

A methodological approach to integrate the vertical stratification of the constructicon in corpus research

Constructions, the basic descriptive units of construction grammar, are organized into networks that range from entirely concrete constructions to fully schematic ones (Fried and Östman 2004: 15–18). However, the vertical stratification of this ‘constructicon’ is only rarely taken into account in corpus studies with a constructionist background (for exceptions, see a.o. Boas 2010, Lenci 2012). We will argue that this leads to several problems, the most fundamental of which we call the Problem of Precedence. This problem asks which level in the constructicon is most relevant in the study of e.g. a particular alternation. To answer this question, we introduce a methodological approach that incorporates the vertical dimension.

To illustrate this approach, we investigate the alternation between the Dutch transitive construction and the *naar*-construction, as in (1)-(2) (cf. Haeseryn et al. 1997: 1168; Broekhuis 2004: 122), based on the Sonar corpus (Oostdijk et al. 2013). We start at the most schematic level and use collocational analyses to identify the verbs *doen* ‘do’ and *kijken* ‘look’ as the most prototypical verbs of the transitive and *naar*-construction. Next, we calculate the semantic proximity of each alternating verb to *doen* ‘do’ and *kijken* ‘look’ in distributional vector space, and use this proximity to predict the choice between both constructions.

We then descend a level in the constructicon and investigate the alternation separately for each alternating verb. Here, we apply the same procedure to the object slot. Finally, we descend even further and investigate the alternation separately for e.g. *zoek (naar) slachtoffer* ‘search victim’ and *zoek (naar) woord* ‘search word’. While we focus on semantic differences, we will control for processing-related and lexical factors.

- (1) *Zakkenrollers zoeken voornamelijk (naar) slachtoffers in het openbaar vervoer,...*
(WR-P-P-G-0000152205.p.4.s.3)
‘Pickpockets primarily search their victims in public transportation,...’
- (2) *Het soort liefde dat (...) deed haar (naar) nieuwe dingen verlangen.*
(WR-P-P-B-0000000132.p.1726.s.15)
‘The kind of love that (...) made her long for new things’

References

- Boas, Hans. 2010. The syntax-lexicon continuum in Construction Grammar. A Case study of English communication verbs. *Belgian Journal of Linguistics* 24(1). 54–82.
- Broekhuis, Hans. 2004. Het voorzetselvoorwerp. *Nederlandse Taalkunde* 9. 31–97.
- Fried, Mirjam and Jan-Ola Östman. 2004. Construction Grammar. A thumbnail sketch. In Mirjam Fried & Jan-Ola Östman (eds.), *Construction Grammar in a cross-language perspective*, 11–86. Amsterdam/Philadelphia: John Benjamins.
- Haeseryn, Walter, Kirsten Romijn, Guido Geerts, Jaap de Rooij and Maarten van den Toorn. 1997. *Algemene Nederlandse Spraakkunst*. Groningen: Nijhoff.
- Lenci, Alessandro. 2012. Argument alternations in Italian verbs: a computational study. In Valentina Bambini, Irene Ricci & Pier Marco Bertinetto (eds.), *Linguaggio e cervello - Semantica. Atti del XLII Congresso Internazionale di Studi della Societ`a di Linguistica Italiana*, 1–26. Rome: Bulzoni.
- 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.