

# Does separability have meaning?

A multimethod study on the semantics of Dutch verb-particle constructions

Ann-Sophie Vrielynck

(F.N.R.S. – Universities of Liège and Antwerp)

# My PhD journey so far...

- Joint PhD between Universities of Liège and Antwerp
- Supervisors: Julien Perrez and Dirk Pijpops
- 2023-2024: ARC project “Grammar from space: how spatial elements become applicatives”
  - WP3: “Diachrony of verb-particle constructions in Dutch”
- 2024-2025: FNRS project
  - “The case of the Dutch prefix *be-*: a multimethodological synchronic and diachronic approach to unravel alternation patterns”
- Thesis: “The applicatives in Dutch”

# The semantics of verb- particle constructions



**Literature**



**Study 1**

Methods

Analysis & results



**Study 2**

Methods

Analysis & results



**Conclusion**

Structure of presentation

# Verb-particle constructions

# Verb-particle constructions

- Non-verbal element + verb
  - e.g., *onder-gaan* ('go down; undergo'),  
*over-lopen* ('overflow; go through')
- Dutch: two types
  1. Separable
  2. Inseparable

# (In)separable verbs

## Separable verbs (SCV)

- *Hij **schakelt** de computer uit.*  
(‘He turns off the computer.’)
- Main stress on preverbal element
- Separation
  - V<sub>1</sub>/V<sub>2</sub> (main clauses)
  - ‘ge’ (past participle)
  - ‘te’ (infinitival particle)
  - modal verbs/auxiliaries

## Inseparable verbs (ICV)

- *Hij **ondergaat** zijn pijn.*  
(‘He undergoes his suffering.’)
- Main stress on base verb
- No separation

# 1) Booij (2002)

## Grammaticalisation process

- Resultative element → SCV → ICV
- Semantic bleaching of preverbal element

## Link separability – meaning

- Separable = concrete, spatial
- Inseparable = abstract, aspectual
- e.g., *doorlopen*

## Counterexamples

### Separable - **abstract**

- Opzoeken

(e.g., *de informatie opzoeken*) ('look up the information')

- Inlopen

(e.g., *je schoenen inlopen*) ('wear in one's shoes')

### Inseparable - **concrete**

- Doorzoeken

(e.g., *het huis doorzoeken*) ('search (through) the house')

- Overspoelen

(e.g., *het land overspoelen*) ('wash over the land')

## 2) Blom (2004, 2005)

### Two diachronic developments

1. Resultative element → SCV
2. Non-resultative element → SCV → ICV

### Link separability – meaning

- Resultative = SCV
  - e.g., *Hij **zaagt** de boom **om**.* ('He **cuts down** the tree.')
  - e.g., *De zon **gaat onder**.* ('The sun **goes down**.')
- Non-resultative = SCV/ICV
  - e.g., *Hij **ondergaat** de pijn.* ('He **undergoes** the pain.')
  - e.g., *Ze **kijkt** de jongen **aan**.* ('She **looks** at the boy.')

However,  
counterexamples

- **Inseparable and resultative**
  - *Onderdrukken* ('suppress')
  - *Onderwerpen* ('subject')
  - *Overtuigen* ('convince')
  - *Overreden* ('persuade')
  - *Overhandigen* ('hand over')

## Research questions

- What is the **semantic difference** between separable and inseparable verbs in contemporary Dutch?
  - 1.1. Do separable verbs have a more **concrete** meaning than inseparable verbs? (Booij 2002)
  - 1.2. Can the distinction separable – inseparable be explained by the distinction between **resultative** and non-resultative? (Blom 2004, 2005)

# Study 1

Corpus-based study

(→ Hypothesis Booij)

## Methods

- All verbs with the particles *door-*, *om-* and *over-*
  - < Van Dale (2022) & Reference File Dutch (van der Vliet 2007)

|              | SCV          | ICV          | Total      |
|--------------|--------------|--------------|------------|
| <i>Door-</i> | 124 (55,36%) | 100 (44,64%) | 224 (100%) |
| <i>Om-</i>   | 129 (64,18%) | 72 (35,82%)  | 201 (100%) |
| <i>Over-</i> | 189 (49,35%) | 194 (50,65%) | 383 (100%) |
| Total        | 442 (54,70%) | 366 (45,30%) | 808 (100%) |

- Link to dataset **Brysaert et al. (2014)**
  - ! limitations
- → **patients** as proxy for verb's meaning (cf. Frasinelli & Schulte im Walde 2019; Pijpops et al. 2021)

## Methods (2)

- 1) Collect corpus data
  - NITenTenzo (SketchEngine) & SoNaR (Oostdijk et al. 2013)
  - Past participles
- 2) Automatic identification of patients
  - SoNaR: Alpino-parses (van Noord 2006)
  - NITenTenzo: Spacy's library *nl\_core\_news\_sm* (Honnibal & Montani 2017)
- 3) Determine concreteness of each patients
  - < Dataset Brysbaert et al. (2014)
  - Removed: pronouns
  - Min. 10 attestations per verb

## Data

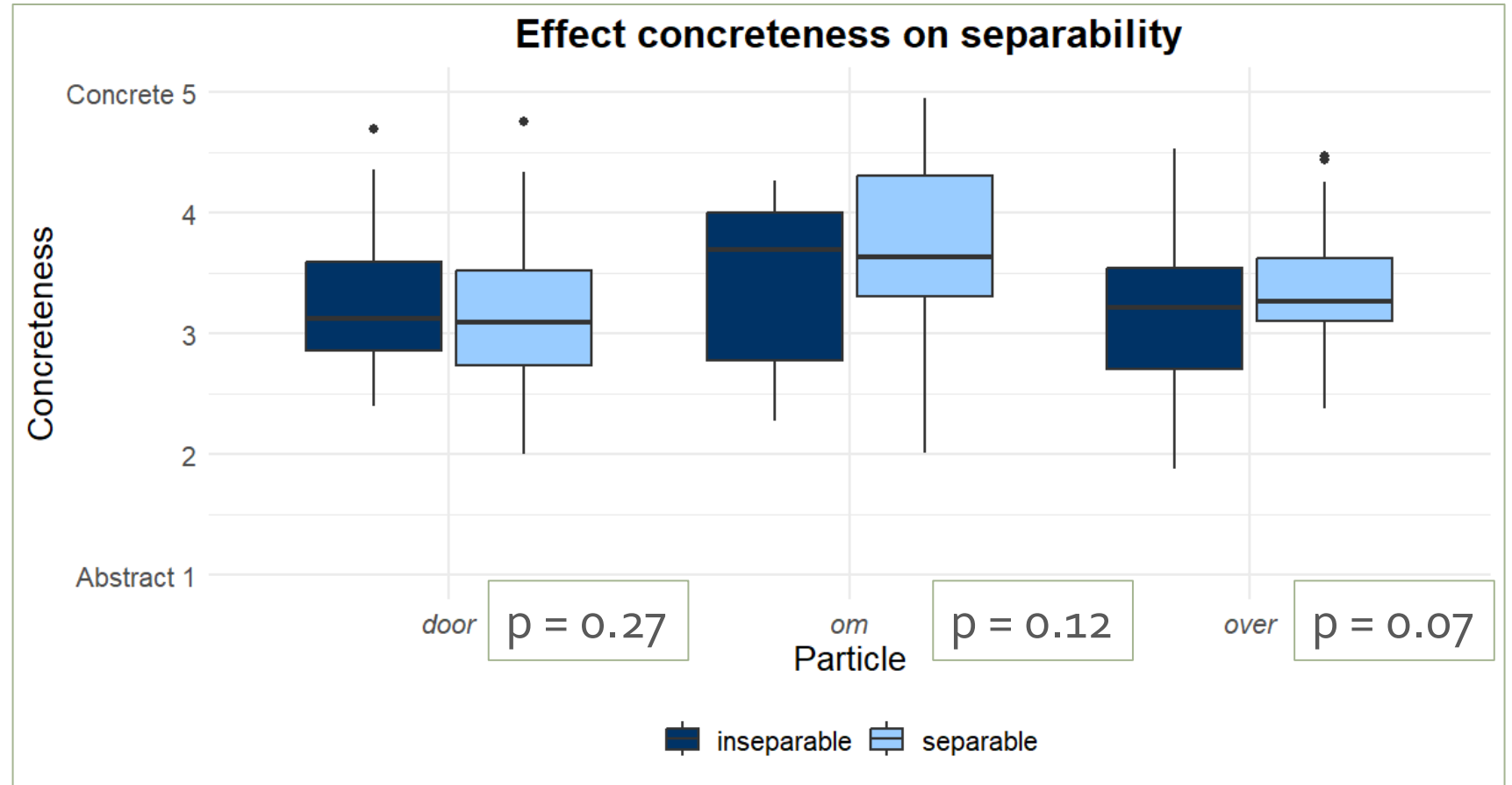
|       | SCV                                       | ICV                                       | Total  |
|-------|---|---|--|
| Door- | 52 verbs (67.53%)<br>15'712 attestations  | 25 verbs (32.47%)<br>9'642 attestations   | 77 verbs (100%)<br><b>25'354 attestations</b>  |
| Om-   | 48 verbs (64.86%)<br>12'944 attestations  | 26 verbs (35.14%)<br>7'455 attestations   | 74 verbs (100%)<br><b>20'399 attestations</b>  |
| Over- | 45 verbs (44.55%)<br>22'077 attestations  | 56 verbs (55.45%)<br>21'501 attestations  | 101 verbs (100%)<br><b>43'578 attestations</b> |
| Total | 144 verbs (57.37%)<br>50'733 attestations | 107 verbs (42.63%)<br>38'598 attestations | 251 verbs (100%)<br><b>89'331 attestations</b> |

# Analysis

- **Binomial logistic regression (R pakket 'glm' (R Core Team (2023))**
  - Average 'concreteness' of each verb
  - `model ← glm(Separability ~ Concreteness, data = d, family = binomial)`

# Results

- Overall p-value = 0.102



- *door- vs. om-/over-*

## Interim conclusion

- Link meaning & separability (cf. Booij 2002)
  - SCV = concrete and spatial
  - ICV = abstract and spectral
- Differences between individual particles

# Study 2

Experimental study

(→ Hypotheses Booij & Blom)

# Methods

- Forced-choice task
- Material
  - Non-verbs (*door-*, *om-* and *over-*) (De Smet 2021; De Smet & Rosseel 2023; ChatGPT)
  - 72 fill-in-the-blank sentences with choice: SCV or ICV (ChatGPT)
    - Resultativity (*zodat* ('so that') vs. *en* ('and'))
    - Concreteness (cf. Brysbaert et al. 2014)



Gedurende 2019 ... we vaak bij opa en oma thuis. (speven)

Gedurende 2019 speefden we vaak bij opa en oma thuis.

Gedurende 2019 spaven we vaak bij opa en oma thuis.

## Methods (2)

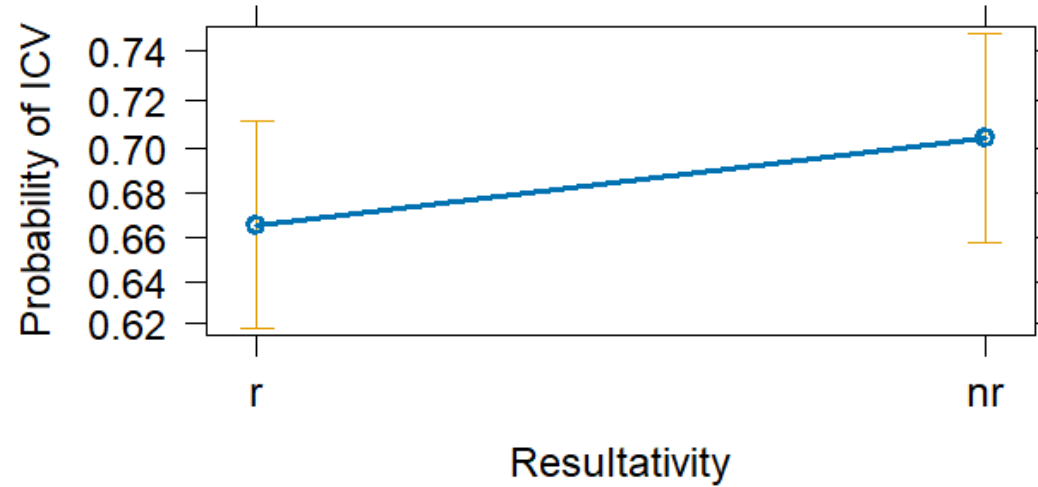
- Participants
  - N = 105
  - 18-25 years old
  - BA2 students of language & literature
  - L1 = Dutch

# Analysis

- Mixed effects regression model per particle
- Factors
  - Concreteness
  - Resultativity
  - Construction
  - Display order
  - Priming
  - Trial number
- Random effects: participant & verb

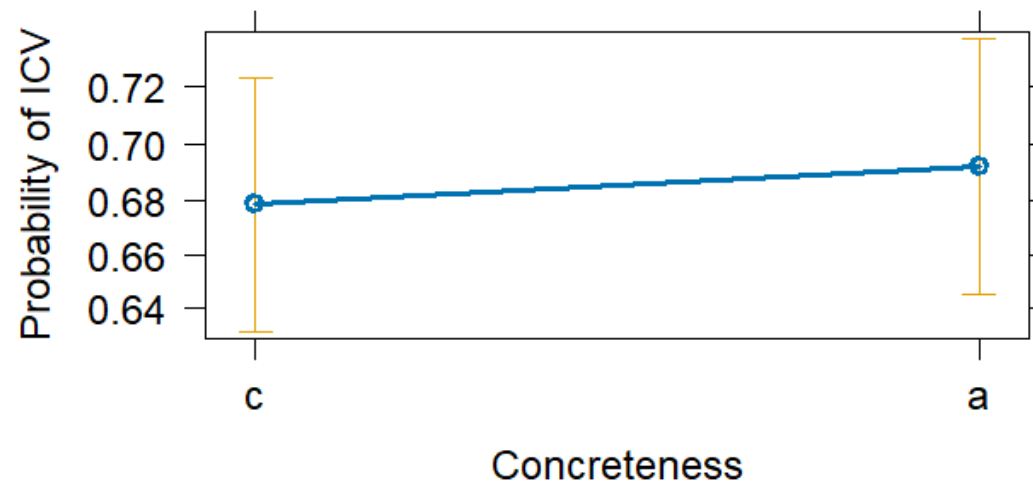
# Results *door-*

## Resultativity - separability effect (*door-*)



$p = 0.05.$

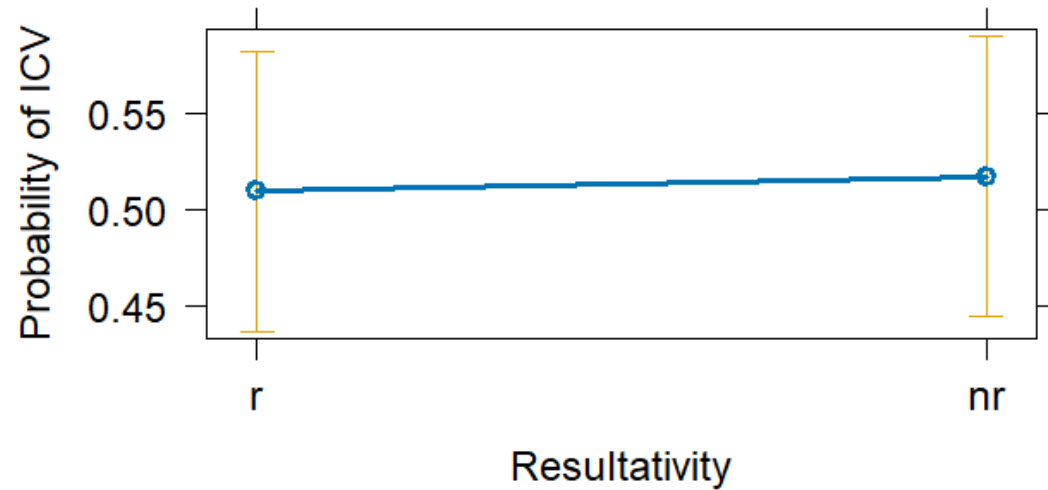
## Concreteness - separability effect (*door-*)



$p = 0.50$

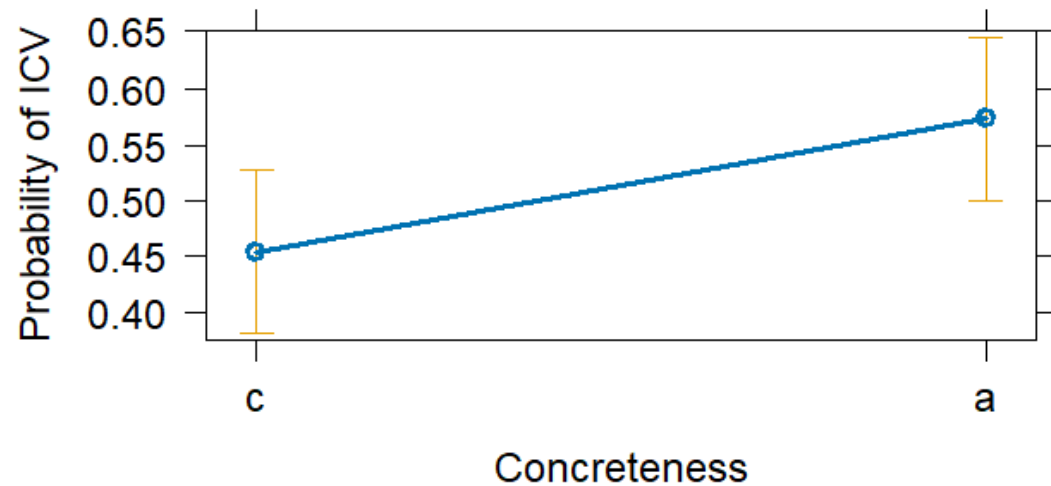
## Results *om-*

### Resultativity - separability effect (*om-*)



$p = 0.73$

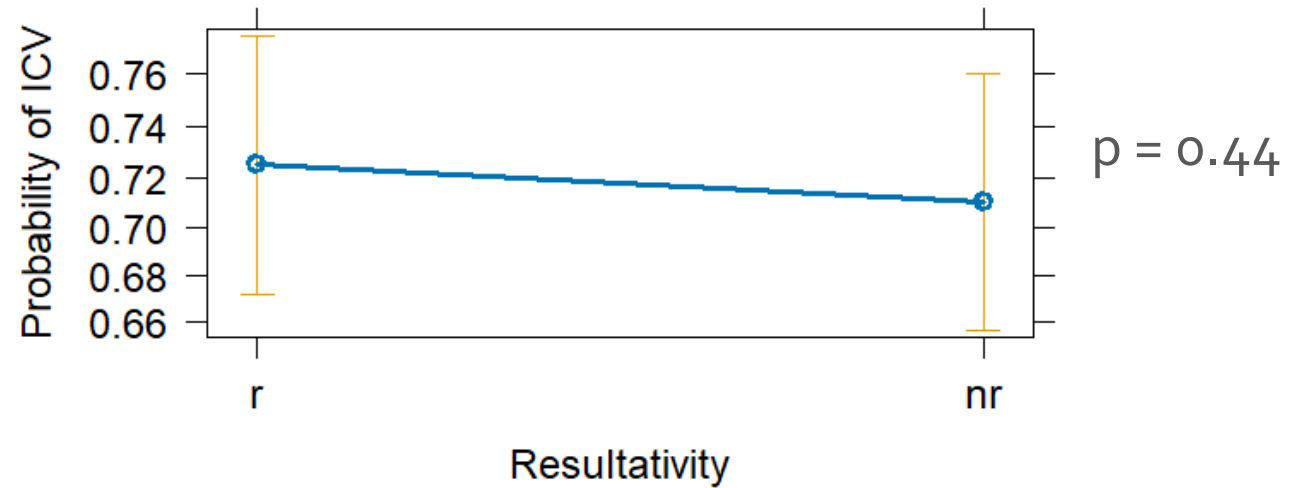
### Concreteness - separability effect (*om-*)



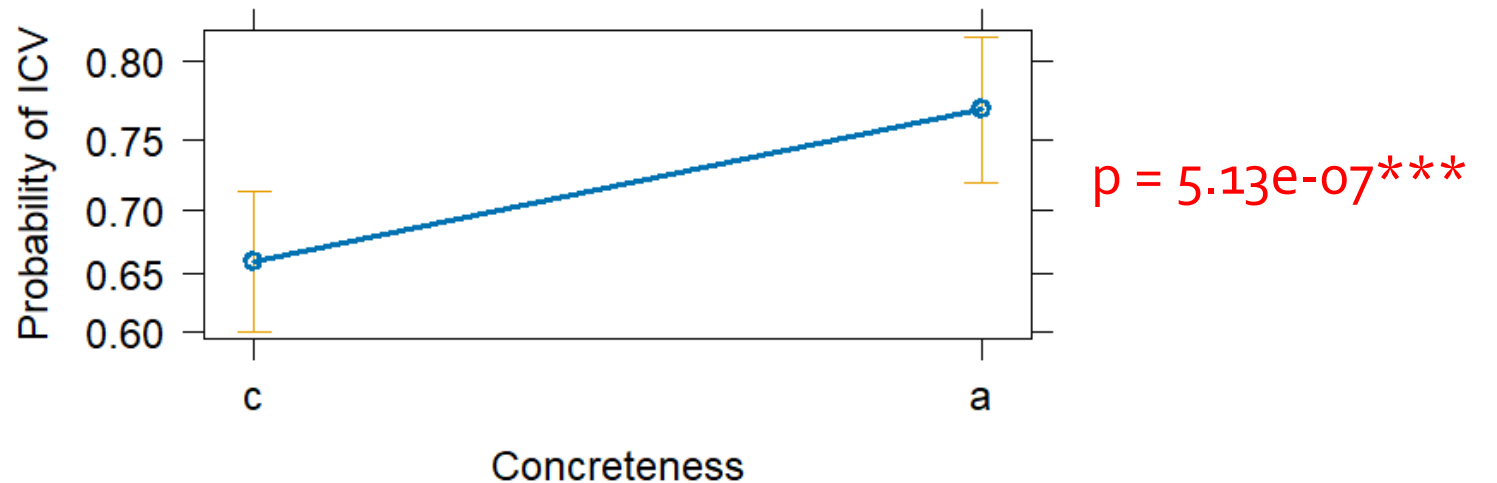
$p = 4.35e-07***$

Results *over-*

### Resultativity - separability effect (over-)



### Concreteness - separability effect (over-)



# Conclusion

## Interpretation of the results

- On a cognitive level: semantic difference between SCV and ICV
- Significant influence of both
  - Concreteness (*om-*, *over-*) (cf. Booij 2002)
  - Resultativity (*door-*) (! marginally significant) (cf. Blom 2004, 2005)
    - However: not shown in corpus study (but limitations!)
- Differences between individual particles
  - *Door-* more salient resultative meaning vs. *om-* and *over-*?

# Thank you!

Questions?

[ann-sophie.vrielynck@uliege.be](mailto:ann-sophie.vrielynck@uliege.be)



## References

- Blom, Corrien (2004). On the diachrony of complex predicates in Dutch: predicative and nonpredicative preverbs. *Journal of Germanic Linguistics* 16. 1–75.
- Blom, Corrien (2005). *Complex predicates in Dutch: synchrony and diachrony*. Utrecht: Netherlands Graduate School of Linguistics.
- Booij, Geert (2002). *The morphology of Dutch*. Oxford: Oxford University Press.
- Brysbaert, Marc, Stevens, Michaël, De Deyne, Simon, Voorspoels, Wouter & Gert Storms (2014). Norms of age of acquisition and concreteness for 30,000 Dutch words. *Acta Psychologica* 150. 80–84.
- Van Dale groot woordenboek van de Nederlandse taal* (2022). Utrecht: Van Dale Uitgevers.
- Frassinelli, Diego & Sabine Schulte im Walde (2019). Distributional interaction of concreteness and abstractness in verb-noun subcategorisation. *Proceedings of the 13th International Conference on Computational Semantics – Short Papers*. 38–43.
- Honnibal, Matthew & Ines Montani (2017). *spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing*.
- [nlTenTen20] Corpus of the Dutch Web 2020, *Sketch Engine*. <https://www.sketchengine.eu/nltenten-dutch-corpus/> (31 mei, 2024)
- van Noord, Gertjan. (2006). At Last Parsing Is Now Operational. In Piet Mertens, Cédric Fairon, Anne Dister & Patrick Watrin (eds.), TALN 2006. *Verbum Ex Machina. Actes de la 13e conférence sur le traitement automatique des langues naturelles*. 20–42. Louvain-la-Neuve: Cental.)

## References (2)

Oostdijk, Nelleke, Reynaert, Martin, Hoste, Véronique & 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.

OpenAI. (2025). *ChatGPT* [Large language model]. chat.openai.com/chat. (Version of March 31, 2025)

Pijpops, Dirk, Speelman, Dirk, Grondelaers, Stef & Freek Van de Velde. 2021. Incorporating the multi-level nature of the construction into hypothesis testing. *Cognitive Linguistics* 32(3). 487–528.

R Core Team (2023). R: A Language and Environment for Statistical Computing\_. R Foundation for Statistical Computing, Vienna, Austria. <<https://www.R-project.org/>>.

De Smet, Isabeau. (2021). *De sterke werkwoorden in het Nederlands. Een diachroon, kwantitatief onderzoek*. KU Leuven [PhD dissertation].

De Smet, Isabeau & Laura, Rosseel. (2023). Who's afraid of homophones? A multimethodological approach to homophony avoidance. *Language and Cognition*. 1–24.

SoNaR-corpus (Version 1.2.1) (2015) [Data set]. Available at the Institute of the Dutch language: <http://hdl.handle.net/10032/tm-a2-h5>

van der Vliet, Hennie (2007). The Referentiebestand Nederlands as a Multi-Purpose Lexical Database. *International Journal of Lexicography* 20(3). 239–257.