguistic Typology, 18(2): 251–288.
Mithun, Marianne. 1995. The relativity of irreality. In Bybee & Fleischman (eds.) 1995, 367–387.
Mithun, Marianne. 1999. The languages of Native North America. Cambridge: Cambridge University Press.
Palmer, F. R. 2001. Mood and modality. 2nd edn. Cambridge: Cambridge University Press.
Queixalós, Francesc. 2024. A Grammar of Sikuani. (LINCOM Studies in Native American Linguistics, 92.) München: Lincom.
Rose, Françoise. 2014. Irrealis and negation in Mojeño Trinitario. En Lev Michael y Tania Granadillo (eds.), Negation in Arawak Languages (pp. 216-240). Leiden: Brill.

Abbreviations

COLL = collective; EVD = evidential; FUT = future; INCL = inclusive; IMP = imperative; IRR = irrealis; ITER = iterative; NEG = negation; OBJ = object; PL= plural; POSS = possessive; POT = potential; PURP = purpose REAL = realis; SBJ = subject; SMLF = semelfactive.

DEFINITENESS IN AVATIME

Dinah Adom Mac-Arthur (Georg-August Universität, Göttingen)

This research presents definiteness in Avatime and argues that the dichotomy of definite markers is unstable in the language. Avatime is a Ghana-Togo-Mountain (GTM) language that belongs to the Kwa group of the larger Niger-Togo phylum. It is a class system language with 7 noun classes marked by affixes. Definiteness is marked by suffixes whose morphological distinctions are determined by the noun class and agreement functions. The markers have a single syllable structure CV or V and participate in ATR harmony with the vowel in the root as shown in (1b) and (1a).

(1) a. ovè
o-ve-è
C_1SG-mouse-C_1SG.DEF
'the mouse'

b. ɔnʊ̀vɔ̀ɛ́
ɔ-nʊ̀vɔ̀-ɛ́
C_1SG-child-C_1SG.DEF
'the child'

According to Schwarz (2009, 2013), there are two types of definite markers cross-linguistically; strong article and weak article. The strong article encodes familiarity and licensed when the definite NP refers to an entity previously mentioned in an ongoing discourse. The weak article is used when the definite NP is unique in the discourse situation.
Following the four major contexts as proposed by Hawkins (1978)[1], the literature on Kwa languages argues that the bifurcation seems to hold. For example, Akan (Owusu 2022) shows a contention between the definite article *no-* used in familiar contexts and the bare noun in unique contexts. In Ga (Renans 2016), *lɛ* is a strong article whiles *nɛɛ* or the bare noun is licensed in unique environments. In Sɛlɛɛ, (Agbetsoamedo 2014, Agbetsoamedo & Duah 2022),a GTM language, the distal demonstrative *nwu* also has definite interpretations and occurs in both familiar and unique contexts. The bare noun is used in relational anaphora.
For the most part, the research on definiteness has focused on languages with one or two articles however, Avatime presents novel data on a language with several definite markers. By using the questionnaire from Duah, Grubic & Renans (2021) based on the four major contexts proposed by Hawkins (1978), the findings show that not only do definite markers in the Avatime occur in both familiar and unique environments, but it is the case that the same marker can occur in both contexts.
In (2a) and (3a), *-le* occurs with the NPs *liwu* 'dress' and *liwo* ' 'sun', respectively. It is evident that, the nouns belong to the same class hence the same marker. The context does not affect the form of the article. In (2b) and (3b), the bare nouns are infelicitous. This is because they strictly have indefinite interpretations.

(2) [Familiar Context: Yesterday, Adzo went to the boutique to buy some clothes. She tells her mother what happened. 'When I went to the boutique...]

[1]The strong article is felicitous in relational anaphora whiles the weak article usually occurs in larger situations(globally unique contexts), immediate situations and part-whole bridging

a. *Ma-mɔ̀* *li-wú.* *Li-wú-lè* *li-pedi.*
1SG.SBJ-see C_3SG-dress. C_3SG-dress-C_3SG.DEF AGR-be.beautiful
'I saw a dress. The dress was beautiful.'

b. *Ma-mɔ̀* *li-wú.* *#Li-wú* *li-pedi.*
1SG.SBJ-see C_3SG-dress. C_3SG-dress AGR-be.beautiful
'I saw a dress. The dress was beautiful.'

(3) [Globally unique Context:Yayra slept and woke up late in the afternoon. The next day she is telling Edem. When I woke up yesterday...]

a. *Li-wò-le* *lìí-kle.*
C_3SG-sun-C_3SG.DEF AGR.IPFV-shine
'The sun was shinning'

b. *#Li-wò* *lìí-kle.*
C_3SG-sun AGR.IPFV-shine
'The sun was shinning'

Avatime shows that the DP in an anaphoric environment is the same DP when the noun appears in a unique environment. Although I do not dismiss Schwarz (2009, 2013) proposal that there is two types of definite markers that differentiate between familiarity and uniqueness cross-linguistically, this research provides evidence that the bipartite notion is unstable in Avatime.

References

Agbetsoamedo, Yvonne. 2014. *Aspects of the grammar of Selee*. University of Stockholm. (Doctoral dissertation).

Agbetsoamedo, Yvonne & Reginald A. Duah. 2022. *Definiteness in Sɛlɛɛ*. A paper presented at the 2nd workshop for Definiteness Across Domain project. University of Postdam.

Duah, Reginald, Mira Grubic & Agata Renans. 2021. The combination of indefinite and definite 'determiners'—a cross-linguistic study. In *Proceedings of sinn und bedeutung*, vol. 25, 238–255.

Hawkins, John A. 1978. *Definiteness and indefiniteness, london, croom helm*. London: Routledge.

Owusu, Augustina P. 2022. *Cross-categorialdefiniteness/familiarity*. Rutgers The State University of New Jersey, School of Graduate Studies. (Doctoral dissertation).

Renans, Agata. 2016. A cross-categorial definite determiner: evidence from ga (kwa). In *Semantics and linguistic theory*, vol. 26, 23–42.

Schwarz, Florian. 2009. *Two types of definites in natural language*. University of Massachusetts Amherst.

Schwarz, Florian. 2013. Two kinds of definites cross-linguistically. *Language and Linguistics Compass* 7(10). 534–559.

Toward a comprehensive typology of nominal TAM

Elia Calligari - University of Pavia

Nominal TAM is the marking of TAM-related values on a nominal head, with effect on its temporal interpretation. It is an understudied topic in typology: the work by Nordlinger & Sadler (2004) was the first attempt to cross-linguistic analysis and inspired reports on this phenomenon in further languages (e.g., Adamou 2011, Aikhenvald 2022, a.o.), but no consensus has been reached on its definition. Tonhauser's (2007, 2008) formalist analyses reject nominal TAM as a category since it implies conflation of nominal and verbal domains notwithstanding the semantic discrepancies between the two classes. Later contributions by Nikolaeva (2015) and Bertinetto (2020), instead, argue for a finer-grained analysis, stressing the similarity between nominal and verbal TAM markers, both encoding all facets of temporal information (i.e. tense, aspect, mood).

Two types of nominal TAM are usually distinguished (Nordlinger & Sadler 2004): *independent nominal TAM* has its scope limited to the noun phrase while clausal temporal information is provided otherwise, as in (1) and (2) below; *propositional nominal TAM* scopes over the whole sentence, cooperating with or substituting for verbal TAM markers in rendering the temporal interpretation of the clause, as in (3), where case contributes to tense expression via agreement.

This ongoing research aims at (i) building a convenience-variety sample through a replicable method, (ii) proposing a comprehensive cross-linguistic description of nominal TAM and of its functional domain, and (iii) exploring the grammaticalization of its markers. Following Miestamo et al.'s (2016) review of variety sample strategies, the Diversity Value technique (Rijkhoff & Bakker 1998, Bakker 2011) was chosen, adopting Ethnologue's (Eberhard et al. 2024) genealogical classification as a basis for calculation. Starting with a pre-determined sample size of 500 languages, we calculate the Diversity Value of all families and extract a set of languages from each in proportion to its diversity value; special rules apply to language isolates, pidgins and creoles. Subsequently, the sample is corrected for areal stratification based on Glottolog's macro-areas (Hammarström & Donohue 2014).

The plan is to rely on a converging evidence workflow (Masini & Mattiola 2019): after searching the grammars of sample languages to track down instances of nominal TAM along with their marking strategies, the second step focuses on a subset of ca. 20 languages with nominal TAM, enabling functional analysis of primary data, i.e., corpora and glossed texts. The final step eventually involves in-depth corpus-based studies on larger corpora of 2 to 4 languages to collect synchronic and diachronic evidence allowing to identify possible paths of grammaticalisation of nominal TAM markers within a source-oriented approach (Cristofaro 2019). These phases will lead to an updated definition of nominal TAM, based on distributional evidence from primary data combined with recent theoretical advancements. Concerning nominal past, a promising comparison may be offered by the markers of discontinuous verbal past investigated in Plungian & van der Auwera (2006), enacting a past reference with the implication that the situation described for the past does not hold for the present, thus resembling the implications of nominal past (compare (1) and (4)). As to nominal future, a deeper analysis by Nikolaeva (2015) show that it encodes values bordering between future tense and irrealis mood, confirming Comrie's (1985) observations on the non-prototypicality of future as a tense. Furthermore, a possible nominal TAM marker in Northern Samoyedic languages has been targeted by a source-oriented pilot study aimed at assessing its grammaticalisation path (Calligari 2024).

Examples [1]

(1) Paraguayan Guaraní (Tupi-Guaraní, Paraguay; Nordlinger & Sadler 2004:781)
O-va-ta *che-róga-**kue**-pe*
3-move-FUT 1SG-house-**PST**-in
'He will move into my **former** house'

(2) Paraguayan Guaraní (Tupi-Guaraní, Paraguay; Nordlinger & Sadler 2004:781)
A-va-va'ekue *hóga-**rã**-pe*
1SG-move-PST 3.house-**FUT**-in
'I moved into his **future** house'

(3) Lardil (Tangkic, Australia; Klokeid 1976:493 in Nordlinger & Sadler 2004:791)

Ngada	*bilaa*	*wu-thur*	*ngimbenthar*	*diin-**kur***	*wangalk-**ur***
1SG.NOM	tomorrow	give-FUT	2SG.**FOBJ**	this-**FOBJ**	boomerang-**FOBJ**

'I'll give you this boomerang tomorrow'

(4) Seychelles Creole (French-based, Seychelles; Michaelis 1993:82 in Plungian & van der Auwera 2006:325)

Mon	***ti***	*vine*
1SG	**DP**	come

'I came/had come' (lit. 'I came, and then went back')

References

Adamou, Evangelia. 2011. Temporal uses of definite articles and demonstratives in Pomak (Slavic, Greece). *Lingua* 121(5). 871–889.

Aikhenvald, Alexandra Y. 2022. Beyond nominal tense: Temporality, aspect, and relevance in Tariana noun phrases. *Studies in Language* 46(1). 40–75.

Bakker, Dik. 2011. Language sampling. In Song, Jae Jung (ed.), *The Oxford handbook of linguistic typology*, 100–127. Oxford: Oxford University Press.

Bertinetto, Pier Marco. 2020. On nominal tense. *Linguistic Typology* 24(2). 311–352.

Calligari, Elia. 2024. *A Corpus-based contribution to the debate on the Northern Samoyedic predestinative suffix*. MA Thesis. Pavia: University of Pavia. https://hdl.handle.net/20.500.14239/26525

Comrie, Bernard. 1985. *Tense*. Cambridge: Cambridge University Press.

Cristofaro, Sonia. 2019. From synchronically-oriented typology to source-oriented typology: Typological universals in diachronic perspective. *CLUB Working Papers in Linguistics* 3. 52–67.

Eberhard, David M. & Simons, Gary F. & Fenning, Charles D. (eds.). 2024. *Ethnologue: Languages of the World*. Twenty-seventh edition. Dallas, TX: SIL International. Online version: http://www.ethnologue.com (last accessed 2025-01-23).

Hammarström, Harald & Donohue, Mark. 2014. Some principles on the use of macro-areas in typological comparison. *Language Dynamics and Change* 4. 167–187.

Klokeid, Terry. 1976. *Topics in Lardil grammar*. Cambridge, MA: MIT (PhD dissertation).

Masini, Francesca & Mattiola, Simone. 2019. Come fare tipologia con categorie non tradizionali? [How to do typology with non-traditional categories?]. *CLUB Working Papers in Linguistics* 3. 282–294.

Michaelis, Susanne. 1993. *Temps et aspect en créole seychellois: valeurs et interferences* [Tense and aspect in Seychelles Creole: values and interferences]. Hamburg: Buske.

[1] Abbreviations: 1 = first person, 2 = second person, 3 = third person, DP = discontinuous past, NOM = nominative, FOBJ = future object, FUT = future, PST = past, SG = singular.

Miestamo, Matti & Bakker, Dik & Arppe, Antti. 2016. Sampling for variety. *Linguistic Typology* 20(2). 233–296.

Nikolaeva, Irina. 2015. On the expression of TAM on nouns: Evidence from Tundra Nenets. *Lingua* 166. 99–126.

Nordlinger, Rachel & Sadler, Louisa. 2004. Nominal Tense in Crosslinguistic Perspective. *Language* 80(4). 776–806.

Plungian, Vladimir A. & van der Auwera, Johan. 2006. Towards a typology of discontinuous past marking. *Language Typology and Universals* 59(4). 317–349.

Rijkhoff, Jan & Bakker, Dik. 1998. Language sampling. *Linguistic Typology* 2. 263–314.

Tonhauser, Judith. 2007. Nominal tense? The meaning of Guaraní nominal temporal markers. *Language* 83(4). 831–869.

Tonhauser, Judith. 2008. Defining Crosslinguistic Categories: The Case of Nominal Tense (Reply to Nordlinger and Sadler). *Language* 84(2). 332–342.

Reviewing the Andoke pre-auxiliary position: neither subject nor focus?

Andoke is a language isolate spoken by some 30 individuals in the Colombian Amazon. Andoke morphosyntax is marked by the presence of an element here labelled 'auxiliary', which occurs in almost all declarative and interrogative sentence types. The Andoke auxiliary usually agrees with the preceding element in noun class – if said element is a noun (1a). Otherwise, the auxiliary resorts to default agreement (1b). The auxiliary is moreover the host of several morphological markers, relating to clause-typing, tense and several epistemic categories.

Previous literature on Andoke (Landaburu 1979; Landaburu 2000; Landaburu 2023 etc.), and in particular on the Andoke auxiliary (Landaburu 1976), provides several explanations on its morphosyntactic function, primarily concerned with the distinction of 'focus' versus 'subject'. Using original data from my own fieldwork, including semi-naturalistic discourse and elicitation, as well as annotating archival data, I show that neither of these labels provide a fitting description of the 'pre-auxiliary position'. Instead, I argue that the pre-auxiliary position cannot be adequately labelled using concepts such as 'focus', 'subject' or 'topic', and must instead be considered a language-particular syntactic position, similar to some Germanic V2 structures.

In particular, when analyzing discourse structure, the pre-auxiliary position is found to more often than not be filled by connectives or adverbs for paragraphs at a time. Focus seems to be instead marked through prosodic prominence. While there is some isomorphism of content interrogative questions and their respective answers (2), in which both the interrogative pronoun and the questioned element appear in pre-auxiliary position, this cannot be attributed to focus. Stimulus-based elicitation shows that the isomorphism between both constructions rather pertains to a desire for syntactic parallelism.

While the results of subjecthood tests are not finalized at this moment, preliminary insights show that 'subject' is not a viable candidate for the pre-auxiliary position either. Setting aside that the majority of pre-auxiliary positions in discourse are filled by non-nominal constituents, it seems that most subjecthood tests that have been already evaluated do not uniformly match the pre-auxiliary constituent. For example, Keenan (1976: 321) lists how subjects "normally express the agent of the action, if there is one", which does not accurately describe Andoke: instead, it seems that when two nominals are overtly expressed in a clause, the noun with the more animate referent is chosen to be in pre-auxiliary position.

Lastly, I show how the information structural concept of 'topic' is not a viable candidate for this position either. Further, I discuss the post-verbal, right-dislocated position within the Andoke clause, which is usually filled by new or contrastive topical constituents to introduce a topic shift – against expectations (Givón 1983: 19; also cf. [3]).

In sum, I argue that the question of 'sujet ou focus?' (en. 'subject or focus?') posed by Landaburu (1976) must be answered by 'neither', and that in the case of Andoke, language-specific descriptive labelling should be preferred over a comparative labelling (cf. Haspelmath 2007).

Examples

(1) a. *dúuɲɛ́kɒ* ***baya*** *nĩ=mã́*^{H} *nõ-hʌʌ́-ʌ*

D.[M] AUX:3M T=with 1SG:COME&DO-speak-FINIT

'I will speak with Dúuɲɛ́kɒ' (stim001:4)

b. *ya-aka* ***bʌ=ɲɛ́ʔ***^{H} *ɲõ=ẽ=ʌ* *yi-yẽ-ĩ-ʌ-i*

3M-alone AUX:DEF=R.PAST home=in=FINIT UP-3M-GO&DO-remain-FINIT

'Alone he stayed at home.' (narr001:6)

(2) a. *híʌ* ***dʌ-tɤ́***^{H} *ha=ɤ́*^{H} *y-ɤ̃-ĩ-i*

what AUX:3INAN:Q-EMPH 2SG=to UP-3COLL-give-FINIT

'What did they give to you?' (narr001:121)

b. *ʌ́ʌ* *dɯ-ʌ-i* ***bʌ*** *o-dó-i* *ʌ=ɲɛ́ʔ*^{H}*=hɤ*

paca[3INAN] grill-ASP?-FINIT AUX:3INAN 1SG-take-FINIT 3INAN=PROX=PRSTV

'I took grilled paca, here!' (narr001:122)

(3) CONTEXT: A lazy man tries to lay eggs with some frogs he was supposed to hunt for his family at home but instead defecates with them. The frogs wrinkle their noses, smelling the odor.

phm phm phm *ka* *bʌ=ɲɛ́ʔH* *páa (..)* *páa (.)*

ID:sniffing CONN AUX:DEF=R.PAST already already

páa *ʌ-pokɤ̃=ka* *bʌ* *ka*

already 3INAN-dawn=CONN AUX:DEF CONN

yẽ-ɤ̃-ẽ *ɤ̃-ẽ-po-i* % ***ɤ̃-pɤko=a***

3M-wife-COLL 3COLL-COME&DO-arrive-FINIT 3COLL-house=in

'[They went:] *phm, phm, phm*; and then already the sun was dawning, and his wife [pl.tant.] arrived, <u>at their house</u>' (narr006)

THEN: At their house, the man's wife notices his absence and his empty hammock [etc.]

References

Givón, T. 1983. *Topic Continuity in Discourse: A quantitative cross-language study*. John Benjamins. https://doi.org/10.1075/tsl.3.

Haspelmath, Martin. 2007. Pre-established categories don't exist: Consequences for language description and typology. 11(1). 119–132. https://doi.org/10.1515/LINGTY.2007.011.

Keenan, Edward L. 1976. Towards a Universal Definition of 'Subject'. In Charles N. Li (ed.), *Subject and Topic*, 303–333. New York: Academic Press.

Landaburu, Jon. 1976. L'énoncé en langue andoke: sujet ou focus? *Amerindia* 1. 19–32.

Landaburu, Jon. 1979. *La langue des Andoke (Amazonie colombienne): grammaire* (Langues et civilisations à tradition orale 36). Paris: Société d'études linguistiques et anthropologiques de France.

Landaburu, Jon. 2000. La Lengua Andoque. In María Stella González de Pérez, María Luisa Rodríguez de Montes & Instituto Caro y Cuervo (eds.), *Lenguas indígenas de Colombia: una visión descriptiva*, 275–288. 1. ed. Santafé de Bogotá: Instituto Caro y Cuervo.

Landaburu, Jon. 2023. Chapter 3: Andoke. In Patience Epps & Lev Michael (eds.), *Language Isolates I: Aikanã to Kandozi-Shapra* (Handbücher Zur Sprach- Und Kommunikationswissenschaft 44), vol. 1, 125–172. De Gruyter. https://doi.org/10.1515/9783110419405-003.

Countability and distributivity in interaction with classifiers in Vaímajã

Languages with classifiers challenge binary count–mass distinctions (Chierchia 2010). A problem presented by well-known classifier languages such as Chinese and Vietnamese is that, while classifiers clearly increase a noun's countability, lexical criteria for comparing count and mass nouns are scarce, obscuring the relationship between countability and classifiers (Cheng & Sybesma 1998; Gil 1996).

In contrast, in Vaímajã (Colombia, East Tukanoan) inanimate nouns belong to three lexically determined number categories: i) *basic singular nouns*, a small closed class linked to spatio-temporal and anatomical semantics, e.g. *jopé* 'door'; ii) *general number nouns*, constituting the majority of noun lexemes, e.g. *jutí* 'textile'; iii) and *mass nouns*, which denote liquids, granular substances, and immaterial concepts, e.g. *okó* 'water'. This allows us to observe the individuating function of classifiers explicitly. I define these categories on the basis of five tests (Table I), exemplified on the next page.

Table I: Properties of lexical number categories

Can the noun…	*jopé* 'door' **BASIC SINGULAR**	*jutí* 'textile' **GENERAL**	*okó* 'water' **MASS**
1) take a pluraliser directly	✓	–	–
2) be modified by a numeral directly	✓	–	–
3) be modified by *paíri* 'big'	✓	✓	–
4) be modified by the distributive quantifier *páɨ́*	✓	✓	–
5) denote singular number with a bare demonstrative	✓	(✓)	–

Basic singular nouns pass all five tests, while mass nouns fail all of them. General number nouns pattern with mass nouns for pluralisation (1) and numeral modification (2), but with basic singular nouns in their ability to be modified by the adjective *paíri* 'big' (3) and the distributive quantifier *páɨ́* 'many' (4). Regarding modification by a bare demonstrative, general number nouns display behaviour unlike both basic singular and mass nouns: unlike mass nouns, general number nouns may be modified by a bare demonstrative, and unlike basic singular nouns, such constructions cannot denote singular reference (5).

Classifiers convert mass and general number nouns into countable, bounded entities by contributing a bounded shape feature to the lexeme, e.g. *okó* 'water' → *oko-rɨ́* {water-CLF:POT} 'pot of water', and from general number nouns, e.g. *jutí* 'textile' → *juti+ro* {textile+CLF:FLAT} 'piece of clothing'. Conversely, mass nouns can be obtained from basic singular and general number nouns by zero derivation. Basic singular nouns are the primary source for "repeaters" (6), classifiers which have (near) identical lexical counterparts (see Seifart 2005: 77–81).

Inanimate nominalisations by *+ri* 'NMLZ' form basic singular nouns by suffixation of a classifier, but they also allow the formation of mass nouns by means of *-ke*/*-je* 'CLF:MASS' (7). Notably, there appears to be no operation that derives general number nouns at all.

These findings highlight the individuating role of classifiers, as it is their boundedness and distributivity that ultimately determine a noun's number category. The emergence of repeaters from basic singular roots suggests a grammaticalization path in which classifiers evolve to contribute boundedness and countability to general number and mass nouns, which lack these features.

Examples (own data)

Tests:

	Basic singular	**General number**	**Mass**
(1)	PLURALISATION		
	a. *jopé+ri* door-PL 'doors'	b. *jutí*(+*ro*)*+*ri* textile+CLF:FLAT+PL 'clothes'	c. *okó*(-*rɨ́*)*+*ri* water-CLF:POT+PL 'pots of water'
(2)	NUMERAL MODIFICATION		
	a. *pɨá día+ri* two river-PL 'two rivers'	b. *pɨá jutí*(+*ro*+*ri*)* two textile+CLF:FLAT+PL 'two pieces of clothing'	c. *pɨá oko*(-*rɨ+ri*)* two water-CLF:POT+PL 'two pots of water'
(3)	DISTRIBUTIVE MODIFIER		
	a. *paʉ vãme* many name 'many names'	b. *paʉ́ pʉ̃ʉ̃* many leaf 'many leaves'	c. *paʉ́ okó*(-*rɨ+ri*)* many water-CLF:POT+PL 'many pots of water'
(4)	STUBs		
	a. *pairi ɨjéro* big mouth 'big mouth'	b. *paíri jutí* big textile 'big shirt/big shirts'	c. **paíri okó* big water (Intended: 'big water')
(5)	UNMARKED DEMONSTRATIVES		
	a. *atí dɨpó* DEM.PROX foot 'this foot'	b. *atí yukɨ́* DEM.PROX tree 'these trees'	c. **atí jĩnírike* DEM.PROX chicha (Intended: 'this chicha')

Repeaters and nominalizations

(6) REPEATER FROM BASIC SINGULAR NOUN

a. *jiká-vejé vejé*
one.INAN-CLF:CHAGRA chagra
'one chagra'

b. *yá-vi ví*
1SG.POSS-CLF:BUILDING house
'my house'

(7) NOMINALISATIONS

a. *wá+ri-ka+ro*
row+NMLZ-INAN+CLF:TREE
'paddle'

b. *wá+ri-ke*
row+NMLZ-MASS
'the act of rowing'

References

Cheng, Lisa Lai-Shen & Sybesma, Rint. 1998. Yi-wan tang, yi-ge Tang: Classifiers and massifiers. *Tsing Hua Journal of Chinese Studies* 3(XXVIII).

Chierchia, Gennaro. 2010. Mass nouns, vagueness and semantic variation. *Synthese* 174(1). 99–149. (doi:10.1007/s11229-009-9686-6)

Gil, David. 1996. Maltese "collective nouns": A typological perspective. *Rivista di Linguistica* 8(1). 53–87.

Seifart, Frank. 2005. *The Structure and Use of Shape-Based Noun Classes in Miraña (North-West Amazon)*. Nijmegen: Radboud Universiteit. (PhD Dissertation.)

Non-canonical reflexive and reciprocal constructions in Basque: new insights into reflexive and reciprocal voices

Kristina Bilbao Hernandez (University of the Basque Country, UPV/EHU)
kristina.bilbao@ehu.eus

Typological research on reflexive and reciprocal constructions has primarily focused on their morphological encoding, distinguishing between nominal and verbal reflexives/reciprocals (Faltz 1985; König & Kokutani 2006; Nedjalkov 2007; Janic et al. 2023). Nominal reflexives/reciprocals typically involve reflexive/reciprocal anaphoric nominals, while verbal reflexives/reciprocals are associated with reflexive/reciprocal voices, marked by verbal voice markers (Kulikov 2011; Zúñiga & Kittilä 2019). However, some inherently reflexive/reciprocal verbs express these meanings without specific morphological encoding (Zúñiga & Kittilä 2019). This pattern is cross-linguistically attested in grooming verbs for reflexivity (Haspelmath 2023) and social interaction verbs for reciprocity (Haspelmath 2007), contrasting with nominal and verbal constructions that rely on morphosyntactic strategies.

This study investigates non-canonical reflexive (1a) and reciprocal (1b) constructions in Basque. Although derived from transitive verbs like *ikusi* 'see', these constructions involve a single argument marked with the absolutive case, triggering absolutive agreement with the intransitive auxiliary *izan* 'be':

(1) a. Ane ispiluan ikusi da.
Ane(ABS) mirror.in see.PFV be.(3ABS)
'Ane has seen herself in the mirror.'

b. Ane eta Miren kalean ikusi dira.
Ane(ABS) and Miren(ABS) street.in see.PFV be.3PL.ABS
'Ane and Miren have seen each other in the street.'

Constructions in (1) lack anaphors or voice markers, setting them apart from nominal and verbal reflexive/reciprocal constructions. Instead, these constructions align with the pattern observed in inherently reflexive/reciprocal verbs such as *dutxatu* 'shower' (2a) and social interaction verbs like *batzartu* 'meet' (2b):

(2) a. Ane dutxatu da.
Ane(ABS) shower.PFV be.(3ABS)
'Ane has taken a shower.'

b. Ane eta Miren batzartu dira.
Ane(ABS) and Miren(ABS) meet.PFV be.3PL. ABS
'Ane and Miren have met.'

Although both (1) and (2) lack morphological encoding of reflexivity/reciprocity, I argue they employ distinct strategies to express those meanings: the former relies on a morphosyntactic strategy, while the latter conveys them lexically. This study demonstrates that, despite the absence of overt reflexive/reciprocal marking, verbs in constructions like (1) are not inherently reflexive/reciprocal, unlike verbs in (2). Furthermore, I propose that constructions like (1) are reflexive/reciprocal voices encoded through the intransitive auxiliary, rather than through a dedicated voice marker.

Three key observations support this claim. First, canonical reflexive/reciprocal constructions with *ikusi*-like verbs (1) are transitive and require anaphors, whereas inherently reflexive/reciprocal verbs (2) cannot occur with anaphors in transitive constructions (Etxepare 2003). Second, *ikusi*-like verbs (1) are not interpreted as reflexive or reciprocal in non-finite clauses without anaphors, whereas grooming verbs and social interaction verbs are. Third, ditransitive verbs can also occur in non-canonical constructions with the intransitive auxiliary. These facts indicate that verbs forming non-canonical reflexives and reciprocals (1) do not express reflexivity/reciprocity lexically. Instead, they require explicit encoding to achieve a reflexive/reciprocal interpretation, which occurs through the intransitive auxiliary in non-canonical constructions.

In conclusion, I argue that these constructions represent reflexive/reciprocal voices in Basque, challenging the conventional view that such voices necessarily involve a verbal morpheme. This study offers new insights into the typology of reflexive and reciprocal constructions, suggesting that reflexive and reciprocal voices may emerge syntactically through auxiliary selection, rather than through morphological voice markers.

References

Etxepare, R. (2003). Reflexives and reciprocals. In J. I. Hualde & J. Ortiz de Urbina (ed.), *A Grammar of Basque* (380–385). Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110895285.363

Faltz, L. M. (1985). *Reflexivization: A study in universal syntax*. New York: Garland.

Haspelmath, M. (2007). Further remarks on reciprocal constructions. In V. Nedjalkov (ed.)., *Reciprocal Constructions*. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.71.74has

Haspelmath, M. (2023). Comparing reflexive constructions in the world's languages. In K. Janic, N. Puddu & M. Haspelmath (ed.)., *Reflexive constructions in the world's languages*. [*Research on Comparative Grammar* 3]. Berlin: Language Science Press.

Janic, K., Puddu, N. & Haspelmath, M. (ed.). (2023). *Reflexive constructions in the world's languages*. [*Research on Comparative Grammar* 3]. Berlin: Language Science Press. DOI: 10.5281/zenodo.7861660

König, E. & Kokutani, S. (2006). Towards a typology of reciprocal constructions: focus on German and Japanese. *Linguistics* 44(2), 271–302. https://doi.org/10.1515/LING.2006.010

Kulikov, L. (2011). Voice typology. In J. J. Song (ed.), *The Oxford Handbook of Linguistic Typology* (368–398). Oxford: Oxford University Press.

Nedjalkov, V. (ed.). (2007). *Reciprocal Constructions*. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.71

Zúñiga, F. & Kittilä, S. (2019). *Grammatical voice*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781316671399

Middle voice in Kannada: Markers in competition

Lennart Chevallier

Keywords: Middle voice, Dravidian, Kannada, Typology

The middle as a third voice category has been associated with various phenomena. A number of authors (e.g. Barber 1975: 18, Klaiman 1991: 92, Kemmer 1993: 4) consider the middle voice to involve *subject-affectedness*. More recent approaches, however, such as Inglese (2021), do not seek a single core meaning, but consider the middle to encode a variety of valency-related functions. For Inglese, in a middle voice system, these middle markers are obligatory with some verbs, i.e., they do not contrast with an active voice and are hence "non-oppositional", while they are non-obligatory with other verbs. In this latter group, they are "oppositional", as they show a functional contrast with the active voice. I follow this definition in my analysis.

Across the Dravidian languages of South India, we see evidence for such a system to varying degrees. According to the etymological dictionary by Burrow & Emeneau (1984, entry 2151), the verb *koḷ-* (also: *koḷḷu*) 'to seize, receive, buy, acquire, etc.' is an auxiliary with reflexive meaning in Tamil, Malayalam, Kannada, Telugu, and other languages. However, upon closer examination, the picture is multi-faceted in that *koḷ-* brings about readings which are not restricted to the reflexive, e.g. in Kannada, *koḷḷu* can serve as a reflexive (1), reciprocal (2), self-benefactive (3), anticausative (4) and passive (5) marker (see examples, next page) and, e.g., obligatorily occurs in *is(a)koḷḷu* 'to take' (Bucher 1899: 55).

In my talk, I will illustrate in which ways the middle marker in Kannada competes with other constructions, showing work in progress on my dissertation. With "competition" I mean that a marker other than the middle marker is also available to encode the same function. This is in line with Inglese (2022), who compares different interaction scenarios of middle markers with other valency-reducing constructions. For example, in Kannada, *koḷḷu* is used as reciprocal marker with the verb *haṃcu* 'to share' resulting in *haṃc-i-koḷḷu* [share-LNK-MM] 'to share with each other'. Reciprocity can, however, also be expressed without *koḷḷu* by the expression *obba-r-ig' obba-ru* [one.HUM-HON-DAT one.HUM-HON] 'to each other' (Sridhar 1990: 124-125). The question remains whether one strategy is unmarked or whether both express slightly different semantics. It is also interesting to note that the primary function of the middle voice in Dravidian appears to be the self-benefactive, unlike e.g. the data discussed in Inglese (2021: 506), but in line with its origin in Dravidian from a self-benefactive construction.

All modern Kannada examples are drawn from two sources: An annotated corpus, which is the result of continuous joint work, and questionnaire-based fieldwork. Data from Middle Kannada are also included to trace this development in Kannada since the 12th century. The annotation is being done in Toolbox and FLEx (SIL) at the morphosyntactic level and a translation of every sentence is provided.

At present, the approach is chiefly descriptive but closely linked to recent typological studies. Future work will also include other aspects of middle marking, e.g., non-valency related functions, and will aim to contribute to a better understanding of the middle voice cross-linguistically.

References

Barber, Elizabeth Jane Wayland. 1975. Voice – beyond the passive. In C. Cogen et al. (eds.), *Berkeley Linguistics Society 1*. 16-24.

Bucher, Johannes. 1899. *A Kannaḍa-English School-Dictionary. Chiefly Based on the Labours of Rev. Dr. F. Kittel*. Mangalore: Basel Mission Book & Tract Depository.

Burrow, Thomas & Murray Barnson Emeneau. 1984. *A Dravidian Etymological Dictionary*. 2nd ed. Oxford: Clarendon Press.

Inglese, Guglielmo. 2021. Towards a typology of middle voice systems. *Linguistic Typology* 26(3), 489-531. https://doi.org/10.1515/lingty-2020-013.

Inglese, Guglielmo. 2022. How do middle voice markers and valency reducing constructions interact? Typological tendencies and diachronic considerations. *Folia Linguistica* 56(2), 239-271. https://doi.org/10.1515/flin-2022-2019.

Kemmer, Suzanne. 1993. *The middle voice*. Amsterdam & Philadelphia: John Benjamins (Typological Studies in Language, 23).

Klaiman, Miriam Holly. 1991. *Grammatical voice*. Cambridge: Cambridge University Press (Cambridge Studies in Linguistics, 59).

Sridhar, Shikaripur Narayanarao. 1990. *Kannada*. London & New York: Routledge (Descriptive Grammars).

Examples

Reflexive

1) *Raːŋi kannaɖi-y-alli avaɭ-ann-e noːɖ-i-koṃɖ-aɭu*

   Rani mirror-EUPH-LOC 3SG.F-ACC-FOC see-LNK-MM.PST-3SG.F.NPRS

   ‘Rani saw herself in the mirror.’

Reciprocal

2) *naːvu caːkoːleːʈ-annu haṃc-i-koṃɖ-evu*

   1PL chocolate-ACC share-LNK-MM.PST-1PL.NPRS

   ‘We shared the chocolate among ourselves.’

Self-benefactive

3) *naːnu id-annu maːɖ-i-koṃɖ-e*

   1SG 3SG.PROX-ACC make-LNK-MM.PST-1SG.NPRS

   ‘I did this for myself.’

Anticausative

4) *baːgilu taːn-aːgi-y-eː mucc-i-koṃɖ-itu*

   door self-ADVLZ-EUPH-FOC close-LNK-MM.PST-3SG.N.PST

   ‘The door closed itself.’

Passive

5) *hosa mane-gaɭu nirmaːŋa-goṃɖ-avu*

   new house-PL construction-MM.PST-3PL.N.NPRS

   ‘The houses were built.’

Rebecca Arkell

Conceptualising and Categorising Body Parts in Cape Verdean Creole: A Semantic Typology

Body part semantics provides valuable insights into how languages categorise human anatomy through cognitive and cultural lenses. While earlier work suggested that hierarchical partonomies are a linguistic universal (Brown, 1976), more recent studies challenge this view, highlighting the need to better understand both universals and variation across languages (Huisman et al., 2021). Languages differ significantly in how they segment, conceptualise, and name body parts (Enfield et al., 2006).

This study investigates how body parts are named and categorised in Badiu, the Santiago variety of Cape Verdean Creole (CVC)—a Portuguese-related creole with substrate influence from Mandinka, Wolof and Temne. It further explores body-related idioms and linguistic expressions, providing insights into the cognitive and cultural models embedded in the language.

A key hypothesis of this study is that variation will emerge in how speakers categorise body parts. Some may distinguish between *arm/hand* and *leg/foot*, aligning more closely with Portuguese, while others may use more general or ambiguous terms like *mó* and *pé* to refer to both, reflecting substrate influence. This expected variation aligns with the view of CVC as part of a continuum, with a *basilect* (furthest from Portuguese) and an *acrolect* (closest to Portuguese) at opposing ends (De Camp, 1971).

Preliminary observations suggest that CVC lacks a general term for 'limb', and that *folgu* ('breath') is not clearly distinct from the act of breathing (Quint, 2000). The presence or absence of specific body part terms is often linked to their cultural significance, as the body plays functional and symbolic roles in people's customs and behaviours (Kraska-Szlenk, 2020). Research on cultural sensorium shows that African and Euro-American models differ in how they structure perception and emotion (Geurts, 2002a, b; Stoller, 1989). While Euro-American models tend to treat states of perception, affect, and disposition as independent categories, many African cultures merge these domains into one single conceptual category (Ameka & Amha, 2022). With its European and African heritage, CVC offers an ideal setting to examine how distinct models of embodiment interact.

An important focus of this study is the semantic extension of the body to express emotions, a widely attested cross-linguistic phenomenon (Enfield & Wierzbicka, 2002). In CVC, the expression d*ja bu da-m ku stángu* ('I am fond of you') literally translates as 'I have given you my stomach', suggesting a conceptual link between the stomach and emotions such as love and affection.

The study employs a field-based approach with approximately 25 native Badiu speakers across age, gender, and socio-economic backgrounds. Elicitation techniques include Enfield's (2006) *Elicitation Guide on Parts of the Body* and van Staden & Majid's (2006) *Body Colouring Task*, as well as targeted prompts to elicit body-based idioms and metaphors relating to emotion and space.

This research contributes to the typology of body part semantics, offering novel insights into both universal and culture-specific cognitive-linguistic models in CVC. It lays the groundwork for future studies on embodiment, semantic extension, and grammaticalisation (Lehmann, 2016). It also opens pathways for comparative analyses with CVC's superstrate (Portuguese) and substrate languages, advancing our understanding of creole evolution and contact-induced change.

References

Ameka, F. K., & Amha, A. (2022). Research on language and culture in Africa. In *Approaches to Language and Culture* (pp. 339–383). De Gruyter. https://doi.org/10.1515/9783110726626-013

Brown, C. H. (1976). General Principles of Human Anatomical Partonomy and Speculations on the Growth of Partonomic Nomenclature. In *Source: American Ethnologist* (Vol. 3, Issue 3).

De Camp, D. (1971). Toward a generative analysis of a post-creole speech continuum. Pidginization and creolization of languages, 347–70.

Enfield, N. J. (2006). Elicitation guide on parts of the body. *Language Sciences*, *28*(2–3), 148–157. https://doi.org/10.1016/j.langsci.2005.11.003

Enfield, N. J., Majid, A., & van Staden, M. (2006). Cross-linguistic categorisation of the body: Introduction. *Language Sciences*, *28*(2–3), 137–147. https://doi.org/10.1016/j.langsci.2005.11.001

Enfield, N. J., & Wierzbicka, A. (2002). Introduction: The body in description of emotion. *Pragmatics & Cognition*, *10*, 1-25.

Geurts, K. L. (2002a). *Culture and the senses. Bodily ways of knowing in an African community*. University of California Press.

Geurts, K. L. (2002b). On rocks, walks, and talks in West Africa. Cultural categories and an anthropology of the senses. *Ethos*, *30*(3), 178–198.

Huisman, John L. A., Roeland van Hout & Asifa Majid. (2021). Patterns of Semantic Variation Differ Across Body Parts: Evidence from the Japonic Languages. Cognitive Linguistics 32(3). 455–486. https://doi.org/10.1515/cog-2020-0079.

Kraska-Szlenk, I. (2020). *Body part terms in conceptualization and language usage*.

Lehmann, C. (2016). *Foundations of body-part grammar*.

Quint, N. (2000). *Grammaire de la langue cap-verdienne: étude descriptive et compréhensive du créole afro-portugais des Iles du Cap-Vert.* L'Harmattan.

Stoller, P. (1989). *Fusion of the worlds. An ethnography of possession among the Songhay of Niger.* University of Chicago Press.

van Staden, M., & Majid, A. (2006). Body colouring task. *Language Sciences*, *28*(2–3), 158–161. https://doi.org/10.1016/j.langsci.2005.11.004

Modelling the Arawakan multilocus classifier systems

This study presents an explorative quantitative analysis of the Arawakan classifier systems, demonstrating how distance matrix modelling can help in understanding complex lexico-grammatical systems such as classifiers.

The Arawakan language family comprises 40 extant languages across South America, clustering in the Amazon. The languages have complex classifier systems, known for being able to occur variously on nouns, verbs, numerals, and nominal modifiers, as shown in (1) by the presence of the same classifier on a numeral, noun, and adjective in Mojeño Trinitario. The classifiers can also be used for different functions, including derivation (2) and qualification (3). For this, they have been referred to as multifunctional (Krasnoukhova, 2012) and multilocus (Dunn & Rose, to appear) classifier systems. Despite the great diversity in the systems, including in the number of classifiers, the classifier forms and the available hosts, and whether such a system exists at all, Dunn (2022) argued that the Arawakan classifier systems had their origin in Proto-Arawakan.

The Arawakan languages are spread across Amazonia, itself highly linguistically diverse in terms of the number of language families (Epps, 2009). There are several linguistic contact areas, which involve the Arawakan languages, such as the Guaporé-Mamoré in the southwestern region and the Caquetá-Putumayo and Vaupés regions in the northwestern Amazon (Epps & Michael, 2017). Classifiers are proposed as a contact feature both within these areas, and also in a proposed macro-area comprising the entire region, on the basis of their distributional and functional properties (Epps & Michael, 2017; Krasnoukhova, 2012). The Arawakan classifiers systems have also been shaped and impacted by this history of regional contact, such as the development of the demonstrative locus in Tariana (Arawakan) from contact with Tukano (Tukanoan) (Aikhenvald, 2012, p. 297).

Distance matrix modelling is a means to quantify the similarity and dissimilarity between two objects. Such modelling of linguistic constructions has previously been used by Arias et al (2022) and Van Gijn et al (2022) to demonstrate how linguistic features and systems such as person marking and TAME have become more similar in unrelated languages due to contact. This paper aims to explore how this methodology can be used to model the distances between the Arawakan classifier systems on semantic, structural, and functional levels. Drawing on data from descriptive grammars of over 20 Arawakan languages, the parameters of classifier locus, construction, function, and meaning are modelled as non-metric distance matrices, allowing for a detailed exploration of the data and the relationships between languages. By drawing the language sample from the southwestern and northwestern Amazonian regions, the modelling is able to both demonstrate the shared familial relationship between the systems and explore the areal relationships and distinctions at the same time.

Building a model of the Arawakan classifier systems with this methodology will create a baseline for further studies, incorporating data on non-Arawakan languages. Forming a comprehensive, 3D picture allows for further research into the Arawakan classifiers with regard to contact and inheritance, two themes that are key to understanding language development and diversity in South America.

Examples

1. Loci in Mojeño Trinitario (Bolivia, Rose, 2024, pp. 48, 75, 95)
    a. *éto-**gi***
    one-**CLF:cyl**
    'one (speaking of, for example, a tree)'
    b. *to* *risa-**gi***
    ART.NH alder-**CLF:cyl**
    'aliso (tree sp.)'
    c. *p-ñi* *'chane* *ñ-etere-g-'o* *p-jo* *'chope-**gie*** *wkugi*
    DEM-M person 3M-jump-CLF:cyl-ACT DEM-NH.SG big-**CLF:cyl** tree
    'The man jumps over the big tree.'
2. Derivation in Apurinã (Brazil, Facundes, 2000, p. 172)
    *xamu-**panhi***
    fire-**CLF:powder**
    'ash'
3. Qualification in Tariana (Brazil, Aikhenvald, 2003, p. 118)
    a. *hinipuku* *han-**ipa***
    garden big-**CLF:big.open.space**
    'big garden (viewed as a large space)'
    b. *hinipuku* *hanu-**puna***
    garden big-**CLF:road**
    'big garden (viewed as a road)'

Abbreviations

ACT	active
ART	article
CLF	classifier
cyl	cylindrical
DEM	demonstrative
M	masculine
NH	nonhuman
SG	singular

References

Aikhenvald, A. Y. (2003). *A Grammar of Tariana, from Northwest Amazonia*. Cambridge University Press.

Aikhenvald, A. Y. (2012). *The Languages of the Amazon*. Oxford University Press.

Arias, L., Emlen, N. Q., Norder, S., Julmi, N., Lemus Serrano, M., Chacon, T., Wiegertjes, J., Howard, A., Azevedo, M. C. B. C., Caine, A., Dunn, S., Stoneking, M., & Van Gijn, R. (2022). Interpreting mismatches between linguistic and genetic patterns among speakers of Tanimuka (Eastern Tukanoan) and Yukuna (Arawakan). *Interface Focus*, *13*.

Dunn, S. (2022). *The Shape of Arawakan Classifiers: An internal source for the Arawakan classifier system(s)* [MA Thesis]. Leiden University.

Dunn, S., & Rose, F. (to appear). Arawak classifiers: A South American multilocus classifier system. In N. O. Schiller & T. Kupisch (Eds.), *The Oxford Handbook of Gender and Classifiers*. Oxford University Press.

Epps, P. (2009). Language classification, language contact, and Amazonian prehistory. *Language and Linguistics Compass*, *3*, 581–606.

Epps, P., & Michael, L. (2017). The Areal Linguistics of Amazonia. In R. Hickey (Ed.), *The Cambridge Handbook of Areal Linguistics* (pp. 934–963). Cambridge University Press.

Facundes, S. da S. (2000). *The Language of the Apurinã people of Brazil: (Maipure/Arawak)* [PhD Thesis]. University of New york at Buffalo.

Krasnoukhova, O. (2012). *The Noun Phrase in the Languages of South America* [PhD Thesis]. Radboud Universiteit Nijmegen.

Rose, F. (2024). *Mojeño Trinitario classifiers: A multilocus and multifunctional system*. Cadernos de Etnolingüística.

Seifart, F. (2007). The Prehistory of Nominal Classification in Witotoan Languages. *International Journal of American Linguistics*, *73*(4), 411–445.

Van Gijn, R., Norder, S., Arias, L., Emlen, N. Q., Azevedo, M. C. B. C., Caine, A., Dunn, S., Howard, A., Julmi, N., Krasnoukhova, O., Stoneking, M., & Wiegertjes, J. (2022). The social lives of isolates (and small language families): The case of the Northwest Amazon. *Interface Focus*, *13*.

Coding asymmetries between plain and axial locations

In this presentation, I introduce new types of coding asymmetry and explain them in terms of frequency of use.

It is well-known in linguistic typology that human languages exhibit a remarkable feature: coding asymmetry universals (e.g. Greenberg 1966; Haspelmath 2021). For example, plural forms tend to be longer than singular forms, accusative forms tend to be longer than nominative forms, and negative forms tend to be longer than affirmative forms, as illustrated in Table 1.

Although many types of coding asymmetries have been reported in the literature (e.g. Haspelmath (2021) lists 25 examples of coding asymmetry pairs), coding asymmetries between plain and axial locations have received little attention. Thus, this presentation addresses such asymmetries.

The present study distinguishes between two kinds of locations: plain locations and axial locations. *Plain locations* refer to general locations, such as 'in the room', whereas *axial locations* are a cover term for intrinsic and relative frames of reference. Axial locations are defined as regions that are spatially separated in an axial manner relative to a ground, such as 'in front of', 'behind', 'above', 'under', and 'beside'. Axial locations have been explored from various perspectives, including semantic typology (Levinson & Wilkins 2006), generative grammar (Svenonius 2006; Cinque 2010), FDG (Mackenzie 2013), and grammaticalization theory (Lehmann 2015; Svorou 1994). However, they have not been systematically investigated from a morphosyntactic typological perspective. The present study aims to fill that gap.

This study conducts exploratory research based on a convenience sample of 40 languages and reports the following two generalizations regarding coding asymmetries:

(1) Strategies used for axial locations are more complex (or longer) than those for general locations.

(2) Strategies used for source relationships in axial locations tend to be more complex than those used for static location and goal relationships in axial locations.

Regarding (1), I have identified several types of strategies used for axial locations, all of which are more complex (or longer) than those used for plain locations, as illustrated in (3) and (4). This finding can be situated within broader patterns of coding asymmetries (Greenberg 1966; Haspelmath 2021). As for (2), in plain locations, source relationships are encoded more complexly than goal relationships (Stolz et al. 2014; Haspelmath 2019). The present study confirms that this generalization also holds for axial locations, as illustrated in (5) and (6).

To explain these two generalizations, I compare several competing frameworks, such as markedness, iconicity, and frequency, and suggest that they are best explained in terms of frequency, as suggested by Haspelmath (2006; 2008; 2021). While markedness and iconicity fail to account for cases where plain and axial locations are coexpressed using the same marker, the frequency-based explanation can account for these instances.

By identifying new types of coding asymmetry, this study contributes to a broader understanding of coding asymmetry universals. Furthermore, by explaining these asymmetries, it advances the discussion on whether markedness, iconicity, or frequency offers the best explanation for these universals.

Table 1: Examples of coding asymmetries (adopted from Haspelmath 2021)

singular	plural	*house – house-s*
nominative	accusative	Russian *ja - menja*
affirmative	negative	*live – don't live*

(4) English
   a. ***in*** *Bordeaux*
   b. ***in front of*** *the building*

(5) Ulwa (Keram; Barlow 2023)
   a. *ma=**in***
      3SG.OBJ=in
      'in it (=the garden)'
   b. *ma=**wan***
      3SG.OBJ=above
      'over it'

(6) English
   a. *My cat goes **under** the table.*
   b. *My cat comes **from under** the table.*

(7) Japanese
   tukue=no *sita=**e**/sita=**kara***
   table=GEN under=ALL/under=ABL
   '(to)/from under the table'

References

Barlow, Russel. 2023. *A grammar of Ulwa*. Berlin: Language Science Press.

Cinque, Guglielmo. 2010. Mapping Spatial PPs: An Introduction. In Guglielmo Cinque & Luigi Rizzi (eds.), *Mapping Spatial PPs: The Cartography of Syntactic Structures, Volume 6*, 3–25. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195393675.003.0001.

Greenberg, Joseph H. 1966. *Language universals, with special reference to feature hierarchies*. The Hague: Mouton. https://doi.org/10.1515/9783110899771.

Haspelmath, Martin. 2006. Against markedness (and what to replace it with). *Journal of linguistics*. Cambridge University Press 42(1). 25–70.

Haspelmath, Martin. 2008. Frequency vs. iconicity in explaining grammatical asymmetries. *Cognitive Linguistics* 19(1). 1–33. https://doi.org/10.1515/COG.2008.001.

Haspelmath, Martin. 2019. Differential place marking and differential object marking. *STUF - Language Typology and Universals* 72(3). 313–334. https://doi.org/doi:10.1515/stuf-2019-0013.

Haspelmath, Martin. 2021. Explaining grammatical coding asymmetries: Form–frequency correspondences and predictability. *Journal of Linguistics* 57(3). 605–633. https://doi.org/10.1017/S0022226720000535.

Lehmann, Christian. 2015. *Thoughts on grammaticalization*. 3rd edn. Berlin: Language Science Press. https://doi.org/10.17169/langsci.b88.99.

Levinson, Stephen C. & David P. Wilkins (eds.). 2006. *Grammars of Space: Explorations in Cognitive Diversity* (Language Culture and Cognition). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511486753.

Mackenzie, J. Lachlan. 2013. Spatial adpositions between lexicon and grammar. In J. Lachlan Mackenzie & Hella Olbertz (eds.), *Casebook in Functional Discourse Grammar* (Studies in Language Companion Series), 67–94. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/slcs.137.04mac.

Stolz, Thomas, Sander Lestrade & Christel Stolz. 2014. *The Crosslinguistics of Zero-Marking of Spatial Relations*. München: De Gruyter (A). https://doi.org/doi:10.1524/9783050065304.

Svenonius, Peter. 2006. The Emergence of Axial Parts. *Nordlyd* 33(1). 49–77. https://doi.org/10.7557/12.85.

Svorou, Soteria. 1994. *The Grammar of Space*. Amsterdam: John Benjamins. https://www.jbe-platform.com/content/books/9789027276575.

Typological Diversity and Areal Convergence in Grammatical Local Case Marking:

A Cross-Genetic Study of the Languages in West Bengal

West Bengal, India, is a linguistically diverse area and home to the languages of four language families: Tibeto-Burman, Dravidian, Austroasiatic, and Indo-Aryan. This paper aims to explore the cross-genetic analysis of grammatical local or spatial cases in four languages of West Bengal. The languages for the present study are Bangla (Indo-Aryan), Santali (Austroasiatic), Dhimal (Tibeto-Burman), and Kurux (Dravidian). Local or spatial cases encode spatial relationships such as location ('at'), destination ('to'), origin ('from'), and path ('through') (Blake 2001, p. 151)'. Despite distinct genetic affiliations, the above-mentioned languages have coexisted in the same geographical area for a long time, providing a unique opportunity to study the interaction between typological variation and contact-induced convergence. The study uncovers both shared patterns and divergent strategies, focusing on the influence of language contact in shaping spatial expressions across language families.

Areal convergence is evident in the functions of local cases in these languages. For example, in these languages, locative markers are used to mark non-canonical objects in some verbs such as 'trust', 'doubt', 'suspect', 'believe' etc., or the ablative case in all four languages extends to psychological predicates, marking experiencers of fear. However, this paper does not confine the analysis solely to areal features in local case marking; it also focuses on examining language-specific contact-induced changes. Case syncretism occurs when different grammatical cases are expressed using the same form. All the languages under study exhibit spatial case syncretism. Some of them are the result of contact. For example, the Indian variety of Dhimal uses the same marker for locative and allative cases; even the Nepalese variety of Dhimal preserves a separate allative marker (/*tʰekapa*/) (King 2008). The Indian variety appears to have lost this distinction, potentially influenced by neighboring Indo-Aryan languages like Bangla and Nepali (Bangla /-*e*/ or /-*te*/; Nepali /-*lai*/).

(1)	*oɟa*	*iskul-**ta***	*ihi*	(2)	*oɟa aro ka*	*ʃiliguri-**ta***	*hani*	*kʰaigʰa*
	he	school-LOC	be		he and i	siliguri-ALL	together	went
	'He is in school.'				'He and I went to Siliguri.'			

This study also examines the influence of other linguistic factors, such as animacy, context, or pragmatic considerations, on the selection and interpretation of local case marking. For example, the Dhimal ablative marker for inanimates (*soŋ*) extends to perlative uses, while the animate ablative (*dosa*) also functions as a comitative marker. Animacy influences case marking here. Animate beings, as socially and cognitively significant entities, are often treated with more grammatical elaboration, leading to a dedicated marker for *dosa*/, which can encode both source and accompaniment. In contrast, inanimate entities are more likely to be treated as passive locations or paths, leading to the overlap between source and path expressed by /*soŋ*/.

This study highlights the intersection of genetic inheritance and language contact in shaping local case systems. The findings contribute to a broader understanding of spatial case typology, particularly in multilingual and contact-rich environments. By examining how these languages structure spatial relations, the study sheds light on the interaction between morphosyntax, semantics, and pragmatics in the linguistic landscape of South Asia.

References

Blake, B.J. (2001). *Case (Second Edition)*. Cambridge, UK: Cambridge University Press.
King, J.T. (2008). A Grammar of Dhimal. Netherlands: LEI Universiteit Leiden

Finding your niche in the complementation system: the distribution of Basque verbal noun complements

Silvie Strauß
Leipzig University

A complementation system consists of a certain inventory of complement clauses (CC) which are distributed over different syntactic and semantic contexts, taking on different functional profiles (cf. Schmidtke-Bode 2014: Chapter 7). Typological studies of complementation systems have revealed cross-linguistic patterns in their organization concerning the degree of syntactic integration of the CCs into the matrix clause (e.g. Givón 1990; Cristofaro 2003; Noonan 2007; Schmidtke-Bode 2014). Zooming in on the organization of the complementation system of an individual language, however, more specific factors may be decisive in dividing the functional space among the CC types available. In this talk I will look at the division of labour in a subpart of the Basque complementation system, namely verbal noun (VN) complements.

VNs in Basque can be inflected in different cases and their case marking has some further consequences for their morphosyntactic properties. Most importantly, only VNs in argument cases can morphologically distinguish between past and non-past (cf. (5) below).

Here the focus will be on the three types of VN complements that are by far most frequent:

1. VNs in argument cases (i.e. absolutive, ergative, dative, instrumental)
2. VNs in the inessive case
3. VNs in the purposive case

Looking at the complementation patterns of about 200 clause-embedding predicates in the Basque Corpus of Contemporary Texts (ETC), the major factor conditioning the use of the different VN complements turns out to be the temporal or aspectual relation between the matrix and the embedded state-of-affairs, as exemplified in examples (1)-(5) and summarized in Table 1. Generally this relation results from the semantics of the matrix verb, although there are a few minimal pairs with them same matrix verbs, too.

(1) prospective
*[irakur-tze-**ko**]* *esan-Ø* *d-i-e-t*
read-NMLZ-**PUR** say-PFV 3SG.ABS-AUX.DITR-3PL.DAT-1SG.ERG
'I told them to read.'

(2) prospective
*[irakur-tze-a-**Ø**]* *erabaki-Ø* *d-u-te*
read-NMLZ-SG-**ABS** decide-PFV 3SG.ABS-AUX.TR-3PL.ERG
'They decided to read.'

(3) imperfective
*[irakur-tze-**n**]* *has-i* *dira*
read-NMLZ-**INE** begin-PFV AUX.ITR.3PL
'They began to read.'

(4) perfective

*[irakur-tze-a-**Ø**]* *lor-tu* *d-u-te*

read-NMLZ-SG-**ABS** manage-PFV 3SG.ABS-AUX.TR-3PL.ERG

'They managed to read (it).'

(5) temporally independent

*[irakur-tze-a-**Ø**/* *irakurr-i* *izan-a-**Ø**]* *espero* *d-u-t*

read-NMLZ-SG-**ABS** read-PFV be-SG-**ABS** hope 3SG.ABS-have-1SG.ERG

'I hope that they read/ have read.'

	purposive	inessive	argument case
prospective	X		X
imperfective		X	
perfective			X
independent			X

Table 1: Temporal/aspectual relations and case.

As Table 1 shows, the different aspectual relations are neatly divided between the CC types. The only overlap is found in the prospective. Here, a secondary factor comes into play: purposive VNs are strongly linked to contexts where the embedded subject is coreferential with the matrix object and appear with subject coreference only in contexts where an absolutive VN is excluded for syntactic reasons.

The mapping between case and aspect is not random: on the functional level, only VNs in argument cases can express tense independently, which is desirable in contexts with independent time reference. On the semantic level, the purposive marker is also found in the prospective participle, whereas the imperfective participle is morphologically a VN in the inessive, a mapping that conforms also to cross-linguistic tendencies (cf. Bybee et al. 1994: 129-133; Dench 2003: 88-89; Epps 2008).

A closer look at the complementation system of an individual language can thus reveal a fine-grained pattern of organization with its members filling specific ecological niches in line with their morphological properties.

References

Bybee, Joan / Perkins, Revere / Pagliuca, William. 1994. *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World.* Chicago/London: University of Chicago Press.

Cristofaro, Sonia. 2003. *Subordination*. Oxford: Oxford University Press.

Noonan, Michael. 2007. "Complementation." In Shopen, Timothy (ed.): *Language Typology and Syntactic Description,* vol. 2. Cambridge: Cambridge University Press. 52-150.

Dench, Alan. 2003. "From purposive/future to present. Shifting temporal categories in the Pilbara languages of north west Western Australia." In Barry J. Blake & Kate Burridge (eds.): *Historical Linguistics 2001*. 87-104.

Egungo Testuen Corpusa (= ETC) [Corpus of Contemporary Texts]. 2021. https://www.ehu.eus/etc/

Epps, Patience. 2008. "From 'wood' to future tense: Nominal origins of the future construction in Hup." *Studies in Language* 32(2): 383-404.

Schmidtke-Bode, Karsten. 2014. *Complement Clauses and Complementation Systems: A Cross-Linguistic Study of Grammatical Organization.* University of Jena, doctoral dissertation.

Intensifier Adjective Constructions in Sintic languages from a typological perspective

Ting Zhang

Shenzhen University & Johannes Gutenberg University of Mainz

Intensifier adjective constructions are a unique type of expressive adjective formation in Sinitic languages, typically combining a monosyllabic core adjective with an expressive morpheme.Their primary function is to intensify the overall depiction and modification conveyed by the core adjective. Constructions like "飞快" (fēi-kuài, 'fly- fast, extremely fast) and "冰寒" (bīng-hán, 'ice cold, extremely cold') explicitly convey intensification and incorporate vivid expressive imagery, whereas adjectives like "天蓝" (tiān-lán, 'sky blue') do not. Many Sinitic dialects exhibit parallel yet distinct forms, reflecting regional variations in construction and function.

These constructions, predominantly found in East Asian languages, exhibit both significant typological relevance and language-specific characteristics. Existing research has primarily concentrated on stative adjectives in modern Sinitic languages, with studies on intensifying adjective constructions being fragmented and largely homogenized, often relying on isolated dialectal data points. Traditionally, these constructions are classified into XA or BA patterns (A = adjective, X/B = non-core component), where the non-core elements have generally been treated as affixes without in-depth analysis. Typological studies on this structure remain scarce, yet research in this area holds considerable value for linguistic typology.

This study systematically classifies intensified adjective constructions into three types based on synchronic and diachronic perspectives. Type A (滚烫 gǔn tàng)retains identifiable original characters, preserving the full lexical meaning of each component, allowing for independent usage. Type B (瘟苦 wēn kǔ)features non-core component that either use original characters or homophones, undergoing partial semantic bleaching while remaining productive within a constrained semantic domain. Type C (稀嫩 xī nèn) consists of non-core component with untraceable etymon, exhibiting complete semantic bleaching and serving solely as intensifiers, freely combining with any core adjective.

From a diachronic perspective, the evolution from Type A to Type C represents a process of inventory fusion. Type A reflects an early stage where two synonymous lexical items are juxtaposed as independent words, connected via their shared semantic features to depict a vivid state. In this stage, meaning is cumulatively stored in distinct cognitive units. With increased usage frequency and language contact, the non-core component gradually loses its original semantic features while retaining an intensifying function. Within dialectal continua and contact zones, these non-core elements undergo functional convergence, becoming interchangeable and indistinguishable in speakers' cognitive processing. Once integrated into the construction, they universally indicate high intensity adjectives, with multiple units merging into a single cognitive storage slot, thereby completing the process of inventory fusion.

Furthermore, this study explores the inventory-internal categories of such constructions and their underlying mechanisms. Visual perception-related adjectives are most commonly stored in the inventory first, followed by other domains. The non-core elements, while originally not contributing to categorical expansion, develop extended functions through trans-categorical correspondence, aligning with the principle of linguistic economy. This study systematically classifies intensifier adjective constructions in Sinitic languages, tracing their typological variation and diachronic evolution through the process of inventory fusion.

References:

Cysouw, M. (2006). DNS Bhat, Pronouns (Oxford Studies in Typology and Linguistic Theory). Oxford: Oxford University Press, 2004. Pp. xii+ 320. *Journal of Linguistics*, *42*(3), 697-703.

Croft, W. (2003). *Typology and universals*. Cambridge University Press.

Schachter, P., & Shopen, T. (1985). Parts-of-speech systems. *Language typology and syntactic description*, *1*, 3-61.

Liu, D. (2014). On the maximum utilization principle of language inventory. Studies of the Chinese Language, 5, 387–401. [刘丹青. (2014). 论语言库藏的物尽其用原则. 中国语文, 5, 387–401.]

Liu, D. (2019). On inventory fission in Chinese. Language Teaching and Linguistic Studies, 5, 34–50. [刘丹青. (2019). 汉语中的库藏裂变. 语言教学与研究, 5, 34–50.]

Liu, D. (2021). The non-independence view of sememes of linguistic elements. Chinese Teaching in the World, 35(2), 147–165. [刘丹青. (2021). 语言单位的义项非独立观. 世界汉语教学, 35(2), 147–165.]

Adjunct clause and complement order in English-Estonian translation

Abstract. I explore the relative position of the adjunct clause in Estonian fiction texts translated from English. Complements tend to be positioned at the end in Estonian clauses, but follow the predicate verb directly in English. This typically leaves the Estonian adjunct to be flanked by the predicate verb and its complements, and the English adjunct to follow the complements. Estonian adjunct clauses tend to precede or follow their matrix clause, but (non-finite) clauses can also occupy the typical adjunct position in the middle (1).

(1) Estonian (Laanem 2021: 6)

O	*V*	*S*	*adjunct*	*complement*	*complement*
[--],	paiska-s	ta	[--] vahele+jää-des	ema-le	näkku
[--],	blow-3SG.PST	s/he	[--] get+caught-CVB	mother-ALL	face[ILL]

‘(*direct speech)*, he yelled to his mother’s face when he was caught [--].’

I investigate whether translated fiction texts contain these flanked adjunct clauses and the circumstances under which they appear, contrasting this with original Estonian literature. I aim to answer the following questions:

1. Where is the adjunct clause located compared to complements in Estonian sentences translated from English? Does this match the source clause word order? Are there different tendencies from original sentences?

2. Under which circumstances is the adjunct clause positioned between the matrix predicate verb and its complements in English-Estonian translation? How does the syntactic form of the adjunct clause affect its position?

The word order in English and Estonian is predominately (although not entirely) determined by different functional layers: in English, word order carries a significance in the semantic meaning of constituents, e.g. the subject precedes the predicate; Estonian follows V2 and the general word order is shaped by information structure (Lindström 2017). Estonian adjunct clauses have been found to fall outside this structure, often following or preceding the matrix clause (Sahkai 1999), but converb clauses, for example, can still be found in the typical adjunct position (Simmul 2020), even though they otherwise often form separate information units (Simmul 2021, 2023). Given that English adjunct clauses rarely appear between the predicate verb and its complements, and the other positions are common in both languages, adjunct clauses translated from English to Estonian could show a stronger bias towards preceding and following the matrix clause compared to original texts.

To find specific answers, I examine translated and original Estonian sentences with adjunct clauses, as well as the English source sentences, taken from 5-6 page excerpts in novels with Estonian releases from 2010 and later, around 20 for each type. For Estonian adjunct clauses, I determine their type, their position relative to matrix clause complements, their length and their semantic function. Contrasting these data sets will highlight differences between English-translational and original Estonian. For English source sentences, I determine the type of the source construction and its position. This allows me to see whether there is a comparable source clause and whether the adjunct clause has shifted in translation. I assess the cases of adjunct clauses appearing between the predicate verb and complements, using clause length, semantic function, source constructions and other context clues to determine why they might have appeared. The results contribute to understanding clause relationships in the Estonian language and in translational language in general.

Sources

Laanem, Tiina 2021. Öised hääled. [Voices in the night.] Tallinn: Kirjastus Pegasus.

Literature

Lindström, Liina 2017. Lause infostruktuur ja sõnajärg. [Information structure and word order in the finite clause]. – *Eesti keele süntaks*, 537–565.

Sahkai, Heete 1999. Eesti verbifraasi sõnajärg. [Estonian verb phrase word order.] – *Keel ja Kirjandus* 1, 24–32.

Simmul, Carl Eric 2020. Süüvides jonni jätmata avastamaks sõnajärjemustreid: des-, mata- ja maks-konverbitarindi sõnajärg. [Examining word order patterns in the -des, mata and -maks converb constructions.] – *Keel ja Kirjandus* 3, 221–242.

Simmul, Carl Eric 2021. des-, mata- ja maks-konverbitarindi inforoll. [Information role of the -des, -mata and -maks converb constructions.] – *Eesti ja soome-ugri keeleteaduse ajakiri = Journal of Estonian and Finno-Ugric Linguistics*, 12 (1), 303–334.

Simmul, Carl Eric 2023. Information structure of converb constructions: Estonian -des, -mata and -maks constructions. – *Folia Linguistica*, 58 (1), 29–77.

Looking outwards and hearing inwards: on SENSORY PATH in a typological perspective

According to Langacker (1991, p. 303), languages attest few, if any, dedicated clause structures to express sensory perception. Instead, such structures tend to be borrowed from other semantic domains, for instance that of space and motion (see Hopper (2004), Slobin (2008), Huumo (2010)). From the perspective of cognitive semantics, Talmy (2000, p.115) suggests a SENSORY PATH schema which states that a directionally encoded perceptual event involves fictive motion between Experiencer and Experienced "in one direction or the other". Therefore, one and the same sensory experience may be directionally conveyed as either oriented from the Experiencer (1) or towards (2) them.

(1) John cast a sharp look at Ann.
(2) Ann's face swam into his view.

In a study on SENSORY PATH in contemporary anglophone fiction, Aquilina (submitted) observes two factors that underlie the choice of one or the other directionality: (i) SENSORY MODALITY (vision being predominantly conceptualized as motion from the Experiencer, while auditory expressions following the opposite pattern) and (ii) ± VOLITIONALITY (for both sensory modalities, volitional perceptual acts being mostly construed as motion from the Experiencer, contrary to non-volitional experiences).

In this paper, I test whether the two aforementioned factors predict the choice of a directionality pattern in 97 languages, distributed across 5 areas (Africa, Australia & New Guinea, Eurasia, North America, South America, Southeast Asia & Oceania) and sampled according to Miestamo et al. (2016, pp. 284-293). Specifically, I focus on argument marking of basic perception predicates varying in SENSORY MODALITY (vision versus hearing) and ± VOLITIONALITY traits. For instance, in Forest Enets (see (3)) a volitional visual experience is construed as oriented from the Experiencer (the stimulus being introduced by a directional preposition), whereas in Mehweb (4) a non-volitional auditory experience is expressed as oriented towards the Experiencer (the latter being marked with lative).

The data has been collected from a typological database BivalTyp (Say et al.2020-) as well as from dictionaries, grammars and personal communication with linguists. The preliminary results suggest that:

(i) languages show a strong tendency toward directional conceptualization of vision and hearing. Idioms which use a transitive pattern for basic perception predicates attest directionality elsewhere: for instance, "(lit.) throw vision" in Yamalero (5).

(ii) as expected, visual modality favors the directionality pattern oriented from the Experiencer, whereas hearing - the opposite one. This finding may be explained by extra-linguistic factors, such as different degrees of control over the stimuli in vision and hearing (see also Enghels 2007). A visual experiencer can choose and change the stimuli by moving eyes, whereas while perceiving sounds, one is less agentive and thus, less prone to be conceptualized as a source of energy. The degree of control over the stimuli might also account for the difference between volitional (Experiencer→) and non-volitional (Experiencer←) experiences.

(iii) the observed asymmetries are not absolute universals. Thus, in Sorani Kurdish (6), a volitional visual experience is conveyed as directed towards the Experiencer, a conceptualization that conflicts with both predictions stated above.

Appendix (examples):

(3) Forest Enets, Uralic (Ovsjannikova 2020)

kasa-jʔ ŋa dʲez seŋiŋa
man-NOM.SG.1SG sky in_the_direction look(IPFV).3SG.S
'My friend is looking at the sky'.

(4) Mehweb, Nakh-Daghestanian (Daniel & Musaeva 2023)

Islam-ize irʁ-uwe le-r muzika
pn - cont(lat) hear-cvb be-npl music
'Islam hears the music.'

(5) Yamalero, Guahiban (Ginebra, personal communication, January 15, 2025)

koneje betsa na-ita-xua-ba-jü kou dipialito
when GOAL.UP.ITV ?-vision-throw-REAL-1SBJ EVD? Ripialito

nam-ta bo-ka-ina kou
way-SG lie-INTR.SG-REAL-VEN.UP EVD?
'I looked up, the Ripialito road was nearby

(6) Sorani Kurdish, Indo-European (Amadeh 2023)

Hîwa xerîk-e çaw le hewr-ek-an de-k-a

Pn busy-be.prs.3sg eye from cloud-def-pl ind-do.prs-3sg

'Hiwa is looking at the clouds.'

References:

Amadeh, M. (2023). Bivalent patterns in Sorani Kurdish. In: Say, Sergey (ed.). *BivalTyp: Typological database of bivalent verbs and their encoding frames.* (Data first published on June 26, 2023; last revised on October 13, 2024.) (Available online at https://www.bivaltyp.info, Accessed on 7 February 2025.)

Aquilina, Y. (submitted). *On SENSORY PATH and other asymmetries in the encoding of vision and hearing: the case of English.*

Daniel, M. & Majsarat, M. (2023). Bivalent patterns in Mehweb. In: Say, Sergey (ed.). *BivalTyp: Typological database of bivalent verbs and their encoding frames.* (Data first published on January 21, 2024; last revised on June 6, 2024.) (Available online at https://www.bivaltyp.info, Accessed on 28 September 2024)

Enghels, R. (2007). *Les modalités de perception visuelle et auditive : Différences conceptuelles et répercussions sémantico-syntaxiques en espagnol et en français.* Max Niemeyer Verlag.

Hopper, R. (2004). *Perception verbs, directional metaphor and point of view in Tokelauan discourse*. Journal of Pragmatics, 36(10), 1741–1760. https://doi.org/10.1016/j.pragma.2004.06.002

Huumo, T. (2010). *Is perception a directional relationship? On directionality and its motivation in Finnish expressions of sensory perception*. Linguistics, 48(1), 49–97. https://doi.org/10.1515/LING.2010.002

Langacker, R. V. (1991). *Foundations of cognitive grammar*. Volume II, Descriptive application. Stanford: Stanford University Press.

Miestamo, M., Bakker, D. & Arppe, A. (2016). *Sampling for variety*. Linguistic Typology, 20(2), 233-296. https://doi.org/10.1515/lingty-2016-0006

Ovsjannikova, M. (2020). Bivalent patterns in Forest Enets. In: Say, Sergey (ed.). *BivalTyp: Typological database of bivalent verbs and their encoding frames.* (Data first published on November 3, 2020; last revised on November 3, 2020.) (Available online at https://www.bivaltyp.info, Accessed on 3 February 2025.)

Say, Sergey (ed.). 2020-. *BivalTyp: Typological database of bivalent verbs and their encoding frames.* (Available online at https://www.bivaltyp.info, Accessed on 27 September 2024.)

Slobin, D. (2008). Relations between Paths of Motion and Paths of Vision: A Crosslinguistic and Developmental Exploration. In Virginia C. Mueller-Gathercole (ed.), *Routes to language: Studies in honor of Melissa Bowerman.* New York: Psychology Press.

Talmy, L. (2000). *Toward a cognitive semantics, Vol. 1. Concept structuring systems.* Cambridge, MA: The MIT Press.

Volitional forms in Mongolic languages

Volitional moods are a group of verbal categories expressing the speaker's wish, desire, or intention (Ammann & van der Auwera, 2004), which often include forms such as the imperative, hortative, jussive, and optative. In this study, functional definitions of these categories are based on Aikhenvald (2010). For the optative, Dobrushina (2011) distinguishes the performative optative, dedicated to blessings and curses, from the desiderative optative, which conveys the speaker's 'powerless wishes'.

This study investigates the modal markers of volitional moods in the Qinghai-Gansu Mongolic languages in China—Eastern Yugur, Mongghul, Mangghuer, Bonan, Dongxiang, and Kangjia. Comparative data from Middle Mongol and common Mongolic languages[1] are included to understand the historical development and synchronic variation of the volitional forms across Mongolic languages. The data include grammars, dictionaries, and historical texts. The Mongolic terms are relabeled using typological terminology from Aikhenvald (2010).

Mongolic languages exhibit a rich inventory of dedicated volitional forms, conditioned by subject person and number (Janhunen, 2012). Conventionally, these forms are grouped as *imperatives*[2] (Janhunen, 2003), and each of them bears a specific name[3]. For instance, [Hortative][4] corresponds to the Mongolic *voluntative* (*-yA*) (examples 1, 2), while [Jussive] is expressed by the *concessive* (*-tU-kA.(y)i*)[5] and *permissive* (*-g*) (examples 3, 4). The *concessive* and *permissive* also express blessings and curses, as [Performative optative] (examples 5, 6). [Apprehensive] is referred to as *dubitative* (*-xU-jA.(y)i*) (examples 7, 8). Irreal wishes of the speaker [Desiderative optative] are expressed by the *desiderative* (*-AAsAi*) or *optative* (*-sU-xA.(y)i*) in Mongolic terms (examples 9, 10). The *desiderative* (**-xA-sU-xA.(y)i > -AAsAi*) evolved from the *optative* (*-sU > -sU-xA.(y)i*) and has replaced it in certain modern languages, such as Khalkha (Janhunen, 2003). However, languages like Buryat still retain the older *optative* (*-sU*), which expresses intentions (examples, 11).

The data shows that in most modern Mongolic languages the [Desiderative optative] (*-sU-xA.(y)i*) can occur with all persons, whereas the older form (*-sU*) is restricted to expressing first person intentions, and the newer form (*-AAsAi*) is primarily used with the third person. The Qinghai-Gansu Mongolic languages have a special *optative* (*-sA*) that is formally identical to their conditional converbs. It remains debated whether *-sA* developed from the conditional converb (**-xA-sU*) through insubordination, represents a cognate of *-sU-xA.(y)i*, or emerged under Turkic influence. Evidence suggests that *-sA* has retained the older function of *-sU* to convey intentions, which is atypical for a conditional. The [Jussive] suffix *-tU-kA.(y)i*, widely attested in other Mongolic languages, has largely disappeared in the region, with only Eastern Yugur preserving it as a lexicalized remnant used in blessings. While both suffixes (*-tU-kA.(y)i* and *-g*) express blessings and curses, available data leaves open the question of whether different evaluative meanings (blessings vs. curses) are more strongly associated with one suffix rather than the other. Unlike the [Desiderative optative],

---

[1] The common Mongolic languages include various groups of dialects such as Khalkha, Oirat, and Buryat.
[2] Poppe (1955) called these forms *vocatives.*
[3] Conventional Mongolic terms are presented in *italics*.
[4] Typological terms are presented in square brackets '[-]'.
[5] The forms are given in the Common Mongolic format following Janhunen (2003). The exact realization may vary across individual languages. Capitalization within a suffix indicates the presence of two synharmonic variants.

-tU-kA.(y)i and *-g* are primarily restricted to third-person subjects, except in cases where a lexicalized device such as *bol-tugai* (*bol 'be; become'*) is present, thus allowing reference to the first or second person.

In summary, Mongolic languages have particularly rich systems that make detailed semantic distinctions. The data shed light on the connections between volitional moods and conditionals and on their diachronic development.

References

Aikhenvald, A. Y. (2010). *Imperatives and Commands*. OUP Oxford.

Ammann, A., & van der Auwera, J. (2004). Complementizer-headed main clauses for volitional moods in the languages of South-Eastern Europe: A Balkanism. In O. Mišeska Tomić (Ed.), *Balkan Syntax and Semantics* (pp. 293-314). John Benjamins Publishing Company. https://doi.org/10.1075/la.67.15com.

Bläsing, U. (2003). Kalmuck. In J. Janhunen (Ed.), *The Mongolic languages* (pp. 229–248). London: Routledge.

Bulchulu [Buluchilaqhu] and Jalsan [Jalsav] (1990) [published 1991] *Jaguv Yuqhur Gala bae Mungqhul Gala* (*Dongbu Yuguyu he Mengguyu*) [Eastern Yugur and Mongolian], Guigaquda.

Chingeltei [Cavggaldai] and Li Keyu [Li Kex vIui] (1988) [published 1991] *Muvgqhur Gala bae Mungqhul Gala* (*Tuzuyu he Mengguyu*) [Mongghul and Mongolian], Guigaquda.

Dobrushina, N. (2011). The optative domain in East Caucasian languages. In G. Authier & T. A. Maĭsak (Eds.), *Tense, aspect, modality and finiteness in East Caucasian languages* (pp. 95-130). Bochum: Brockmeyer Verlag.

Janhunen, J. A. (2012). *Mongolian*. John Benjamins Publishing Company.

Janhunen, J. (Ed.). (2003). *The Mongolic Languages (Routledge Language Family Series)*. Taylor & Francis.

Junast [Zhaonasitu] (1981) *Dongbu Yuguyu Jianzhi* [A brief outline of Eastern Yugur]. Beijing: Minzu Chubanshe [The Ethnic Publishing House].

Poppe, N. (1955). *Introduction to Mongolian Comparative Studies*. Helsinki: Suomalais-Ugrilainen Seura.

Skribnik, E. (2003). Buryat. In J. Janhunen (Ed.), *The Mongolic languages* (pp. 102–128). London: Routledge.

Appendix

VOL - Voluntative, CONC - Concessive
PERM - Permissive, DES - Desiderative
OPT - Optative, DUB - Dubitative
FPRT - Future participle, EMPH - Emphatic marker

[Hortative] - Eastern Yugur (Junast, 1981)

(1) bə dʒen mede-**ja**
1 self know-VOL
'Let me make the decision myself!'

(2) Buda denjəŋdə χaldala jawə-**ja**
1.pl.incl film see go-VOL
'Let's go see a movie!'

[Jussive] - Kalmuck (Bläsing, 2006)
(3) umsh-**txa**
read-CONC
'he/they should read; he/they must read'

[Jussive] - Eastern Yugur (Junast, 1981)
(4) mulas naad-**gani**!
children play-PERM
'Let the children play!'

[Performative Optative] - Eastern Yugur (Bulchulu & Jalsan, 1991)
(5) Møŋke seigqan nasti, eŋke seigqan dʒirɑalti
long beautiful be…years.old safe beautiful happiness

su:-qi:n beleg bəl-qə bəl-**tɔgwɑi**!
have-FPRT gift be-FPRT be-CONC
'Wish you a long life, good health and happiness!'

(6) tʃə ʃəke la sein su:-**gane**
you big EMPH good have-PERM
'May you be infinitely happy!'

[Apprehensive] - Khalkha (Janhunen, 2012)
(7) mart-**oodzai**
forget-DUB
'May s/he/you not forget!'

[Apprehensive] - Mongghul (Chingeltei & Li, 1991)
(8) nohui jau-**gujee**
dog bite-DUB
'The dog may bite!' (or: 'let it not happen that the dog bites!', 'I hope the dog will not bite').

[Desiderative Optative] - Khalkha (Janhunen, 2012)
(9) en' eubel tzas yix or-**aosai**
this winter snow big enter-DES
'I hope it will snow a lot this winter!'

[Desiderative Optative] - Eastern Yugur (Bulchulu & Jalsan, 1991)
(10) teŋger dʒasən oro-**so**
sky snow fall-OPT
'It would be great if it snowed!'

[Intentions] - Buryat (Skribnik, 2006)
(11) unsha-**huu**-bdi
read-OPT-1.pl.excl
'We shall read (without you).'

Postverbial construction with a deictic motion verb *kel-* 'come' in Western Yugur

Zhencao Zhong
Johannes Gutenberg-Universität Mainz; Minzu University of China

This research investigates the functions of the postverbial construction (PVC) with the deictic motion verb *kel-* 'come' in the Western Yugur language and discusses their possible origins.

A postverbial construction refers to an analytic grammatical unit made up of a main verb (V1) expressing lexical meaning and an auxiliary verb (V2) expressing actionality or other grammatical concepts, which are connected by a converbial marker (Johanson 2021, 597).

Western Yugur is one of the native languages spoken by the Yugur people in northwestern China. It is an endangered Turkic language with about 2,000 speakers, spoken in the Gansu-Qinghai linguistic area (Zhong 2019, 4). It is in close contact with Mandarin Chinese, Amdo Tibetan, Khalkha Mongolian, and Eastern Yugur. Existing research claims that the language is heavily influenced by its neighbouring languages, showing contact-induced changes such as copied phonemes from Mandarin Chinese and some loss of morphosyntactic elements shared with other modern Turkic languages (Chen 2004). However, reports about morphosyntactic copies in the language are relatively scarce.

The Western Yugur data used for analysis come from a 3,213-sentence corpus, based on naturalistic conversation. Data for comparison are either elicited or taken from existing literature.

There are 363 sentences with a PVC in the corpus, 90 of which contain a PVC with *kel-*. This research describes the function of the PVC with the deictic motion verb *kel-* 'come' in Western Yugur and briefly compares these functions with similar constructions in both genetically related and contact languages.

Four different functions of the PVC with *kel-* are identified: (1) cislocative, (2) continuity of action, (3) resultative, and (4) discourse marker (see examples [1]–[4]). Functions (1) and (2) are shared with most Turkic languages, e.g., Modern Uyghur, Salar, and Tuvan. Functions (3) and (4) are not reported in other Turkic languages but are found in Chinese {V1 + lái 'come'}, as seen in examples [5]–[6]. Additionally, a similar resultative function is found in the Santa language, a Mongolic language in the Gansu-Qinghai area (see example [7]).

pər avaka kʰərək kettə
[1] *pər avaka kʰər-**ək** **kel**-tə*
one old.man get.in-CVB POSTV.come-COP
'An old man came in.' (ybe202408170101.mp4, 00:51:16)

antaqa jozək kiɣen
[2] *antaqa joz-**ək** **kel**-ɣen*
that pass-CVB POSTV.come-PTCP
'That is how we lived.' (ybe202408170101.mp4, 00:42:04)

am pʰaqʰasa pʰaqʰak kelomastə sa
[3] *am pʰaqʰa-sa pʰaqʰa-**k** **kel**-o-mas-tə sa*
now poop-COND poop-CVB POSTV.come-CONT-AOR.NEG-COP PRT
'When we try to poop, we can't get it out.' (ybe202408170101.mp4, 00:44:45)

tej tej øryenkenta tʰiʂ ʂəŋtʂʰɛntuita nakə am sa
[4] *tej tej øryen-ken-ta tʰiʂ ʂəŋtʂʰɛntui-ta nakə am sa*
just just study-PTCP-LOC DM production.team-LOC that now PRT

ti-kelse mal mut^{h}oŋ ɕyeɕiaoto a
ti-kel-se *mal* *mut^{h}oŋ* *ɕyeɕiao-to* *a*
say-POSTV.come-COND livestock schepherd.child school-COP that

ʂʐɣe k^{h}ərəptə
ʂʐ-ɣe *k^{h}ər-əp-tə*
inside-DAT get.in-CVB-COP
'When I just started school, I came to the school for children of shepherds in the production team.' (ybe202408170101.mp4, 00:40:35)

他/她 听 来 了
[5] *t^{h}a^{44}* ***t^{h}iŋ^{21} lɛ^{53}*** *liɔ^{21}*
s/he listen come:PFV PRT
'S/he understood it.' (fieldnotes)

说起来…… 还 一箱 葡萄 落 车上 了
[6] ***shuōqǐlái*** *hái* *yìxiāng* *pútao* *là* *chēshàng* *le*
DM still one.box grape left in.the.car PRT
'Speaking of which, there's still a box of grapes left in the car.' (cf. BLCU Corpus Center, (Xun et al. 2016))

alima baolu-dʐɯ irə wo
[7] *alima* *baolu-****dʐɯ irə*** *wo*
fruit ripe-CVB come PFV
'The fruit rope.' (cf. Liu (2009: 146))

I propose that the resultative and quotative uses were copied from Chinese and that the Western Yugur language has accommodated them with an unbalanced manner. The resultative function is more widely adopted and spread, whereas the quotative function is still in its early stages of entering Western Yugur morphosyntax.

Abbreviations

AOR	aorist
COND	conditional
CONT	continuous
COP	copular
CVB	converb
DAT	dative
DM	discourse marker
LOC	locative
NEG	negative
PFV	perfective
PTCP	participle
POSTV	V2 in a postverbial construction
PRT	particle
PVC	postverbial construction

References

Chen, Zongzhen. 2004. *Xībù Yùgùyǔ Yánjīu [Research into the Western Yugur Language]*. Beijing: China Minzu Photographic Art Press.

Johanson, Lars. 2021. *Turkic*. Cambridge Language Surveys. Cambridge: Cambridge University Press.

Liu, Zhaoxiong. 2009. 'Dōngxiāngyǔ Jiǎnzhì [A Brief Description of the Dongxiang Language]'. In *Zhōngguó Shǎoshù Mínzú Yǔyán Jiǎnzhì Cóngshū [A Brief Description of the Chinese Ethnic Languages]*, revised, 6:107–75. Beijing: Minzu Press.

Xun, Endong, Gaoqi Rao, Xiaoyue Xiao, and Jiaojiao Zang. 2016. 'The Construction of the BCC Corpus in the Age of Big Data'. *Corpus Linguistics* 3 (1): 93–109, 118.

Zhong, Yarjis Xueqing. 2019. 'Rescuing a Language from Extinction: Documentation and Practical Steps for the Revitalisation of (Western) Yugur'. Doctoral dissertation, Canberra: The Australian National University.