

Nicole Grégoire

Untangling Multiword Expressions

A study on the representation and variation of Dutch multiword expressions

The central topic of this dissertation is the phenomenon *Multiword Expressions*. In this dissertation, Multiword Expressions (MWEs) are defined as combinations of words that have linguistic properties not predictable from the individual components or the normal way they are combined. MWEs are studied from multiple perspectives.

The first part of this study describes the design, implementation and population of DuELME, a Dutch Electronic Lexicon of Multiword Expressions. DuELME has been created out of a need for a large number of lexical descriptions of Dutch MWEs organized in such a way that they can be used in a wide variety of different grammatical frameworks, approaches to MWEs and their implementations. The result is a resource that contains lexical descriptions of over 5,000 MWEs, and that can be integrated into Dutch NLP systems with a minimal amount of manual effort.

A corpus-based study of the variation potential of MWEs has been carried out in the second part of this dissertation. Central in this part is the Idiom Variation Potential Hypothesis, which postulates that the presence of certain properties of the idiom parts are responsible for their in principle unlimited variation potential. The hypothesis has been tested for 25 idioms taken from DuELME using both corpus data and panel judgements as empirical material.

Untangling Multiword Expressions is of interest to those working in the domain of linguistics, corpus linguistics, and to those who are concerned with the creation of linguistic resources.

Netherlands
Graduate
School of
Linguistics

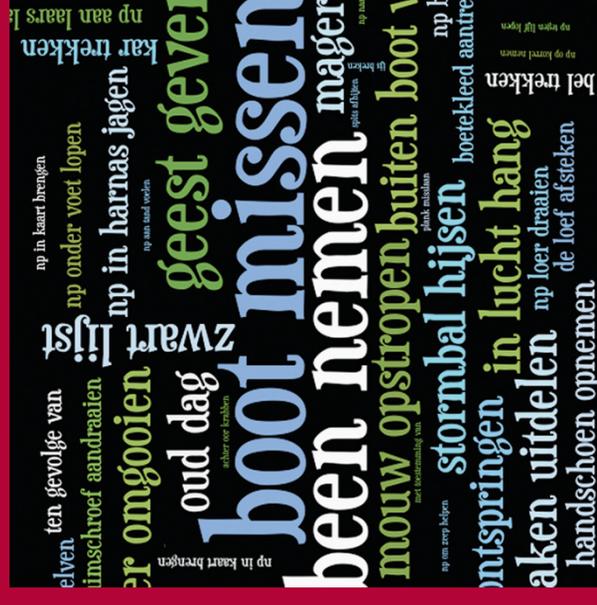
: LOT

ISBN 978-94-6093-010-2

Nicole Grégoire

Untangling Multiword Expressions

A study on the representation and variation of Dutch multiword expressions



Netherlands
Graduate
School of
Linguistics

: LOT

Landelijke Onderzoekschool Taalwetenschap

