

Reduction rate and lexical-distributional characteristics affect the perception of Estonian spontaneous speech

Kaidi Lõo¹, Pärtel Lippus¹, Fabian Tomaschek², Benjamin V. Tucker^{2,3}

¹ University of Tartu, ² University of Tübingen, ³ University of Alberta

kaidi.loo@ut.ee

Spontaneous speech can be characterized by a large amount of variation in pronunciation, including vowel reduction, as well as assimilation and deletion of segments (see e.g., Ernestus & Warner 2011; Tucker & Ernestus 2016). Previous research has indicated that the more reduced words are, the more difficult it is to comprehend them. Critically, this process is particularly difficult when words are isolated and presented out of context (Ernestus et al. 2002, Arnhold et al. 2017). However, not much is known on how the recognition of reduced speech is influenced by the lexical-distributional characteristics of words.

The aim of the current study is to investigate the perception of spontaneous speech in Estonian, a morphologically rich Finno-Ugric language. More specifically, we were interested in two research questions. 1) Is it also the case for Estonian that the perception of less reduced forms is easier than that of more reduced (quantified as the difference in segments between the SAMPA transcription of actual pronunciation and the dictionary form)? 2) How does reduction interact with lexical measures such as surface frequency and inflectional paradigm size (i.e., the number of attested inflected forms given an inflectional paradigm in the spontaneous speech corpus (Lippus et al. 2020, see also Lõo et al. 2018).

To investigate these questions, we conducted an online word perception experiment with 125 native speakers of Estonian (92 females and 33 males, age range 18-65 years) and 1000 randomly selected inflected nouns (excluding foreign, compound, abbreviated words etc) from the Estonian spontaneous speech corpus (Lippus et al. 2020). The stimuli were divided into four lists, and each participant listened to 250 stimuli. We followed the design devised by Arnold et al. (2017) in which listeners were first asked to indicate whether they recognised the word and then type in what they heard, thereby identification speed and accuracy were measured. After the experiment, participants filled out a short background questionnaire.

The overall identification accuracy was 54.7%. The data was analysed using generalized additive mixed effects models (Wood, 2017). Preliminary results indicate that less reduced ($\beta=-1.68$, $z=-14.83$, $p<0.001$), more frequent word ($\beta= 0.87$, $z=5.31$, $p<0.001$) and with more segments ($\beta= 0.36$, $z=7.62$, $p<0.001$) as well as with fewer inflectional paradigm members ($\beta=-1.18$, $z=-65.95$, $p<0.001$) were recognized most accurately. Furthermore, reduction rate interacted with surface frequency in such a way that frequent less reduced words were particularly easy to correctly identify ($Chi.sq.=17.02$, $p=0.002$).

In summary, research on morphologically complex lesser-studied languages such as Estonian complements the field of speech perception with typological diversity. Our results show that the perception of Estonian inflected words is affected by a variety of complex word-level factors. The current analysis explored in particular the relationship between acoustic and lexical characteristics of spontaneous speech (cf. Ernestus & Baayen, 2007). Further analyses will be conducted to investigate whether perception is different depending on whether the reduction occurred in the stem versus in the affix as well as a more in-depth phonetic analysis of the stimuli.

References

- Arnold, D., Tomaschek, F., Sering, K., Lopez, F., & Baayen, R. H. (2017). Words from spontaneous conversational speech can be recognized with human-like accuracy by an error-driven learning algorithm that discriminates between meanings straight from smart acoustic features, bypassing the phoneme as a recognition unit. *PloS one*, *12*(4), e0174623.
- Ernestus, M. T. C., & Baayen, R. H. (2007). The comprehension of acoustically reduced morphologically complex words: the roles of deletion, duration, and frequency of occurrence. *Proceedings of the 16th International Congress of Phonetic Sciences*, Saarbrücken, pp. 773-776.
- Ernestus, M., Baayen, H., & Schreuder, R. (2002). The recognition of reduced word forms. *Brain and language*, *81*(1-3), 162-173.
- Ernestus, M., & Warner, N. (2011). An introduction to reduced pronunciation variants. *Journal of Phonetics*, *39*(S1), 253-260.
- Lippus, P., Tuisk, T., Ots, N., Teras, P. (2020). Phonetic corpus of Estonian spontaneous speech v.1.0.6 [online]. Available at: <https://doi.org/10.15155/1-00-0000-0000-0012BL>.
- Lõo, K., Järvikivi, J., Tomaschek, F., Tucker, B. V., and Baayen, R. H. (2018). Production of Estonian case-inflected nouns shows whole-word frequency and paradigmatic effects. *Morphology*, *28*:71–97.
- Tucker, B. V., & Ernestus, M. (2016). Why we need to investigate casual speech to truly understand language production, processing and the mental lexicon. *The Mental Lexicon*, *11*(3), 375-400.
- Wood, S. N. (2017). *Generalized additive models: an introduction with R*. CRC press.