

LOT
586

Jelke Bloem

Processing verb clusters

Processing verb clusters

In theory, linguists' lives are becoming easier due to the increasing availability of digital texts that can serve as data for linguistic research. The internet is full of language data and its volume is increasing at a faster rate than any linguist could keep up with. Yet this data is only used to a limited extent, because it is not always clear how we can find, among this sea of data, the utterances and constructions that are of particular interest to us linguists. When we do try to do this, it results in such a large amount of data that any further processing will have to be automatic, using models that can generalize over large volumes of data.

This dissertation describes the word order variation that exists in Dutch two-verb clusters, such as *gezien hebben* 'seen have' or *hebben gezien* 'have seen' by applying quantitative research methods to larger text collections. The results show that the processing complexity of the utterance plays an important role in the choice between these two word orders. The dissertation consists of eight articles covering historical and synchronic aspects of verb cluster order variation, with a focus on new, computational research methods that can be applied to the problem. Chapters two, three and six mainly concern the use of digital tools for linguistic research, while the chapters four, five, seven, eight and nine apply such tools to research questions concerning verb cluster variation.

ISBN 978-94-6093-371-4

DOI <https://dx.medra.org/10.48273/LOT0586>

Jelke Bloem

Processing verb clusters



Netherlands Graduate School of Linguistics
Landelijke Onderzoekschool Taalwetenschap



UNIVERSITY OF AMSTERDAM

Amsterdam Center for Language and Communication