

IS REDUNDANCY USEFUL IN LANGUAGE? AGENT-RECIPIENT DISAMBIGUATION IN ENGLISH AND DUTCH

EVA ZEHENTNER^{*1} and DIRK PIJPOPS²

^{*}Corresponding Author: eva.zehentner@es.uzh.ch

¹English Department, University of Zurich, Zurich, Switzerland

²Department of Modern Languages, University of Liège, Liège, Belgium

This paper discusses the competing evolutionary motivations of efficiency versus robustness in language processing and learning (MacWhinney et al. 2014), both from a typological and diachronic perspective. Specifically, we assess the potential benefits or costs of redundancy in morphosyntactic marking of participant roles, comparing and testing two opposing hypotheses:

On the one hand, following the most crucial tenet in usage-based linguistics that language use affects – or even determines – grammar (Bybee 2010), we assume that language is organised in a way that facilitates efficient usage (e.g. Gibson et al. 2019). On this account, redundant marking should be dispreferred. Well-known typological ‘trade-off’ distributions and diachronic trajectories between word order and morphological case marking seem to support this point (Fedzechkina et al. 2017). Furthermore, prepositional marking is often only applied in contexts where it comes with some added processing benefit (cf. Pijpops et al. 2018 on the impact of complexity on Dutch transitive object marking, or Tal et al. 2020, Levshina 2021 on ambiguity/atypicality in differential object marking).

On the other hand, however, we pursue Van de Velde's (2014) argument that a certain amount of redundancy – or rather, ‘degenerate’ marking (involving many-to-many relationships) – is in fact beneficial from a usage perspective: redundancy constitutes an indispensable component of any degenerative Complex Adaptive System, and thus also of language (Steels 2000; Beckner et al. 2009). Such redundancy/degeneracy comes with two important advantages, viz. robustness and evolvability: most importantly for the present paper, the former entails that redundant marking offers protection against information loss in the noisy language channel, even though it may be less efficient. Redundancy is furthermore assumed to increase learnability, particularly in more complex situations (e.g. Tal et al. 2021).

Our case study to assess the plausibility of what we call the ‘strict-efficiency’ versus the ‘robustness’ account is participant role marking in ditransitive clauses in Present Day Dutch and English, for a comparative perspective, as well as historical English for a diachronic view. More precisely, we investigate the interaction between strategies used to distinguish agents and recipients in transfer-events, e.g. with verbs of giving as in (1) and (2).

- (1) They_{AGENT} give a book to the student_{RECIPIENT}.
- (2) Ze _{AGENT} geven een boek aan de student_{RECIPIENT}.

Since both agents and recipients in ditransitive clauses are prototypically animate (sentient) and volitional (e.g. Newman 1998; Naess 2007; Haspelmath 2015), disambiguating these roles based on semantic-pragmatic information is usually difficult if not impossible. Morpho-syntactic cues are hence indispensable in determining ‘who gave what to whom’. Among the strategies language users have at their disposal are (i) constituent order (e.g. SVO in Present Day English), (ii) case marking/ formal differentiation (e.g. subject vs object pronoun forms in PDE), (iii) subject-verb agreement, and (iv) prepositional marking. Employing multiple strategies at the same time constitutes redundant marking; for example, in (1) all four disambiguation strategies are given. Meanwhile in (3), none are used, resulting in an ambiguous sentence.

- (3) Mijn baas kan je niet zomaar een uitbrander geven.
 ‘You can’t just give my boss a telling-off’ or ‘My boss can’t just give you a telling off.’

In our study, we make use of the *Sonar Corpus of Written Dutch* (Oostdijk et al. 2013), a pre-compiled dataset of ditransitives from the *ICE-GB* (Röthlisberger 2018) and the *Penn Parsed Corpus of Middle English* (PPCME2; Kroch et al. 2000). Instances of ditransitive clauses with *give* are extracted from the corpora, and coded for the strategies instantiated by them. Following the ‘strict-efficiency’ account, we then expect language users to prefer employing a single strategy for each instance. By contrast, based on the degeneracy/ robustness account, we anticipate sentences that simultaneously instantiate multiple strategies to be most common, and cases where only one strategy is at work to be rare. Our results indicate that even though the precise strategies and their disambiguation power differ between Dutch and English, both languages show substantial redundancy to be the default. Still, redundancy seems to operate within limits, with four-fold strategy use being rare, and two simultaneous strategies being most common. Our diachronic results are in line with this conclusion: We find that English appears to have moved towards more redundant marking over time, but that after a short period of ‘exuberant’ redundancy, double redundancy is settled on as the norm.

In a final step, we assess the question of whether redundant marking is particularly frequent in complex environments, here measured as sentence length in words (excluding the subject and object arguments of the respective ditransitive patterns). Our findings are again mixed: for Dutch and historical English, complexity emerges as an influential predictor; in Present Day English, however, no significant effect can be observed. We interpret this outcome of our study in light of the differing degrees of variability of strategies in the languages/ stages.

References

- Beckner, C., Blythe, R., Bybee, J., Christiansen, M., Croft, W., Ellis, N., Holland, J., Ke, J., Larsen-Freeman, D., & Schoenemann, T. (2009). Language is a Complex Adaptive System: Position paper. *Language Learning*, 59(1), 1-26.
- Bybee, J. (2010). *Language, usage and cognition*. Cambridge: Cambridge University Press.
- Gibson, E., Futrell, R., Piantadosi, S.P., Dautriche, I., Mahowald, K., Bergen, L. & Levy, R. (2019). How efficiency shapes human language. *Trends in Cognitive Sciences*, 23(5), 389-407.
- Fedzechkina, M., Newport, E., & Jaeger, F. (2017). Balancing effort and information transmission during language acquisition: Evidence from word order and case marking. *Cognitive Science*, 41(2), 416-446.
- Haspelmath, M. (2015). Ditransitive constructions. *Annual Review of Linguistics* 1, 19-41.
- Kroch, A., Taylor, A., & Santorini, B. (2000-). The Penn-Helsinki Parsed Corpus of Middle English (PPCME2). Second edition, release 4.
- MacWhinney, B., Malchukov, A., & Moravcsik, E. (eds.). (2014). *Competing motivations in grammar and usage*. Oxford: Oxford University Press.
- Naess, A. (2007). *Prototypical transitivity*. Amsterdam: Benjamins.
- Newman, J. (1998). Recipients and ‘give’ constructions. In W. van Langendonck and W. Van Belle (Eds.), *The dative, Vol. 2: Theoretical and contrastive studies* (pp. 1-28). Amsterdam: Benjamins.
- Oostdijk, N., Reynaert, M., Hoste, V., & Schuurman, I. (2013). The construction of a 500-million-word reference corpus of contemporary written Dutch. In P. Spyns and J. Odijk (Eds.), *Essential speech and language technology for Dutch, Theory and applications of Natural Language Processing* (pp. 219-247). Heidelberg: Springer.
- Pijpops, D., Speelman, D., Grondelaers, S., & Van de Velde, F. (2018). Comparing explanations for the Complexity Principle. Evidence from argument realization. *Language and Cognition*, 10(3), 514-543.
- Röthlisberger, M. (2018). *The dative dataset of World Englishes*. KU Leuven.
- Steels, L. (2000). Language as a Complex Adaptive System. In M. Schoenauer, K. Deb, G. Rudolph, X. Yao, E. Lutton, J. Merelo and H.-P. Schwefel (Eds.), *Proceedings of PPSN VI* (pp. 17-26). Berlin: Springer.
- Tal, S., E. Grossman, H. Rohde & Arnon, I. (2021). Speakers use more redundant referents with language learners: Evidence for communicatively-efficient referential choice. Preprint. *PsyArXiv*.
- Tal, S., Smith, K., Culbertson, J., Grossman, E., & Arnon, I. (2020). The impact of information structure on the emergence of differential object marking: An experimental study. Preprint. *PsyArXiv*.
- Van de Velde, F. (2014). Degeneracy: the maintenance of constructional networks. In R. Boogaart, T. Colleman and G. Rutten (Eds.), *Extending the scope of Construction Grammar* (pp. 141-179). Berlin: Mouton de Gruyter.