

Is language efficient or redundant? How language users distinguish the agent from the recipient in English and Dutch

Eva Zehentner¹ & Dirk Pijpops²

¹ University of Zurich, eva.zehentner@es.uzh.ch; ² University of Liège, dirk.pijpops@uliege.be

Keywords: Redundancy, corpus, efficiency, argument marking, robustness

Language is systemically redundant, meaning it often features several strategies to perform the same function (Van de Velde 2014). For instance, to form the past tense, Dutch may use ablaut, as in *bid ~ bad* 'pray' or a *-de* suffix, as in *graaf ~ graafde* 'dugged'. However, such systemic redundancy does not entail syntagmatic redundancy. That is, it is not necessarily the case that the various strategies are used in one and the same utterance. In fact, the occurrence of double forms such as Dutch *begin ~ begon*, lit. 'begin ~ began-ed', is highly exceptional (De Smet 2021). The reason seems obvious: redundant marking is superfluous, and would unnecessarily burden production processing (Sinnemäki 2009; Kurumada & Jaeger 2015; Leufkens 2015). Syntagmatic redundancy should accordingly be avoided for reasons of efficiency. This may be done either directly by language users, or by grammar evolving so that various strategies strictly complement one another, and do not overlap. However, it has also been argued that syntagmatic redundancy is actually useful and therefore common, because (i) it enhances the robustness of the linguistic signal against information loss (Fedzechkina et al. 2012; Levshina 2021), and (ii) it increases learnability (Sloutsky & Robinson 2013; Tal & Arnon 2022).

We test these competing accounts by investigating agent-recipient disambiguation in English and Dutch. These languages may employ the same four morphosyntactic strategies to distinguish agents from recipients, viz.

- (i) Constituent order, e.g. *The student gives the lecturer a book.*
- (ii) Nominal marking, e.g. *The student gives **us** a book.*
- (iii) Verbal agreement, e.g. *The student gives the lecturer a book.*
- (iv) Prepositional marking, e.g. *The student gives a book **to** the lecturer.*

Under an efficiency-focused account, we expect combinations of these strategies to be rare, and single marking to be common. Under an account focused on redundancy as useful, we conversely expect multiple marking to be the default. We investigate this case study using both Present Day corpus data from Dutch (Oostdijk et al. 2013) and English (Röthlisberger 2018). While both languages employ the same strategies, they crucially do so in different ways. Dutch generally allows more leeway to the individual language producers to decide whether or not to apply a certain strategy. By contrast, in English, this choice is determined to a larger degree by grammar, with only prepositional marking really being optional for the language producer. In addition to this synchronic comparison, we also investigate the usage of these strategies throughout the development of English, using the *Penn Helsinki Parsed Corpus of Middle English* (Kroch et al. 2000). We find that despite substantial differences in the specific application, redundancy seems to be the default in all languages and stages; however, it importantly seems to operate within efficiency constraints.

Finally, we also discuss our results in terms of different approaches to efficiency, viz. the question whether our results support redundancy as (non-)efficient for production, (non-)efficient for comprehension, or whether applying redundancy only in specific environments may in fact represent an efficient trade-off between production and comprehension.

References

- Fedzechkina, Maryia, Florian Jaeger and Elissa Newport. 2012. Language learners restructure their input to facilitate efficient communication. *Proceedings of the National Academy of Sciences* 109(44). 17897–17902
- Kroch, Anthony, Ann Taylor and Beatrice Santorini. 2000. The Penn-Helsinki Parsed Corpus of Middle English (PPCME2). Department of Linguistics, University of Pennsylvania.
- Kurumada, Chigusa and Florian Jaeger. 2015. Communicative efficiency in language production: Optional case-marking in Japanese. *Journal of Memory and Language* 83. 152–178.
- Leufkens, Sterre. 2015. *Transparency in language: A typological study*. Utrecht: LOT.
- Levshina, Natalia. 2021. Cross-Linguistic Trade-Offs and Causal Relationships Between Cues to Grammatical Subject and Object, and the Problem of Efficiency-Related Explanations. *Frontiers in psychology* 12. 648200.

- Oostdijk, Nelleke, Martin Reynaert, Véronique Hoste and Ineke Schuurman. 2013. The Construction of a 500-Million-Word Reference Corpus of Contemporary Written Dutch. In Peter Spyns & Jan Odijk (eds.), *Essential Speech and Language Technology for Dutch, Theory and Applications of Natural Language Processing*, 219–247. Heidelberg: Springer.
- Röthlisberger, Melanie. 2018. The dative dataset of World Englishes. KU Leuven. <https://doi.org/10.5281/zenodo.2553357>.
- Sinnemäki, Kaius. 2009. Complexity in core argument marking and population size. In Geoffrey Sampson, David Gil & Peter Trudgill (eds.), *Language complexity as an evolving variable*, 126–140. (Oxford Studies in the Evolution of Language). United Kingdom: Oxford University Press.
- Sloutsky, Vladimir and Christopher Robinson. 2013. Redundancy matters: Flexible learning of multiple contingencies in infants. *Cognition* 126(2). 156–164.
- Tal, Shira and Inbal Arnon. 2022. Redundancy can benefit learning: Evidence from word order and case marking. *Cognition* 224. 105055.
- Van de Velde, Freek. 2014. Degeneracy: the maintenance of constructional networks. In Ronny Boogaart, Timothy Coleman & Gijsbert Rutten (eds.), *Extending the Scope of Construction Grammar*, vol. 1, 141–179. Berlin: Mouton de Gruyter.