

Kasse ja koere: Use of nonstandard inflectional forms among adults speaking Estonian

Mari Aigro, University of Tartu

Annika Kängsepp, University of Tartu

Morphological overgeneralizations entail situations in which a nonstandard form is produced due to the speaker applying a rule to an instance with which it is not conventionally used. Children are among the most widely discussed overgeneralizers, often producing forms such as **toosin sulle mõmmi* (~ *tõin sulle mõmmi* ‘I brought you a teddy bear’). This phenomenon of using nonstandard stems or morphological formatives is widely assumed to recede from speech production during childhood, with uniform use of standard forms being expected by adulthood (Clark 2009, Blything et al 2014).

However, this does not seem to be the case. Adults seem to incorporate overgeneralizations of similar types in their speech, with corpora showcasing instances such as *näen tänaval kasse ja *koere* (~ *koeri/koerasid*, ‘I see cats and dogs in the street’). While very little existing research investigates this phenomenon, speakers’ occasional divergence from language norms is evident from the fact that in Estonian a wide range of resources exist for adult speakers with the aim of preventing such use instances (implicating the existence of non-standard production as grounds for these resources), which even include a phone hotline (*Keelenõuanne, Eesti keele instituut*).

From the perspective of analogy-based (Eddington 2000, Blevins 2016) and schema-based (Langacker 2000) usage accounts, such nonstandard use instances are not only easily explained but also expected. This further emphasizes the gap in terms of how little is known about the extent and internal motivators behind adult non-standard language use. We know that semantic compatibility facilitates the degree to which speakers are tolerant towards non-standard uses, meaning they do not perceive everything deemed incorrect by the language standard as equally incorrect (Ambridge et al 2008). We also know that when asked to produce forms for nonce words, speakers are affected by related frequency values, as well as formative productivity and phonological neighbourhood sizes (the number of similar lexemes in their inventory, Dąbrowska 2008).

Present research is based on a dataset of almost 1,000 nonstandard nominal and pronominal inflection forms in the Estonian National Corpus 2021 (2.4 billion tokens, Koppel & Kallas 2022). We focus on the proportion with which a nonstandard instance is used compared to its standard counterpart (e.g., the percentage which **koere* makes up of all uses in that nominal cell, also including *koeri* and *koerasid* in ‘dogs.PAR.PL’). We investigate which morphological, distributional, and lexical factors affect this proportion, essentially leading to more frequent nonstandard inflectional uses.

References

- Ambridge, Ben, Julian M. Pine, Caroline F. Rowland, and Chris R. Young, ‘The Effect of Verb Semantic Class and Verb Frequency (Entrenchment) on Children’s and Adults’ Graded Judgements of Argument-Structure Overgeneralization Errors’, *Cognition*, 106.1 (2008), pp. 87–129, doi:[10.1016/j.cognition.2006.12.015](https://doi.org/10.1016/j.cognition.2006.12.015)
- Blevins, James P., *Word and Paradigm Morphology* (Oxford University Press, 2016), doi:[10.1093/acprof:oso/9780199593545.001.0001](https://doi.org/10.1093/acprof:oso/9780199593545.001.0001)
- Blything, Ryan P., Ben Ambridge, and Elena V. M. Lieven, ‘Children Use Statistics and Semantics in the Retreat from Overgeneralization’, ed. by Mark Aronoff, *PLoS ONE*, 9.10 (2014), p. e110009, doi:[10.1371/journal.pone.0110009](https://doi.org/10.1371/journal.pone.0110009)
- Clark, Eve, *First Language Acquisition*, 2nd edn (Cambridge University Press, 2009)
- Dąbrowska, Ewa, ‘The Effects of Frequency and Neighbourhood Density on Adult Speakers’ Productivity with Polish Case Inflections: An Empirical Test of Usage-Based Approaches to Morphology’, *Journal of Memory and Language*, 58.4 (2008), pp. 931–51, doi:[10.1016/j.jml.2007.11.005](https://doi.org/10.1016/j.jml.2007.11.005)
- Eddington, David, ‘Spanish Stress Assignment within the Analogical Modeling of Language’, *Language*, 76.1 (2000), p. 92, doi:[10.2307/417394](https://doi.org/10.2307/417394)
- Koppel, Kristina, and Jelena Kallas, ‘Eesti Keele Ühendkorpuste Sari 2013–2021: Mahukaim Eestikeelsete Digitekstide Kogu’, *Eesti Rakenduslingvistika Ühingu Aastaraamat Estonian Papers in Applied Linguistics*, 18 (2022), pp. 207–28, doi:[10.5128/ERYa18.12](https://doi.org/10.5128/ERYa18.12)
- Langacker, Ronald W., ‘A Dynamic Usage-Based Model’, in *Usage-Based Models of Language*, ed. by Michael Barlow and Suzanne Kemmer (CSLI Publications, 2000), pp. 1–63