MEDAL

Methodological Excellence in Data-Driven Approaches to Linguistics

**Final Conference**

Institute of Estonian and General Linguistics
University of Tartu
9 - 10 October 2025

# BOOK OF ABSTRACTS

Funded by the European Union

UK Research and Innovation

# Table of Contents

## Plenary talks



## Presentations

***Alphabetical by first author's surname***

## Integrating phylogenetic, typological, and neurobiological perspectives on language

*Balthasar Bickel*
*University of Zurich*

A core property of the human language faculty is its intrinsic dynamic. Unlike the communication systems of other great apes, language is subject to relentless change, yielding massive typological diversity in its structure. Phylogenetic methods have revolutionized the way this dynamic can be understood, efficiently separating systematic biases from random fluctuations in linguistic evolution. Yet results are limited because we have only access to linguistic evolution since the Neolithic, a period characterized by unusually high densities and contact between human populations. This makes it unclear whether biases reflect only recent history or properties of language at the level of the entire species over its 300-500 ky history. These issues can be resolved to some extent by convergent evidence from experimentally testable mechanisms that drive biases in observed evolution, such as a preference for a certain structure (e.g. agent-first sentences) over another (e.g. agent-last sentences). Biases scale up to the species level if they are grounded in (neuro)biological mechanisms that fully persist under maximally diverse conditions, in particular if they persist under adverse conditions, where for example current usage frequencies in a language are at odds with the proposed mechanism. In my talk I will review recent work along these lines, highlighting the methodological developments that are needed for fully integrating the phylogenetic-typological perspective on language with the neuroscience of language comprehension and production.

## Advances in experimental techniques for testing links between human language and cognition

*Jennifer Culbertson*
*University of Edinburgh*

Human languages exhibit striking variation. At the same time, certain linguistic patterns crop up again and again, while others seem to be extremely rare. What these tantalising observations tell us about human language is one of the most contentious questions in linguistics. Do similarities between languages reflect a special capacity for language that has evolved only in humans? Do they reflect more general features of the human mind, potentially shared with our ancestors? Are they just down to accidents of history? Traditionally, linguists have argued for one or another of these answers based on limited sources of evidence. For example, it is common to base claims on small samples of languages, case studies of how a handful of languages change over time, or examples of how individual languages are learned. In this talk, I highlight a growing body of research using a different approach: artificial language experiments. I show how this approach can be used to bring crucial empirical evidence to bear on how language is shaped (or not!) by the human linguistic and cognitive system.

**From lexicon to eye-tracking: methodological pathways to understanding the encoding of speed in language**

*Piia Taremaa*
*University of Tartu*

This lecture presents a multifaceted exploration of the linguistic encoding of speed within the motion domain, using case studies from Estonian. The investigation is guided by two primary questions: how is speed expressed in motion descriptions, and do these linguistic realisations of speed reflect grounded cognition? To address these questions, the presentation introduces a selection of methodological approaches, each contributing a unique perspective on the expression of speed in language. For example, lexical analysis – combining dictionary searches with corpus-based techniques – reveals imbalances in the frequency and variety of expressions for fast versus slow motion. Behavioural profiling within corpus data uncovers patterns in clause structure and biases toward spatial or manner encoding. Complementing these analyses, rating experiments are used to assess speakers' intuitions about word meanings along the speed dimension. In a more dynamic setting, elicited language production using visual stimuli reflects the influence of speed on both the content and delivery of motion descriptions. Additionally, reaction time measurements offer insights into the cognitive processing of fast versus slow motion terms, while eye-tracking data sheds light on how attention is distributed during language comprehension involving speed cues. Overall, the presentation emphasises the complementary nature of these methodologies and argues for the benefits of multi-methodological research designs. It advocates for converging evidence as a cornerstone of robust linguistic analysis.

Authors:

Aigro, M. Taremaa, P. Pilvik, M.-L. Lõo, K. Mackiewicz, J. Chuprina, A. Milin, P. Vihman, V.-A., Divjak, D.

## Native-like eye movements: Investigating L1 and L2 processing of pronominal subjects in Estonian

Many languages enable speakers to encode subject person information twice, meaning that, in addition to mandatory verbal morphology, subject person information may also occur as a separate pronoun, as in (1) in Estonian.

(1) *Ma* *armasta-n* *küpsise-i-d.*
**1SG** love-PRS.**1SG** cookie-PL-PAR
'I love cookies.'

However, such languages vary in terms of the extent to which speakers use pronominals to express subject information. In addition, while certain fundamental cognitive factors are likely to encourage the occurrence of pronouns cross-linguistically, pronoun use in different languages is likely to also be regulated by language-specific contextual conditions.

This study uses eye-tracking to compare the ways in which Estonian L1 speakers and Russian speakers using Estonian as L2 process subject pronouns or the lack thereof in Estonian. Each participant was shown written stimuli adapted from spoken corpora transcriptions. Spoken data was preferred as pronoun use may be subject to editing in written language, which may affect the contexts in which pronouns are used and omitted.

First, stimuli only involve contexts which feature pronouns significantly more frequently in Estonian than in Russian (1SG, 2PL). This enables us to investigate a possible L2 effect in processing overt pronouns, namely one where explicit pronouns would hinder L2 reading measures in the given stimuli more than L1 reading measures.

Second, half of the stimuli each participant saw included manipulation, meaning that either an originally occurring pronoun had been deleted, or one had been provided where originally there was none. This allows us to first establish eye-tracking with naturalistic stimuli as a fruitful method for tracking the conditions facilitating syntactic variation in language, by showing that L1 speakers' reading measures are sensitive to manipulations. Furthermore, it also allows us to investigate whether L1 speakers are more sensitive to pronoun manipulation than L2 speakers, as the latter's sensitivity to the contextual conditions facilitating pronoun use in Estonian can be expected to be reduced.

In addition, we collected various proficiency indicators from all L2 participants to see if both hypothesised effects – L2 speakers being more sensitive to pronoun use, but less sensitive to manipulation of pronoun use – decrease with increased L2 proficiency. If this was the case, it would demonstrate that being more native-like in their linguistic usage patterns also means that their eye movements and the processing patterns they imply become more like those of L1 speakers.

# The Potential of Animated Videos in The Development of English Tense-aspect Usage for Saudi EFL Learners

Sarah Alshahrani, Petar Milin, Dagmar Divjak

University of Birmingham

**Keywords:** tense-aspect acquisition, incidental learning, animated videos, classroom intervention

Implicit learning is the acquisition of knowledge within a natural process without a conscious awareness of the task of learning while explicit learning tends to be a more conscious process in which the learner makes judgments of hypotheses to find/acquire the structure (Ellis & Rebuschat, 2015). Although explicit instruction raises rule awareness, it doesn't always lead to full mastery, particularly in complex areas like tense and aspect. In Saudi EFL classrooms, this reliance on explicit instruction has contributed to learners' ongoing difficulties in mastering these grammatical features.

This study examines the role of animated videos in tense-aspect incidental learning among Saudi EFL learners, comparing the effectiveness of enhanced versus non-enhanced input. Sixty intermediate-level female students were divided into three groups. Control Group **CG** (N=20) learned through reading dialogues that presented tense and aspect structures in the form of written texts. Treatment Group 1 **TG1** (N=20) watched enhanced animated videos where grammatical features were visually highlighted using colour coding. Treatment Group 2 **TG2** (N=20) viewed non-enhanced animated videos, where target structures were naturally integrated into the narrative without visual emphasis. Learners' knowledge was assessed using a forced-choice task, along with grammar judgment and error correction tasks, before and after the intervention

A Generalized Linear Mixed-Effects Regression analysis showed a statistically significant improvement in **TG2** ($p = 0.000007$). However, **TG1** didn't show significant changes ($p = 0.3$), and **CG** showed changes but not highly significant ($p=0.02$). Our study suggests that instructional interventions should carefully consider input saliency and cognitive load in incidental learning. It challenges assumptions about the benefits of enhanced input in SLA, emphasizing the role of naturalistic exposure in grammar learning but also the value of traditional grammar teaching.

**References:**

Ellis, N. C., & Rebuschat, P. (2015). Implicit AND explicit language learning: Their dynamic interface and complexity. In (Vol. 48, pp. 3-24). John Benjamins Publishing Company. https://doi.org/10.1075/sibil.48.01ell

## Metaphors without Borders : Multilingual Approaches to Metaphor Identification

*Bhumika Bhattacharya and Indranil Dutta*

Humans use metaphors to convey complicated concepts in more concrete domains embedded in the knowledge of their world. In the conceptual metaphor theory which is regarded as the main linguistic and cognitive theory behind metaphors (Lakoff and Johnson, 1980; Lakoff, 1993), it is understood that metaphors are primarily a mental activity, and language is merely a side effect of these conceptual metaphors. From this theory it is extrapolated that metaphors are agnostic in regards to syntactic structure, that is, a conceptual mapping can be expressed through any syntax that the speaker uses. This is a fundamental challenge for computing: how to operationalize a deeply cognitive phenomenon in systems that process primarily statistical regularities of language. For the past two decades, syntactic and semantic heuristics have dominated computational metaphor detection, a system using lexical features, part-of-speech tags, and dependency relations. While these approaches yielded moderate success, they do not capture the generalizability and flexibility of metaphoric cognition, especially with low-resource languages lacking in annotated corpora.

This poster examines whether multilingual training can improve the detection of metaphors, which are cognitive and thus language-agnostic phenomena and try to postulate that training in multiple languages enhances model performance in each language relative to monolingual training, building on recent developments in metaphor detection that combine transformer-based models with linguistic theories like the Metaphor Identification Procedure (MIP). To ensure reproducibility, I will replicate recent benchmark experiments using publicly accessible code and datasets. All pre-processing scripts, trained models, and code will be publicly available. I will also assess transfer performance on a small manually annotated Hindi dataset (adapted from existing parallel corpora) and refine multilingual transformer models (mBERT, XLM-R) on English benchmark datasets. Using the Metaphor Identification Procedure (MIP), I will then generate a pilot dataset of about 500 Hindi metaphors.

This research aims to highlight metaphor detection as a cognitively universal task and argue that metaphor detection can benefit from cross-lingual conceptual mappings. Robust models built with the aim of accurate metaphor detection will enhance downstream applications such as sentiment analysis and misinformation detection. By examining cross-lingual architectures, this research will not only contribute to computational metaphor processing but also attempt to build more generalizable models.

## Acquisition of negation in Mongolian using the Sketch Acquisition approach

*Dorjderem Byambasuren, Shanley Allen and Dorjderem Byambasuren*

Negation structures are typically assumed to be one of the complex constructions in child language (Tagliani, Vender, and Melloni, 2022). Although the acquisition of negation has been extensively studied in a number of different languages such as English (Klima and Bellugi, 1966), German (Hummer et al., 1993), Japanese (McNeill & McNeill, 1973), and Korean (Choi and Zubin, 1985), very little is known about how children acquire negation in morphologically complex languages. Mongolian is an agglutinative language that allows a variety of negative markers, including negative particles, auxiliaries, and suffixes. In this study, we examine the development of negation structures in Mongolian-speaking children aged 1,8-4;0 and their caregivers by analyzing six hours of spontaneous naturalistic speech data to determine the trajectory of development. The data from 10 children, living in urban and rural communities of Mongolia, were processed following the guidelines of the Sketch Acquisition Manual (Defina et al., 2023). We found that negation constructions were present from 1;8 and increased with age. Negation particles were acquired first, followed by auxiliary verbs and grammatical morphemes, consistent with the finding that analytic forms emerge earlier than synthetic forms. Children used postverbal negation with suffixes more frequently than preverbal negation particles, consistent with the use of caregiver input. Finally, the development of negation functions followed the sequence from rejection to non-existence to denial, which aligns with patterns observed in other languages. These results shed light on Mongolian acquisition in general and in relation to caregiver speech and contribute to the understanding of how negation is acquired in agglutinative and morphologically complex languages.

**Multimodal Synchronization in Children's Peer Conversations: A Data-Driven Analysis of Motion and Language**

Understanding synchronization in communication, both verbal and non-verbal, is crucial for gaining deeper developmental insights in preschool children. Verbal synchronization involves adjusting speech patterns to align with conversation partners, while non-verbal synchronization includes coordinating physical movements like gestures or spatial distance. The development of synchronization in conversation among peers is a multimodal process, intertwining linguistic and non-linguistic modalities. This study explores the emergence of conversational synchronization by examining both motion and language in a peer-talk corpus of 60 Hebrew-speaking children, divided into four age groups (3;0–3;6, 3;6–4;0, 4;0–5;0, 5;0–6;0, five triads per group). Approximately 15 hours of synchronized audio and video recordings were processed using AI-driven tools, enabling detailed tracking of the children's physical proximity and their referring expressions—ranging from zero reference, to pronouns, to full lexical nouns with various morphosyntactic features.

Quantitative analyses reveal that when children are closer together, they favor less explicit referential forms, whereas greater distance brings more explicit lexical choices to the fore. Notably, specific morphosyntactic constructions (such as zero reference in past-tense verbs or second-person imperative forms) show varying correlations with distance at different ages. These patterns often take on U-shaped or inverted-U trajectories: both younger and older children link their language choices to spatial distance, whereas intermediate-age groups exhibit weaker correlations. Furthermore, the variance in distance itself follows a U-shaped curve across ages—peaking around 3;0, dipping between 3;6 and 5;0, then rising again by 6;0—indicating an initial period of broad spatial exploration, a phase of relative convergence, and a return to more variable physical positioning. By contrast, variance in linguistic features generally decreases over time, suggesting a tightening range of referential strategies in the face of changing social dynamics. Taken together, these findings align with a psychologically motivated, cognitively grounded perspective on information processing, indicating that as children refine their discourse skills, they continue to adapt how they manage proximity—albeit in an increasingly structured manner.

Methodologically, this work illustrates the power of integrating motion energy analysis with corpus-based annotation to unravel the multimodal fabric of early peer interaction. By documenting how children's physical and linguistic behaviors evolve in tandem, this study illuminates not only the emergence but also the progressive development of conversational synchronization. The findings highlight both stable patterns and renewed exploratory phases as children navigate the complex trade-offs between referencing efficiency and social engagement. Moreover, our results emphasize that children's capacity to coordinate bodily motion with referential precision strengthens over time, offering new insights into how synchronized interaction evolves in parallel with communicative skills.

## Negative Markers Variation in Cangin Languages: A Comparison between Paloor, Noon, and Saafi

*Anna Marie Diagne*

This presentation offers a comparative description of verbal negation structures in three Cangin languages spoken in Senegal: Paloor, Noon, and Saafi. The study focuses on identifying negative morphemes, their syntactic placement, and differencies observed between these closely related languages. Drawing on oral corpus data and existing grammatical descriptions, the analysis highlights both morphosyntactic convergences and divergences among them, while contributing to a broader documentation of negation systems in Atlantic languages.

**Negation in Turkish Sign Language: A Corpus-Based Analysis of Manual and Nonmanual Markers**

This study investigates the frequency and distribution of negation markers (both manual and nonmanual) in Turkish Sign Language (TİD). Negation, a universal linguistic phenomenon, demonstrates significant variation across signed languages (SLs), revealing unique typological patterns. For example, some SLs mainly employ nonmanual markers (e.g., headshake), while others have been shown to rely mainly on manual negators. TİD has traditionally been classified as being the latter type, i.e., a manual-dominant SL (Zeshan, 2006). The present study explores this classification from the perspective of corpus linguistic research.

For this study, data from a subset of an existing corpus was used (Dikyuva et al., 2017). The data is composed of video recordings of 66 participants involved in semi-structured conversations, recorded using two cross-positioned cameras that more accurately capture not only the manual signs but also the nonmanual aspects of their communication. The data was then annotated using ELAN software, focusing on tier-based GLOSS analysis (Crasborn & Sloetjes, 2008). The negation-related annotations identified 16 different manual signs, and 13 nonmanual markers (NMMs). The NMMs include head movements, brow movements, and mouth gestures. In the majority of negation statements (83.91% of all annotated negation statements), manual and nonmanual forms were used simultaneously, requiring a modified coding system to more efficiently code them. This modified coding system allowed us to do a more in-depth examination of the combined use of manual and nonmanual markers.

Using the previously annotated TİD corpus, we coded 1,418 negative clauses and identified three categories: a) NM marking only (81 / 1,418 cases: 5.72%); b) manual marking only (148 / 1,418: 10.37%); and c) combined use of manual and NM markers (1,189 / 1,418 cases: 83.91%). We will discuss the four most commonly found manual negators: the basic negators DEĞİL-1 (NOT-1) and DEĞİL -2 (NOT-2); the negative existential YOK (NOT-EXIST); and the negative adverbial HİÇ (NEVER). We will also discuss the use of NMMs with these manual negation forms. Our findings clearly show that TİD negation is mostly expressed by using a combination of both manual negators and NMMs.

While our analyses seem to underscore TİD's classification as a manual-dominant SL, it is also clear that the majority of cases still included NMMs. The frequent co-occurrence of NMMs with manual negators further suggests that TİD features a hybrid system (Makaroğlu, 2021), thus challenging the traditional dichotomy. Our functional analysis of the role of NMMs in TİD negation thus sheds new light on discussions regarding the typology of TİD.

Crasborn, O., & Sloetjes, H. (2008). *Enhanced ELAN functionality for sign language coropora*.
Dikyuva, H., Makaroğlu, B., & Arik, E. (2017). *Turkish Sign Language Grammar*. Ministry of Family and Social Policies, Ankara: Fersa Ofset.
Makaroğlu, B. (2021). A Corpus-Based Typology of Negation Strategies in Turkish Sign Language. *Dilbilim Araştırmaları Dergisi*, *2021*, 111-147. https://doi.org/10.18492/dad.853176
Zeshan, U. (Ed.). (2006). *Interrogative and negative constructions in sign languages*. Ishara Press.

## Exploring the neural signature of implicit morphological language learning abilities: the role of resting state beta waves in mastering subject-verb agreement in a morphologically complex language

*Dagmar Divjak, Petar Milin and Kiran Pandurang Phalke*

While most EEG studies on language have analysed time data, there is evidence from the analysis of frequency data suggesting that it is in fact resting state alpha waves that accurately predict language learning abilities (Prat et al. 2016). However, the tests administered in Prat et al. (2016) essentially equate language learning with vocabulary learning. While lexical development is no doubt crucial for L2 development, for many languages, the challenges inherent in lexical development are eclipsed by those inherent in morphology and syntax.

To start bridging this knowledge gap, we repeated the controlled language learning task presented in Ez-zizi et al. (2024) while recording cortical brain activity via 64 electrodes using an ANT Neuro mobile kit with a 500Hz sampling rate, both during a 5-minute resting period before the experiment and during training. We recruited 30 monolingual participants with British English as their first language, aged 18 or over, neurotypical, without learning disabilities and with normal or corrected-to-normal vision. Participants were taught Subject-Verb agreement via an implicit pseudo-artificial language learning task modelled on the plural past tense in Polish which boasts virile and non-virile past tense endings. First, participants were taught the Polish labels of the different animal and human characters used in the learning task. During the learning task, participants were presented with audio-visual scenes representing an action performed by a group of human and/or animal characters. After training participants were tested on unseen character combinations carrying out familiar actions, and they were asked to indicate whether a given past tense form accurately described a scene.

The 64-channel raw EEG data collected during five-minutes of eyes-closed resting-state at the start of the experiment was processed to obtain power in different frequency bands using EEGLAB toolbox for MATLAB (v. R2023b). Data from 10 participants had to be discarded before data analysis, 8 because of poor EEG data quality (recordings with fluctuating or high impedance values) and 2 because of software malfunction (for one of the memory tasks). The raw power calculated over certain electrodes in frontal (F3, F4, AF3, AF4, F7, F8), central (FC5, FC6), parietal (P7, P8), temporal (T7, T8) and occipital (O1, O2) brain regions was averaged and converted to logarithmic values. These values correlate with accuracy scores from the language learning task: while the multiple-comparison corrected Pearson coefficients of correlation were not significant, the most significant correlation before the correction was obtained for mid-beta band (15-17.5 Hz) over frontal (F7) and central (FC6) electrodes. Additionally, raw power over electrodes was converted to laterality coefficients between pairs of left and right hemispheric electrodes. When these measures were correlated with language learning task accuracy scores, it was again observed that significant correlation existed before multiple-comparison correction, between frontal (F3-F4) electrodes in the beta frequency range.

Overall, our EEG results point in the direction of beta frequency band over frontal electrodes being most useful in predicting the language learning task scores, with lower resting-state beta band power correlating with overall higher test scores. We will discuss these findings and their implications for the study of language learning.

**Selected references**

Ez-zizi, A., Divjak, D., & Milin, P. (2024). Error-correction mechanisms in language learning: Modeling individuals. Language Learning, 74(1), 41-77.

Prat, C. S., Yamasaki, B. L., Kluender, R. A., & Stocco, A. (2016). Resting-state qEEG predicts rate of second language learning in adults. Brain and language, 157, 44-50.

**The role of the face in language emergence**

The primary mode of (early) human communication is face-to-face interaction, involving multimodal and meta-communicative signals such as facial expressions (Vigliocco et al., 2014). It is widely recognized that such facial signals support communication in existing languages by e.g. facilitating the prediction of upcoming utterances, repair and backchanneling (Micklos & Woensdregt, 2023; Nota et al., 2021). These kinds of processes likely also play a role in the pragmatic inference relevant to coordinating on novel linguistic variants (Healey et al., 2007; Micklos et al., 2018). However, communication experiments that simulate the process of coordinating and converging on novel linguistic variants often overlook the role of such signals. Even more strikingly, many of these experiments are not conducted in a face-to-face context.

To address this gap and explore the role of facial signals in language emergence, we combine methods from research on multimodal communication with a traditional communication game paradigm commonly used in language evolution experiments. We manipulate facial visibility in a between-subjects design (34 dyads total) yielding two conditions: face visible and face invisible. Dyads repeatedly communicate about a set of twelve abstract shapes (adapted from Macuch Silva et al. (2020)) while being video and audio recorded, with roles of director and matcher switching every trial: the director invents and writes a novel label for the target shape, and the matcher guesses what the target shape is from an array of shapes upon seeing the label. Writing (and speaking) using existing language was prohibited during the task. Participants played the game, using the same shapes, over four consecutive rounds.

We focus on two measures and their progression over time across conditions: (1) communicative success, capturing accuracy of a guess, and (2) convergence, capturing how similarly participants in a dyad label a shape. We have also conducted a preliminary analysis of response times for guessing and labeling. Results provide insight into how the presence of facial signals influences communication and coordination when building a novel communication system, contributing to a deeper understanding of the face's role in language emergence.

Participants who do not see each other lack access to facial signals that support pragmatic inference, for example by signalling feedback (Hömke et al., 2018, in press), but also confidence regarding the label or guess (Nölle et al., 2023; Swerts & Krahmer, 2005). We predicted that a lack of this information would hinder communicative success and convergence. However, we also acknowledge the possibility that participants could adapt to the absence of facial signals in ways that facilitate communication, such as using more iconic labels (Roberts et al., 2015; Tamariz et al., 2018) or meta-communicative vocalizations.

Results show both communicative success and convergence increasing significantly over rounds, but no significant main or interaction effect of condition (see fig. 1 and 2). However, the face invisible condition showed significantly more variance with respect to convergence: pairs who could see each other converged more uniformly, while levels of convergence were more varied across pairs who could not. Furthermore, a preliminary analysis of guessing and naming times revealed that both were slower overall when participants interacted face-to-face. Moreover, this effect on guessing times interacts with communicative success and convergence: trials with incorrect guesses or the use of labels that were less converged on were much slower when the face was visible compared to invisible.

A highly plausible explanation for the longer response times is the increased processing load that comes with participants attending to facial signals, especially when they are less confident in their guess. In addition, the greater variation in convergence in the face invisible condition suggests that emerging communication systems are more susceptible to variation when facial signals are unavailable. Despite these observations, the overall similar levels of communicative success and convergence between conditions suggest that facial signals are not indispensable for success or convergence. We speculate that participants can adaptively overcome the absence of facial signals through alternative means, such as using more iconicity or meta-communicative vocalizations. Taken together, these findings suggest that while facial signals can influence early communication, they are not essential for building a successful communication system, at least in the written medium. To get a more nuanced understanding of how the availability of facial signals affects communicative success and convergence, we will explore the role of meta-communicative vocalizations and iconicity of the labels as potential compensatory strategies, while also examining the role of facial expressivity in future research.

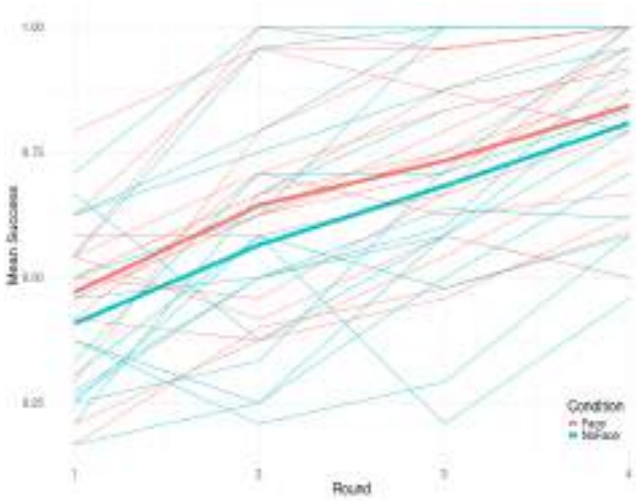



Figure 1: Mean proportion of successful trials per round for each condition in bold, with dyads visualized separately.

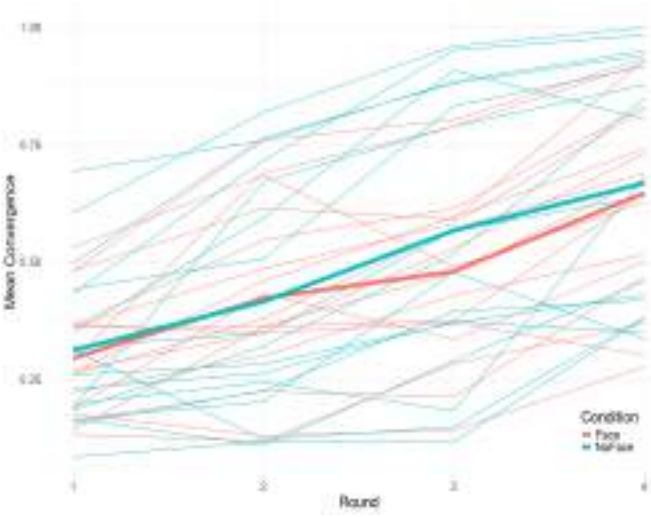



Figure 2: Mean convergence per round for each condition in bold, with dyads visualized separately.

**References**

Healey, P. G. T., Swoboda, N., Umata, I., & King, J. (2007). Graphical Language Games: Interactional Constraints on Representational Form. *Cognitive Science*, *31*(2), 285–309. https://doi.org/10.1080/15326900701221363

Hömke, P., Holler, J., & Levinson, S. C. (2018). Eye blinks are perceived as communicative signals in human face-to-face interaction. *PLOS ONE*, *13*(12), e0208030. https://doi.org/10.1371/journal.pone.0208030

Hömke, P., Levinson, S. C., Emmendorfer, A. K., & Holler, J. (in press). Eyebrow movements as signals of communicative problems in human face-to-face interaction. *Royal Society Open Science*. https://pure.mpg.de/pubman/faces/ViewItemOverviewPage.jsp?itemId=item_3389878

Macuch Silva, V., Holler, J., Ozyurek, A., & Roberts, S. G. (2020). Multimodality and the origin of a novel communication system in face-to-face interaction. *Royal Society Open Science*, *7*(1), 182056. https://doi.org/10.1098/rsos.182056

Micklos, A., Silva, V. M., & Fay, N. (2018). The prevalence of repair in studies of language evolution. *Proceedings of the 12th International Conference on the Evolution of Language (Evolang12)*. The Evolution of Language. Proceedings of the 12th International Conference on the Evolution of Language (Evolang12). https://doi.org/10.12775/3991-1.075

Micklos, A., & Woensdregt, M. (2023). Cognitive and Interactive Mechanisms for Mutual Understanding in Conversation. In *Oxford Research Encyclopedia of Communication*. https://doi.org/10.1093/acrefore/9780190228613.013.134

Nölle, J., Wu, Y., Arias, P., Yang, Y., Garrod, O. G. B., Schyns, P. G., & Jack, R. E. (2023). *Multimodal markers of confidence and doubt: Inferring Feeling of Knowing from facial movements*.

Nota, N., Trujillo, J. P., & Holler, J. (2021). Facial Signals and Social Actions in Multimodal Face-to-Face Interaction. *Brain Sciences*, *11*(8), Article 8. https://doi.org/10.3390/brainsci11081017

Roberts, G., Lewandowski, J., & Galantucci, B. (2015). How communication changes when we cannot mime the world: Experimental evidence for the effect of iconicity on combinatoriality. *Cognition*, *141*, 52–66. https://doi.org/10.1016/j.cognition.2015.04.001

Swerts, M., & Krahmer, E. (2005). Audiovisual prosody and feeling of knowing. *Journal of Memory and Language*, *53*(1), 81–94. https://doi.org/10.1016/j.jml.2005.02.003

Tamariz, M., Roberts, S. G., Martínez, J. I., & Santiago, J. (2018). The Interactive Origin of Iconicity. *Cognitive Science*, *42*(1), 334–349. https://doi.org/10.1111/cogs.12497

Vigliocco, G., Perniss, P., & Vinson, D. (2014). Language as a multimodal phenomenon: Implications for language learning, processing and evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1651), 20130292. https://doi.org/10.1098/rstb.2013.0292

**A Dataset of Scalar Estonian Document-level Subjectivity Evaluation**

Karl Gustav Gailit, Kadri Muischnek, Kairit Sirts

Subjectivity analysis is a well-established task in Natural Language Processing, where subjectivity describes a text that consists of the author's thoughts and opinions, meanwhile objectivity describes a text that contains facts and represents reality. The most common approach to the task of subjectivity analysis is analyzing the subjectivity of individual sentences, such as the popular subjectivity dataset SUBJ (Pang & Lee 2004), which consists of sentences from movie reviews (labeled as subjective) and plot summaries (labeled as objective). Subjectivity analysis can also be performed on the document level, such as detecting whether newspaper articles are subjective (such as opinion pieces) or objective (such as news) (Toprak & Gurevych 2009).

To analyze subjectivity there first need to be datasets to analyze subjectivity on and there are two methods for creating these datasets. The first is using human annotators to assign the texts subjectivity labels, while the second is using the source of the text as the label – for instance, labeling movie reviews as subjective and plot summaries as objective. However, the latter method is much less accurate and therefore unreliable. (Přibáň & Steinberger 2022)

While existing datasets generally label subjectivity dichotomically, as either objective or subjective, humans experience subjectivity more like a scale, where one end is total objectivity and the other subjectivity. While Likert scales have been previously used in subjectivity analysis (Wiebe 2000; Escouflaire, Descampe & Fairon 2024), we propose having annotators instead score the subjectivity of texts on a full numerical scale between 0 and 100, where 0 describes a completely objective text and 100 a completely subjective text. Using this method, our goal is to create a usable and versatile dataset for Estonian document-level subjectivity analysis.

The poster presentation shall describe the results of this project, with focus on the correlation between the annotators. We will also explore various other aspects, such as correlation between the annotators scores and aspects of the text itself, including text length and genre. Additionally, there will be an overview of how annotations vary over time, even with a single annotator.

Sources:

Escouflaire, Louis, Antonin Descampe, and Cédrick Fairon. "Unveiling Subjectivity in Press Discourse: A Statistical and Qualitative Study of Manually Annotated Articles." *Discours* 34. 2024.

Pang, Bo, and Lillian Lee. "A sentimental education: sentiment analysis using subjectivity summarization based on minimum cuts." *Proceedings of ACL*. 2004.

Přibáň, Pavel, and Josef Steinberger. "Czech Dataset for Cross-lingual Subjectivity Classification." *Proceedings of the Thirteenth Language Resources and Evaluation Conference*. 2022.

Toprak, Cigdem, and Iryna Gurevych. "Document level subjectivity classification experiments in deft'09 challenge." *Actes du cinquième DÉfi Fouille de Textes*: 91. 2009.

Wiebe, Janyce M. "Learning Subjective Adjectives from Corpora." 2000.

## Combining EEG and eye tracking: from study design to data analysis

*Maria den Hartog and Andreas Wölfle*

Linguists have various methods at their disposal to investigate language processing, including electroencephalography (EEG) and eye tracking (ET). Event-related potentials (ERPs) recorded using scalp EEG measure the brain activity which occurs following a (language) stimulus. In ET, cognitive processes are inferred from eye movements. Eye movements are generally regarded as detrimental to ERP measurements, because the electrical activity generated by eye movements obscures the electrical activity stemming from the brain (Luck, 2005). To circumvent this problem, words are often presented using a rapid serial visual presentation (RSVP) paradigm for ERP recording. However, eye movements naturally occur during normal reading and the onset of a fixation on a word means that word can start to be processed in the brain. This principle has been fruitfully utilized to compute fixation-related potentials (FRPs; Baccino & Manunta, 2005; Degno & Liversedge, 2020). FRPs are ERPs where the "event" used for the computation of the ERP is the fixation on a target word.

We illustrate the methodological challenges and opportunities of FRPs via a study examining the effect pronouns of address in German (den Hartog et al., 2024). This study centers around the question: does the choice for a formal ('Sie' you; V) or informal ('du' you; T) pronoun of address affect how short emotional narratives are processed? We consider all the methodological steps taken to answer this research question, from the study's design, to its practical implementation and the data analysis pipeline required to synchronize the EEG and ET data streams. We also provide instructions on how to construct and analyze a combined EEG and ET experiment using conventional EEG and ET software: SR Research Experiment Builder and Data Viewer, BrainVision Recorder, MathWorks MATLAB with the FieldTrip toolbox (Oostenveld et al., 2011), and R.

Starting with study design, the combination of EEG and ET allows more naturalistic stimuli to be used than regular ERP studies might employ. Participants in our study read short narratives in which positive or negative emotions are conveyed, with each story presented in full on a computer screen. We share our considerations regarding the physical properties of the stimuli, which can affect low-level processing. We also show how to create a marker in the EEG recording at the start of stimulus presentation to synchronize the EEG and ET signals, and how to create markers in the EEG recording contingent on fixations to a word to use in the analysis later on. Moving to the practical implementation of the experiment, we show the physical setup of our experiment and the challenges we faced connecting the EEG and ET systems in the lab. Next, we discuss our analysis pipeline. Data analysis starts with the extraction of the ET data in Data Viewer and its integration into the EEG data in MATLAB. We discuss how to assess the degree of synchronization between the EEG and ET systems and how the ET data can be used in the EEG analysis to handle eye movement artifacts. Finally, we discuss how the resulting FRPs can be analyzed using cluster-based permutation testing in MATLAB, or using parametric methods by exporting the data to R.

### References

Baccino, T., & Manunta, Y. (2005). Eye-Fixation-Related Potentials: Insight into Parafoveal Processing. Journal of Psychophysiology, 19(3), 204–215. https://doi.org/10.1027/0269-8803.19.3.204

Degno, F., & Liversedge, S. P. (2020). Eye Movements and Fixation-Related Potentials in

Reading: A Review. Vision, 4(1), 11. https://doi.org/10.3390/vision4010011

den Hartog, M., Wöfle, A., & de Hoop, H. (2024). Du and Sie—Processing German forms of address in an emotional narrative [OSF Preregistration]. https://osf.io/3zgcf/

Luck, S. J. (2005). An introduction to the event-related potential technique. MIT Press. https://mitpress.mit.edu/9780262621960/an-introduction-to-the-event-related-potential-technique/

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. Computational Intelligence and Neuroscience, 2011, 1–9. https://doi.org/10.1155/2011/156869

**From motion to marking: an artificial language learning study on differential argument marking**

Many languages exhibit Differential Argument Marking (DAM), most commonly in the form of Differential Object Marking (DOM) and less frequently Differential Subject Marking (DSM) (Bossong, 1985; Aissen, 2003; Witzlack-Makarevich & Serzant, 2018). Accounts grounded in communicative efficiency propose that languages expend more coding in less expected configurations—e.g., atypical pairings of roles and referents—thereby increasing predictability (Gibson et al., 2019; Haspelmath, 2019, 2021a,b; Levshina, 2021). Yet such accounts alone do not explain why DOM is cross-linguistically more common than DSM (Holmes & Vihman, 2025). We argue that the asymmetry reflects the interaction of (i) a general bias toward accusative alignment and agent-centered construal (Comrie, 2013; Longenbaugh & Polinsky, 2017) and (ii) a metaphorical extension from motion to transitive events, whereby *goals* more readily map onto patients and *sources* onto agents (Bárány, 2018; Motamedi et al., 2021).

We test these predictions using an artificial language learning paradigm with directional and transitive constructions. During training, goal/source markers appeared only in directional clauses; transitive clauses were unmarked with flexible word order (50% SOV, 50% OSV). In a subsequent comprehension test, participants saw novel transitive sentences in which one argument was unexpectedly marked: in one group with the *goal* marker, in another with the *source* marker. Native speakers of Thai (analytic, neutral alignment) and Estonian (rich morphology, accusative alignment) were tested to probe cross-linguistic robustness (Bisang, 1991; Iwasaki & Ingkaphirom, 2007; Kont, 1963; Tamm, 2004).

Results showed an overall bias to interpret marked arguments as *patients*, consistent with an accusative-leaning preference. Critically, only the *goal*-marking condition exhibited a positive relationship between accurate learning of the locative function and patient interpretations in marked transitives, suggesting that goals extend more readily to patient-marking than sources to agent-marking. Exposure to goal-marked transitives also increased SOV interpretations of unmarked transitives, indicating an interaction between role marking and unmarked word-order preferences.

These findings provide experimental support for asymmetric extensibility from motion to argument marking and help explain the greater prevalence of DOM relative to DSM. By linking predictability-based marking to alignment biases and to the semantics of locatives, the study offers preliminary evidence for why object-marking systems are more likely to arise and stabilize cross-linguistically than subject-marking systems.

1

**References**

Aissen, J. (2003). Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory*, 21(3), 435–483.

Bárány, A. (2018). DOM and dative case. *Glossa: A Journal of General Linguistics*, 3(1), 97.

Bisang, W. (1991). Verb serialisation, grammaticalisation, and attractor positions in Chinese, Hmong, Vietnamese, Thai and Khmer. In H. Seiler & W. Premper (Eds.), *Partizipation: das sprachliche Erfassen von Sachverhalten* (pp. 509–562). Tübingen: Narr.

Bossong, G. (1985). *Differentielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Gunter Narr. doi:10.5281/zenodo.4697660

Comrie, B. (2013). Alignment of Case Marking of Full Noun Phrases. In M. S. Dryer & M. Haspelmath (Eds.), *WALS Online* (v2020.4). Zenodo.

Gibson, E., Futrell, R., Piantadosi, S. T., Dautriche, I., Mahowald, K., Bergen, L., & Levy, R. (2019). How efficiency shapes human language. *Trends in Cognitive Sciences*, 23(5), 389–407. doi:10.1016/j.tics.2019.02.003

Haspelmath, M. (2019). Differential place marking and differential object marking. *Language Typology and Universals*, 72(3), 313–334.

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

Haspelmath, M. (2021b). Role–reference associations and the explanation of argument coding splits. *Linguistics*, 59(1), 123–174.

Holmes, J., & Vihman, V. (2025). Typicality biases in interpreting unmarked sentences: an artificial language learning experiment on differential argument marking. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 47.

Iwasaki, S., & Ingkaphirom, P. (2007). *A reference grammar of Thai*. Cambridge: Cambridge University Press.

Kont, K. (1963). Käändsõnaline objekt läänemeresoome keeltes [Declinable objects in Finnic languages]. In *Keele ja Kirjanduse Instituudi uurimused IX*. Tallinn: Eesti NSV Teaduste Akadeemia.

Levshina, N. (2021). Communicative efficiency and differential case marking: A reverse-engineering approach. *Linguistics Vanguard*, 7(s3).

Longenbaugh, N., & Polinsky, M. (2017). Experimental approaches to ergative languages. In J. Coon, D. Massam, & L. Travis (Eds.), *The Oxford Handbook of Ergativity* (pp. 709–734). Oxford: Oxford University Press.

Motamedi, Y., Smith, K., Schouwstra, M., Culbertson, J., & Kirby, S. (2021). The emergence of systematic argument distinctions in artificial sign languages. *Journal of Language Evolution*, 6(2), 77–98.

Tamm, A. (2004). Relations between Estonian verbs, aspect, and case. PhD dissertation, Eötvös Loránd University, Budapest.

Witzlack-Makarevich, A., & Serzant, I. A. (2018). Differential argument marking: Patterns of variation. In I. A. Serzant & A. Witzlack-Makarevich (Eds.), *Diachrony of differential argument marking* (pp. 1–40). Berlin: Language Science Press.

## Collostructional analysis of the Finnish simile construction using different association measures

*Jenni Ihanus*

My talk will deal with the application of collostructional analysis (e.g. Stefanowitsch & Gries 2003) to the study of the Finnish simile construction [X tekee jotakin kuin Y] 'X does something like Y'. The aim is to identify the verbs that Finnish speakers are most likely to associate with this construction, and collostructional analysis is a promising tool for corpus-based assessment of the entrenchment of the relationship between a verb and the target construction (see e.g. Gries, Hampe & Schönefeld 2005; Stefanowitsch & Flach 2017). However, there are several methodological options for conducting the analysis, and my talk will focus on how the choice of the association measure (AM) affects the results of the analysis (on this topic, see e.g. Wiechmann 2008). The AM is used to obtain the ranking of the verbs.

The AMs I compare are the widely used log-likelihood value, which gives weight to the frequency of co-occurrence of a construction and a verb, and the log odds ratio, which measures association more clearly (Gries 2022). My data consists of about 2,000 instances of the simile construction [X tekee jotakin kuin Y] from a corpus of journalistic texts (Yleisradio 2025). I consider both established ("nukun kuin tukki" 'I sleep like a log') and new similes ("algoritmi toimii kuin anopin kakkuresepti" 'the algorithm works like a mother-in-law's cake recipe') as instances of the target construction.

The results of the collostructional analysis differ considerably when different AMs are used. When the chosen AM is the log-likelihood value, the verbs most strongly associated with the simile construction are verbs that express the behaviour or action of the subject (e.g. "käyttäytyä" 'behave', "toimia" 'act' 'work'). These verbs are also the most frequent verbs in my data. The high frequency is at least partly explained by the fact that the verbs "käyttäytyä" and "toimia" are generic and polysemous in meaning, which makes these verbs suitable for a wide range of purposes. On the other hand, when the chosen AM is the log odds ratio, the verbs most strongly associated with the target construction are quite specific in meaning, rarely used in journalistic texts and often part of established Finnish similes. These verbs include, for example, frequentative verbs such as "putkahdella" ('to repeatedly crop up') and verbs such as "sihistä" 'hiss' and "vapista" 'shake'.

Although both AMs produce interesting results, it is impossible to judge, without other evidence, which one is better at predicting the verbs that language users associate with the simile construction. Presumably, different types of frequency information contribute to the entrenchment of linguistic items, and different factors, such as the salience of the linguistic expression, are involved in this process (e.g. Divjak 2019: §7; Stefanowitsch & Flach 2017). The next step is to use the results of the collostructional analysis of the simile construction in further research, for example in production tasks aimed at language users. This seems to be fruitful if attention is paid to how the chosen AM affects the results of the analysis.

### References

Divjak, D. 2019. Frequency in language. Memory, attention and learning. Cambridge University Press.

Gries, S. T. 2022. What do (some of) our association measures measure (most)? Association? Journal of Second Language Studies 5(1), 1–33.

Gries, S. T., Hampe, B. & Schönefeld, D. 2005. Converging evidence. Bringing together experimental and corpus data on the association of verbs and constructions. Cognitive Linguistics 16(4), 635–676.

Stefanowitsch, A. & Flach, S. 2017. The corpus-based perspective on entrenchment. In Hans-Jörg Schmid (ed.), Entrenchment and the psychology of language learning. How we reorganize and adapt linguistic knowledge, 101–127. De Gruyter.

Stefanowitsch, A. & Gries, S. T. 2003. Collostructions. Investigating the interaction of words and constructions. International Journal of Corpus Linguistics 8(2), 209–243.

Wiechmann, D. 2008. On the computation of collostruction strength. Testing measures of association as expressions of lexical bias. Corpus Linguistics and Linguistic Theory 4(2), 253–290.

Yleisradio 2025. Yle Finnish News Archive 2011-2021, Korp [data set]. Kielipankki. http://urn.fi/urn:nbn:fi:lb-2022032203.

**Developmental Trajectories in Speech-Brain Coherence: Investigating its continuum and correlation to language acquisition**

One of the most fascinating recent discoveries in neurolinguistics is that neuronal oscillations synchronize with external signals like speech. In early child development, this is crucial for better understanding language acquisition milestones, such as the ability to segment continuous speech into smaller linguistic units[1]. Studies on infants show that infant-directed speech amplitude modulations are particularly strong around certain frequency ranges, corresponding to stressed syllables and syllables. This early entrainment suggests that infants are able to neurally track speech, thus supporting early language development. Indeed, the initial literature in this field indicates a link between speech-brain-coherence and early language acquisition[3]. This speech-brain-coherence facilitates temporal alignment of the brain to specific frequencies. Such alignment enables maximizing attention to linguistic features that aid in speech segmentation[4, 5]. This might suggest that infants with greater tracking may have an early advantage for language acquisition. However, the evidence for this linking hypothesis is to date poor and, crucially, hardly ever investigated through a systematic and longitudinal approach.

Herein, we longitudinally assess whether individual neural tracking of speech trajectories relates to individual differences in later language skills. Two key questions are asked: (i) How does neural tracking change in early brain development (6, 9, 12 months of age)? (ii) How does the speech-brain-coherence relate to concurrent language acquisition and to later language growth (measured in the second year of life)? We expect that at different ages, different rates might drive tracking mechanisms. Moreover, those children with better tracking abilities across early life will show greater language growth/skills in later language measures.

We recently completed EEG data collection (N = 129) from monolingual Dutch infants while they listen to child-directed stories at three different timepoints in the first year of life (6, 9, 12 months). Additionally, we collected early CDI language development data in the second year of life (12, 18, 24 months). Speech-brain coherence data will be longitudinally compared between the three sessions and regressed against language skills in their second year of life. All collected data will be analysed in the upcoming months.

**References**

1. Giraud, A. L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature neuroscience, 15*(4), 511-517.
2. Nencheva, M. L., & Lew-Williams, C. (2022). Understanding why infant-directed speech supports learning: A dynamic attention perspective. *Developmental Review*, 66, 101047.
3. Menn, K. H., Ward, E. K., Braukmann, R., Van den Boomen, C., Buitelaar, J., Hunnius, S., & Snijders, T. M. (2022). Neural tracking in infancy predicts language development in children with and without family history of autism. *Neurobiology of Language, 3*(3), 495-514.
4. Meyer, L. (2018). The neural oscillations of speech processing and language comprehension: state of the art and emerging mechanisms. European Journal of Neuroscience, 48(7), 2609-2621.
5. Junge, C., Cutler, A., & Hagoort, P. (2014). Successful Word Recognition by 10-Month-Olds Given Continuous Speech Both at Initial Exposure and Test. Infancy, 19(2), 179-193. https://doi.org/https://doi.org/10.1111/infa.12040

**When does the pro drop? A case study of pro-drop in Macedonian**

In this paper I examine how and when pro-drop occurs in Macedonian (South Slavic). Macedonian is considered a 'pro-drop' language, but thus far there has been no empirical study on the conditions under which subject expression does or does not occur. Macedonian has case-marked pronouns and verbal endings marked for person and number. The pronouns are considered 'optional', at least in a semantic sense. This is illustrated in example 1.

1) (Jas) oda-m doma.
1SG.NOM go.IPFV-PRS.1SG home
'I am going home.'

To see how often 'pro-drop' actually occurs, I conducted a corpus study of the Macedonian narrative corpus of the SpeechReporting database (Jordanoska 2023). The total length comprises about 2 hours and it contains about 15.000 words. To make the data comparable to the ones used by Piivik et al., (submitted), the dataset was restricted to verbs with a first or second person ending. This resulted in 731 verbs, the precise distribution of person and number of which is shown in Figure 1.

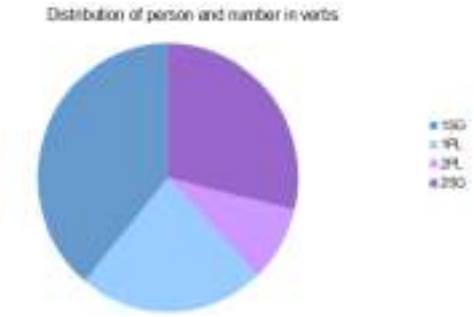



*Figure 1: Distribution of verbal endings*

As shown in Table 1, of those verbs, only 139 (19%) had an overt subject pronoun, whereas in 592 verb clauses (81%) the pronoun was 'dropped'.

| | |
|---|---|
| pronoun | 139 (19%) |
| no pronoun | 592 (81%) |
| total | 731 (100%) |

Table 1: Percentage of overt pronoun vs dropped pronoun

Thus, the 'dropped' subject is the default, and the real question is: When is the subject pronoun realized? A usual suspect (as suggested in Hyams 1989, Friedman 1993, Gundel et al. 1993, and Mitkovska, & Bužarovska 2018) for subject expression in Macedonian is information structure. Friedman (1993:291) suggests that subject pronouns are realized when the subject is

either a focus or a shifted topic. However, when we look at the discourse conditions under which subjects are expressed in the corpus, we see that there has to be more to this story than information structure. Consider example 2.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 2) | šo | ќe | me | jadi-š? | gleda-š | deka | **jas** | | |
| | why | FUT | 1SG.ACC | eat-PRS.2SG | see-PRS.2G | that | 1SG.NOM | | |
| | dojdo-v | | tebe | da | te | kana-m | | za | |
| | come-AOR.1SG | | 2SG.ACC | to | 2SG.ACC | invite-PRS.1SG | | for | |
| | kum | | | | | | | | |
| | godfather | | | | | | | | |

‘Why would you eat me? Can’t you see that I have come to invite you to be a godfather?’ (Macedonian_volkot_kum.30-1)

In (2) the pronoun *jas* ‘I’ is neither new or contrastive information (focus), nor is it a shifted topic, as the referent is already given by *me* ‘me’ in the preceding sentence.

Thus, we have to look at other factors as well. Following the coding scheme proposed by Pilvik et al. (submitted), the Macedonian data in this paper has been coded for clause type, valency of the verb, and tense of the verb. The preliminary results seem to indicate that clause type and valency play a role in the expression of subject pronouns, whereas tense does not.

First, let’s look at clause type. While 246 verbs (33,6%) occurred in subordinate clauses and 485 (66,3%) in main ones across the board, in the verbs with overt pronouns the distribution was more skewed towards main clauses: of the 139 overt pronouns, 111 (79,8%) occurred in main clauses, whereas only 28 (20,1%) occurred in subordinate ones. Across the board there were more transitive verbs (466 or 63,7%), than intransitive verbs (265 or 36,3%). However, when we look at the verbs with overt subject pronouns, we see that 63 (45,3%) of the overt pronouns occurred with transitive verbs, and 76 (54,7%) with intransitive ones. Finally, let’s look at tense. Across the board, 108 (14,8%) verbs were in the past tense versus 623 (85,2%) in the present tense. Of the verbs with a pronoun, 27 (19,4%) were in the past tense, 112 (80,6%) in the present. This seems roughly the same distribution as the distribution of verbs across the board.

Compared to the findings in Pilvik et al. (submitted) and Wilbur (this conference), we first of all see that the expression of subject pronouns varies greatly across languages and genres. Pilvik et al. (submitted) found 72% overt subject pronouns for Estonian, 59% for Russian, 11% for Finnish and 5% for Polish. Wilbur’s (2014) narrative corpus of Pite Saami contained 38% overt pronouns. Pilvik et al. suggest an influence of genre rather than language family, thus a direction for future research would be to compare the data in this paper to data from a Macedonian subtitle corpus. The data suggests that clause type plays a role in Macedonian subject expression, something that Piivik et al found for Estonian and Polish as well. Tense, on the other hand, which does not seem to play a role in Macedonian subject expression, does play a role in Estonian and Russian. Thus, we can conclude from this preliminary data that different factors play a role in the expression of subject pronouns.

**References**

Friedman, V. (1993) *Macedonian*. In: Comrie B and Corbett GG (eds) The Slavonic languages (pp. 249–305). London/New York: Routledge.
Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 274-307.
Hyams, N. (1988) A principles and parameters approach to the study of child language. *Papers and Reports on Child Language Development 27*: 153–61
Jordanoska, I. (2023). A narrative corpus of Macedonian. In Nikitina, Tatiana, Ekaterina Aplonova, Izabela Jordanoska, Ekaterina Biteeva, Abbie Hantgan-Sonko, Guillaume Guitang, Olga Kuznetsova, Rebecca Paterson, Elena Perekhvalskaya & Lacina Silué (eds.) The SpeechReporting Corpus: Discourse Reporting in Storytelling. CNRS-LLACAN & LACITO, http://discoursereporting.huma-num.fr/index.html
"When less is not more: A comparative corpus study of optional subject pronoun expression in Finnic and Slavonic".
Wilbur, J. (2014). *A grammar of Pite Saami* (p. 295). Language Science Press.

**Distribution of global and local cues in Russian clauses with SVO, OVS and S-drop constructions**

In natural communication, speakers balance between communicative success and production ease, which the message encoding reflects (Hörberg & Sjons, 2023). Packaging information asymmetrically (e.g., *given* vs. *new,* often associated with *topic* and *focus*, Laleko, 2022) is a strategy so as using cues helping the addressee build expectations about the sentence structure and its elements' roles. In event descriptions, global discourse (e.g., noun phrase (NP) givenness) and local (e.g., animacy) cues help assign event participants' roles, encoded grammatically in subjects and objects (Foley, 2007), but the cues' strengths differ. The usage-based account predicts that the differences can be inferred from the cues' distributions in a corpus. This allowed Hörberg (2018) to explore distributional differences of NP properties in Swedish subject-verb-object (SVO), object-verb-subject (OVS), and passive clauses. Replicating the finding that fronted objects are often discourse prominent, topical, and given, he also showed that they signal topic shifts and contrasts. Such *new*, informationally 'heavy' NPs tend to go with informationally 'light' verbs and prominent subjects, showing that speakers prefer to keep the discourse information flow steady (Hörberg, 2018; Fenk-Oczlon, 2001). Does the pattern hold for other languages, e.g., Russian, whose word order is highly pragmatically motivated (Laleko, 2022)?

In Russian OVS clauses, objects are most often given/topical and subjects are new. Fronted *new* objects have also been found, but their function is debated (Slioussar & Makarchuk, 2022). Other factors (e.g., verb semantics) need more attention. Exploring several factors behind the SOV use in a web corpus, Slioussar and Makarchuk (2022) found that pronominal, not lexical, objects were fronted significantly more often. Objects' givenness (accessibility, assessed via NP modifiers) and focus type made SOV significantly more probable, but only objects' givenness was analysed, and other patterns require examining. To date, there is no systematic research on distributional differences of factors (cues) for particular Russian word order patterns (Slioussar & Makarchuk, 2022) and other relevant constructions, e.g., dropped subjects (S-drop) in transitive clauses require to be highly retrievable (i.e., given/topical) (Bizzarri, 2015), but to what extent does object's givenness/prominence play a role?

In this talk, I present the results of an ongoing study on the distribution of discourse (givenness, pronominalisation, person, definiteness, focus particles) and local (animacy, case, verb semantics, aspect, tense) cues in Russian SVO, OVS, and S-drop clauses (e.g., *Меня устраивал график*, I.ACC approve-of.PST.M.SG schedule.NOM.M.SG and *Илью Ильича обидели*, Ilya-Ilyich.ACC.SG offend.PST.PL). Givenness is assessed within the paradigm combining Prince (1981), Ariel (1990), and Gundel et al. (1993) (Hörberg, 2018). The balanced corpus 'SynTagRus' of over 1,000,000 tokens (Droganova et al., 2018) is used. I expect to find that objects are given/topical/prominent, and subjects new/focused to a higher degree in OVS than SVO clauses, but fronted *new* objects, used for topic shifts or contrasts, with prominent post-verbal subjects might also be frequent. Dropped subjects are likely to be given/topical/prominent and combine with new/focused objects to a higher degree than non-dropped ones. Such findings would be in line with the more general crosslinguistic pattern that fronted objects are more often given than new but also with the discourse information flow balance.

References

Ariel, M. (1990). *Accessing Noun-Phrase Antecedents*. London: Routledge.

Bizzarri, C. (2015). Russian as a partial pro-drop language. *Annali di Ca'Foscari. Serie occidentale*, *49*, 335–362.

Droganova, K., Lyashevskaya, O., & Zeman, D. (2018). Data Conversion and Consistency of Monolingual Corpora: Russian UD Treebanks. In *Proceedings of the 17th International Workshop on Treebanks and Linguistic Theories (TLT 2018)*, December 13–14, 2018, Oslo University, Norway (No. 155, pp. 52-65). Linköping University Electronic Press.

Fenk-Oczlon, G. (2001). Familiarity, information flow, and linguistic form. In Joan Bybee & Paul J. Hopper (eds.), *Frequency and the emergence of linguistic structure*, 431–448. Amsterdam: John Benjamins.

Foley, W. A. (2007). A typology of information packaging in the clause. In T. Shopen (Ed.), *Language Typology and Syntactic Description: Volume 1: Clause Structure*. 2nd ed., Vol. 1, 362–446. Cambridge University Press. https://doi.org/10.1017/CBO9780511619427.007

Gundel, J.K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language 69*(2). 274–307. https://doi.org/10.2307/416535

Hörberg, T. (2018). Functional motivations behind direct object fronting in written Swedish: A corpus-distributional account. *Glossa: A Journal of General Linguistics*, *3*(1), Article 1. https://doi.org/10.5334/gjgl.502

Hörberg, T., & Sjons, J. (2023). Speakers balance their use of cues to grammatical functions in informative discourse contexts. *Language, Cognition and Neuroscience*, *38*(2), 175–196. https://doi.org/10.1080/23273798.2022.2102667

Laleko, O. (2022). Word order and information structure in heritage and L2 Russian: Focus and unaccusativity effects in subject inversion. *International Journal of Bilingualism*, *26*(6), 749–766. https://doi.org/10.1177/13670069211063674

Prince, E. F. (1981). Toward a Taxonomy of Given-New Information. In P. Cole (Ed.), *Radical Pragmatics*, 223–255. New York: Academic Press.

Slioussar, N., & Makarchuk, I. (2022). SOV in Russian: A corpus study. *Journal of Slavic Linguistics*, *30*, 1–14. https://doi.org/10.1353/jsl.2022.a923076

MultiplEYE: Standardizing Multilingual Eye-Tracking for Global Reading Research

The MultiplEYE project is an international research initiative designed to create a standardized eye-tracking dataset to advance reading research across multiple languages. A central goal of MultiplEYE is to assemble an open-access, multilingual dataset that enables cross-linguistic comparisons by addressing traditional research limitations in eye-movement studies. This presentation will outline the project's progress and achievements in developing a unified experimental procedure and stimuli corpus. This includes the development of reading materials, formulation of comprehension questions, and selection of supplementary cognitive tests. The network adheres to rigorous data collection standards, meta-data documentation, and the creation of a custom, open-source preprocessing pipeline.

MultiplEYE distinguishes itself by balancing consistency and diversity. Standardization across labs ensures that cross-linguistic comparisons are valid. At the same time, diversity in languages, and participant backgrounds increases the generalizability of the findings. The project covers a broad linguistic spectrum, including languages from the Indo-European, Uralic, Turkic, Sino-Tibetan, Semitic, and Eskaleut families, thus filling a critical gap in multilingual eye-tracking research.

Ultimately, the project not only aims to provide a valuable resource for reading research but also to establish best practices for open research data in the field of eye-tracking. By standardizing data collection, documentation, and sharing processes, MultiplEYE enhances reproducibility and reusability of eye-movement data, contributing to the advancement of reading research.

The MultiplEYE network actively welcomes more labs to join this collaborative international effort, thus fostering global cooperation in reading research across languages and cultures.

# Genderless, Yet Not Neutral: Investigating Gender Bias with a Mixed-Method Approach in Estonian

Elisabeth Kaukonen

Keywords: *ECR, gender, Estonian, corpus linguistics, experimental linguistics*

Feminist linguistics posits that language both reflects and reproduces gender-related social practices, including gender stereotypes (e.g., Cameron 1992; Eckert & McConnell-Ginet 1995; Talbot 1998; Heilman 2001; Mills 2008). While many studies have explored this link between language, stereotypes and the biases these stereotypes may evoke (Oakhill et al. 2005; Gygax et al. 2008; Garnham & Yakovlev 2015; Molinaro et al. 2016), they have largely focused on grammatical or natural gender languages, leaving grammatically genderless languages underexplored.

This presentation addresses this research gap, by investigating gender stereotypes and bias in grammatically genderless languages, using Estonian as a case study and presenting the key findings of my doctoral dissertation. The study employs two different methodological approaches, combining corpus linguistics with a quasi-experimental study. First, data from Estonian sports news corpora from 2020 and web subcorpora of the Estonian National Corpus 2023 were analyzed, utilizing corpus linguistics and corpus-assisted discourse studies (Partington et al. 2013; Stefanowitsch 2020). The findings reveal how Estonian gender-marked vocabulary conveys societal roles associated with men and women. Second, a Likert-scale survey was conducted, to investigate gender perception of Estonian occupational titles. The survey results, based on the responses from 581 participants, complement the findings from the corpus study and illustrate how gender-marked language influences the gender perception of native speakers. This mixed-method approach highlights how gender stereotypes manifest in a grammatically genderless language, challenging the assumption that such languages are more gender-neutral (Prewitt-Freilino 2012). Furthermore, the presentation will underscore material for future research by outlining potential experiment designs, such as the sentence evaluation task (see Oakhill et al. 2005; Gygax et al. 2008; Hammond-Thrasher & Järvikivi 2023), that can provide deeper insights into the perception of gender.

Ultimately, this research underscores that grammatically genderless languages, like Estonian, can exhibit gender bias comparable to gendered languages. By addressing these issues, this research contributes to a broader understanding of how gender operates across different languages and cultures.

**References**

Cameron, Deborah 1992. *Feminism and Linguistic Theory*. London: MacMillan Press Ltd. https://doi.org/10.1007/978-1-349-22334-3

Eckert, Penelope and Sally McConnell-Ginet. 2003. *Language and Gender.* Cambridge: Cambridge University Press

Garnham, A., and Yuri Yakovlev. 2015. The Interaction of Morphological and Stereotypical Gender Information in Russian. *Frontiers in Psychology* 16. https://doi.org/10.3389/fpsyg.2015.01720

Gygax, Pascal M., Ute Gabriel, Oriane Sarrasin, Alan Garnham and Jane Oakhill. 2008. Generically intended, but specifically interpreted: When beauticians, musicians, and mechanics are all men. *Language and Cognitive Progresses* 23(3). 464–485. https://doi.org/10.1080/01690960701702035

Hammond-Thrasher, Stephanie and Juhani Järvikivi. 2023. Rating gender stereotype violations: The effects of personality and politics. *Frontiers in Communication* 8. doi.org/10.3389/fcomm.2023.1050662

Heilman, Madeline E. 2001. Description and Prescription: How Gender Stereotypes Prevent Women's Ascent Up the Organizational Ladder. *Journal of Social Issues* 57(4), 657–674. https://doi.org/10.1111/0022-4537.00234

Mills, Sara. 2008. *Language and Sexism*. Cambridge: Cambridge University Press

Molinaro, Nicola, Jui-Ju Su and Manuel Carreiras. 2016. Stereotypes override grammar: Social knowledge in sentence comprehension. *Brain and Language* 155–156, 36–43. https://doi.org/10.1016/j.bandl.2016.03.002

Partington, Alan, Alison Duguid and Charlotte Taylor. 2013. *Patterns and Meanings in Discourse. Theory and Practice in Corpus Assisted Discourse Studies* (CADS) (Studies in Corpus Linguistics). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Prewitt-Freilino, Jennifer L., Andrew T. Caswell, and Emmi K. Laakso. 2012. The Gendering of Language: A Comparison of Gender Equality in Countries with Gendered, Natural Gender, and Genderless Languages. *Sex Roles* 66(3–4), 268–281. https://doi.org/10.1007/s11199-011-0083-5

Stefanowitsch, Anatol. 2020. *Corpus Linguistics. A guide to the methodology*. Berlin: Language Science Press.

Talbot, Mary. 1998. *Language and Gender. An Introduction*. Malden: Blackwell Publishers Inc.

## Investigating shared syntax in Spanish-English bilinguals and bilingual cognitive models: Code-switching increases cross-language structural priming

After hearing a grammatical structure in one language, bilinguals become more likely to produce that structure in their other language. This cross-language structural priming phenomenon is commonly interpreted as evidence for shared syntax in bilinguals [1,2]. However, bilinguals mix languages in ways that do not always follow the script of typical cross-language structural priming experiments [3]. They not only switch languages between, but also within sentences. We hypothesize that code-switching in the prime sentence increases implicit learning of shared syntax, leading to stronger cross-language priming compared to single-language primes. We test for this effect in an implicit learning model of priming [4] and in Spanish-English bilinguals from the US.

We conducted four simulated structural priming experiments, using instances of the Spanish-English Bilingual Dual-path model [5], which was previously used to simulate both code-switching [5] and cross-language structural priming [6]. Following [6], we trained model instances on artificial versions of Spanish and English, to be used as simulated participants. In a structural priming experiment, these simulated participants were presented with Spanish active or passive primes before producing English transitives. The primes either had an English (code-switched) determiner and noun (Examples a,b) or noun only (c,d), or were entirely in Spanish (e). Code-switches were at the beginning (a,c) or the end of sentences (b,d).

Examples
(a) “**the boy** empuja el juguete”
(b) “el niño empuja **the toy**”
(c) “el **boy** empuja el juguete”
(d) “el niño empuja el **toy**”
(e) “el niño empuja el juguete”

Only with a code-switched determiner and noun at the beginning (a), structural priming was increased compared to entirely Spanish primes (visualized by the graph on the left in Figure 1) as evidenced by a significant positive interaction between code-switch condition and priming (Est. = 0.18, p = < 0.001) in our mixed effects analysis. We therefore used code-switched primes of type (a) in our follow-up experiment with human participants.

We tested for the interaction between code-switching and prime structure in 190 Spanish-English bilinguals in the US. In a pre-registered online experiment (https://aspredicted.org/WQG_Z9S), participants wrote 60 English picture descriptions after hearing active or passive primes that were entirely Spanish or code-switched. As predicted by the model, the experiment revealed that structural priming in participants was stronger after code-switched compared to entirely Spanish primes (visualized by the graph on the right in Figure 1) as evidenced by a significant positive interaction between code-switch condition and priming (Est. = 0.15, p = 0.002) in the mixed effects analysis. Together, these results suggest that processing code-switches in a prime sentence can result in increased prediction error, which in turn can lead to increased implicit learning of shared syntactic representations, resulting in stronger cross-language structural priming. These results also demonstrate that the Bilingual Dual-path model can be used to predict novel psycholinguistic effects in human participants.

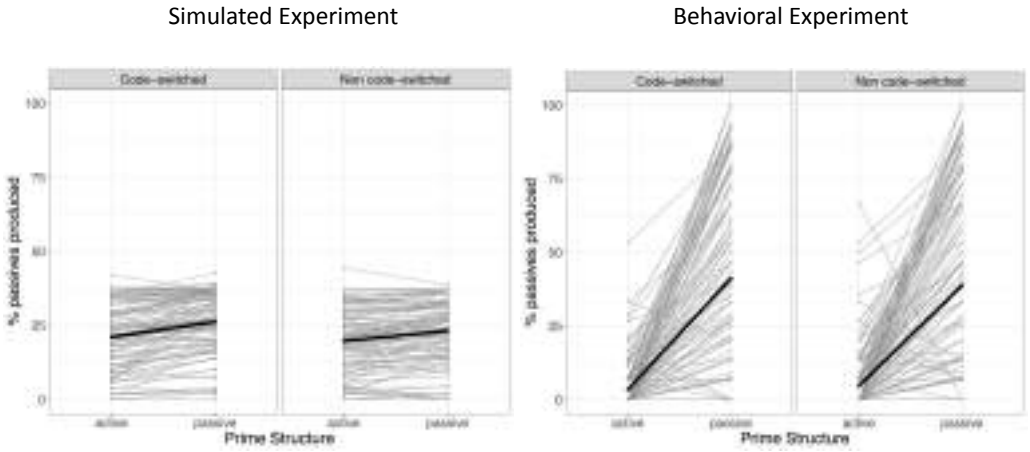



**Fig. 1** Results from the Simulated Experiment, graph on the left, and Behavioral Experiment, graph on the right, with code-switched determiner and noun at beginning of prime sentence, in the left panels, see Example (a), compared to non-code-switched primes, in the right panels. The thick black lines visualize the priming effect across all analyzed trials by connecting the percentage of passive responses after active primes to the percentage of passive responses after passive primes. The thin grey lines show the same for each individual (simulated) participant.

**References**

[1] Hartsuiker, R. J., Pickering, M. J., & Veltkamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English bilinguals. *Psychological science*, *15*(6), 409-414.
[2] Van Gompel, R. P., & Arai, M. (2018). Structural priming in bilinguals. *Bilingualism: Language and Cognition*, *21*(3), 448-455.
[3] Kootstra, G. J., & Rossi, E. (2017). Moving beyond the priming of single-language sentences: A proposal for a comprehensive model to account for linguistic representation in bilinguals. *Behavioral and Brain Sciences*, *40*.
[4] Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological review*, *113*(2), 234.
[5] Tsoukala, C., Broersma, M., Van den Bosch, A., & Frank, S. L. (2021). Simulating code-switching using a neural network model of bilingual sentence production. *Computational Brain & Behavior*, *4*(1), 87-100.
[6] Khoe, Y. H., Tsoukala, C., Kootstra, G. J., & Frank, S. L. (2023). Is structural priming between different languages a learning effect? Modelling priming as error-driven implicit learning. *Language, Cognition and Neuroscience*, 1-21.
[7] Kootstra, G. J., Van Hell, J. G., & Dijkstra, T. (2010). Syntactic alignment and shared word order in code-switched sentence production: Evidence from bilingual monologue and dialogue. *Journal of Memory and Language*, *63*(2), 210-231.
[8] Goldrick, M., Putnam, M., & Schwarz, L. (2016). Coactivation in bilingual grammars: A computational account of code mixing. *Bilingualism: Language and Cognition*, *19*(5), 857-876.

## Gender Differences in Spoken Estonian: Exploring Politeness Markers

*Sofia Kriuchkova*

As social norms often differ by gender, it is expected that men and women perceive politeness differently. Observations suggest that women tend to be more polite, as femininity is traditionally associated with softness and compliance. Women also often make more effort to participate in conversations. (Fishman, 1983) Key studies of the differences in politeness strategies between men and women demonstrate that, compared to men, women are more likely to use pragmatic particles, negative-politeness strategies, the subjunctive mood, a higher number of compliments and apologies, and facilitative tag questions (e.g. Berryman-Fink & Wilcox, 1980; Edelsky, 1976). Later quantitative studies in English and other languages confirm these findings (e.g. Gauthier, 2017; Senkina et al., 2017; Siregar et al., 2023). In Estonian linguistics, attention so far has focused primarily on the study of politeness markers and their use in various contexts (e.g. Haavel, 2019; Lindström, 2009; Pajusalu et al., 2010). However, it has not yet been studied how gender norms influence the selection of politeness markers. The aim of the planned study is to identify and describe gender differences in the use of politeness markers in the spoken language of native Estonian speakers, as well as to compare the findings with the results established for other languages. Along with gender, age will be taken into account as an additional factor. Special attention will be paid to teenage speech, as it is assumed that gender differences are less pronounced in this group than among adult speakers. The data for the study will be drawn from multiple corpora, including teenage speech and adults' everyday dialogues. The methodology will involve corpus analysis with automated text processing in Python and statistical evaluation. The poster will present the research design, methodological approach, and expected outcomes. Although the analysis has not yet been completed, the project aims to contribute to Estonian linguistics and provide a basis for further research in variational linguistics.

**References:**

Berryman, C. L. & Wilcox. J. R. 1980. Attitudes toward male and female speech: Experiments on the effects of sex-typical language. Western Journal of Speech Communication 44(1). 50–59.

Edelsky, C. 1976. Subjective reactions to sex-linked language. The Journal of Social Psychology 99(1). 97–104.

Fishman, P. M. "Interaction: The work women do." Feminist research methods. Routledge, 2019. 224-237.

Gauthier, M. 2017. Age, gender, fuck, and Twitter: A sociolinguistic analysis of swear words in a corpus of British tweets. PhD diss., Université Lumière Lyon 2.

Haavel, Aive., 2019. Ametiasutuse kliendikirjade viisakus ja tundetoon. Tallinn: Tallinna Ülikool, Humanitaarteaduste instituut. Magistritöö

Lindström, L., 2009. Kõnelejale ja kuulajale viitamise vältimise strateegiaid eesti keeles. Emakeele Seltsi aastaraamat, (55), 88-118.

Pajusalu, R., Vihman V., Klaas, B., & Pajusalu K., 2010. Eestlaste ja venelaste suhtluskäitumine: sina, teie ja keegi veel. Eesti Rakenduslingvistika Ühingu aastaraamat, 6, 207-224.

Senkina, D. 2018. How Mothers and Fathers Address their Sons and Daughters in Russian and Vice Versa: A Quantitative Study. Higher School of Economics Research Paper No. WP BRP, 69.

Siregar, L. S., Manurung, L. W., Silitonga, H. H., & Naibaho, T. 2023. Refusal Speech Act Used By Male And Female Sellers At Pasar Sambu Medan. Innovative: Journal Of Social Science Research, 3(2), 4648-4657.

## Creating The Penitentiary Document Corpus Treebank for Late Medieval Latin

Morpho-syntactic analysis forms the basis of many linguistics research topics be they the analysis of single linguistics features or the study of entire texts. The aims of this study are twofold: Firstly, this study describes the creation of a gold standard annotated treebank of the late medieval penitentiary documents. The penitentiary handled the petitions of the members of the Catholic church regarding special licences, dispensations and absolutions. (Salonen & Schmugge, 2009) The documents from four regions, i.e. Medieval Sweden, Norway, England and Wales have been published as modern editions. These editions have been turned into a structured machine readable xml-database and published as the PeDoCo corpus in the Fin-Clarin database (Kupari et al. 2024, forthcoming). From this corpus a set of 1,200 tokens was taken to serve as a sufficient sample to explore the performance of different state of the art models available (Kupari et al. 2024). We describe in maticulous detail the process of creating this treebank sample set especially in regards to the syntactic level. Secondly, we expand the annotation task to the much larger above mentioned 200,000 token PeDoCo dataset using the CoNLL-U format for quantitative analysis..

The PeDoCo texts present several specific challenges for morpho-syntactic parsing and NLP tasks more broadly. First and foremost, it is essential to recognize that the original texts contain expanded abbreviations and editorial restorations, which can complicate automatic processing. Additionally, the pronounced formulaicity of these texts further affects their linguistic structure. See Illustration 1 for an example with grammatical analysis of such recurring formulaic expressions in the corpus, highlighting the difficulty of the qualitative grammatical analysis. The corpus is characterized by a high degree of ungrammaticality, encompassing a full spectrum of linguistic variation, from conventional usage to severe deviations. This includes scribal errors, typographical mistakes, confusion arising from the use of formulaic language, and exceptionally long sentence structures. These features do not reflect colloquial Latin but rather represent genuine human errors in the transmission of text.

The PeDoCo corpus thus represents a highly specific genre: a dataset that includes mistakes and nonsensical elements as an inherent feature. This stands in stark contrast to Classical Latin textbook texts, which have been emended over centuries. Instead, the PeDoCo corpus captures the linguistic reality of the papal penitentiary, offering insight into the challenges and inconsistencies that characterized its documentary practices. The careful description of these phenomena sheds light into the difficulties of using the Universal Dependencies framework and the need for discussions on best practices.

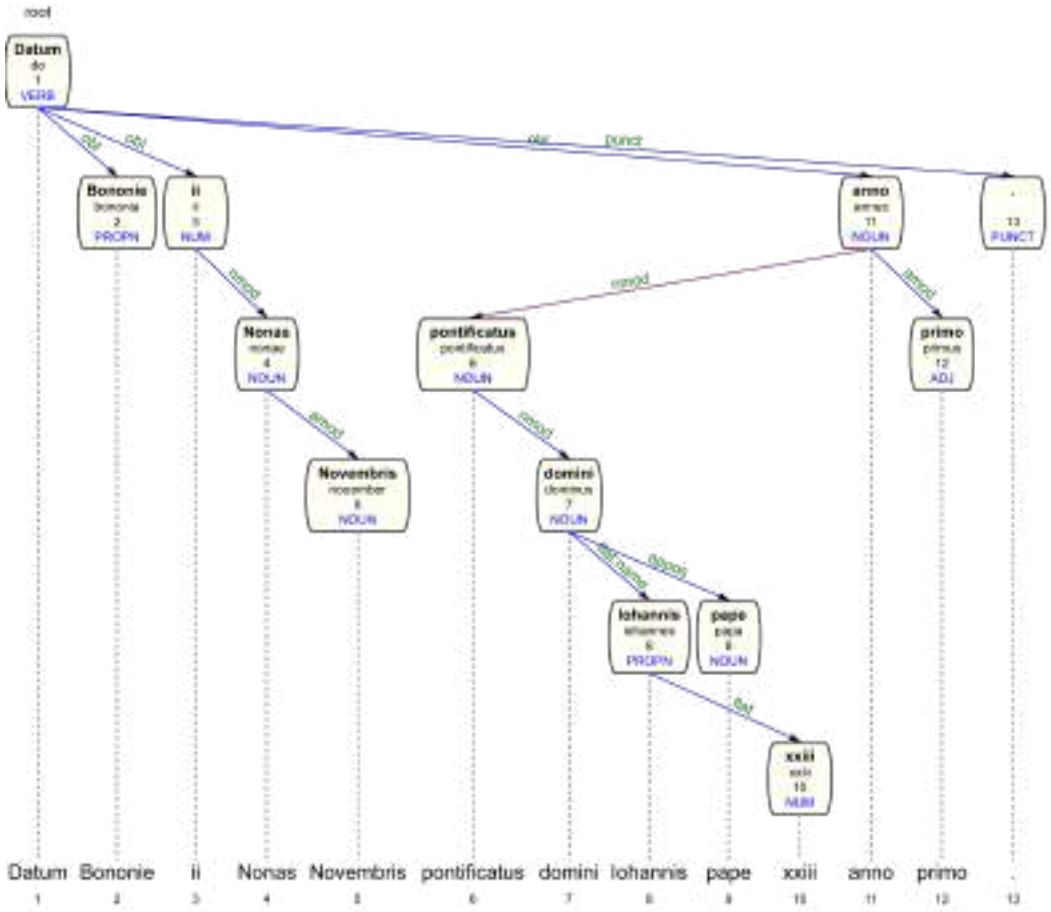



Illustration 1. Visualization of a Gold Standard annotated tree (*AP* doc. 3, 4.11.1410, Bologna, case of Petrus Holmstani) https://github.com/Orange-OpenSource/conllueditor

| | FORM | LEMMA | UPOS | XPOS | FEATS | HEAD | DEPREL | DEPS | MISC |
|---|---|---|---|---|---|---|---|---|---|
| 1 | Datum | do | VERB | _ | _\|Aspect=Perf\|Degree=Pos\|Neu | 0 | root | _ | _ |
| 2 | Bononie | bononia | PROPN | _ | _\|Fem=\|Loc=\|Sing= | 1 | obl | _ | _ |
| 3 | ii | ii | NUM | _ | _= | 1 | obl | _ | _ |
| 4 | Nonas | nonae | NOUN | _ | _\|Acc=\|Fem=\|Plur= | 2 | nmod | _ | _ |
| 5 | Novembris | november | NOUN | _ | _\|Gen=\|Masc=\|Sing= | 4 | amod | _ | _ |
| 6 | pontificatus | pontificatus | NOUN | _ | _\|Gen=\|Masc=\|Sing= | 11 | nmod | _ | _ |
| 7 | domini | dominus | NOUN | _ | _\|Gen=\|Masc=\|Sing= | 6 | nmod | _ | _ |
| 8 | Iohannis | iohannes | PROPN | _ | _\|Gen=\|Masc=\|Sing= | 7 | flat:name | _ | _ |
| 9 | pape | papa | NOUN | _ | _\|Gen=\|Masc=\|Sing= | 7 | appos | _ | _ |
| 10 | xxiii | xxiii | NUM | _ | _= | 8 | flat | _ | _ |
| 11 | anno | annus | NOUN | _ | _\|Abl=\|Masc=\|Sing= | 1 | obl | _ | _ |
| 12 | primo | primus | ADJ | _ | _\|Abl=\|Degree=Pos\|Masc=\|Sing= | 11 | amod | _ | _ |
| 13 | . | . | PUNCT | _ | _= | 1 | punct | _ | _ |

Illustration 2. A visualization of a sample sentence in table format using the CoNLL-U editor. The treebank uses the Universal Dependencies formalism.

## References

Clarke, Peter D., and Patrick N. R. Zutshi. 2013. "*Supplications from England and Wales in the registers of the Apostolic Penitentiary, 1410-1503. Volumes I-III"*, 1410-1464 Canterbury and York Society, Vol. CIII. Woodbridge, Suffolk: Canterbury / York Society.

Hanna-Mari Kristiina Kupari, Erik Henriksson, Veronika Laippala, and Jenna Kanerva. 2024. Improving Latin Dependency Parsing by Combining Treebanks and Predictions. In *Proceedings of the 4th International Conference on Natural Language Processing for Digital Humanities*, pages 216–228, Miami, USA. Association for Computational Linguistics.

Jørgensen, Torstein, and Gastone Saletnich. 2004. "*Synder og pavemakt: botsbrev fra Den Norske Kirkeprovins og Suderøyene til Pavestolen, 1438-1531"*. Diplomatarium Poenitentiariae Norvegicum. Misjonshøgskolens forlag.

Risberg, Sara, and Kirsi Salonen. 2008. *"Appendix: Acta Pontificum Suecica. 2, Acta Poenitentiare: Auctoritate papae: the church province of Uppsala and the apostolic penitentiary 1410-1526."* Diplomatarium Suecanum. Stockholm: Riksarkivet.

Salonen, Kirsi, and Ludwig Schmugge. *A Sip from the "Well of Grace": Medieval Texts from the Apostolic Penitentiary*. Catholic University of America Press, 2009. *JSTOR*.

## Variation in the Case Forms of Indefinite Pronouns: A Multi-Method Approach Combining Corpus Studies and Eye-Tracking

*Annika Kängsepp*

Variation in the case forms of the Estonian indefinite pronouns keegi ‘someone’, miski ‘something’, kumbki ‘either’, and ükski ‘none’ has been documented for over a century, both in written language (Rull 1917) and dialects (Saareste 1955). This variation concerns the placement of the -gi/-ki, which can occur after the case ending (e.g., kelle-le=gi someone-ALL=CLIT ‘to someone’), before the case ending (e.g., kellegi-le someone.INDF-ALL), between two case endings (e.g., kelle-le-gi-le someone-ALL-INDF-ALL), or both before and after the case ending (e.g., kellegi-le=gi someone.INDF-ALL=CLIT) (Saareste 1923, 1936). The variation has a strong dialectal background: forms where -gi/-ki is placed after the case ending have historically been prevalent only in Southern and Northeast Estonia (Saareste 1955). Despite this, in standard Estonian, the normative placement of -gi/-ki occurs only after the case ending.

Given that this variation has been systematically studied only in written Estonian (Pant 2018, 2020), this presentation explores the variation using written and spoken data, with a combination of corpus studies and eye-tracking. The aim of corpus studies was to provide an overview of the extent of the variation and describe the factors influencing the choice between the normative and non-normative forms. Since written language limits the factors that can be examined, studying spoken language was necessary to understand the impact of sociolinguistic factors and prosody. While the corpus studies provide insights into the production of forms, eye-tracking offers a complementary view into real-time processing of varying forms by tracking eye movements such as fixations and regressions during reading (e.g., Holmqvist et al., 2011). Taking a broader perspective, the eye-tracking experiment explores how normativity and frequency affect the reading of morphologically varied forms.

The corpus study of written language utilized data from the Estonian National Corpus 2021 (2.4 billion words; Koppel & Kallas 2022). For the corpus study of spoken language, data was drawn from Estonian Public Broadcasting’s Radio Corpus (109 million words; Lippus et al., 2023a) and the Estonian Podcast Corpus (85 million words; Lippus et al., 2023b). The experimental stimuli were based on natural examples from the written corpora and further manipulated to obtain comparable context length and structure.

The results from the corpus studies indicate that in written Estonian, -gi/-ki is predominantly placed after the case ending (78.6%), whereas in spoken Estonian, it occurs at almost equal frequency after (54.2%) and before (43.3%) the case ending. In written language genre, and whether the text had been edited, was the strongest factor influencing variation. In spoken language, speech rate emerged as the most influential factor, with faster speech favoring the placement of -gi/-ki before or between two case endings. (Kängsepp 2024, 2025) Regarding processing, the hypothesis for the ongoing eye-tracking study posits that there are significant differences between how normative and non-normative forms are read, especially regarding late measures (see e.g., Godfroid 2020).

The presentation exemplifies how combining methods provides deeper insight into language variation and offers a more comprehensive understanding of both the production and processing of morphologically complex forms.

This work was supported by the National Program: Estonian Language and Culture in the Digital Age project “Morphosyntactic variation in Estonian” (EKKD-TA2).

## References

Godfroid, Aline. 2020. Eye Tracking in Second Language Acquisition and Bilingualism. A Research Synthesis and Methodological Guide. New York: Routledge.

Holmqvist, Kenneth, Marcus Nyström, Richard Andersson, Richard Dewhurst, Halszka Jarodzka & Joost van de Weijer. 2011. Eye tracking: A comprehensive guide to methods and measures. Oxford University Press.

Koppel, Kristina & Jelena Kallas. 2022. Eesti keele ühendkorpuste sari 2013–2021: mahukaim eesti keele kogu. [Estonian National Corpus 2013–2021: The largest collection of Estonian language data]. Eesti Rakenduslingvistika Ühingu aastaraamat = Estonian papers in applied linguistics 18. 207–228. https://doi.org/10.5128/ERYa18.12.

Kängsepp, Annika. 2024. Käändevormide varieerumine ja seda mõjutavad tegurid kirjalikus keeles indefiniitpronoomeni keegi näitel. [The variation in the case forms of the indefinite pronoun keegi 'someone' in written Estonian]. Keel ja Kirjandus 11. 1016–1037. https://doi.org/10.54013/kk803a3.

Kängsepp, Annika. 2025. Indefiniitpronoomenite keegi ja miski käändevormide varieerumine suulises keeles. [Variation in the case forms of the indefinite pronouns keegi 'someone' and miski 'something' in spoken Estonian]. Eesti Rakenduslingvistika Ühingu aastaraamat = Estonian papers in applied linguistics 21. 121–139. https://doi.org/10.5128/ERYa21.07.

Lippus, Pärtel, Tanel Alumäe, Siim Orasmaa, Katrin Tsepelina & Liina Lindström. 2023a. Eesti Rahvusringhäälingu raadiosaadete korpus. [Estonian Public Broadcasting's Radio Corpus]. https://doi.org/10.23673/re-441.

Lippus, Pärtel, Tanel Alumäe, Siim Orasmaa, Maarja-Liisa Pilvik & Liina Lindström. 2023b. Eesti taskuhäälingukorpus. [Estonian Podcast Corpus]. https://doi.org/10.23673/re-445.

Pant, Annika. 2018. Asesõnade keegi, miski, kumbki käändevormide varieerumine eesti kirjakeeles. [The variation of case forms in pronouns keegi, miski, kumbki in written Estonian]. Bakalaureusetöö. http://hdl.handle.net/10062/60560.

Pant, Annika. 2020. Pronoomenite keegi, miski, kumbki, ükski käändevormide kasutus tänapäeva eesti keeles. [The variation of case forms usage of pronouns keegi, miski, kumbki, ükski in modern Estonian]. Magistritöö. http://hdl.handle.net/10062/68143.

Rull, Aado. 1917. Kaaslõpp -gi ja ta kidurad kaimud. [The clitic -gi and its lesser relatives]. Eesti Kirjandus 2. 82–88.

Saareste, Albert. 1923. Kellegile, mingit, kuskil. [To someone, some, somewhere]. Eesti Keel 4. 116–117.

Saareste, Andrus. 1936. Eesti õigekeelsuse päevaküsimustest. [On Current Issues in Estonian Orthography]. Eesti Kirjandus 12. 548–556.

Saareste, Andrus. 1955. Petit atlas des parlers estoniens. Väike eesti murdeatlas. [A Small Estonian Dialect Atlas]. Uppsala: Almqvist & Wiksells.

## Combining different methods to comprehend Estonian L2 vowel perception and production

Previous studies (Leppik et al., 2023) have shown that Spanish L1 learners of Estonian find it hard to differentiate between Estonian mid vowels. While in production, the learners merge /ø/ and /ɤ/ into an ambiguous mid vowel and often confuse it with /o/, in perception, they identify /ø/ most often as /ɤ/ and /e/, and only in 5% of the cases as /o/. This discrepancy between perception and production suggests that the orthography might have an influence on the production of L2 learners. Furthermore, several studies on other languages have noted the interference of orthography in L2 production (e.g., Escudero and Wanrooij, 2010; Young-Scholten et al., 2002). The vowels of Estonian that are new for Spanish L1 learners are presented in orthography with modified Latin letters *ä, ö, õ, ü*. Thus, the learners might simply ignore the diacritics above the letters and merge the unfamiliar vowel categories with the vowels that are represented with unmodified characters (*a, u, o*).

A Visual World Paradigm eye tracking study is being carried out to investigate the effect of orthography on vowel perception. Eye tracking data gives a possibility to follow the process that precedes the final decision. The study combines pictures and printed words which are presented in quadruplets, where every quadruplet contains a target (e.g., *söök* 'food'), orthographic competitor (e.g., *sool* 'salt'), phonetic competitor (e.g., *seen* 'mushroom') and unrelated distractor (e.g., *pall* 'ball'). The target and competitors always start with the same consonant. Half of the stimuli are presented as pictures and the other half as printed words. During the experiment the participant hears a sentence, e.g., '*Vali pilt, kus on köök'* ('Choose a picture where there is a kitchen') and has to click on the correct picture (or the printed word).

The paper will present preliminary results from the eye tracking study and compare the previous results of production and perception with the new data. We believe that combining data from these three different methods helps to comprehend better the difficulties of Estonian L2 perception and production and identify aspects that need more focus in Estonian L2 pronunciation teaching.

### References

Escudero, P., & Wanrooij, K. (2010). The effect of L1 orthography on non-native vowel perception. Language and speech, 53(3), 343-365.

Leppik, K., Lippus, P., & Asu, E. L. (2023). The perception and production of Estonian vowels and vocalic quantity contrasts by Spanish L1 learners. Ampersand, 11, 100147. DOI: 10.1016/j.amper.2023.100147.

Young-Scholten, M. (2002). Orthographic input in L2 phonological development. An integrated view of language development: Papers in honor of Henning Wode.

# The Emergence of Kinship Terminology: an Information-theoretic Agent-based Approach

Every society uses a communicative system to refer to members of their family, or *kin*. A kinship system consists of a set of words that are used to refer to the position that a family member occupies in the family tree; for instance, mother, sister or uncle. Evidently, different languages use different word forms for the same family member (e.g. the English word *grandmother* is *amona* in Basque). However, cross-linguistic variability exists also at the semantic level, such that different languages may group family members into different categories. For instance, English or Dutch use gender-based distinction for grandfather and grandmother, but –unlike Chinese or Vietnamese– there is no distinction with regards to maternal or paternal side (Murdock, 1970).

Kemp and Regier (2012) observe that this variation is, in fact, constrained by two principles: first, a pressure for *informativeness*, such that communication is possible without too much ambiguity; and second, a pressure for simplicity that limits the amount of semantic categories, such that the complexity of the system is bounded. The authors observe that languages exhibit a (near-)optimal trade-off between these two opposing principles. While similar observations have been reported in other semantic domains (Carr et al., 2020; Zaslavsky et al., 2018), the exact source of these pressures is not fully understood, hence such studies used computational simulations to leverage the influence of different factors like learnability, communicative needs or channel capacity.

To address this, we propose the following methodological approach. We present an agent-based computational model that simulates the emergence of a communication protocol (or 'language') between two neural network agents. We focus on the domain of kinship, hence agents are trained to develop a communication system to refer to members of their kin. Our agents do not have any prior linguistic knowledge; therefore, in order to communicate, they send signals by choosing symbols ('words') which are initially meaningless. Over the course of training, we reward interactions that lead to successful communication, hence over time agents develop a system of signal-meaning associations that allow them to refer to different members of their kin. Notably, we do not guide the agents to specific signal choices, thus their behavior emerges from the combination of learning model and environment. As in Zaslavsky et al. (2018), we use an information-theoretic framework to analyze the evolving trajectory of the languages that our agents develop during training. Our approach, however, extends the mentioned framework and adapts it to the study of kinship, which –unlike color– is a discrete domain.

Our agents learn to communicate successfully in the kinship referential game, although –as can be seen in Figure 1–the accuracy varies depending on the channel capacity (i.e. the number of different words that the an agent can potentially choose from). Notably, endowing the agents with enough words provides them with sufficiently large search space, but this does not entail that agents make use of every possible word; in fact, Figure 2 suggests that the opposite is the case. As can be noted, the complexity of the emerged languages (which reflects how agents effectively attribute the vocabulary to kinship members, such that higher complexity entails more fine-grained semantic categories) increases during training, but the final complexity is close to the theoretical optimal curve. At the same time, informativeness also increases during training, resulting in near-zero information loss in the final language. Our work shows that domain-general learning –as provided by our neural network agents– combined with a referential objective and sufficient channel capacity are plausible pressures behind the (near-)optimal kinship systems found in the world's languages.

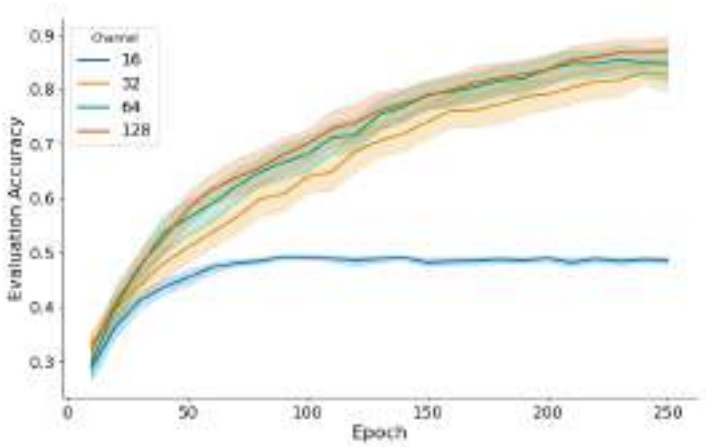
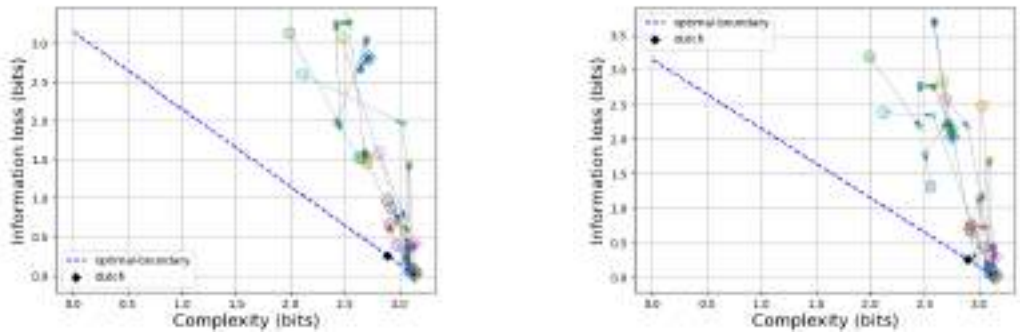



Figure 1: Communication accuracy of agents playing the kinship referential game, over the course of learning epochs.




(a) Ego: Alice

(b) Ego: Bob

Figure 2: Complexity and information loss of emerged languages (for channel capacity 64). The complexity of Dutch is shown as a black diamond, as an example of a natural language. For the emerged languages, we show the complexity and information loss during learning. The initial state is shown with a circle, and the final state is shown with a diamond. Different trajectories represent different initializations of the model. The blue dotted line represents the theoretical optimal curve. The area below the curve is theoretically unachievable. Since some kinships vary based on the gender of the ego, we present a model for a female ego (Alice) and a model for a male ego (Bob).

## References

Carr, J. W., Smith, K., Culbertson, J., and Kirby, S. (2020). Simplicity and informativeness in semantic category systems. *Cognition*, 202.

Kemp, C. and Regier, T. (2012). Kinship categories across languages reflect general communicative principles. *Science*, 336:1049–1054.

Murdock, G. P. (1970). Kin term patterns and their distribution. *Ethnology*, 9(2):165–208.

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

# Multimodal Negation: Gesture and Prosody in Estonian and Turkish

Pärtel Lippus[a], Eva Liina Asu[a], Maarja-Liisa Pilvik[a], Liina Lindström[a], Sevilay Sengul Yildiz[b], Hatice Zora[c], Asli Ozyurek[c,d]

Institute of Estonian and General Linguistics, University of Tartu, Estonia[a]
Donders Centre for Cognition, Radboud University, Nijmegen, The Netherlands[b]
Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands[c]
Centre for Language Studies, Radboud University, Nijmegen, The Netherlands[d]

The aim of this paper is to investigate the relationship between speech and gesture in negation phrases in two typologically different languages: Estonian and Turkish. More specifically, the alignment of negation gestures with the prosodically prominent syllables will be studied.

Earlier multimodal studies on different languages (e.g. English [1], French [2], Indonesian [3]) have shown that negation can be associated with recurrent gestures. While these gestures primarily have different communicative functions than the beat gestures, they are still often aligned with prosodically prominent units in speech [4].

In Estonian, standard (clausal) negation is asymmetric [5]: the negative verb form is different from the affirmative form (negation particle *ei* + different connegative verb stem, e.g. *Ma ei jookse* 'I not run.CNG'). The negation particle always immediately precedes the verb. In speech, the verb receives a pitch accent and the negation particle is therefore deaccented. In Turkish, verbal negation is expressed through suffixation, and the negation suffix -mA is added directly to the verb stem (verb stem + negation suffix -mA, e.g., *Ben koşmam* 'I run.NEG.1SG'). While word-level prosodic prominence typically falls on the final syllable of a word in Turkish, some non-accentable morphemes, like the negation suffix, can shift it to the preceding syllable [6].

The Estonian data consists of 20 video recordings of dialogues (à 30 min) between 40 speakers taken from the Phonetic Corpus of Estonian Spontaneous Speech [7]. The Turkish data consists of spontaneous dialogues including 87 different speakers taken from YouTube. The speech was annotated in Praat [8] (words, sounds, syllables, morphological categories) and the gestures were manually annotated in ELAN [9]. The prosodic analysis focused on duration, pitch, and intensity relative to negation markers.

The Estonian data contained 2056 negation phrases of which 475 (23%) were associated with a gesture (339 head shakes, 136 hand gestures) and the Turkish data contained 211 negation phrases of which 179 (85%) were accompanied by a gesture (72 backward hand flips, 19 head tilts, and 9 eyebrow movements — all occurring in isolation — and 79 combined gestures involving hand and head or hand and eyebrows). Both in Turkish and Estonian the majority (79-87%) of the gestures preceded the negation marker. While we expected the gestures to be aligned with the accented syllable, the onset of hand gestures in Estonian preceded the accented syllable by an average of 497 ms and that of head gestures by 542 ms. In Turkish, when aligned with the negation marker, the onset of the gesture stroke preceded it by 648 ms. However, when aligned with the accented syllable that precedes the negation marker, the lead time was reduced to 504 ms — a pattern strikingly similar to what we found in Estonian.

## References


1. S. Harrison, 'The organisation of kinesic ensembles associated with negation', *GEST*, vol. 14, no. 2, pp. 117– 140, Dec. 2014, doi: 10.1075/gest.14.2.01har.
2. S. Harrison and P. Larrivée, 'Morphosyntactic Correlates of Gestures: A Gesture Associated with Negation in French and Its Organisation with Speech', in *Negation and Polarity: Experimental Perspectives*, vol. 1, P. Larrivée and C. Lee, Eds., in Language, Cognition, and Mind, vol. 1. , Cham: Springer International Publishing, 2016, pp. 75–94. doi: 10.1007/978-3-319-17464-8_4.
3. P. Siahaan and G. P. Wijaya Rajeg, "Multimodal language use in Indonesian: Recurrent gestures associated with negation," p. 769328, 2023, doi: 10.17617/2.3527196.
4. P. Wagner, Z. Malisz, and S. Kopp, 'Gesture and speech in interaction: An overview', *Speech Communication*, vol. 57, pp. 209–232, Feb. 2014, doi: 10.1016/j.specom.2013.09.008.
5. M. Miestamo, *Standard Negation: The Negation of Declarative Verbal Main Clauses in a Typological Perspective*. Mouton de Gruyter, 2005. doi: 10.1515/9783110197631.
6. Kabak, B., & Zora, H. (In press). Psycholinguistics and Turkish: Prosodic representations and processing. In Johanson, Lars (ed.). *Encyclopedia of Turkic Languages and Linguistics*. Brill.
7. P. Lippus, K. Aare, A. Malmi, T. Tuisk, and P. Teras, 'Phonetic Corpus of Estonian Spontaneous Speech v1.3'. Institute of Estonian and General Linguistics, University of Tartu, Oct. 20, 2023. doi: 10.23673/RE-438.
8. P. Boersma and D. Weenink, 'Praat: doing phonetics by computer'. Feb. 27, 2021. [Online]. Available: http://www.praat.org
9. 'ELAN [Computer software]'. Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, 2023. [Online]. Available: https://archive.mpi.nl/tla/elan

**Pre-attentional Processing of Estonian Quantity Stimuli in Russian-Estonian Bilinguals**

Estonian prosody features a three-way quantity distinction (Q1, Q2, Q3) based on the duration of the first (stressed) syllable, the duration of the second syllable, and the pitch cue (Lippus et al., 2011). Previous behavioral studies indicate that native Estonian speakers use both durational and pitch cues to distinguish between Q2 and Q3, whereas Russian native speakers rely primarily on durational differences (Meister & Meister, 2011). EEG experimental studies also confirmed that perception of the quantity stimuli varied not only based on physical characteristics but also according to the participant's native language, confirming that the meaning of a word influences the early automatic processing of phonological features (Kask et al., 2021). The current study aimed to investigate how Russian-Estonian bilinguals differ in perception of Estonian quantity, depending on the age at which they began acquiring Estonian (age of acquisition) and their level of Estonian.

The study employed a two-part experimental design: 1) a web-based self-report questionnaire collecting demographic data, musicality, handedness, language skills, Estonian language exposure, use and learning experience; 2) laboratory-based EEG measurements using the mismatch negativity (MMN) paradigm (Näätänen et al., 2004) to assess the early neural responses to Estonian quantity distinctions.

Stimuli consisted of 2 minimal triplets (sada 'hundred' vs. saada 'send' vs. saada 'get'; pole 'is not' vs. poole 'half' gen vs. poole 'toward'), recorded by male and female native Estonian speakers. Eight stimuli series included natural speech samples, re-synthesized versions manipulating pitch and duration, and a control condition with non-linguistic stimuli. There were 1 standard and 3 deviant stimuli in each series. EEG data were recorded from 64 scalp electrodes, with additional electrodes tracking eye movements and facial muscle activity. Pre- and post-experimental tests included critical flicker frequency test (Simonson & Brožek, 1952), lexical proficiency task (LexEst, Lõo et al, 2024), audiometry, and self-reported fatigue questionnaire. Participants watched a silent movie while stimuli were presented to headphones.

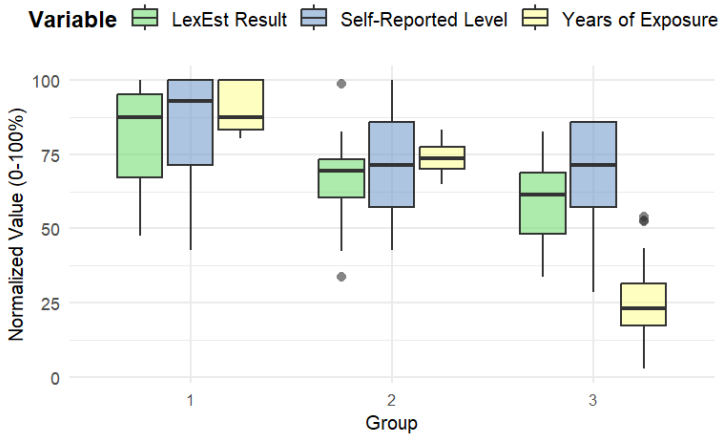
A total of 73 Russian-Estonian bilinguals participated in the study, divided into three groups based on their age of acquisition (AoA) of Estonian: bilingual first language acquisition (BFLA) group (AoA 0-4 years) included 24 participants, mean age 23.9 years (SD = 6.6), early second language acquisition (SLA) group (AoA 5-8 years) included 26 participants, mean age 26.2 years (SD = 6.4) and late second language acquisition (SLA) group (AoA 15-34 years) included 23 participants, mean age 29.3 years (SD = 6.3).

EEG data were preprocessed using BrainVision Analyzer 2.1, with standardization and artifact correction applied. Cleaned data were averaged within each subject for each event type (standard and deviant). Further analyses were conducted in R, focusing on frontal regions of interest (AF3, F1, F3, F5, FC1, FC3; AF4, F2, F4, F6, FC2, FC4). A repeated-measures ANOVA was performed to examine MMN differences across participant groups and conditions.

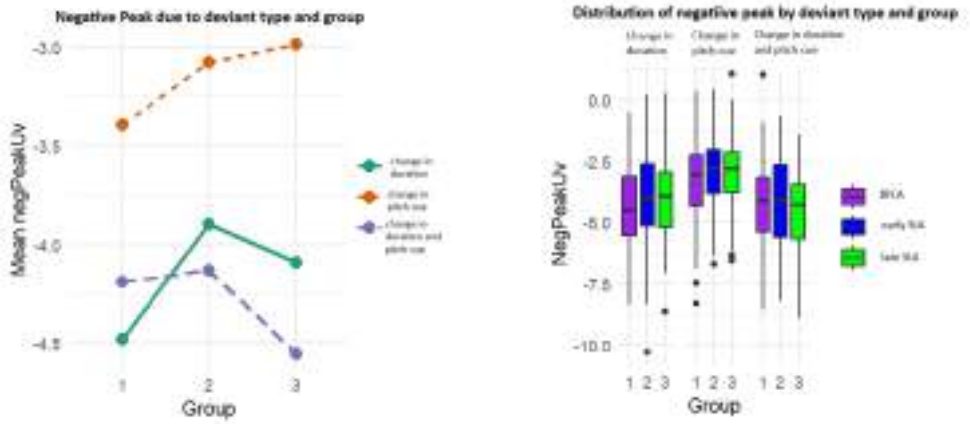
Preliminary results indicate that early exposure to language enhances pre-attentional processing of language-specific stimuli. These findings align with previous research suggesting that linguistic background influences early and automatic auditory processing due to brain plasticity.

## Figures

Figure 1. Distribution of LexEst results, self-reported level of Estonian and years of exposure to Estonian (% of life) by group (1 – BFLA, 2 – early SLA, 3 – late SLA).




Figures 2. MMN response for series using 'saada' with manipulations in change and duration of the first vowel.




## References

Kask, L., Põldver, N., Lippus, P., & Kreegipuu, K. (2021). Perceptual Asymmetries and Auditory Processing of Estonian Quantities. *Frontiers in Human Neuroscience, 15*. https://doi.org/10.3389/fnhum.2021.612617

Lippus, P., Pajusalu, K., & Allik, J. (2011). The role of pitch cue in the perception of the Estonian long quantity. In S. Frota, G. Elordieta, & P. Prieto (Eds.), Prosodic Categories: Production, Perception and Comprehension (pp. 231–242). Springer Netherlands.

Lõo K., Bleive B., Leppik K., Malmi A. (2024). LexEst. https://sisu.ut.ee/lexest

Meister, L., & Meister, E. (2011). Perception of the short vs. Long phonological category in Estonian by native and non-native listeners. *Journal of Phonetics*, *39*(2), 212–224. https://doi.org/10.1016/j.wocn.2011.01.005

Näätänen, R., Pakarinen, S., Rinne, T., & Takegata, R. (2004). The mismatch negativity (MMN): towards the optimal paradigm. Clinical Neurophysiology, 115, 140-144.

## The Estonian Auditory Lexical Decision database: expanding methodological diversity in psycholinguistics

*Kaidi Lõo, Anton Malmi and Benjamin V. Tucker*

Large-scale lexical decision databases are important for understanding word recognition, yet the few current auditory lexical decision datasets (Dutch: Ernestus & Cutler, 2015; French: Ferrand et al., 2018, and English: Tucker et al., 2019) focus on rather morphological simple languages. To address this gap, we introduce the Estonian Auditory Lexical Decision Database (EALD), a dataset with responses from 412 native Estonian speakers (aged 18-72 years; 321 females, 85 males, 10 non-binary) to 8800 Estonian words and pseudowords, collected online using the OpenSesame software and the JATOS server Mindprobe (Lange et al., 2015). The stimuli entail monomorphemic, inflected, derived, and compound Estonian words. Compared to the languages mentioned above, Estonian has a more complex agglutinative and fusional case-marking system, resulting in large morphological paradigms that significantly influence word recognition processes (Lõo et al., 2018). In addition to behavioral responses, we provide auditory and lexical information, enabling researchers to analyze the effects of syllable and morphological structure, phonotactics, word frequency, and both orthographic and phonological neighborhood density on auditory word recognition.

Our preliminary results using regression analyses show that reaction times are faster with increasing written word frequency and inflectional paradigm size, and decreasing phonological neighborhood density. Data shows a nonlinear relationship between listeners' age and reaction time. Reaction times increase between 18 and 30 years decrease between 30 and 50 years, and increase again after 50 years. Additionally, time of day influences performance, with slower responses recorded at night time.

By making all our auditory, behavioral, and lexical data of EALD available in OSF, we aim to enhance the current methodological diversity in psycholinguistics, providing researchers with a valuable tool to explore morphological processing and spoken word recognition in Estonian.

**Beyond Typology: Individual Differences in Second Language Learning and Academic Achievement among Russian-Speaking University Students in Estonia**

**Authors:** J., Mackiewicz[1], V. Vihman[2], P. Milin[1]

[1]University of Birmingham, [2]University of Tartu

**Abstract**

This study builds on prior research examining the role of typological distance in second language (L2) acquisition among university students from diverse linguistic backgrounds (Mackiewicz, under review). The main findings indicated that slower linguistic progress among Chinese learners, compared to their European peers—whose first languages are typologically closer to English—was largely attributable to differences in initial proficiency and the extent of exposure to English. More succinctly, the findings suggest that, at higher levels of L2 proficiency, the amount and nature of learners' prior experience with the target language may exert a stronger influence on developmental trajectories than structural distance alone.

To explore this hypothesis further, we examine Russian L1 speakers' proficiency in Estonian, a language that is also typologically distant from their native tongue. Unlike international students in the UK, however, Russian-L1 learners of L2 Estonian are typically born and raised in the country, rather than migrating there. The broader *situated environment* in which their L2 development occurs is unique in several respects. First, active proficiency in Estonian among non-native speakers remains relatively low, with only around 41% reporting functional fluency (Kaldur et al., 2023). With regards to Russian L1 university students, they have reported feeling underprepared for studying in Estonian-based curricula, despite having studied in Estonian schools aimed at raising L2 Estonian proficiency (Klaas-Lang et al., 2022) and often fall behind in academic performance, which in turn affects societal integration (Erss, 2023). Finally, the urgency surrounding Estonia's ongoing education reform—which aims to establish Estonian as the sole language of instruction in all schools by 2030 (Popova, 2023)—creates additional pressure on L2 acquisition. Together, these factors constitute a complex and dynamic environment in which Estonian is acquired as a second language, offering a distinctive lens through which to examine L2 development more broadly.

Drawing on a sample of university students who met the B2 language entry requirements of the Common European Framework of Reference for Languages, we examine how individual differences relate to L2 proficiency—specifically vocabulary knowledge and reading skills—and to self-reported academic achievement. By focusing on a linguistically minoritised student population with diverse language learning histories—ranging from early immersion in Estonian to late-start, classroom-based instruction—this study offers novel insights into the complex factors shaping language development and academic outcomes in second language higher education contexts.

Erss, M. (2023). Comparing student agency in an ethnically and culturally segregated society: How Estonian and Russian speaking adolescents achieve agency in school. *Pedagogy, Culture & Society*, *33*(2), 439–461.

Kaldur, K., Pertšjonok, N., Toomik, K., Khrapunenko, M., Kivistik, K., Aavik, A.-L., & Jurkov, K. (2023). *The organisation of Estonian language learning in restricted language environments: Creating sustainable and effective integrated solutions*. Balti Uuringute Instituut. URL: https://www.ibs.ee/en/publications/the-organization-of-estonian-

language-learning-in-limited-language-environments-challenges-and-sustainable-solutions

Klaas-Lang, B., Koreinik, K., & Rozenvalde, K. (2023). *From school to university: What features of communication culture do students with Russian as a dominant language notice? Journal of Estonian and Finno-Ugric Linguistics*, *14*(2), 187–211.

Mackiewicz, J. (under review). Are they ready? Rethinking linguistic preparedness of international students in English-medium Higher Education.

Popova, Z. (2023). *CONFIRMED: 2030 IS THE End-Date of The Russian-Language Education in Estonia*. FUEN Policy Paper. URL: https://fuen.org/assets/upload/editor/docs/doc_PuAbUeyh_TheEstonianExperiment.pdf

**How do children acquire consonants with multiple lingual gestures?**

Anton Malmi (University of Tartu, Estonia), Claire Nance (Lancaster University, UK)

As children grow, the size and proportions of their vocal anatomy also evolve [1]. Depending on the child, adult-like vocal anatomy proportions are reached anywhere between the ages of 7–18 years [2]. A child's tongue is proportionally shorter than an adult's tongue. Also, the larynx, which is connected to the tongue, is proportionally higher up, bringing the posterior part of the tongue a lot closer to the hard palate [3]. Cross-linguistic studies show that children acquire different speech sounds in a hierarchical order dependent on these physiological changes. Specifically, consonant acquisition continues up to around 6 years old, and consonants requiring multiple articulatory gestures are reported to be acquired late, for example, /dʒ/, /kʷʰ/, /ɹ/ [4].

Our study investigates how children negotiate the constraints of their developing anatomy and the need to produce intelligible speech. We focus on consonants with multiple articulatory gestures, using secondary palatalisation as a case study. Due to the relative size of the hard palate and tongue, child speech is reportedly produced with a palatal quality, and audible palatalisation is produced early on [5], [6]. But, producing adult-like phonemic secondary palatalisation requires precise coordination of multiple tongue gestures, which is reported to be challenging for younger children [7]. We aim to investigate 1) if children develop a different articulatory strategy for palatalisation compared to adults and 2) how tongue gestures necessary for secondary articulations are refined across childhood.

In this pilot study, we use ultrasound tongue imaging to investigate children's tongue gesture development for phonemically palatalised /lʲ nʲ sʲ tʲ/ and non-palatalised consonants /l n s t/ in Estonian. We have collected ultrasound tongue images and acoustic recordings of eight first-language Estonian-speaking children (four aged 3–4 years, four aged 6–7) and eight first-language Estonian-speaking adults aged 32–46. The recordings were made in Articulate Assistant Advanced (AAA) v221.4.1. The participants were asked to repeat a pre-recorded word that they heard through the laptop speakers. The experiment consisted of 24 single words. This includes eight minimal pairs for the palatalisation contrast (16 words total) and a further eight words with a palatalised consonant where no corresponding non-palatalised minimal pair exists. Each word was repeated five times. We recorded 120 tokens from each participant, making the total number of tokens to 1920.

The recordings were made using a Telemed Micrus ultrasound machine, a 20mm radius convex probe at 3MHz frequency, ~90Hz frame rate, and a stabilisation headset [8] and lapel microphone. Splines were fitted to the ultrasound data using DeepLabCut in AAA [9]. Acoustic data were labelled in Praat [10] and then reimported into AAA for defining acoustic landmarks in the analysis. Ultrasound spline coordinates were rotated to each speaker's occlusal plane [11] and extracted at the acoustic midpoint for further analysis.

The analysis is ongoing. We are using two methods to quantify and compare tongue shapes: Modified Curvature Index (MCI) and the Number of Inflections (NINFL). MCI calculates normalised curvature of the tongue shape and shows how curled up or stretched out the tongue is [12], [13]. NINFL calculates the number of times a curve changes from concave to convex [14]. We will fit linear mixed-effects models to the MCI and ordinal mixed-effects models on the NINFL data to study the main effects and interactions of age, consonant, and palatalisation.

Following the results from a similar study with Gaelic children [15], we expect Estonian children's tongue shapes for palatalised consonants to be flatter than adults' due to the relatively larger size of the tongue compared to the hard palate. We also predict that the tongue shapes go from fewer to more inflections with age as children acquire advanced control of their tongue. We discuss the implications of these results for normative data in clinical practice and for models of the acquisition of speech-motor control [16].

[1] H. K. Vorperian, R. D. Kent, M. J. Lindstrom, C. M. Kalina, L. R. Gentry, and B. S. Yandell, “Development of vocal tract length during early childhood: A magnetic resonance imaging study,” *J. Acoust. Soc. Am.*, vol. 117, no. 1, pp. 338–350, Jan. 2005, doi: 10.1121/1.1835958.
[2] H. K. Vorperian, “Anatomic development of the vocal tract structures as visualized by magnetic resonance imaging,” The University of Wisconsin, 2000.
[3] D. Abakarova, S. Fuchs, and A. Noiray, “Developmental Changes in Coarticulation Degree Relate to Differences in Articulatory Patterns: An Empirically Grounded Modeling Approach,” *J. Speech Lang. Hear. Res. JSLHR*, vol. 65, no. 9, pp. 3276–3299, Sep. 2022, doi: 10.1044/2022_JSLHR-21-00212.
[4] S. McLeod and K. Crowe, “Children’s Consonant Acquisition in 27 Languages: A Cross-Linguistic Review,” *Am. J. Speech Lang. Pathol.*, vol. 27, no. 4, pp. 1546–1571, Nov. 2018, doi: 10.1044/2018_AJSLP-17-0100.
[5] M. Vihman and V.-A. Vihman, “From first words to segments: A case study in phonological development,” in *Experience, Variation, and Generalization: Learning a first language*, E. V. Clark and I. Arnon, Eds., in Trends in Language Acquisition Research. , Amsterdam: John Benjamins Publishing Company, 2011.
[6] N. Zharkova, “Strategies in the acquisition of segments and syllables in Russian-speaking children.,” *Dev. Paths Phonol. Acquis.*, vol. 2, 2004.
[7] M. Denny and R. S. McGowan, “Implications of Peripheral Muscular and Anatomical Development for the Acquisition of Lingual Control for Speech Production: A Review,” *Folia Phoniatr. Logop.*, vol. 64, no. 3, pp. 105–115, May 2012, doi: 10.1159/000338611.
[8] L. Spreafico, M. Pucher, and A. Matosova, “UltraFit: A Speaker-friendly Headset for Ultrasound Recordings in Speech Science,” presented at the Proc. Interspeech 2018, 2018, pp. 1517–1520. doi: 10.21437/Interspeech.2018-995.
[9] A. Wrench and J. Balch-Tomes, “Beyond the Edge: Markerless Pose Estimation of Speech Articulators from Ultrasound and Camera Images Using DeepLabCut,” *Sensors*, vol. 22, no. 3, Art. no. 3, Jan. 2022, doi: 10.3390/s22031133.
[10] P. Boersma and D. Weenink, “Praat: doing phonetics by computer [Computer program].” 2025.
[11] J. M. Scobbie, E. Lawson, S. Cowen, J. Cleland, and A. A. Wrench, “A common co-ordinate system for mid-sagittal articulatory measurement,” *QMU CASL Work. Pap. WP-20*, Jun. 2011, Accessed: Mar. 06, 2025. [Online]. Available: https://eresearch.qmu.ac.uk/handle/20.500.12289/3597
[12] K. M. Dawson, M. K. Tiede, and D. H. Whalen, “Methods for quantifying tongue shape and complexity using ultrasound imaging,” *Clin. Linguist. Phon.*, vol. 30, no. 3–5, pp. 328–344, May 2016, doi: 10.3109/02699206.2015.1099164.
[13] M. Dokova, E. Sugden, G. Cartney, S. Schaeffler, and J. Cleland, “Tongue Shape Complexity in Children With and Without Speech Sound Disorders,” *J. Speech Lang. Hear. Res.*, vol. 66, no. 7, pp. 2161–2183, 2023.
[14] J. L. Preston, P. McCabe, M. Tiede, and D. H. Whalen, “Tongue shapes for rhotics in school-age children with and without residual speech errors,” *Clin. Linguist. Phon.*, vol. 33, no. 4, pp. 334–348, Apr. 2019, doi: 10.1080/02699206.2018.1517190.
[15] C. Nance and S. Kirkham, “Acquiring tongue shape complexity in consonants with multiple articulations,” in *Proceedings of the 5th International Symposium on Applied Phonetics*, Tartu, Estonia, 2024.
[16] S. Tilsen, “Selection and coordination: The articulatory basis for the emergence of phonological structure,” *J. Phon.*, vol. 55, pp. 53–77, Mar. 2016, doi: 10.1016/j.wocn.2015.11.005.

## How similar are language skills among family members in early childhood? A systematic review of nuclear family associations

*Magda Matetovici, Hans-Fredrik Sunde, Sergio Miguel Pereira Soares, Selim Sametoglu, Elsje van Bergen and Caroline Rowland*

Studies that show positive associations between speech input from parents and the language skills of young children prompt us to assume that differences in language abilities are directly caused by how and how much parents talk. However, parents' input also reflects their language skills, which are subject to genetic as well as environmental influences, thus could affect the language skills of children through both or either of these pathways. Studies rarely correlate parent and child language skills during childhood, either because they do not consider that parent input might depend on their skills, or because they equate input with parents' language skills. We argue that parent input and parent language skills are related but separate constructs, and that it is important to know the effect of both in order to understand the intergenerational transmission of language skills and the reasons why children differ in their language abilities. As a first step in this direction, in this paper we present a systematic review that investigates the similarity between the language skills of children and their parents, as well as between the parents themselves; the latter correlation informs our interpretation of the genetic influence of the two parents on their children (if two parents are more alike than we would expect by chance, we also expect a larger correlation between parents and children). We performed a pre-registered systematic search in PsycINFO, SCOPUS, Web of Science, ERIC and ProQuest Dissertations and Theses of English language articles. We present correlations between the language skills of parents and their children and between the language skills of the two parents in their native language(s) using a language test or assessment. We document the direction and magnitude of these associations, summarize the explanations these studies provide for their findings and identify potential moderators. Lastly, we assess the methodological quality of studies included in our review and document potential differences between high and low quality papers.

## The development of speaking skills in B1-level Estonian L2 courses

*Katrin Mikk*

Speaking is an important but difficult skill for second/foreign language learners to acquire (Thornbury 2005). Even learners who pass standardized language tests may still struggle with everyday spoken communication. While speaking skills are most effectively developed in the language environment (Doehler, Eskildsen 2022), they can also be successfully developed in the language classroom. In this context, teachers play a crucial role in selecting appropriate and engaging learning activities (Bittner-Collins, 2012) and increasing the amount of speaking in class. Like other language skills, speaking can be acquired or learned both implicitly and explicitly. In my study, I investigate how the language learning methods used in language courses influence the development of learners' speaking skills. To explore this, I observed nine Estonian L2 group lessons (three lessons per group). Each course lasted 120 hours (approximately three months). For data collection, I recorded class activities and captured learners' speaking performances at both the beginning and end of the course. Additionally, I developed an assessment tool to measure and evaluate learners' speaking development. This instrument offers a more comprehensive approach compared to earlier tools for assessing speaking skills. The assessment tool is used to evaluate learners' speaking development across several dimensions, including willingness and initiative to speak, collaboration with a partner, fluency, pronunciation, vocabulary, and grammatical accuracy. As this study is part of my doctoral dissertation, in the future, I will compare the results from each group, considering the activities and language learning methods used in the lessons, in order to answer the following questions: 1. How much time was spent on speaking activities in class, and what was the amount of speaking for each student? 2. What is the correlation between classroom activities and the development of learners' speaking skills?

In my presentation, I will present findings obtained through the assessment tool, which show changes in different components of learners' speaking skill development over the duration of the course.

**References:**

Bittner-Collins, T. (2012). Improving the Speaking Skill through a Controlled-Learning Environment for 2nd year students of Traducción Inglés-Español. Universidad de Las Américas.

Doehler, S. & Eskildsen, S. (2022). Emergent L2 grammar in and for social interaction: introduction to special issue. The Modern Language Journal 106, 3–22.

Thornbury, S. (2005). How to Teach Speaking. Harlow: Longman.

**What experimental evidence can tell us about homesign creation**

Deaf and hard of hearing children (DHH) born to a hearing, non-signing family often innovate their own gestural systems known as homesigns. These are of great interest to language development researchers because they exhibit certain fundamental properties of natural languages such as recursion, combinatoriality, and argument structure (Flaherty et al., 2021) despite their creators not having access to a (fully) usable linguistic model. The standard assumption in homesign research is that caregivers unknowingly provide their children with input by producing gestures alongside their speech. These have been shown to lack linguistic structure; thus, it is further assumed that the children systematize the inconsistent input they were exposed to, innovating linguistic structure. Although this assumption has been discussed at length from a theoretical standpoint, it has never been, to my knowledge, discussed or examined from a neurocognitive perspective.

The present research is an original review that synthesizes experimental evidence from a variety of methods including event-related fMRI, artificial language learning, and silent gesture aimed at investigating the internal processes that would be necessary for the standard assumption to be possible. The findings of the reviewed experiments can be classified into three categories: (i) the linguistic dimension of gestures (ii) the effects of cross-modal reorganization following congenital deafness, and (iii) children's increased capacity to regularize inconsistencies in their input and innovate linguistic structure.

The nature of gestures, particularly when produced simultaneously with speech, has been the subject of a large body of research (see Özyürek, 2014 for a comprehensive review) where they are often argued to supplement the semantic, pragmatic, and syntactic content of spoken utterances. As such, their processing interacts with that of speech and largely takes place in the language areas of the brain, specifically, in the inferior frontal and superior temporal gyri (IFG and STG), which correspond to Broca's and Wernicke's areas, respectively (Wolf et al., 2017). While it is known that hearing children also make use of gestures during acquisition (Özçalışkan & Dimitrova, 2013), homesigners do not have access to the spoken language information that gestures complement, they only have access to the gestures. Could they be able to derive more information from these gestures than their hearing peers and, if yes, what enables them to do so?

Sensory loss induces an adaptive mechanism of the brain known as cross-modal reorganization, where areas designated to the processing of the lost modality are repurposed, leading to enhanced performance in the remaining modalities (Bonna et al., 2021). Crucially, this is the result of sensory loss itself, not of individual experience, e.g., sign language exposure in childhood (Hall et al., 2019), which might in fact improve spoken language outcomes in DHH children with cochlear implants (Davidson et al., 2014). In the congenitally deaf, the primary auditory cortex is repurposed, engaging in the processing of visual stimuli, resulting in both attentional and perceptual enhancements, specifically in the peripheral visual field (Scott et al., 2014). fMRI studies have shown that non-linguistically structured manual communication (gestures) activate brain areas that are traditionally involved in the processing of language (e.g., STG) to a *significantly greater extent in early deaf individuals* than in hearing individuals (Simon et al., 2020). This suggests that gestures can be processed as linguistic material, especially in the early deaf, where cortical activation patterns during the processing of meaningful gestures (emblems) have been shown to be near-identical to those during the processing of American Sign Language (ASL) signs (Husain et al., 2009).

Regarding the ability to innovate structure, a common experimental method is silent gesture (e.g., Bohn et al., 2019; Özçalışkan et al., 2016). Bohn et al. (2019) found that children spontaneously created core linguistic properties, including conventionality and compositionality when asked to describe a picture scene without using spoken language. Compositional gestures improved comprehension compared to holistic ones and subtle differences were ignored with participants copying each other's gestures and converging on one form, demonstrating an increased capacity for regularization (Bohn et al., 2019). Experiments employing different methodologies, such as miniature artificial languages (Hudson Kam & Newport, 2009) and pattern recognition and rearrangement tasks (Kempe et al., 2015) report comparable findings regarding children's capacity for regularizing inconsistent input and converging on forms that are more easily transmissible and reproducible. In comparison, adults are more likely to reproduce inconsistencies until the number of alternate forms is too great or they occur in low frequency (Hudson Kam & Newport, 2009).

This project highlights not only the necessity of interdisciplinarity in research, but also of the use of novel experimental methodologies in scientific inquiry. Technological advances have allowed us to obtain evidence for hypotheses that form the foundation of a major current within the field of linguistics (see Finkl et al., 2020 for experimental evidence of a 'core language network' separated from externalization channels). Similarly, the collection of evidence presented here has provided support for claims that have been central to homesign research for at least the past 20 years (see Goldin-Meadow, 2005) without having been explored to their full extent.

References

Bonna, K., Finc, K., Zimmerman, M., Bola, M., Mostowski, P., Szul, M., Rutkowski, P., Duch, W., Marchewka, A., Jednoróg, K., & Szwed, M. (2021). Early deafness leads to re-shaping of functional connectivity beyond the auditory cortex. *Brain Imaging and Behavior, 15*, 1469-1482.

Davidson, K., Lillo-Martin, D., & Chen Pichler, D. (2014). Spoken English language development among native signing children with cochlear implants. *J Deaf Stud Deaf Educ*. *18*(2), 238-250.

Finkl, T., Hahne, A., Friederici, A. D., Gerber, J., Mürbe, D., & Anwander, A. (2020). Language without speech: Segregating distinct circuits in the human brain. *Cerebral Cortex*, *30*, 812-823.

Flaherty, M., Hunsicker, D. & Goldin-Meadow, S. (2021). Structural biases that children bring to language learning: A cross-cultural look at gestural input to homesign. *Cognition, 211*, 1-11.

Goldin-Meadow, S. (2005). *The Resilience of Language: What gesture creation in deaf children can tell us about how all children learn language*. Psychology Press.

Hall, M. L., Hall, W. C., & Caselli, N. K. (2019). Deaf children need language, not (just) speech. *First Language*, *39*(4), 367-395.

Hudson Kam, C. L. & Newport, E. L. (2009). Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology*, *59*(1), 30-66.

Husain, F. T., Patkin, D. J., Thai-Van, H., Braun, A. R., & Horwitz, B. (2009). Distinguishing the processing of gestures from signs in deaf individuals: An fMRI study. *Brain Research*, (1276), 140-150.

Kempe, V., Gauvrit, N., & Forsyth, D. (2015). Structure emerges faster during cultural transmission in children than in adults. *Cognition*, *136*, 247-254.

Özçalışkan, Ş. & Dimitrova, N. (2013). How gesture input provides a helping hand to language development. *Seminars in Speech and Language*, *34*(4), 227-236.

Özçalışkan, Ş., Lucero, C., & Goldin-Meadow, S., (2016). Does language shape silent gesture? *Cognition*, *148*, 10-18.

Özyürek, A. (2014). Hearing and seeing meaning in speech and gesture: insight from brain and behaviour. *Phil. Trans. R. Soc. B*, (369):20130296, 1-10.

Scott, G. D. Karns, C. M., Dow, M. W., Stevens, C., & Neville, H. J. (2014). Enhanced peripheral visual processing in congenitally deaf humans is supported by multiple brain regions, including primary auditory cortex. *Frontiers in Human Neuroscience, 8*:177, 1-9.

Simon, M., Lazzouni, L., Campbell, E., Delcenserie, A., Muise-Hennessey, A., Newman, A. J., Champoux, F., & Lepore, F. (2020). Enhancement of visual biological motion recognition in early-deaf individuals: Functional and behavioral correlates. *PLoS ONE, 15*(8), 1-17.

Wolf, D., Rekittke, L. M., Mittelberg, I., Klasen, M., & Mathiak, K. (2017). Perceived conventionality in co-speech gestures involves the fronto-temporal language network. *Frontiers in Human Neuroscience*, *11*:573.

**Grammar description of Estonian Sign Language: Methodological Approaches**

**Goal.** The paper describes a number of methodological approaches to developing a grammar description for Estonian Sign Language, one of the sign languages used by the Estonian Deaf communities.

**Questions.** Two main questions concern the grammar description and the design of the described grammar of Estonian Sign Language: What are the methodological challenges and solutions for the development of the grammar sketch in Estonian Sign Language? How can the grammar sketch of Estonian Sign Language be designed?

**Findings.** The paper explores six issues that emerged during the process of grammar description: (a) the definition of grammar description; (b) possible cross-linguistic and cross-modal effects in the description process; (c) the role of linguistic knowledge of Estonian Sign Language users during the process of data collection and analysis; (d) the process of reviewing technical glossaries regarding linguistic concepts and grammar design; and (e) the design of the layout; (f) collaboration with deaf and academic communities during the process.

**Discussion.** These issues are explored within the framework of community-based empowerment research (cf. Czaykowska-Higgins, 2009; Yamada, 2007; Rice, 2018) to preserve and explore the linguistic diversity of Estonian Sign Language in Estonia. (according to Hale, 1992; Kadanya, 2006). The paper provides a starting point for discussing the role of grammar description, including language resources and language documentation as a toolkit.

Keywords: Estonian Sign Language; grammar description; language resources; terminological development; community-based research

**References**

Czaykowska-Higgins, E. (n.d.). Research Models, Community Engagement, and Linguistic Fieldwork: Reflections on Working within Canadian Indigenous Communities. *Language Documentation & Conservation*, *3*(1), 15–50.

Rice, K. (2018). Collaborative research: Visions and realities. In T. B. Shannon & J. Carmen (Eds.), Insights from Practices in Community-Based Research (pp. 13-37). De Gruyter Mouton. https://doi.org/doi:10.1515/9783110527018-002

Yamada, R.-M. (2007). Collaborative Linguistic Fieldwork: Practical Application of the Empowerment Model. Language Documentation & Conservation, 1(2), 257-282.

Hale, K., Craig, C., England, N., Laverne Masayesva, J., Krauss, M., Watahomigie, L., & Yamamoto, A. (1992). Endangered Language. *Language* 68(1), 1-42.

Kadanya, J. L. (2006). Writing grammars for the community. *Studies in Language* 30(2), 253-257. https://doi.org/https://doi.org/10.1075/sl.30.2.04kad

**Assessing scalar meaning: a first exploratory study on some Italian scalar-additive focus particles**

Sentences with scalar-additive *even*-like particles convey three layers of meaning. They entail the corresponding sentences without the particle (*assertion*); they presuppose that at least one of the associate's [1] focus alternatives satisfies the predicate (*additive presupposition*), and that these alternative(s) are somewhat ordered (*scalar presupposition*) [2], [3], [4]. When *even*-like particles interact with negation, both the assertion and the scalar and additive presuppositions get reversed. In Italian, the closest correspondent of *even* is held to be *persino* [5]. However, other particles can trigger similar scalar and additive presuppositions. For example, *pure-neppure* and *anche-neanche* [6], [7], [8]. At present, little experimental work has been carried out on Italian *even*-like operators, though. To our knowledge, only one study was conducted on the processing of *persino* in adults [9]. Hence, it is yet to be determined whether Italian adult speakers use the various scalar-additive focus operators of their language interchangeably, or whether somewhat of a division of labour holds among them.

To verify this, we designed a multiple-choice cloze test targeting the operators *persino-persino...non, pure-neppure* and *anche-neanche*. We first presented participants with a drawing and a short story and then asked them to fill in a concluding sentence by choosing one of these six particles provided as alternatives (Figure 1). Following [10], we ideated the pictures and stories in order to make participants build expectations about which of the characters shown was the most or least likely to carry out the action described – an action which could eventually be either successfully accomplished or failed by all. Via employment of the multiple choice cloze-test methodology, we could observe how the rate of particle selection varied depending on the character focalised in the concluding sentences and on the outcome of the stories. Importantly, then, this paradigm allowed us to collect more intuitive responses concerning the acceptability of the targeted particles in the various conditions compared to other well-established methodologies involving more explicit judgements or ratings. Via within-subject manipulation of Polarity (*positive-negative*) and Character (*likely-middle*), our experiment consisted of 80 (48 experimental) trials, which could appear under four possible conditions.

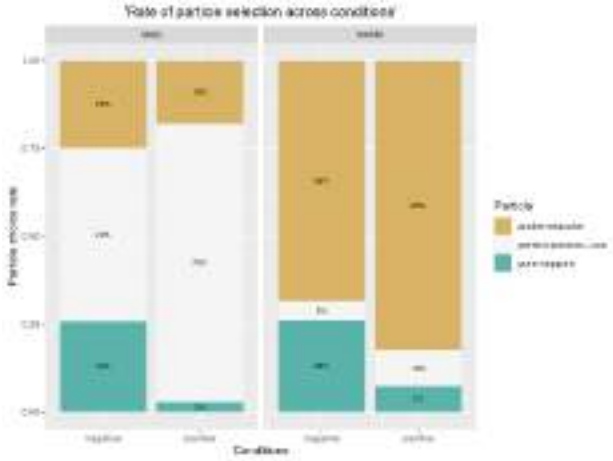
We analysed the responses of 82 monolingual Italian adults (39 females; age $M$=30) via multinomial logistic regressions on Particle Type, with Polarity and Character as fixed effects in interaction and Participant as random intercept. The model returned a significant interaction between Polarity and Character ($\chi^2$=64.20, $p$<.001), as well as a significant main effect of both Character ($\chi^2$=375.03, $p$<.001) and Polarity ($\chi^2$=15.50, $p$<.001). As shown in Figure 2, *persino-persino…non* were the most selected alternatives in both positive and negative contexts with focus on the likely character (79% and 49%, respectively), which seems to indicate that these were the alternatives mostly associated with the expression of the scalar presupposition. Interestingly, the choice rate of *persino…non* in negative contexts was lower than that of *persino* in positive ones. *Pure-neppure* were rather chosen more in negative (26%) than in positive contexts (3%) in association with the likely character. In negative ones, *neppure* was chosen at an equal rate when coupled with the likely (26%) and the middle character (26%), and its selection rate on the likely was almost on a par with that of *neanche* (25%). All this seems to indicate that also *anche*, *neanche* and *neppure* – though not *pure* – are compatible with a scalar interpretation. Last, the fact that *anche-neanche* were the most selected alternatives in association with the middle character (*anche:* 83%; *neanche:* 68%) appears to suggest that these were felt as the 'less scalar' among the alternatives provided.

All in all, these data not only seem to point out that, in Italian, some scalar-additive focus operators are associated with the expression of the scalar presupposition more than others. They also show that some of the positive-negative pairs investigated – particularly, *persino-persino…non* and *pure-neppure* – are not always employed in an exact complementary way as their positive counterparts – a fact certainly worth scrutinizing in further detail.

## Figures

*Figure 1* – An example of experimental trial (translation in italics)


*Figure 2* – Proportion of particle selection across conditions




## References

[1] Krifka, M.: 1998, Additive particles under stress. *Proceedings of Semantics and Linguistic Theory VIII*, 111-129.
[2] Horn, L. (1969). A presuppositional analysis of only and even. *Chicago Linguistic Society (CLS) 5*, 98-107.
[3] Karttunen, L. & Peters, S. (1979). Conventional implicature. In C-K. Oh & D.A. Dinneen (Eds.), *Syntax and Semantics: Presupposition* (Vol. 11, pp. 1-56). Academic Press.
[4] Rooth, M. (1985). *Association with focus*. PhD dissertation. University of Massachusetts at Amherst.
[5] Gast, V. & van der Auwera, J. (2011). Scalar additive operators in the languages of Europe. *Language 87*(1), 2-54.
[6] Tovena, L.M. (2006). Dealing with alternatives. *Proceedings of Sinn und Bedeutung 10*(2), 373-388.
[7] Tovena, L. M. (2005). Discourse and addition. *Proceedings of the Workshop Discourse Domains and Information Structure ESSLLI 2005*, 47-56.
[8] Mari, A. & Tovena, L.M. (2006). A unified account for the additive and scalar uses of Italian neppure. In C. Nishida & J-P.Y. Montreuil (Eds.), *New Perspectives on Romance Linguistics: Morphology, Syntax, Semantics, and Pragmatics*. (Vol. 1, pp. 187-200). John Benjamins Publishing Company.
[9] Bello Viruega, I. & Nadal, L. (2023). *Processing scalarity: an experimental study on the Italian focus operator perfino*. [Manuscript submitted for publication].
[10] Kim, S. (2011). *Focus particles at syntactic, semantic, and pragmatic interfaces: the acquisition of only and even in English.* PhD dissertation. University of Hawaii.

## Sign Language Users in Estonia: Language Demography Survey

**Goal and background**. This paper provides insights into the linguistic landscape of sign language users in Estonia through a series of selected qualitative and quantitative methods. While there is no comprehensive statistical data on the linguistic diversity of sign languages in Estonia, the most recent census provides only basic data: 1,464 sign language users in Estonia, with 754 people using sign language as their first language (Lass, Vassiljeva, & Randoja, 2023), without any indication of their linguistic background and other sociolinguistic variations.

**Questions**. The paper investigates two questions concerning sign languages and their users equally: (a) what patterns can be identified in the linguistic landscape of sign language users in Estonia and (b) what are the linguistic profiles of sign language users in Estonia?

**Methodology.** The mixed method of data collection includes two quantitative and one qualitative method. The data on linguistic demography and language attitudes/choice of sign language users in Estonia were collected through an online questionnaire-based survey, while the language profile of sign language users was collected through personal interviews as an exploratory method. The collection of statistical data was complemented by on-site sessions in Tallinn, Narva, Tartu, Võru, Pärnu, Kuressaare (N=149), while the language portrait interviews took place in Tartu, Narva and Tallinn (N=15). The use of mixed methods also provides indirect data on the linguistic vitality of Estonian Sign Language through responses on language attitudes and choices (Pauwels, 2016).

**Findings.** The paper presents findings in four categories: (a) the patterns and distribution of sign languages, especially Estonian Sign Language, and other languages in Estonia; (b) the role of sign languages and languages in the lives of sign language users; (c) the general language attitudes towards sign languages and languages in Estonia; (d) the language choices of Estonian Sign Language users in three selected settings.

**Discussion**. The paper discusses the novelty of the applied methods as a triangulation method, beyond the fact that it is the first time the language demography survey in Estonia is available in Estonian Sign Language for its users. The results provide invaluable insights into (a) the linguistic landscape of the sign language communities in Estonia; (b) the distribution of Estonian Sign Language users along sociolinguistic variables; (c) the language knowledge of sign language users; (d) the language attitudes and choices of sign language users with respect to the languages in their immediate environment. In addition, the data are very useful for examining the diversity of backgrounds and needs of sign language users. Beyond practical applications, this research has broader implications for language policy, curriculum development and communication accessibility in Estonia, serving as a crucial tool for

shaping language resources, education and social planning (Verdoodt, 2017, Moreno-Fernández, 2023; Lass et al., 2023).

Keywords: Estonian Sign Language; sign multilingualism; language demography; language portrait; language attitude and choice

**References:**

Laiapea, V., Miljan, M., Toom, R., & Sutrop, U. (2002). *Eesti viipekeele seisundikirjeldus*. Haridusministeeriumi ja Eesti Keele Instituudi. https://dspace.ut.ee/server/api/core/bitstreams/a50951c3-b0cf-46a9-a9b1-b4ad03e67a97/content

Lass, H., Vassiljeva, T., & Randoja, K. (2023). *Eesti rahvastik on tunduvalt mitmekesisem kui 10 aastat tagasi*. https://rahvaloendus.ee/et/uudised/eesti-rahvastik-tunduvalt-mitmekesisem-kui-10-aastat-tagasi

Moreno-Fernández, F. (2023). *Language Demography* (1st ed.). Routledge. https://doi.org/10.4324/9781003327349

Pauwels, A. (2016). Linguistic demography: Census surveys. (Chapter 3). In *Language Maintenance and Shift* (pp. 35–47). Cambridge University Press.

## Language Resources in Estonian Sign Language Research

**Introduction.** Language resources include data, tools and standards that are essential for documenting, preserving and sharing sign language data in deaf communities, not only for research purposes, but also for educational and cultural reasons.

**Issues.** Recent developments, in particular the lexicographic work on Estonian Sign Language carried out by the Estonian Language Institute, are available as a language resource toolkit for different user groups. At the same time, a network of language resources for Estonian Sign Language that focuses on the development of sustainable and usable tools needs to be established. Moreover, more efforts needs to be made to establish discourse in Deaf and academic communities about the role of Estonian Sign Language resources.

**Proposal.** The paper explores and examines the role and functions of language resources in the research of Estonian Sign Language with the aim of revising the precedent works: (a) Visuolab (Rathmann et al., 2024) for documenting and sharing the sign language data, (b) Signbank 2.0 (Quadros et al., 2024) for organising lexicographically the signs of Estonian Sign Language, (c) transcription and annotation conventions for creating the linguistic data on sign language data, (d) management, storage and sharing of sign language data; (e) the role of metadata in the data management of sign language data.

**Discussion.** The paper discusses the role of sign language resources within the system of Estonian language resources and examines sustainable and effective ways to maintain and share these resources among the Deaf and academic communities in light of the F.A.I.R. and C.A.R.E. principles (Wilkinson et al. 2016; Caroll et al. 2020) as tools serving Digital Humanities 2.0 (Davidson, 2008). It also explores how the software and toolkits are reviewed and verified in collaboration with stakeholders from the Deaf and academic communities.

Keywords: Estonian Sign Language; language resources; Signbank 2.0; Visuolab

**References**

Carroll, S. R., Garba, I., Figueroa-Rodríguez, O. L., Holbrook, J., Lovett, R., Materechera, S., Parsons, M., Raseroka, K., Rodriguez-Lonebear, D., Rowe, R., Sara, R., Walker, J. D., Anderson, J., & Hudson, M. (2020). The CARE Principles for Indigenous Data Governance. *Data Science Journal 19*(1), 43. https://doi.org/10.5334/dsj-2020-043

Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3**, 160018 (2016). https://doi.org/10.1038/sdata.2016.18

Rathmann, C., Quadros, R. M. de, Geissler, T., Peters, C., Fernandes, F., Loio, M. P., & Franca, D. (2024). VisuoLab: Building a sign language multilingual, multimodal and multifunctional platform. In E., Efthimiou, S-E., Fotinea, T., Hanke, J. A., Hochgesang, J., Mesch, & M., Schulder (Eds.) *Proceedings of the LREC-COLING 2024 11th Workshop on the Representation and Processing of Sign Languages: Evaluation of Sign Language Resources* (pp. 282-289) https://aclanthology.org/2024.signlang-1.32.pdf

Quadros, R. M. de, Rathmann, C., Romanek, P. Z., Fernandes, F., & Condé, S. (2024). Signbank 2.0 of Sign Languages: Easy to Administer, Easy to Use, Easy to Share. In E., Efthimiou, S-E., Fotinea, T., Hanke, J. A., Hochgesang, J., Mesch, & M., Schulder (Eds.) Proceedings of the LREC-COLING 2024 11th Workshop on the Representation and Processing of Sign Languages: Evaluation of Sign Language Resources (pp. 302-314) https://aclanthology.org/2024.signlang-1.34.pdf

Davidson, C. N. (2008). Humanities 2.0: Promise, Perils, Predictions. In *Publications of the Modern Language Association of America*. 123(3):707–717.

**Comparing human and LLM predictions of English 'actually' tokens in natural speech**

This study compares large language models (LLMs) and human participants in their ability to predict a pragmatically-sensitive word meaning, specifically the English discourse marker *actually.* Pragmatic meanings convey beliefs, emotions, and discourse goals, requiring sensitivity to real-world context. Humans acquire language through social interaction and direct experience, while LLMs learn from massive textual datasets without direct exposure to real-world situations. This fundamental difference in how words are learned raises questions about the extent to which LLMs 'understand' context-dependent pragmatic meaning.

We conducted a series of cloze tests using naturally-occurring spoken corpus examples containing *actually* (extracted from the Providence corpus[1] in CHILDES[2], varying both the type and quantity of linguistic and behavioral context provided. We created six distinct conditions ranging from minimal context (target utterance only) to rich contexts which include preceding and following utterances as well as behavioral descriptions. Participants—both human (recruited via Amazon Mechanical Turk[3]) and LLMs (BERT[4], RoBERTa[5], ELECTRA[6], GPT-3.5[7], GPT-4-Turbo[8], and GPT-4o[9])—were tasked with predicting the missing word.

Fig. 1 presents the results of the cloze task. GPT-4-Turbo consistently achieved the highest accuracy, even surpassing human participants in several conditions, while RoBERTa also performed near human levels. For both LLMs and humans, accuracy increased with more linguistic context (especially preceding context) and was unaffected or even negatively affected by the presence of behavioral descriptions. However, a closer analysis showed that LLMs and humans excelled on different subsets of the data: inter-predictor agreement, measured by Krippendorff's alpha[10], demonstrated that human participants aligned closely with one another, while LLMs showed high internal consistency but low agreement with human judgments (Fig. 2). This suggests humans and LLMs are using fundamentally distinct mechanisms for predicting this meaning.

Our study involves a number of methodological innovations. Our use of spoken corpus data (albeit represented as text) for our cloze test items ensures that our results offer a high degree of ecological validity. Furthermore, we included descriptions of actual co-occurring behaviors in order to provide a representation of the physical and social context surrounding the tokens. The fact that behavioral descriptions failed to improve accuracy suggests behavioral information is not effectively represented by text, highlighting a challenge in integrating multi-modal contextual information in word prediction tasks. Finally, our study measures both accuracy as well as inter-predictor agreement between LLMs and humans; the two-pronged approach allowed us to discover important differences between LLMs and humans that are masked by similarities in overall performance.

Our study contributes to the growing discourse on the interpretive capabilities of LLMs in the pragmatic domain. By combining real corpus data, varied contextual conditions including behavioural information, and multiple measures of success, our methods revealed a nuanced picture of the strengths and limitations of LLMs on pragmatic language tasks.

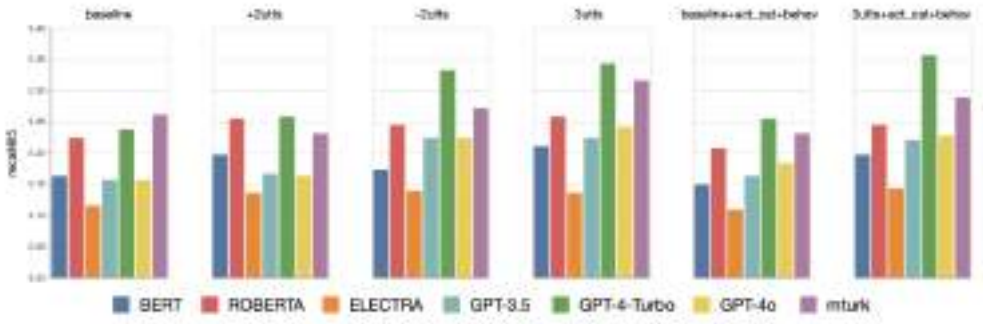



**Fig. 1.** Proportion of accurate guesses (guesses containing 'actually') for each model and human participants, in each condition. Baseline = target utterance only; +2utts = target + following utterance' -2utts = target + preceding utterance; 3utts = target, preceding and following utterance; baseline+act_cat+behav = target, activity category (e.g. 'eating lunch') and description of accompanying behaviors (e.g. 'child picks up spoon'); 3utts+act_cat+behav = target, preceding and following utterance, activity category, behavioral description

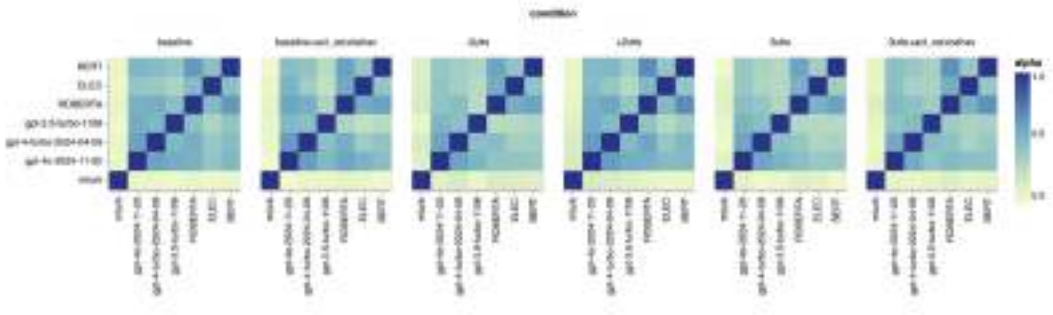


**Fig. 2.** Inter-annotator agreement (using Krippendorff's alpha) between different models and human participants

## References

[1] Demuth, K., Culbertson, J. & Alter, J. (2006). Word-minimality, epenthesis, and coda licensing in the acquisition of English. Language & Speech, 49(2), 137-174.
[2] Rose, Y., & MacWhinney, B. (2014). The PhonBank Project: Data and software-assisted methods for the study of phonology and phonological development. In J. Durand J., U. Gut, & G. Kristoffersen (Eds.), The Oxford Handbook of Corpus Phonology (pp. 380-401). Oxford University Press.
[3] Kevin Crowston. (2012). Amazon mechanical turk: A research tool for organizations and information systems scholars. In Shaping the Future of ICT Research. Methods and Approaches: IFIP WG 8.2, Working Conference, Tampa, FL, USA, December 13-14, 2012. Proceedings, pages 210–221. Springer.
[4] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.
[5] Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Man- dar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. (2019). Roberta: A robustly optimized bert pretraining approach. arXiv preprint arXiv:1907.11692.
[6] Kevin Clark, Minh-Thang Luong, Quoc V Le, and Christopher D Manning. (2020). Electra: Pre-training text encoders as discriminators rather than generators. arXiv preprint arXiv:2003.10555.
[7],[8] Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, et al. (2023). Gpt-4 technical report. arXiv preprint arXiv:2303.08774.
[9] Aaron Hurst, Adam Lerer, Adam P Goucher, Adam Perelman, Aditya Ramesh, Aidan Clark, AJ Os- trow, Akila Welihinda, Alan Hayes, Alec Radford, et al. (2024). Gpt-4o system card. arXiv preprint. arXiv:2410.21276.
[10] Klaus Krippendorff. (2018). Content analysis: An intro- duction to its methodology. Sage publications.

## Using machine translation to study contact-induced language change

*Heete Sahkai*

Language contact can induce language change, including typological change (Thomason 2001, 2003, Heine & Kuteva 2005, Matras 2009). Many languages in the world are currently subject to an unprecedented scale of contact with English as a global language (see e.g. Zenner 2023). This scale of contact is expected to give rise to interferences and contact-induced change in the languages undergoing such contact, raising the question how to discover and study these potential changes. The goal of this talk is to explore a method that allows both to test specific hypotheses and to mine for potential changes in a data-driven way. The method relies on translation as an instance of language production in a contact situation (Kranich 2014, Koetze 2020). Translated texts are thus expected to document the interferences that can arise in a particular contact situation: "While translating a text from a source language (SL) to a target language (TL), the bilingual individual must activate his/her competence in both these languages. The product of this process can exhibit an impact of features of the SL on the target text (TT)." (Kranich 2014:97)

There are potentially many ways of studying source language interferences in translated texts, for instance, by comparing corpora of original and translated texts. This talk proposes a method relying on translation engines that have been trained on parallel corpora of source language texts and their translations into the target language. The method consists in using machine translation to translate original target language texts into the source language and then back to the target language, and comparing the original texts with their backtranslations. Grammatical differences between the original texts and their backtranslations can potentially result from source language interferences that are present in the training data. This method allows for a more controlled comparison of original and translated texts than simply comparing corpora of distinct original and translated texts since the two conditions are otherwise more equal.

The talk will report a pilot study that compared head-dependent orders in 100 sentences from Estonian original fiction works extracted from the Estonian National Corpus 2023 (ENC2023, Koppel et al. 2023) and their machine-translated backtranslations from English, generated with the MTee translation engine (Bergmanis et al. 2022). More specifically, the study focused on head-dependent orders in non-finite VPs complementing nouns. Estonian is characterised by a variation in the order of heads and dependents in several contexts where English only allows head-initial order. It can thus be hypothesised that contact with English will increase the proportion of head-initial orders in Estonian. The results of the pilot study supported this hypothesis as the original sentences contained significantly less head-initial VPs than their backtranslations. The results were validated by comparing head-dependent orders in a set of 478 original and 478 translated sentences extracted from the fiction sub-corpus of the ENC2023. Again, the translated sentences contained significantly more head-initial VPs than the original sentences.

### References

Bergmanis, T., M. Pinnis, R. Rozis, J. Šlapiņš, V. Šics, B. Bernāne, G. Pužulis, E. Titomers, A. Tättar, T. Purason, H.-A. Kuulmets, A. Luhtaru, L. Rätsep, M. Tars, A. Laumets-Tättar, & M. Fishel. 2022. MTee: Open Machine Translation Platform for Estonian Government. In Proceedings of the 23rd Annual Conference of the European Association for Machine Translation, pages 309–310, Ghent, Belgium. European Association for Machine Translation.

Heine, B. & Kuteva, T. (2005). Language Contact and Grammatical Change. Cambridge University Press. https://doi.org/10.1017/CBO9780511614132

Koetze, H.. 2020. Translation, contact linguistics and cognition, in Fabio Alves and Arnt Lykke Jakobsen (eds.) The Routledge Handbook of Translation and Cognition ed.. Abingdon: Routledge, 113-132. https://www.routledgehandbooks.com/doi/10.4324/9781315178127-9

Koppel, K.; Kallas, J.; Jürviste, M.; Kaljumäe, H. (2023). Eesti keele ühendkorpus 2023. Lexical Computing Ltd. / Eesti Keele Instituut.

Kranich, S.. 2014. Translations as a locus of language contact. In J. House (Ed.), Translation: A multidisciplinary approach (pp. 96– 109). New York: Palgrave Macmillan.

Matras, Yaron. 2009. Language Contact (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.

Thomason, S. 2001. Contact-induced typological change, in Haspelmath, M., König, E., Oesterreicher, W., & Raible, W. (Eds.). Language Typology and Language Universals 2. Teilband. Berlin, Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110194265, 1640-1648.

Thomason, S. G. (2003). Contact as a Source of Language Change. In B. D. Joseph & R. D. Janda (Eds.), The Handbook of Historical Linguistics (1st ed., pp. 687–712). Wiley. https://doi.org/10.1002/9780470756393.ch23

Zenner, E. (2023). Recent Impact of English on Other Germanic Languages. Oxford Research Encyclopedia of Linguistics. Retrieved 22 Feb. 2025, from https://oxfordre.com/linguistics/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-1046.

## Online Computerized Adaptive Tests of Children's Vocabulary Development in Dutch

*Selim Sametoğlu, Magda Matetovici, Marieke van den Akker and Caroline Rowland*

Measuring children's early language skills is vital for understanding their cognitive, social, and behavioral development. The MacArthur-Bates Communicative Development Inventories (CDI) are widely used parent-report tools to assess young children's language but, their length imposes significant time burdens on participants. Shorter fixed forms have been introduced but remain inefficient, as all items must still be administered, including those irrelevant to the child's developmental level.

Computerized adaptive testing (CAT) offers a promising alternative. By dynamically selecting test items based on prior responses, CAT minimizes the number of items needed while maintaining measurement precision. Previous studies on American English, Mexican Spanish, and other languages have shown that CAT significantly reduces completion time while retaining strong correlations with full CDI scores.

Here we report on the development of a CAT version of the Dutch CDI (NL-CDI-CAT) based on the data from N = 1650 12 to 30-month-old Dutch-learning children living in the Netherlands. Using item-response theory (IRT) modeling, we identified suitable models for vocabulary comprehension and production, and conducted CAT simulations to optimize design parameters. We then collected new data (N = 60) using both CAT and traditional CDIs, together with language subscales of the Dutch-Bayley's Infant Development Scale (Bayley-III-NL) to evaluate test-retest reliability and validity (convergent and concurrent) of NL-CDI-CAT.

By reducing participant burden, the Dutch CDI-CAT offers an efficient alternative to the assessment of language development in Dutch-speaking children. This can potentially facilitate early identification and interventions for language development-related disorders.

**Distributive or collective: Which interpretation is more accessible?**

This study examines how Italian-speaking children and adults interpret a specific type of ambiguous sentence – transitive sentences with plural subjects and an indefinite object. By comparing children's performance across two linguistic domains, production and comprehension, we explore the implications of our findings for semantic representation and language development.

**Introduction** Ambiguous sentences with plural subjects can be understood as involving individual actions or as collective group actions. For example, sentences such as "the girls are carrying a ladder" can mean that each girl has her own ladder (*distributive* interpretation) or that they are all carrying a single ladder together (*collective* interpretation). Semantic theories suggest that the distributive interpretation is more complex, as it requires a semantic operator ($D$=λPλx∀y[y ≤ x ∧ Atom(y) → P(y)]) that applies a property P to each atomic subset y of a sum x (1,2,3). Across different languages and tasks, adults tend to favor collective interpretations of ambiguous sentences (4,5,6,7), whereas children's preferences vary, with studies reporting no interpretation preferences or even a tendency toward distributive readings – thus questioning the assumption of their greater complexity (6,7,8). In Italian, only one study has explored this topic (7), finding that children accept both readings in a Truth Value Judgment Task (TVJT) until the age of 10/11.

**The study** To our knowledge, no study has yet explored the production domain and thus combined evidence from production and comprehension. We investigated how Italian-speaking children and adults interpret ambiguous sentences and how they produce linguistic expressions to disambiguate them, comparing their performance across tasks from a developmental perspective. We included in our sample 31 preschoolers (16 females; age $M$=5.4, range=5;3–6;3), 44 second graders (22 females; age $M$=7.5, range=6;11–8;5) and 32 adults (15 females; age $M$=34.3). We employed similar stimuli in both tasks, that is, scenes with two pictures representing either a distributive or collective interpretation of actions familiar to young children and ambiguous between the two interpretations, such as *carrying a ladder* (Figure 1). In the production study, participants described each picture, and responses were coded as marked if including distributive (e.g., *ognuno*, "each") or collective markers (e.g., *insieme*, "together"), or unmarked. In the comprehension task, each trial was paired with a sentence with definite plurals in subject position, either unmarked (and thus ambiguous) (e.g., *le ragazze portano una scala¸* "the girls are carrying a ladder"), marked for distributivity (e.g., *le ragazze portano* ***ognuna*** *una scala¸* "the girls are carrying a ladder **each**") or marked for collectivity (e.g., *le ragazze portano* ***insieme*** *una scala¸* "the girls are carrying a ladder **together**"). Participants selected the picture that best matched the sentence. To prevent the comprehension task (where linguistic markers were presented with corresponding pictures) from influencing the production task, we administered it first in a separate session.

**Results** In the comprehension task, the distributive and collective conditions served as a baseline to ensure that children distinguished between the two readings, as confirmed by their high accuracy scores (preschoolers: $M_{DISTR}$=85%, $M_{COLL}$=95%; second graders: $M_{DISTR}$=93%, $M_{COLL}$=98%). In the unmarked condition, our main focus, a Kruskal-Wallis test revealed no significant difference between age groups (p=0.65): all groups rarely selected the distributive picture ($M_{PRESCH}$=15.05%; $M_{SECONDGR}$=12.88%; $M_{ADULTS}$=18.23%) (Figure 2). In the production task (Figure 3), mixed models of logistic regression showed that overall adults produced more linguistic markers than both groups of children (p<0.0001) and second graders more than preschoolers (p<0.0001). Moreover, adults and second graders produced more distributive than collective markers (p=0.001; p<0.0001), while preschoolers did not show a significant difference between the two (p=0.099).

**Discussion** In the comprehension domain, all three age groups favored collective interpretations of ambiguous sentences. While this was anticipated for adults, the same tendency in young children was unexpected: this is the first evidence of a collective preference for definite plural subjects in preschoolers, supporting the semantic hypothesis that distributive interpretations are not the default meaning of ambiguous sentences. The production data are compatible with this conclusion, as distributive descriptions were marked more frequently than collective ones by older children and adults. However, our results on comprehension contrast with previous ones (6,7,8), likely due to differences in experimental paradigms (e.g., TVTJs, picture selection tasks, act-out tasks) or stimuli (e.g., transitive sentences with numerals in subject position) or cross-linguistic differences (9). Finally, despite their similar comprehension patterns, younger children differed markedly from older children and adults in production, rarely using disambiguating markers. Further research is needed to establish at what stage of the concept-to-form mapping they encountered difficulties – recognizing the need for disambiguation or retrieving the appropriate linguistic markers.

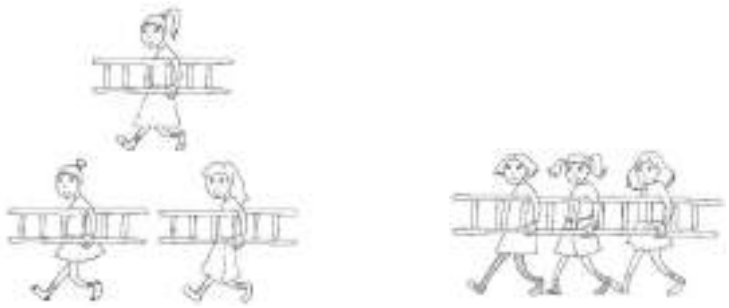


**Figure 1:** Example trial: distributive and collective pictures representing *girls carrying a ladder*.

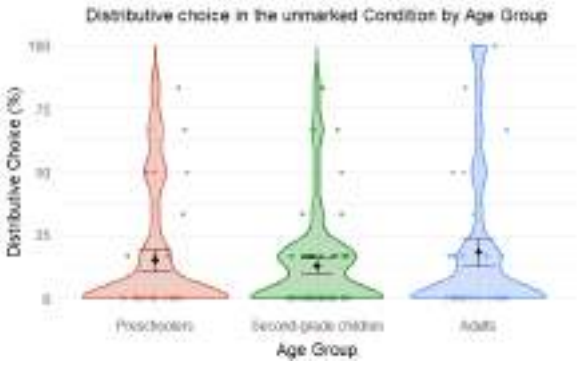



**Figure 2:** Percentages of distributive choice to match unmarked sentences by Age Group (the error bars refer to the standard error).

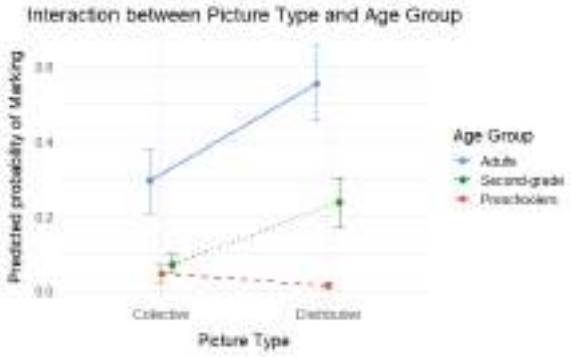



**Figure 3:** Probability of Linguistic Marking predicted by the interaction between Picture Type and Age Group (the bars display the standard error).

## References


(1) Link, G. (1983). The logical analysis of plurals and mass terms: A lattice-theoretic approach. In P. Portner, & B. H. Partee (Eds.), *Formal Semantics - the Essential Readings*, (pp. 127–147). Blackwell.

(2) Link, G. (1987). *Generalized Quantifiers and Plurals* (pp. 151–180). Springer Netherlands, Dordrecht.

(3) Champollion, L. (2016). Covert distributivity in algebraic event semantics. *Semantics and Pragmatics, 9*.

(4) Frazier, L., Pacht, J. M., & Rayner, K. (1999). Taking on semantic commitments, ii: collective versus distributive readings. *Cognition*, *70*:87–104.

(5) Dotlačil, J., & Brasoveanu, A. (2021). The representation and processing of distributivity and collectivity: ambiguity vs. underspecification. *Glossa: A journal of general linguistics*, *6*:6.

(6) Grinstead, J., Padilla-Reyes, R., & Nieves-Rivera, M. (2021). A Collective-Distributive Pragmatic Scale and the Developing Lexicon. *Language Learning and Development, 17*:1, 48-66.

(7) Pagliarini, E., Fiorin, G., & Dotlačil, J. (2012). The acquisition of distributivity in pluralities. In A. K. Biller, E. Y. Chung, & A. E, Kimball (Eds.), *Proceedings of the 36th annual Boston University Conference on Language Development* (pp. 387–399). Cascadilla Press.

(8) Syrett, K., & Musolino, J. (2013). Collectivity, distributivity, and the interpretation of plural numerical expressions in child and adult language. *Language acquisition*, *20*.

(9) Bosnić, A., & Spenader, J. (2019). Acquisition path of distributive markers in Serbian and Dutch: Evidence from an act-out task. In M. M., Brown, & B. Dailey (Eds.), *Proceedings of the 43rd Boston University Conference on Language Development* (pp. 94–108). Cascadilla Press.

Iconicity is a core property of spoken and signed languages (Perniss et al., 2010). The meanings of some iconic signs can be guessed without prior linguistic knowledge at least by some hearing participants (Ortega et al. 2019 but see Sevcikova, Sehyr & Emmorey 2019), although language knowledge considerably boosts the perception of iconicity (Occhino et al. 2017). Iconicity also may drive the evolution of grammatical structure in emerging sign languages (Aronoff et al., 2005; Sandler et al. 2011). Many sign languages show great cross-linguistic similarities in their lexicons and grammars due to overlap in iconic motivations and mappings (e.g. Currie et al. 2002, Börstell et al. 2016). Many signing communities have a large proportion of signers with variable access to sign languages and a range of sign language acquisition experiences and it has been suggested that this may impact sign language structure (e.g., Schembri et al., 2018), perhaps resulting in a greater proportion of systematically iconic and highly learnable structures. Nevertheless, the bulk of studies on the pervasiveness of iconicity in sign languages have focussed on the lexicon.

In the SignMorph Project, we are designing novel ways to try to answer questions about the nature and role of iconicity in British Sign Language (BSL) grammar. In a first set of studies, we are interested in exploring iconicity in BSL morphology by testing hearing non-signers on whether they correctly identify the meaning of morphological modifications in BSL, using both (1) a forced-choice guessing task and an (2) open response task on Prolific. We presented 100 hearing participants without sign language knowledge with pairs of signed stimuli: a base/unmodified form (or 'citation form') of a sign and the same sign form with morphological modification. Morphological modifications of verbs included directionality (modification of the base sign's initial and final location), plural marking (sweep vs. repeated movement added to the base sign), and aspect marking (fast vs. slow reduplication of the base sign), as well as non-manual modification of verbs and adjectives (puffed cheeks vs. tongue protrusion co-produced with the base sign). Combining descriptive (Figure 1) and inferential statistics, our data suggests that the meaning of BSL morphology is guessable in the forced choice task even without language knowledge. A Bayesian binomial regression analysis (random intercepts per item and participant) suggests that the accuracy of hearing non-signers in picking the correct response was reliably much better than chance for all four categories of morphological modification. In the open response task with 20 participants, however, we found lower levels of accuracy overall. Together, these findings suggest that iconicity in sign language morphology is partly accessible without sign language knowledge due to shared human cognition. Showing that hearing non-signers can access iconicity in morphological structures similar to how they are able to use it on the lexical level emphasizes the resilience of iconicity in sign languages and highlights the importance of the core property of iconicity in language emergence and evolution.

We are now designing a second set of studies looking at the relationship between iconicity and learnability in BSL morphology. Following Smith (2024), we will show participants (hearing non-signers) images of animals performing movements (e.g., a giraffe moving horizontally) accompanied by videos of BSL descriptions consisting of a lexical item followed by a classifier construction depicting both the referent and the movement. Our stimuli consist of 6 animals performing 3 movements (18 stimuli total). We will test how accurately and how quickly non-signers are able to produce the BSL morphological constructions used to describe these scenes in at least two conditions: an iconic BSL condition pairing scenes with the relevant iconic BSL descriptions and a counter-iconic condition randomly pairing BSL descriptions to scenes. We will run the experiment on Prolific Academic, collecting data from 40 participants per condition, and our analysis will follow our pre-registration. Because of the iconic properties of sign language morphology, we expect participants in the BSL condition to produce classifier constructions that are more accurate and faster than participants in the counter-iconic condition.

Together these two sets of studies represent innovative combinations of existing experimental methods to investigate the iconicity and learnability of BSL morphology.

**Are each and all the same? The development of universal quantifiers in Dutch**

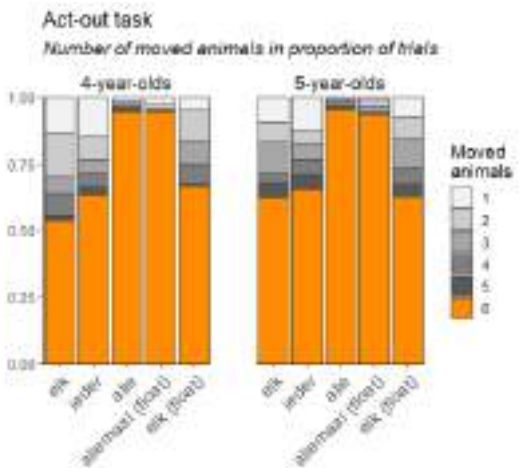
Most languages have multiple universal quantifiers (e.g., *all*, *each*), which specify slightly different meanings (e.g., Gil, 1995). In English, for example, *each pig flew to the moon* usually means that each pig individually flew to the moon, while *all pigs flew to the moon* can also mean that all pigs flew to the moon together. This difference is due to quantifier *distributivity* (e.g., Champollion, 2017): *Each*, a distributive quantifier, forces the separate application of the predicate to the individuals in the quantified set, while *all*, a non-distributive quantifier, also allows the collective application of the predicate to the set as-a-whole. Children learning their first language must figure out that there are multiple terms that express universal quantification, and that these terms differ in distributivity. What is the trajectory of this development? Do children initially treat all universal quantifiers the same and later develop sensitivity to distributivity (e.g., Brooks & Braine, 1998; Syrett, 2019)? Or do they learn that distributive quantifiers specify distributivity from the outset?

We conducted a Dutch experiment, in which we tested children's command of the non-distributive universal quantifier *alle* and the distributive universal quantifiers *elk* and *ieder* (which are synonymous). For exploratory purposes, we also included floating quantifiers, which are not adjacent to the noun they modify (e.g., *The pigs have all flown to the moon*), using floating *elk* and floating *allemaal* (the floating form of *alle*). Our experiment consisted of an act-out task and a truth-value judgment task, which were administered in separate sessions. In the act-out task, children moved toy animals based on prompts like *Nu gaan alle varkens naar de wei* ("Now, all pigs go to the field"). In the truth-value judgment task, children evaluated sentences like *Hier staan alle varkens in de wei* ("Here, all pigs are in the field") against matching and non-matching pictures. In both tasks, we varied the quantifier in the prompt.

Seventy-six children participated in our experiment (age: 4;0-5;9). Data collection of the act-out task is complete (Figure 1). These data show that the non-distributive quantifiers (*alle* and floating *allemaal*) are more often interpreted as universal quantifiers than the non-distributive quantifiers — a pattern confirmed by mixed-effects model analyses (Tables 1-2). Data processing for the truth-value judgement task is still ongoing, so we have not yet conducted inferential analysis for this task. Descriptively, however, the preliminary results of this task corroborate the pattern in the act-out task: When the picture *matched* the sentence, non-distributive quantifiers are accepted *more* often. When the picture *mismatched* the sentence, non-distributive quantifiers are accepted *less* often. Moreover, both tasks show no clear differences between determiner and floating quantifiers or any strong age effects. This suggests that children acquire *alle* and floating *allemaal* as universal quantifiers before age four, while the acquisition of *elk* and *ieder* is protracted.

Our results suggest that children acquire non-distributive universal quantifiers before distributive ones. One explanation is that children may lack semantics of the distributive *elk* and *ieder*, which are low-frequent words (anonymous, *in prep*). Another explanation is that children may understand *elk* and *ieder* as distributive quantifiers and struggle with their interpretation in the context of our experiment. To test the latter hypothesis, we will conduct a follow-up experiment with an act-out task (in which children give hay bales to individual animals, using prompts *Nu krijgt elke koe hooi* 'Each cow gets hay') and a truth-value judgment task (in which children evaluate sentences like *Hier heeft elke koe hooi* 'Each cow has hay'). In this experiment, the context involves a one-to-one correspondence between the individuals in the quantified set ('cows') and the predicated property ('hay'), and thus promotes a distributive reading. If children understand that *elk* and *ieder* are distributive, they may understand these terms better in the context of this task. We start data collection soon, so we can present these results at the MEDAL conference.

Figure 1: Act-out task results




*Note*. In the act-out task, children were given six toy animals. On each trial, they were asked to move these toy animals to a field, with prompts like 'Nu gaat elk varken naar de wei' (Now, each pig goes to the field). These results show that the children moved all six toy animals when prompted with the non-distributive quantifiers ‘alle’ or ‘allemaal (float)’, even though the other quantifiers (which are all distributive) also express universal quantification. Moreover, there is no clear descriptive effect of age. Here, we represent age as categorical for ease of illustration.

Table 1: Logit mixed-effect model analyses, act-out task

| | Df | $\chi^2$ | p |
|---|---|---|---|
| Analysis 1: alle, elk, and ieder conditions | | | |
| Quantifier | 2 | 19.90 | < 0.001 |
| Age | 1 | 2.04 | 253 |
| Quantifier x Age | 2 | 1.02 | 602 |
| Analysis 2: alle, allemaal (floating), elk, and elk (floating) conditions | | | |
| Quantifier | 3 | 9.78 | 20 |
| Age | 1 | 2.15 | 143 |
| Quantifier x Age | 3 | 0.41 | 939 |

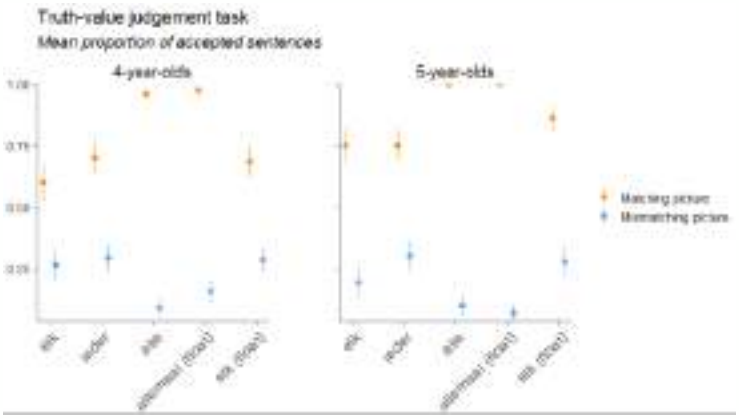
*Note*. These outcomes report model mixed-effect model comparisons, assessed using $\chi^2$ tests on the log-likelihood of different models in which fixed effects were removed one by one. Age was a continuous variable, centred around the mean, and Quantifier was sum-coded. The random-effects structure included random slopes for Quantifier by Participant and Item as well as random intercepts by Participant and Item (the most maximal model that converged). We first analysed the data in the *alle*, *elk* and *ieder* conditions to assess differences between the determiner quantifiers. Then, we analysed the data in the *alle*, *allemaal* (floating), *elk*, and *elk* (floating) conditions to assess differences between determiner and floating quantifiers.

Table 2: Pairwise comparisons Quantifier, act-out task.

| Comparison | SE | z-ratio | p-value |
|---|---|---|---|
| Analysis 1: alle, elk, and ieder conditions | | | |
| alle-elk | 2.69 | 4.00 | < 0.001 |
| alle-ieder | 2.62 | 3.86 | < 0.001 |
| elk-ieder | 0.36 | -1.16 | 248 |
| Analysis 2: alle, allemaal (floating), elk, and elk (floating) conditions | | | |
| alle-allemaal (floating) | 2.79 | 0.66 | 509 |
| elk-elk (floating) | 0.58 | -1.74 | 163 |
| alle-elk | 2.57 | 2.83 | 27 |
| alle-elk (floating) | 2.55 | 2.46 | 69 |
| elk-allemaal (floating) | 2.26 | 2.41 | 69 |
| allemaal (floating)-elk (floating) | 2.24 | 1.98 | 143 |

*Note*. The pairwise comparisons were carried out using estimated marginal means. *p*-values were adjusted using the Holm method.

Figure 2: Preliminary truth-value judgement task results




*Note*. In the truth-value judgement task, children were asked to evaluate sentences like ‘Hier staat elk varken in de wei’ (“Here, each cow is in the field”) against matching or mismatching pictures. These preliminary results suggest that children’s accuracy is higher in the all and allemaal (float) conditions compared to the other conditions (which involve non-distributive quantifiers): They accept these sentences more often when the picture is matching and less often when the picture is mismatching.

**References**

Brooks, P. J., & Braine, M. D. (1996). What do children know about the universal quantifiers all and each?. *Cognition*, *60*(3), 235-268.

Champollion, L. (2017). Parts of a whole: Distributivity as a bridge between aspect and measurement (Vol. 66). Oxford University Press.

Gil, D. (1995). Universal quantifiers and distributivity. In *Quantification in natural languages* (pp. 321-362). Dordrecht: Springer Netherlands.

Syrett, K. (2019). Distributivity. In C. Cummins & N. Katsos (Eds.), The Oxford Handbook of Experimental Semantics and Pragmatics (1st ed., pp. 143–155). Oxford University Press.

The success of Neural Language Models on syntactic island effects is not universal: strong *wh*-island sensitivity in English but not in Dutch

**Introduction.** A much-debated question in linguistics is how humans acquire grammatical knowledge: are we born with a language-specific learning capacity, or can we learn language from input alone? The recent introduction of neural language models (NLMs) can greatly influence this debate. NLMs learn solely from their input in combination with their inductive biases, and thus without any built-in linguistic representations. If these networks can learn specific grammatical phenomena, this suggests that these phenomena can, in principle, be learned without language-specific learning biases. Recent research has looked at the learnability of one of the most studied phenomena in experimental syntax, *syntactic island effects* (see example in Table 1), to investigate whether NLMs are sensitive to island violations (e.g., Wilcox et al. 2024). Syntactic islands (e.g., *wh*-phrases) are structures in which formation of filler-gap dependencies is blocked, and are an ideal test bed here because they rarely occur in training data and NLMs do not have built-in linguistic knowledge to fall back on. Research has mostly shown successful results: NLMs seem able to model island effects in English, but the correspondence between model and human behavior on islands is only assumed. In existing research, the behaviors of NLMs are almost never compared to human data and are almost exclusively researched in English, which makes it difficult to claim that NLMs can model island effects in ways that are comparable to humans. The present study addresses these gaps by incorporating actual data from human experiments and by looking beyond English.

**Methods.** We make two important improvements on earlier work. First, we present an NLM and human participants with the same sentences (Table 1). By collecting both model-assigned sentence probabilities and participant acceptability judgments, we directly compare whether the model represents island sensitivity similarly to humans. Second, we take this approach beyond English and compare NLM and human behavior in both English and Dutch, since the languages, though related, differ in their word order (SVO vs. SOV). Moreover, we are currently running similar experiments in Turkish, a language that is more morphologically complex than English and Dutch and has a flexible word order (we expect to have these results at the conference).

**Results.** The results for English and Dutch are shown in Figure 1. Figure 1 (top) shows that the strong *wh*-island sensitivity of NLMs in English is replicated and that this sensitivity is comparable to that of English participants: the NLM and the participants show the same patterns in their results. The same cannot be said for Dutch, however (see bottom of Figure 1). While the Dutch participants showed a strong sensitivity to *wh*-island violations, with patterns comparable to the English participants, the sensitivity of the Dutch model was not statistically significant (although in the right direction).

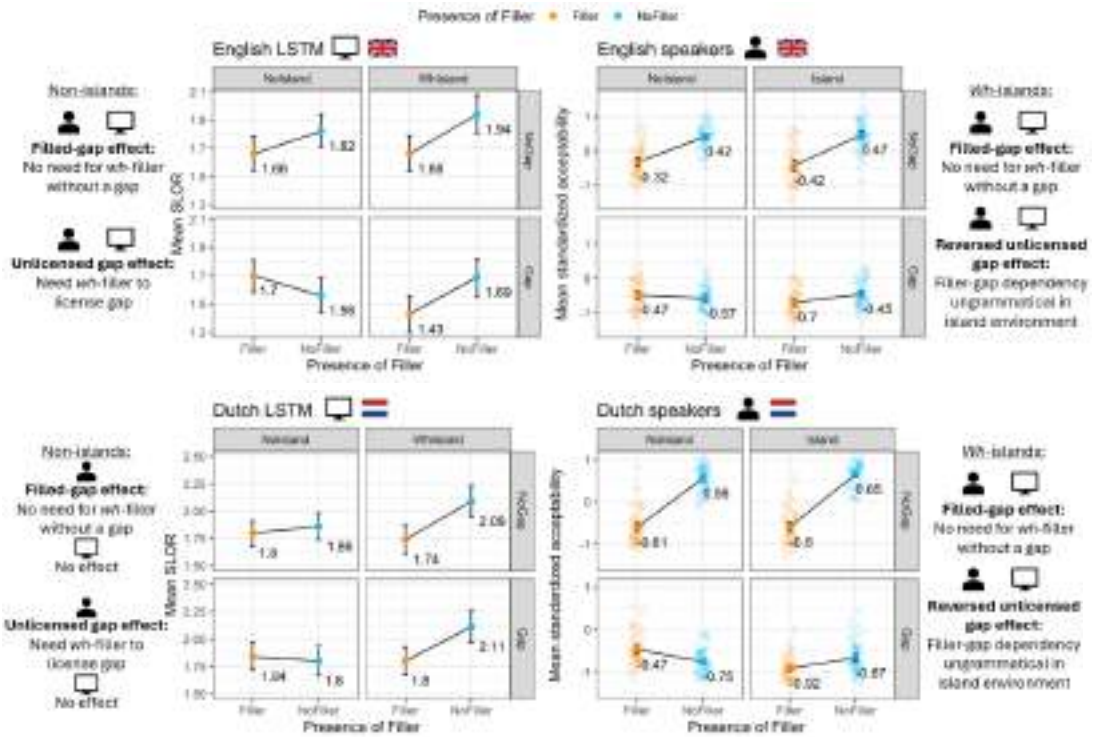
**Conclusion.** NLMs are not successful in all languages (yet) (e.g., Dutch), so more cross-linguistic research is necessary before NLMs can be claimed to bear on the human capacity for grammar learning.

**Table 1.** *Example sentences used in the Dutch (NL) and English (EN) experiments, crossing the factors* PRESENCE OF GAP *('cookies' vs. _ (a gap); indicated in orange), and* PRESENCE OF FILLER *('that' vs. 'what'; indicated in blue) in non-islands and wh-islands (indicated in red between curly brackets).*

| Gap? | Filler? | | Example sentence in non-island and ***wh*-island** configuration |
|---|---|---|---|
| No | No | **NL** | *Ik weet dat jij {denkt dat/betwijfelt of} de bakker koekjes maakt in de bakkerij.* |
| | | | I know that you {think that/doubt if} the baker cookies makes in the bakery |
| | | **EN** | I know that you {think that/doubt whether} the baker makes cookies in the bakery. |
| No | Yes | **NL** | **Ik weet wat jij {denkt dat/betwijfelt of} de bakker koekjes maakt in de bakkerij.* |
| | | | I know what you {think that/doubt if} the baker cookies makes in the bakery |
| | | **EN** | *I know what you {think that/doubt whether} the baker makes cookies in the bakery. |
| Yes | No | **NL** | **Ik weet dat jij {denkt dat/betwijfelt of} de bakker _ maakt in de bakkerij.* |
| | | | I know that you {think that/doubt if} the baker GAP makes in the bakery |
| | | **EN** | *I know that you {think that/doubt whether} the baker makes _ in the bakery. |
| Yes | Yes | **NL** | *Ik weet wat jij {denkt dat/*betwijfelt of} de bakker _ maakt in de bakkerij.* |
| | | | I know what you {think that/doubt if} the baker GAP makes in the bakery |
| | | **EN** | I know what you {think that/*doubt whether} the baker makes _ in the bakery. |

*Note*. The Dutch and English sentences only differ in the object-verb order in the embedded sentence (*koekjes maakt* vs. 'makes cookies').

**Figure 1.** *Mean standardized acceptability judgements (right plot) and mean Syntactic Log-Odds Ratio value (i.e., frequency- and length-corrected surprisal; left plot) for every combination of* PRESENCE OF GAP *and* PRESENCE OF FILLER *within non-islands (top and bottom left) and* wh*-islands (top and bottom right) for English (top plots) and Dutch (bottom plots). Error bars represent 95% confidence intervals.*




## References

Wilcox, E.G., Futrell, R., & Levy, R. (2024). Using Computational Models to Test Syntactic Learnability. Linguistic Inquiry, 55(4), 805-848. https://doi.org/10.1162/ling_a_00491

**Narrative Arcs in Computational Literary Studies: Methods and Perspectives with Examples from the Field of Computational Drama Analysis**

The concept of the narrative arc fundamentally shapes how we understand stories, yet it lacks a clear and consistent definition. In computational literary studies (CLS), different methodological approaches uncover different arcs. For this reason, I argue that, despite the general and vague use of the term, it is important to examine different methodologies in detail and in relation to each other.

There are two main strands of research within CLS that attempt to quantify the narrative arc: those focusing on linguistic-semantic changes and those examining formal changes. Linguistic approaches can themselves be further subdivided. One line of inquiry investigates thematic shifts in a text using methods such as word or document embeddings and topic modeling. These techniques can reveal how rapidly a text transitions between topics, the semantic distance of those shifts, and the degree of circularity in the narrative structure (Toubia 2021, Schmidt 2015).

Another line of semantic analysis involves tracking the temporal distribution of predefined word groups—as opposed to those generated through text mining techniques such as topic modeling. One of the most popular and widely used methods for identifying narrative arcs is the analysis of changes of sentiment in texts. This approach is based on the assumption that fluctuations in emotional tone across a narrative correlate with its structure—revealing plot arcs through patterns of rising and falling sentiment, which could indicate changes in narrative tension and the fate of the heroes. At the same time, the distribution of other word groups can also be examined in order to analyze the development of the plot.

It is also possible to focus on formal changes in a text. Prose fiction has received relatively little attention from this perspective so far. However, examples can also be found here, mainly in the trends observed in sentence length. In contrast, computational drama analysis often examines the formal structure of plays mostly through the study of character networks (see Fischer et al. 2017, Szemes and Nagy 2025, Szemes 2025).

This poster not only reviews these methodological strands but also presents findings from my own research in computational drama analysis. By combining sentiment analysis with the temporal study of character networks, I show how narrative arcs can be detected in dramatic texts. The results demonstrate that such arcs can illuminate key structural and thematic aspects of narratives, providing new perspectives for interpretation. Moreover, they also reveal meaningful genre-level differences and similarities.

**Bibliography**

Boyd L., Ryan, Kate G. Blackburn, and James W. Pennebaker. ‘The Narrative Arc: Revealing Core Narrative Structures through Text Analysis’. *Science Advances* 6, 32 (2020).

Elkins, Katherine, *The Shape of Stories*, Cambridge UP, 2022.

Frank Fischer et al., ’Network Dynamics, Plot Analysis. Approaching the Progressive Structuration of Literary Text’, *DH2017* (2017) https://dh2017.adho.org/abstracts/071/071.pdf.

Jockers, Matthew. 'Introduction to the Syuzhet Package' *CRAN* (2020), https://cran.r-project.org/web/packages/syuzhet/vignettes/syuzhet-vignette.html.

Reagan, Andrew J., Lewis Mitchell, Dilan Kiley, Christopher M Danfort, and Peter Sheridan Dodds. 'The Emotional Arcs of Stories Are Dominated by Six Basic Shapes'. *EPJ Data Sience* (2016). https://doi.org/10.1140/epjds/s13688-016-0093-1.

Schöch, Christoph, 'Topic Modeling Genre: An Exploration of French Classical and Enlightenment Drama', *DHQ* 11 (2017), 2. https://www.digitalhumanities.org/dhq/vol/11/2/000291/000291.html.

Szemes, Botond, Mihály Nagy, 'Temporal Aspects of Structural Differences in Dramatic Genre", *Zeitschrift für digitale Geisteswissenschaften* 10 (2025). DOI: 10.17175/2025_007

Szemes Botond, 'Mapping the Dramaturgical Arc: Computational Approaches to Shakespeare' in: *Conference Reader: Second Workshop on Computational Drama Analysis (Berlin, 03.09.2025)*: https://zenodo.org/records/16814035

**Word learning of metaphors in space and time domains**

Children struggle to learn temporal word meanings that do not refer to observable referents.[1] As an underlying mechanism, polysemy could facilitate learning such abstract meanings by mapping them onto familiar meanings in a related and concrete domain.[2-4] Prior work on how word meanings are learned and metaphorically extended revealed that children learned the meaning of a novel word better in the space domain than in the sound domain. Once they learn the meaning of the novel word, they can extend the meaning to the other domain equally well.[4] However, these domains do not rely on shared conceptual structures to the same extent as do space and time.[5] To better understand the learning mechanism, we investigated the differences in learning and extending word meanings as a function of the domain (space, time). Here, we ask to what extent shared conceptual structures between space and time scaffold how spatial and temporal meanings are learned and extended.

In Study 1, our preregistered sample consisted of Turkish speaking 4-5-year-olds ($n$ = 40, $M_{age}$ = 4.90) and 6-to-7-year-olds ($n$ = 38, $M_{age}$ = 4.94). Photos of eight objects and videos of eight events with short vs. long versions were used to teach a novel word. We manipulated the direction of metaphorical extension and randomly assigned children to space-to-time or time-to-space directions. During word training, children were taught either the spatial meaning of the novel word "*femi"* (meaning long) with long vs. short objects in the space domain or the temporal meaning of the novel word with long vs. short events in the time domain. We tested to what extent children could learn the meaning of the novel word in the trained domain and extend the meaning to an untrained domain.

For word learning, we tested whether abstract temporal meanings would be more difficult to learn than spatial meanings. A *glmer* model revealed a fixed effect of Domain (space, time; β = -3.106, *SE* = 0.485, $p < .001$) and interaction between Domain and Age (4-5-year-olds, 6-to-7-year-olds; β = -1.974, *SE* = 0.937, $p = .035$; Fig 1A), indicating that children were more accurate in space domain than time, but this difference was larger for 4- to 5-year-olds than 6- to 7-year-olds. In fact, 4-to 5-year-olds performed at chance level in the time domain ($p = .130$). For word extension, we tested whether word extension reflects the dominant direction in language, such that extending the meaning from space to time would be easier than from time to space. A *glmer* revealed a fixed effect of Age (β = -0.860, *SE* = 0.356, $p = .016$; Fig 1B). There was no significant difference in Direction (space-to-time, time-to-space) or an interaction between Direction and Age. However, only 6- to 7-year-olds in the space-to-time direction performed above chance level ($p < .001$), indicating that although there was no difference, children struggled to extend the meaning to an untrained domain.

In Study 2, we reduced the cognitive load in temporal discrimination to test whether the difficulty of learning the temporal meaning is related to the differences in spatial and temporal processing. To do so, we presented video recordings of events simultaneously in the time domain. Our preregistered sample consisted of Turkish speaking 4- to 5-year-olds ($n$ = 40, $M_{age}$=5.17) and 6- to 7-year- olds ($n$ = 40, $M_{age}$=7.11). For word learning, A *glmer* model revealed a fixed effect of Domain (β = -1.631, *SE* = 0.635, $p = .010$) and Age (β = -1.630, *SE* = 0.612, $p = .007$; Fig 2A). Importantly, both age groups performed above chance level in both domains. For word extension, A *glmer* revealed a fixed effect of Age (β = -2.697, *SE* = 0.637, $p < .001$; Fig 2B). There was no significant difference in Direction or an interaction between Direction and Age. However, 4- to 5-year-olds still performed at chance level both in space-to-time ($p = .056$) and time-to-space directions ($p = .051$).

In both studies, children learned the spatial meaning of the novel word better than the temporal meaning. Reducing the cognitive load of temporal discrimination supported children in learning and extending the temporal meaning, but it did not eliminate the difference between spatial and temporal meaning or developmental differences between age groups. Once 6- to 7-year-olds learned the meaning, they could extend it to the untrained domain equally well. These results suggest that polysemy supports learning abstract word meanings by mapping them onto familiar meanings in a related and concrete domain.[2-4]

## Figures

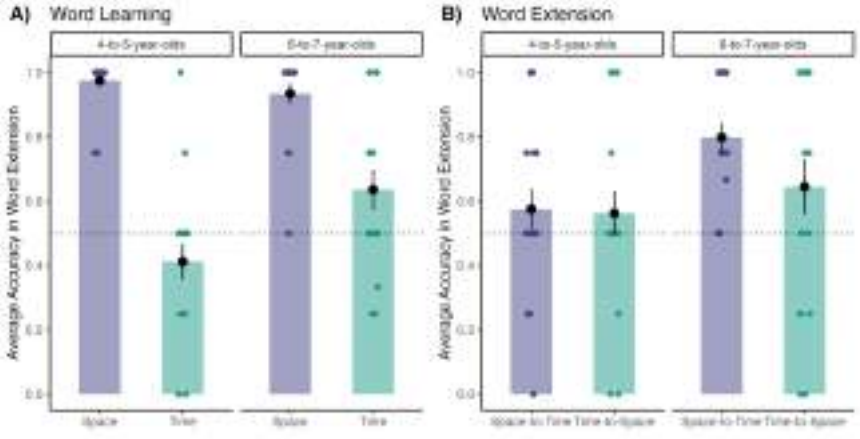



Figure 1. Mean Accuracy for (A) Word Learning in Space and Time Domains and (B) Word Extension in Space-to-Time and Time-to-Space Directions in Study 1

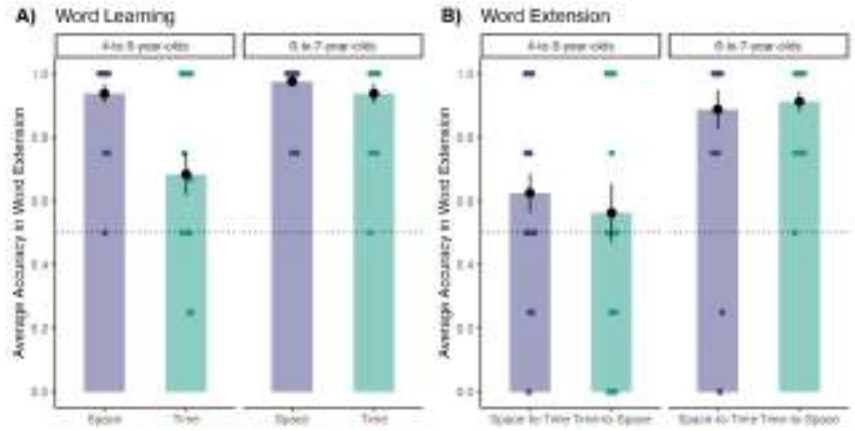



Figure 2. Mean Accuracy for (A) Word Learning in Space and Time Domains and (B) Word Extension in Space-to-Time and Time-to-Space Directions in Study 2

## Reference

1. Tillman, K. A., & Barner, D. (2015). Learning the language of time: Children's acquisition of duration words. *Cognitive Psychology*, *78*, 57–77. https://doi.org/10.1016/j.cogpsych.2015.03.001
2. Srinivasan, M., & Rabagliati, H. (2015). How concepts and conventions structure the lexicon: Cross-linguistic evidence from polysemy. *Lingua, 157*, 124-152. https://doi.org/10.1016/j.lingua.2014.12.004
3. Clark, H. H. (1973). Space, time, semantics and the child. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language*. Academic Press.
4. Starr, A., Cirolia, A. J., Tillman, K. A., & Srinivasan, M. (2020). Spatial metaphor facilitates word learning. *Child Development*, *92*(3), e329-e342. https://doi.org/10.1111/cdev.13477
5. Srinivasan, M., & Carey, S. (2010). The long and the short of it: On the nature and origin of functional overlap between representations of space and time. *Cognition*, *116*(2), 217–241. https://doi.org/10.1016/j.cognition.2010.05.005

**What hides behind the covert expression of perception in traditional storytelling: A contrastive corpus-based study**

Languages vary widely in the way they encode sensory information and in how systematically they do it for particular senses. Perceptual experiences can be expressed lexically by perception verbs (Aikhenvald & Storch 2013; San Roque et al. 2015, inter alia), 'depictive' devices such as ideophones and onomatopoeia (Dingemanse 2011), and demonstratives (Evans & Wilkins 2000), or grammatically through sensory evidentials (Aikhenvald 2004). In addition, the expression of perception can remain implicit in the narrative context without overt reference to the specific sense. For instance, the verb 'laugh' in (1) implies the perception of the sound by characters in the story. In a similar manner, explicit mention of indexes such as footprints or traces, as in (2), suggests their perception by a character, involving vision. Often, covert expression of perception can be found in speech and thought, attributed to characters in the story, as in (3).

(1) *bé* *gbógló* ***lɔ̄gɔ́*** *bé* *bɔ̀lè* *é* *gnù*
then hyena laugh then bird DEF fly
'Then the Hyena **laughed**, and the bird flew away.' Wan (Mande; Côte d'Ivoire)

(2) *è* *zō* *è* *nɛ̄* *ē* ***cɛ̰́*** ***bī*** *è* *glā* ***tɑ̃́íí***
3SG came 3SG child DEF foot trace DEF took until
'He went [and] followed his child's **footprints**.' Wan (Mande; Côte d'Ivoire)

(3) *Ou,* *komənsəj,* *s'ajbə* *kojkaʔa* *taharıabiʔ* *ŋanuəmənɨ*
INTERJ INTERJ seven idol:AUG now real:ADV
*hütəðüt'üŋ* *d'üðɨ* *d'ermənɨ* *təniʔıarıaiʔ* *kən'it'ürübəðatəʔ.*
real:ADV period:GEN middle:PROLAT so:LIM:ADV cut:DRV:PASS:INFER:3PL.R
'(He climbed the mountain.) "Ah, it's rather interesting. Here are seven idols cut in the middle of their bodies."' Nganasan (Uralic; Northern Siberia)

A corpus-based study of narratives in five understudied languages from three linguistic areas (Chuvash, Nganasan, Selkup, Udihe, and Wan) shows that overt expression of perception by lexical means is predominant for vision, while auditory perception more often remains implied in the context (AUTHORS submitted). In this study, we explore possible reasons for this asymmetry in the expression of auditory and visual perception, focusing specifically on narrative factors.

Covert expression of perception tends to occur in contexts associated with the internal focalization of perceptual experience, conceptualized as a primitive contextual model 'I came – I saw – I understood – I said: WOW!' (Skribnik 2023: 240). Such focalization often precedes an important narrative part, or opens a new scene, as in (3). We use a combination of criteria for narrative analysis from Labov (1972) and for thematic unity from van Dijk (1981) to annotate a corpus of narratives for episode boundaries and check whether the type of perceptual expression correlates with its position in the narrative structure.

Our preliminary results show that overt lexical encoding of visual perception often occurs at the beginning of new episodes, while covert expression precedes particularly important and often surprising narrative developments. Visual but not auditory perception is often evoked in referring to events and characters in previous narrative parts. Such references are more likely to be described explicitly but their presence does not correlate with a particular position in the narrative structure. Thus, we also identify cases in which overt reference to perception may happen independently of its position in the narrative structure but may be caused by the additional functions of highlighting important narrative parts through their more frequent mention in different contexts.

More generally, our study illustrates how new methods for data annotation and analysis can be used for quantitative cross-linguistic comparison of corpus data from understudied languages.

**References**

Aikhenvald, A. Y. 2004. *Evidentiality*. Oxford: Oxford University Press.

Aikhenvald, A. Y. & A. Storch. 2013. Linguistic expression of perception and cognition: A typological glimpse. In: AIkhenvald, A. Y. & A. Storch (eds.), *Perception and Cognition in Language and Culture*, 1–46. Leiden: Brill.

Dijk, van T. 1981. Episodes as units of discourse analysis. In: Tannen, D. (ed.), *Analyzing discourse: Text and talk*, 177–195. Georgetown: Georgetown University Press.

Dingemanse, M. 2011. *The Meaning and Use of Ideophones in Siwu.* Nijmegen: Max Planck Institute for Psycholinguistics.

Evans, N. & D. Wilkins. 2000. In the mind's ear: The semantic extensions of perception verbs in Australian languages. *Language* 76(3), 546–592.

Labov, W. 1972. *Language in the inner city: Studies in the Black English vernacular*. Philadelphia: University of Pennsylvania Press.

San Roque, L., K. H. Kendrick, E. Norcliffe, P. Brown, R. Defina, M. Dingemanse, T. Dirksmeyer, N.J. Enfield, S. Floyd, J. Hammond, G. Rossi, S. Tufvesson, S. van Putten & A. Majid. 2015. Vision verbs dominate in conversation across cultures, but the ranking of non-visual verbs varies. *Cognitive Linguistics* 26, 31–60.

Skribnik, E. 2023. Reading Siberian folklore: Miratives, pre-mirative contexts and "Hero's Journey". In: Aikhenvald, A. Y., R. L. Bradshaw, L. Ciucci & P. Wangdi (eds.), *Celebrating Indigenous Voice,* 237–261. Berlin & Boston: Mouton de Gruyter.

Teenagers' use of new quotatives in spoken Estonian: A combined quantitative and qualitative analysis

**Abstract**
European languages are known for having developed 'new quotative' strategies consisting of etymologically non-reportative markers like similative/comparative adpositions (e.g. Eng. *like*, Ru. *tipa* 'like'), demonstratives (Ger. *so*), and quantifiers (Eng. *all*, Swe. *ba(ra)* 'just, only') (Buchstaller & van Alphen 2012: xiv). Colloquial Estonian shows evidence of a similar development in both online communication (Teptiuk 2019: 273–277) and spoken language (Kängsepp et al. 2022), employing lexical sources found in other European languages (e.g. similative *nagu* in [1]). Additionally, a lexical source not previously discussed is also used in quotative constructions, namely, the indefinite pronoun *mingi* 'some [kind of]' as shown in (2).

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| (1) | *ja* | *siis* | ***ma*** | ***ol-i-n*** | ***nagu*** | oh my god | *nii* | creepy |
| | and | then | 1SG | be-PST-1SG | like | INTERJ | so | creepy |
| | 'and then **I was like** oh my god, so creepy' | | | | | | | |
| (2) | *õhtu-l* | *läh-me* | | ***mingi*** | *tere* | *Maarika* | | |
| | evening-ADE | go-PRS.1PL | | some.kind.of | hello | PN | | |
| | 'in the evening, we could go [and be] like hi, Maarika!' | | | | | | | |

Although new quotative strategies in colloquial written Estonian have been investigated in previous studies, there are no systematic attempts to approach this topic in spoken Estonian. This study addresses this gap and describes quotative strategies used in colloquial spoken Estonian, with a focus on the quotative use of the similative marker *nagu* 'like' and the indefinite pronoun *mingi* 'some [kind of]'. We use data from the Estonian teenagers' spoken corpus, consisting in self-recorded oral conversations between two or more speakers aged 10–18 (Koreinik et al., 2023). We use two data samples: the first includes all instances of reported discourse observed in ten representative recordings (470 examples, ~8 hours). The second sample includes all uses of *mingi* and *nagu* from the spoken corpus (79 transcribed recordings, total duration ~59 hours) and taking a random sample balanced for age and gender of the speakers for both markers. This yielded 1068 observations of *nagu* and 1083 observations of *mingi*, of which we found 8% and 6% quotative uses, respectively.

We first present an overview of quotative constructions observed in the first data sample, including conventional strategies, new quotatives, and 'defenestration' where the quotative construction is formally unrepresented (Spronck 2017). Second, we use the two samples and discuss quotative uses of *nagu* and *mingi*. In terms of form, we focus on the morphosyntactic contexts of each marker and their combination with other markers within a quotative strategy. Functionally, we investigate how these markers are used with reported discourse: in quotations and self-quotations, introducing reported thought, non-verbal demonstrations, as well as reported speech. Finally, we discuss the value of a combined approach based on both quantitative and qualitative analysis.

**References**

Koreinik, K., Mandel, A., Pilvik, M.-L., Praakli, K., Vihman, V.-A. 2023. Outsourcing teenage language: A participatory approach for exploring speech and text messaging. *Linguistics Vanguard.* Doi: 10.1515/lingvan-2021-0152

Kängsepp, A., Mandel, A., Pilvik, M.-L., Tenisson, A., Vihman V. 2022. 'Parasite word or conversational tool? The use of *nagu* 'like' and *mingi* 'something' in teen talk.' Presented on 13.10.2022, Tartu, Estonia.

Spronck, S. 2017. Defenestration: deconstructing the frame-in relation in Ungarinyin, *Jrnl of Pragmatics* 114, 104–133.

Teptiuk D. 2019. *Quotative indexes in Finno-Ugric (Komi, Udmurt, Hungarian, Finnish and Estonian),* University of Tartu dissertation.

## Abstract

**Challenges in the Qualitative Analysis of Texts: A Comparison of Five AI Models for Evaluating Affective Tonality in University Students' Self-Reported Attitudes Toward AI Applications**

*Andrus Tins*

Following the public release of ChatGPT on November 30, 2022, public discourse in Estonia became notably polarized in its response to artificial intelligence (AI). Preliminary studies suggest that this polarization began to soften within six months as the general public became more accustomed to the once-novel concept of AI technologies. Nevertheless, by spring 2023, the comment sections of major Estonian news outlets continued to reflect a pronounced divide between pro- and anti-AI sentiments (Tins, 2023). At that time, AI remained a relatively novel concept for majority of the general population, and its rapid adoption provoked skepticism and emotionally charged reactions.

The present study builds on this context by examining affective responses to AI among university students in Estonia. Designed as an experimental investigation, the study explores the potential of AI-assisted tools for the qualitative interpretation of textual data within the humanities.

The dataset comprises responses from 82 full-time undergraduate students (median age: 21) enrolled in cultural and creative disciplines at an Estonian university, collected in autumn 2024. Participants were asked to reflect on whether and how they had used AI applications for personal benefit, and were encouraged to elaborate more broadly on their experiences and opinions.

The data were analyzed using five model–scale combinations: OpenAI's GPT-4o and GPT-3.5 models, each applied with the PANAS (Positive and Negative Affect Schedule; Watson, Clark, & Tellegen, 1988) and VAD (Valence–Arousal–Dominance; cf. Bradley & Lang, 1994; Mohammad, 2018) affective frameworks; and the Estonian-language classifier tartuNLP/EstBERT128-sentiment (Tanvir et al., 2021). To provide a comparative reference, two human coders independently analyzed the same dataset using the exact same instructions given to the GPT models (including definitions of the PANAS and VAD scales, numerical labels, and scoring rules).

The findings suggest that students remain affectively divided in their attitudes toward AI, as indicated by the relatively low incidence of neutral sentiment classifications. Notable differences emerged between models. The EstBERT sentiment classifier, for example, identified approximately two-thirds of the texts as expressing a *negative* net sentiment. In contrast, the GPT-3.5-based models—using both the PANAS and VAD frameworks—classified about two-thirds of the texts as conveying a *positive* net emotional

tone. While the GPT models are not specifically designed for sentiment classification in Estonian language, this study aimed to explore their analytical performance under highly structured and consistent prompting conditions.

**References**

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The Self-Assessment Manikin and the Semantic Differential. Journal of Behavior Therapy and Experimental Psychiatry, 25(1), 49–59.

Mohammad, S. M. (2018). Obtaining reliable human ratings of valence, arousal, and dominance for 20,000 English words. In Proceedings of the Annual Conference of the Association for Computational Linguistics.

Tanvir, H., Kittask, C., Eiche, S., & Sirts, K. (2021). EstBERT: A Pretrained Language-Specific BERT for Estonian. In Proceedings of the 23rd Nordic Conference on Computational Linguistics (NoDaLiDa) (pp. 11–19).

Tins, A. (2023). Narrating Artificial Intelligence (AI) in 2023: An Estonian Case Study on AI Lore. Yearbook of Balkan and Baltic Studies. 253-282. 10.7592/YBBS6.12.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. Journal of Personality and Social Psychology, 54(6), 1063–1070. DOI: 10.1037/0022-3514.54.6.1063

## Heritage Estonian use in Estonian-Norwegian bilingual children: a closer look at narratives

*Adele Vaks*

Heritage language is a language that is acquired at home, in the context of a different societal language. As such, heritage language research offers a fascinating window into the relationship between language exposure and language use. This talk explores the patterns of language exposure and use in child bilinguals acquiring Estonian as a heritage language in Norway.

It is both intuitive and well attested that language exposure has a significant effect on language acquisition. However, the relationship between the two is not necessarily straightforward, especially in multilingual contexts. A nuanced examination of language exposure is needed to explain language outcomes. Even closely connected measures like age of onset of acquisition and length of exposure can have distinct effects on language competence (Smolander et al., 2021). It is also well attested that both the quantity (Ågren et al., 2014; Thordardottir, 2019; Unsworth, 2013), and quality (Daskalaki et al., 2020; Hoff & Core, 2013) of exposure play a significant role.

The sample in this study includes 23 children aged 5–7 years. Sixteen of them have two Estonian-speaking parents, while seven have one Norwegian and one Estonian parent. All children completed three tasks from the Litmus toolkit (Armon-Lotem et al., 2015), in both Estonian and Norwegian: the Sentence Repetition Task (SRT), the Cross-Linguistic Lexical Tasks (CLTs), and the Multilingual Assessment Instrument for Narratives (MAIN). Parents detailed their daily language environments in a customised Q-BEx questionnaire (De Cat et al., 2022).

The talk presents an overview of the lexical and morphological diversity in bilingual children's Estonian narratives, asking whether measures of linguistic diversity in this type of task correlate with scores from more structured, targeted language tests (SRT and CLTs). Correlations between diversity measures in language use, and language exposure measures (amount of current exposure, amount of cumulative exposure, a richness score indicating quality of exposure) are also investigated. By comparing the results from different methods for collecting language data, we get a better understanding of what each method is best suited for and how using them together might deepen our understanding of the children's language use and competence.

**Sources:**

Armon-Lotem, S., de Jong, J., & Meir, N. (Eds.). (2015). Assessing Multilingual Children: Disentangling Bilingualism from Language Impairment (Vol. 13). Multilingual Matters. https://www.multilingual-matters.com/page/detail/?k=9781783093120

Daskalaki, E., Blom, E., Chondrogianni, V., & Paradis, J. (2020). Effects of parental input quality in child heritage language acquisition. Journal of Child Language, 47(4), 709–736. https://doi.org/10.1017/S0305000919000850

De Cat, C., Kašćelan, D., Prevost, P., Serratrice, L., Tuller, L., & Unsworth, S. (2022). Quantifying Bilingual EXperience (Q-Bex): Questionnaire manual and documentation. https://doi.org/10.17605/OSF.IO/V7EC8

Hoff, E., & Core, C. (2013). Input and Language Development in Bilingually Developing Children. Seminars in Speech and Language, 34(4), 215–226. https://doi.org/10.1055/s-0033-1353448

Smolander, S., Laasonen, M., Arkkila, E., Lahti-Nuuttila, P., & Kunnari, S. (2021). L2 vocabulary acquisition of early sequentially bilingual children with TD and DLD affected differently by exposure and age of onset. International Journal of Language & Communication Disorders, 56(1), 72–89. https://doi.org/10.1111/1460-6984.12583

Thordardottir, E. (2019). Amount trumps timing in bilingual vocabulary acquisition: Effects of input in simultaneous and sequential school-age bilinguals. International Journal of Bilingualism, 23(1), 236–255. https://doi.org/10.1177/1367006917722418

Unsworth, S. (2013). Assessing the role of current and CUMULATIVE exposure in simultaneous bilingual acquisition: The case of Dutch gender. Bilingualism: Language and Cognition, 16(1), 86–110.

Ågren, M., Granfeldt, J., & Thomas, A. (2014). Combined effects of age of onset and input on the development of different grammatical structures: A study of simultaneous and successive bilingual acquisition of French. Linguistic Approaches to Bilingualism, 4(4), 462–493. https://doi.org/10.1075/lab.4.4.03agr

**The relationship between lexicon, morphology and syntax in English and Estonian**

*Caroline Rowland, Seamus Donnelly, Piia Taremaa, Ada Urm, Adele Vaks, Virve Vihman and Tiia Tulviste*

A long-standing question in language development – which also touches on fundamental questions in linguistic theory – is the nature of the relationship between early lexical and grammatical knowledge. The very strong correlation between the two has led some to argue that lexical and syntactic knowledge may be inseparable (e.g. Bates & Goodman 1997), consistent with usage-based theories that eschew a distinction between the two systems (e.g. Goldberg 1995, Lieven 2016). However, little research has explicitly examined whether early lexical and syntactic knowledge are statistically separable and, if so, whether morphology patterns with the lexicon or syntax. Two underappreciated methodological challenges are present in such research. First, the relationship between lexical, morphological and grammatical knowledge may change during development. Second, non-linear mappings between true and observed scores on the language measurement scales we use could lead to spurious multidimensionality. Donnelly et al. (2025) addressed these challenges by using data from several time points and a statistical method robust to such non-linear mappings. We apply their analysis to two languages with contrasting morphological systems, to test for effects on the relationship between the three domains.

Donnelly and colleagues examined item-level vocabulary and syntax data on American English from the Wordbank database (data collected using Communicative Development Inventories, CDIs), using Item Response Theory. They also analysed longitudinal corpora of American and British English, using the same CDI scale at 18/19, 21, 24 and 30 months. In both studies, they found evidence that, while there is a very strong relationship between vocabulary and syntax knowledge in early language development, the two are clearly separable from about 18 months. In this talk, based on research conducted in the MEDAL project, we extend the analysis to morphology, and compare the results from English and Estonian, a language with a much more extensive and complex morphological system. We aimed to determine (a) whether the same patterns hold for morphology when we compare cross-linguistic data, and (b) whether morphology aligns with lexical or syntactic development. Our analyses show that the data from both languages are best explained with a three-factor model.

Although we expected differing effects for morphology in the two languages, the models show greater similarities in acquisition across languages, and suggest multidimensionality in both. Together with the earlier results, the evidence points to the existence of three separable constructs. The three-factor model suggests that at least somewhat different sources underlie the individual differences in lexical, syntactic and morphological acquisition in both English and Estonian. We interpret this to support the idea that the development of lexical, morphological and syntactic systems involve separable, but partially overlapping cognitive processes. We will discuss the implications for theories of early language acquisition and linguistic structure.

**References**

Bates, E. & Goodman, J. C. (1997). On the Inseparability of Grammar and the Lexicon: Evidence from Acquisition, Aphasia and Real-time Processing. Language and Cognitive Processes, 12(5–6), 507–584.

Donnelly, S., Kidd, E., Verkuilen, J. & Rowland, C. (2025) The separability of early vocab-

ulary and grammar knowledge. Journal of Memory and Language, 141.

Goldberg, A. E. (1995). Constructions: A construction grammar approach to argument structure. Chicago: University of Chicago Press.

Lieven, E. (2016). Usage-based approaches to language development: Where do we go from here? Language and Cognition, 8(3), 346-368.

Learning from learning: Sources of evidence about transparent vowels in Finnic vowel harmony

**Background** Almost every language with vowel harmony (VH) has vowels that do not actively harmonize: some vowels are *transparent* (acting as neither targets nor triggers), some are *opaque* (acting as triggers but not targets), and some vacillate in their behaviour, depending e.g. on their location in the word. Multiple such patterns are found in those Finnic languages with progressive [+/-back] vowel harmony (Kiparsky & Pajusalu, 2003; Fejes et al., 2024), as summarized in Table 1, but several crucial patterns are found in quite under-studied languages, with scarce existing literature. In this work, we use two intertwined data-driven approaches – both corpus analyses and learning simulations – to investigate the nature and limits of transparency. Especially by studying how VH transparency can be learned, what can we learn about transparency itself?

**Empirical questions** In the typical case, Finnic backness harmony spreads left-to-right from the root-initial stressed vowel to subsequent ones, often including suffixes. But what if the word begins with one (or more) transparent vowels? Several languages allow both front-harmonic (1) and back-harmonic (2) spans to begin later in the word*;* spans of multiple transparent vowels are found at least in Finnish (Suomi et al, 2008) and Seto (Kiparsky and Pajusalu, 2001). From the learner's perspective: how much data demonstrates these patterns, especially the possibility of front-transparent + back-harmonic sequences?

| (1) Transparent + **front vowel** spans | | | (2) Transparent + **back vowel** spans | |
|---|---|---|---|---|
| *pitä-mä*<br>hold-SPN | Votic (Markus & Rozhanskiy, 2022:333) | | *pehk**o-a***<br>bush-PART | Ingrian (Markus & Rozhanskiy, 2014:244) |
| *etelä*<br>south | Finnish (adapted from Suomi et al, 2003) | | *sisik**u**nn**a***<br>entrails-GEN | Seto (adapted from Eesti Keele Instituut, n.d.) |

Another unknown is how harmonic behaviour beyond an initial span of transparent vowels depends on morphological structure. Karlsson (2018) states that in Finnish, stems without back vowels must take front vowel suffixes, even if the stem's front vowels are *all* transparent; this means that transparent-back sequences are possible within roots (3) but not across a root+suffix boundary (4). To what extent this generalizes to other Finnic languages, and how frequently either structure occurs in learners' inputs, are not clear.

| (3) Root-internal transparent + **back vowel** | | | (4) Transparent Root vowels + **front suffix** | |
|---|---|---|---|---|
| *klinikk**a***<br>clinic | Finnish (adapted from Suomi et al, 2008) | | *venee-ssä*<br>boat-INE | Finnish (adapted from Karlsson, 2018) |

**Analyses** To shed light on these questions, we first analyze existing corpus data from Seto, Karelian, Livonian, Votic, and Kihnu Estonian (Lindström et al., 2022; Boyko et al., 2022) with the aim of summarizing any post-transparent VH patterns. One pilot study focuses on a minority Estonian dialect spoken on Kihnu Island (Lindström et al., 2022), revealing skews in the input that learners get about transparent vowels and backness harmony. For example: of the 4780 non-compound Kihnu wordtypes in the corpus, *only five* provide explicit information about whether a sequence of transparent vowels in the three initial syllables can precede front- vs back-harmonic vowels (see bolded cells in Table 2).

Our second approach uses algorithmic simulations of VH acquisition (as in Vesik, in prep) to shed light on what kinds of patterns are likely to be learnable from the available input data. For instance: if the vast majority of a learner's input data consists of relatively short roots, then the developing grammar may only need to encode vowel harmony emanating from the first two or so vowels. As preliminary support for this idea, we extracted from the Kihnu Estonian corpus all monophthong-only lemmas, and removed remaining suffixes (e.g. *-ma*, supine/infinitive; *-mine*, nominalizer; *-gi*, emphatic). Of the resulting 1461 unique stems, a full 1183 (81%) have only one or two syllables. Overall, our learning simulations (drawing from this and other minority language corpora) demonstrate that combining methodologies can provide deeper insight into what kinds of linguistic patterns can be learned from realistic exposure to a language.

| Transparency pattern | Sample languages | Front + transparent | Back + transparent |
|---|---|---|---|
| *i* and *e* both fully transparent | Finnish (also Karelian, Ingrian, Veps) | *pitä-nyt*<br>has kept | *luke-nut*<br>has read |
| *i* fully transp; *e* transp only in non-final stem syllables | Votic | *läsi-mä*<br>be.sick-SPN | *релла-ma*<br>play-SPN |
| *i* fully transparent | North Seto | *`üt'li-vä*<br>say-SPN | *`lõppi-va*<br>end-SPN |
| *i* fully transp; *e* transp only in initial stem syllables | South Seto | *üt's'indä*<br>alone | *emakkõnõ*<br>mother. |

**Table 1**. Data from Karlsson, 2018; Markus & Rozhanskiy, 2022; Lindström et al., 2022

| n | # of words with n syllables | # of which contain transparent Vs in all syllables before the last | # of which have a final V that is… | | |
|---|---|---|---|---|---|
| | | | front | back | transparent |
| 3 | 914 | 125 (e.g., *kiviga*) | 37 | 22 | 66 |
| 4 | 230 | 22 (e.g., *vedeleväd*) | 3 | 2 | 17 |
| 5 | 18 | 2 (e.g., *inimesele*) | 0 | 0 | 2 |

**Table 2**. Kihnu Estonian monophthong-only words with initial transparent vowels (summarized from Lindström et al., 2022).

**References**


Boyko, T., Zaitseva, N., Krizhanovsky, N., Krizhanovsky, A., Novak, I., Pellinen, N., & Rodionova, A. (2022). The Open Corpus of the Veps and Karelian Languages: Overview and Applications. KnE Social Sciences. doi: 10.18502/kss.v7i3.10419.

Eesti Keele Instituut. (n.d.). *Seto sõnastik* [Seto dictionary]. Retrieved February 16, 2025, from https://arhiiv.eki.ee/dict/setosonastik/.

Fejes, L., Siptár, P., & Vago, R.M. (2024). Vowel Harmony in Uralic Languages. In van der Hulst, H. and Ritter, N. A. (Eds.), *The Oxford handbook of vowel harmony*. Oxford University Press.

Karlsson, F. (2018). *Finnish: A comprehensive grammar*. Taylor and Francis.

Kiparsky, P., & Pajusalu, K. (2001). *Seto vowel harmony and the typology of disharmony*. MS. https://web.stanford.edu/~kiparsky/Papers/seto.pdf

Kiparsky, P., & Pajusalu, K. (2003). Towards a typology of disharmony. *The Linguistic Review*, 20, 217-241.

Lindström, L., Todesk, T., & Pilvik, M.L. (2022). Corpus of Estonian Dialects. Institute of Estonian and General Linguistics, University of Tartu. https://datadoi.ee/handle/33/492.

Markus, E., & Rozhanskiy, F. (2014). Comitative and terminative in Votic and Lower Luga Ingrian. *Linguistica Uralica, 50*(4), 241-257.

Markus, E., & Rozhanskiy, F. (2022). Votic. In Bakró-Nagy, M., Laakso, J., & Skribnik, E. (Eds.), *The Oxford guide to the Uralic languages*. Oxford University Press.

Pajusalu, K. (2022). Seto South Estonian. In Bakró-Nagy, M., Laakso, J., & Skribnik, E. (Eds.), *The Oxford guide to the Uralic languages*. Oxford University Press.

Suomi, K., Toivanen, J., & Ylitalo, R. (2008). *Finnish Sound Structure: Phonetics, phonology, phonotactics and prosody*. (Studia Humaniora Ouluensia 9). Oulu: University of Oulu.

Vesik, K. (in prep). *Gradual Error-Driven Learning and Typology of Finnic Vowel Interactions*. Ph.D. thesis, University of British Columbia.

**Approximating language attitudes in a sociolinguistic decision experiment**

The decision to use a particular language in interaction depends on various contextual, intra- and interpersonal factors. Yet, observing a particular exchange as a singular event allows us to control for varying contextual and interpersonal factors to experiment with the intrapersonal domain, which is commonly investigated through introspection, surveys, and interviews in sociolinguistic research. This paper outlines an experimental approach that is currently in preparation, in which the participants' revealed preferences for languages shall be measured quantitatively.

The starting point for this experiment is the assumption borrowed from economics and game theory that each decision is linked to a comparison of the expected utility (or benefit) of all possible options. The design of the utility function can be as broad and complex as necessary and does not imply a reduction of sociolinguistic detail. A focus on a singular event helps to further reduce complexity to only those factors which are relevant in the given setting or in the interaction within a set of interlocutors. Successful or unsuccessful communication will provide reference points for the individual to adjust their expected utility for each possible language choice. In the present case, the main focus lies on materialistic gains and a psychic costs/benefits, with the latter functioning as a penalty or reward for the use of a given language within each participant. If two interlocutors share the same benefit of a successful interaction, they still might have different utility depending on their attitude towards the language used.

The planned experiment will have multilingual participants solve a series of language-based tasks collaboratively in pairs, where their success determines their shared measurable reward at the end of each task. After establishing a base-line for each participant's productivity on the tasks, as well as after every round of tasks, participants are given feedback on their expected utility in the form of points per task, points per minute, or correct answers per task. In a second step, incentives are offered to change the language regime by rewarding or penalising the choice of one of the experiment's languages. After informing participants of their success rates, the incentives are altered and participants given the opportunity to accept the newly offered conditions or retain the old ones. The psychic costs and rewards can be approximated through the difference of the actual success rate and the highest or lowest accepted incentive after several repetitions.

The presentation will illustrate the general concept of the experimental approach, expected results, and their link to sociolinguistic research on language attitudes. It will, furthermore, discuss advantages and drawbacks in experiment design. Since this experiment is still in preparation, it would greatly benefit from being discussed at the conference with the option to developing into a collaborative project.

*ABSTRACT* for the MEDAL Final Conference, Tartu, 9-10 September 2025

**Subject expression in Pite Saami**

Joshua Wilbur • University of Tartu
joshua.wilbur@ut.ee

Pite Saami is a critically endangered Uralic language spoken by a handful of individuals from Arjeplog municipality in Sweden and adjacent areas of Norway. It has a limited amount of language resources available for linguistic study; most of these resources are linguistically annotated transcriptions of spoken language (often accompanied by audio and even video recordings). Pite Saami is an accusative language, and the grammatical subject of Pite Saami clauses is in the nominative case. Finite verbs in Pite Saami inflect for various morphological categories, including – crucially – person and number of the grammatical subject (person categories are: first, second, third; number categories are: singular, dual (for highly animate referents), plural) (Wilbur 2014). The grammatical subject may be realized as a full noun phrase, as in the first half of (1), or as a pronoun, as in (2); however, it is not always realized at all, as in (3) or in the latter half of (1).

(1) *jus gusa da-jt ulli, dä burri da-jt*
if cow\NOM.PL that-ACC.PL reach\3PL.PRS then eat\3PL.PRS that-ACC.PL
'if the cows reach it, then they eat it' [pit080924.221]

(2) *mån bådá-v Áhkábákte-s*
1SG.NOM come-1SG.PRS Á.-ELAT.SG
'I am coming from Áhkábákkte' [pit080924.004]

(3) *daga-jme lastbijla-jt ja daggári-jt*
make-1PL.PST truck-ACC.PL and such-ACC.PL
'we made trucks and such things' [pit080924.067]

In this talk, I will present the initial results of an on-going investigation into the factors determining whether the subject is realized in Pite Saami. The initial results at the time of submitting this abstract are based on a pilot study using a 22 minute conversation between two speakers (the source of the examples above); the recording is from 2008 and contains 2693 tokens and 694 utterances.[1] In order to make the study comparable with other GoLD MEDAL studies on subject realization in Estonian, Finnish, Polish and Russian ([anonymous] manuscript), I have restricted the data set to include only utterances with a finite verb in first or second person, in indicative mood; however, I also included negated clauses because the Pite Saami negation verb is a fully inflecting finite verb, including morphology referencing the person and number of the grammatical subject.

The resulting dataset consists of 194 finite verbs from this recording. Of these, 38% (n = 74) have an overt subject pronoun, while the other 62% (n = 120) lack any overt, realized subject. These initial results indicate that Pite Saami may differ significantly from similar, spoken language data assessed for Estonian (a related Uralic language) and Russian, which showed 72% and 59% expressed pronominal subjects, respectively, while also differing from subtitle data for Finnish and Polish, with 11% and 5% expressed subjects, respectively ([anonymous] manuscript). However, the Pite Saami data is still being coded and analyzed, so other various factors which may affect subject realization in Pite Saami, (e.g., person; tense; distance to the nearest previous subject reference; potentially affirmative vs. negation) will be presented in this talk.

[1] Note that these are spontaneous, spoken language utterances that correspond at least as much to intonation units as they do to discourse-level syntactic units typically referred to as "clauses".

**Abbreviations**

| | |
|---|---|
| 1 | first person |
| 3 | third person |
| acc | accusative case |
| elat | elative case |
| nom | nominative case |
| pl | plural number |
| prs | present tense |
| pst | past tense |
| sg | singular number |

**Bibliography**

[anonymous] (manuscript). “When less is not more: A comparative corpus study of optional subject pronoun expression in Finnic and Slavonic”. In.

Wilbur, Joshua (2014). *A grammar of Pite Saami*. Studies in Diversity Linguistics 5. Berlin: Language Science Press.

**Udmurt discourse particles in response utterances and beyond: results from corpus and acceptability judgment task**

In this talk, we discuss various methods we applied when researching syntactic and pragmatic differences between Udmurt (Permic, Uralic) discourse particles *ja*, *ma*, and *nu*. These elements stand out in the rich system of Udmurt discourse particles in that they can only precede structures they modify. In most cases, they occur sentence-initially, unlike other discourse particles in Udmurt. The meaning of *ja*, *ma*, and *nu* can roughly be rendered by the English sentence-initial *well*:

(1)[S1] *ta-ze* *ug* *ju.* [S2] ***ja***, *odig-ze* *ju=ni!*
this-ACC.3SG.POSS NEG.NPST.1SG drink PTCL one-ACC.3SG.POSS drink=PTCL
'[S1] I'm not drinking this one. [S2] Well, drink one [shot]!' (Personal materials)

**Hypothesis.** Based on limited corpus data, we expect that there are differences between the three particles in terms of their context of use, which can be seen in response contexts of different kinds.

**Data and method.** The data for our analysis comes from:
1) acceptability ratings for constructed examples set up using the conversational board technique (Wiltschko 2021), conducted in Standard Udmurt (6 speakers);
2) traditional fieldwork acceptability judgments from three speakers of Beserman (traditionally classified as a dialect of Udmurt);
3) manually annotated samples of contexts with these particles from the Udmurt and Beserman corpora.

The combinations of values from the following two parameters were assessed in the acceptability judgment task.1. Type of initiative utterance (9 different speech acts). 2. Type of response containing a particle: stand-alone particle as a positive response (‘yes’/’okay’); particle preceding another positive response word; particle preceding a positive response word followed by a contradiction; particle preceding a negative response.

The acceptability scale had five points presented descriptively. It was accompanied by an additional comment field for possible paraphrases and notes about (im)politeness and prosodic integration.

Corpus examples were annotated by the following parameters: sentence type of the particle-containing utterance, dialogic/monologic use, sentence type of the initiative utterance (if any), prosodic integration, position of the particle inside the turn, and co-occurrence with other particles.

The **results** of the acceptability ratings showed different mapping of the three particles along the two parameters. The differences in the acceptability of the particles across various types of response utterances, supported by the corpus evidence, led us to the conclusion that the particles play different roles in discourse topic management: *ja* indicates topic closure, *nu* indicates topic continuity, and *ma* marks the issue under discussion as trivial or not of big importance.

In our talk, we will first focus on the motivation for using acceptability ratings in the study of discourse particles. Second, we will discuss the interpretation of the results of the acceptability ratings, considering the inter-speaker agreement level. Finally, we’ll show how the results of the acceptability ratings correspond to corpus data.

**References:**

Arkhangelskiy, Timofey, Maria Usacheva (eds.). 2025. *Beserman Multimedia Corpus.* Available online at https://beserman.web-corpora.net/index_en.html.

Arkhangelskiy, Timofey, Maria Medvedeva. 2025. *Corpus of Standard Udmurt.* Available online at udmurt.web-corpora.net.

Wiltschko, M. 2021. *The grammar of interactional language.* Cambridge University Press.

MEDAL

Methodological Excellence in Data-Driven Approaches to Linguistics

**Final Conference** Book of Abstracts
9 - 10 October 2025
Institute of Estonian and General Linguistics
University of Tartu
medal.ut.ee

Horizon Europe project 101079429
UKRI project 10047684

Funded by the European Union

UK Research and Innovation