

Diachronic corpus study on Dutch compound verbs: a pilot study

Ann-Sophie Vrielynck (University of Liège and Antwerp)

Supervisor: Julien Perrez (University of Liège)

Co-supervisor: Dirk Pijpops (University of Antwerp)



Introduction

Compound verbs

- Non-verbal element (spatial) + verb
 - E.g., *onder-gaan* ('go down'), *onder-gaan* ('undergo'), *uit-schakelen* ('turn off')
- Applicative function
- Traditional categorisation
 - Separable
 - Inseparable

Separable compound verb (SCV)

- E.g., "De zon **gaat onder**." ('The sun goes down.')
- Main stress on particle
- Separation of the preverbal element and the base verb
 - V1/V2 position in main clauses
 - Past participle marker 'g(h)e-'
 - Infinitival marker 'te'
 - Middle Dutch negation marker 'ne' or 'en'

Inseparable compound verb (ICV)

- E.g., "Hij **ondergaat** zijn lijden." ('He undergoes his suffering.')
- Main stress on base verb
- Remains a compound verb
 - In V1/V2 position in main clauses
 - After the infinitival marker 'te'

Diachronic development (1): i.a. Booij 2002

- Grammaticalisation:

Resultative phrases → separable particles → inseparable prefixes

- Semantic bleaching of the preverbal element
- E.g., door-lopen/door-lopen (Booij 2002: 218)
 - het bos door-lopen (SCV) -> 'to walk through the forest'
 - de cursus door-lopen (ICV) -> 'to complete the course'

But...

- Possible counterexamples (cf. Blom 2005: 71-72)

1) SCVs without concrete, spatial meanings

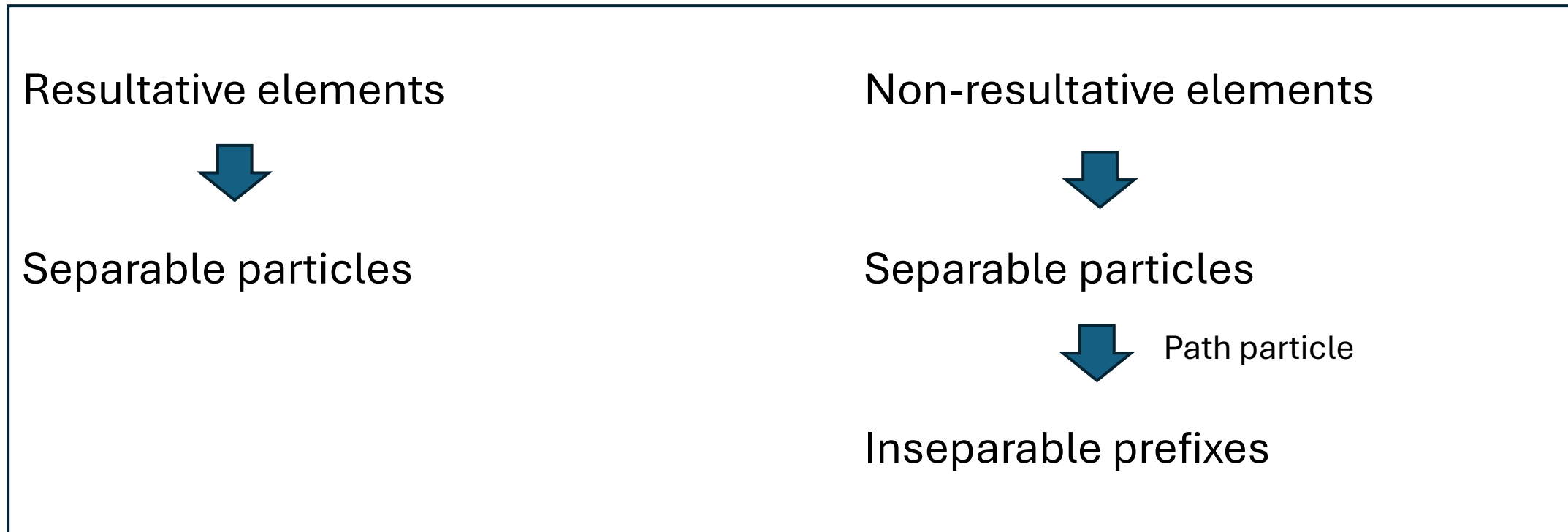
- "De informatie **opzoeken**" (up-search) 'to look up the information'
- "Je schoenen **inlopen**" (in-walk) 'to wear in one's shoes'

2) ICVs with spatial meanings

- "Het huis **doorzoeken**" (through-search) 'to search the house'
 - -> the searching goes through the house
- "Het huis **omgeven**" (around-give) 'to surround the house'
 - -> something is around the house

Diachronic development (2): Blom (2004, 2005)

- Grammaticalisation:



Diachronic development (2): Blom (2004, 2005)

- Hypotheses Blom:
 - Separable before inseparable (= Booij 2002)
 - Distinction resultative/non-resultative
 - Resultative = preverbal element expresses a result that affects the direct object (transitive verb) or the subject (intransitive verb)
 - Semantic change (loss resultative meaning) as necessary condition for evolution separable/inseparable
- Hypotheses confirmed with Middle Dutch data (CD Middle Dutch 1998)

But...

- Possible counterexamples: resultative and inseparable
 - *Onderwerpen* ('subject')
 - *Onderdrukken* ('suppress')
 - *Overtuigen* ('convince')
 - *Overreden* ('persuade')
 - *Overhandigen* ('hand over')
- « not productive and not systematically used with inseparable verbs » (Blom 2004: 40, 43)

Diachronic development (3): van der Horst (2008)

- Two independent developments

Adverb + verb



Separable verb

Adverb + verb



Inseparable verb

Research questions

Research questions

- 1) Can we validate the hypotheses proposed by Blom based on new, different corpus data?
 - a) Is there a distinction between resultative and non-resultative elements in the development of compound verbs?
 - b) Did the inseparable verbs evolve from the separable ones?
 - c) Did a semantic shift (towards a holistic meaning) precede the structural change from separable to inseparable?
- 2) How can we explain the exceptions to Blom's hypotheses (2005)?

Methodology

Corpus data

1) **The EDGeS Diachronic Bible Corpus** (Bouma et al. 2020)

- 13 Dutch Bible texts (1360-1939)
- 9,126,714 tokens

2) **Dbnl (The Digital Library of Dutch Literature)** (Van Olmen 2019)

- Literary texts (16th-20th century)
- 9,766,040 tokens

Data-extraction

- Selection of verbs
 - Exceptions to the hypotheses of Blom
 - *Onderwerpen* ('subject')
 - *Onderdrukken* ('suppress')
 - *Overtuigen* ('convince')
 - *Overreden* ('persuade')
 - *Overhandigen* ('hand over')
 - *Ondergaan* ('go down, undergo')
- Regular expressions

Data cleaning

- Manually removed:
 - E.g., *Overreden* used as past participle of *overrijden* ('run over')
 - E.g., *Onderwerp* used as a noun ('the subject')

Data

Verb	Number of attestations
<i>Onderdrukken</i> ('suppress')	236
<i>Onderwerpen</i> ('subject')	668
<i>Ondergaan</i> ('go down, undergo')	706
<i>Overreden</i> ('persuade')	145
<i>Overhandigen</i> ('hand over')	101
<i>Overtuigen</i> ('convince')	828

Annotation

- 1) Separability: separable, inseparable, ambiguous
- 2) Meaning of the verb
 - Resultativity: resultative or non-resultative
 - Semantics (cf. Middle Dutch Dictionary & Dictionary of the Dutch Language)
 - Compositional (C) versus holistic (H)
- 4) Period
 - 1300-1500 = Middle Dutch
 - 1501-1700 = Early Modern Dutch
 - 1701-1900 = Modern Dutch
 - 1901-2000 = Contemporary Dutch

Example of annotation

- "**Ende doen de Sonne onder ghegaen was, [...]**" (EDGeS corpus, nl_1560_Biestkens-allverses.txt) (‘And when the sun has gone down’)
 - 1) SCV (*ghe-*)
 - 2) Meaning
 - Resultative (-> the sun is down)
 - = The setting of celestial bodies
 - Compositional (*onder* ‘down’ + *gaan* ‘go’)
 - 3) Period
 - 1560 -> Early Modern Dutch

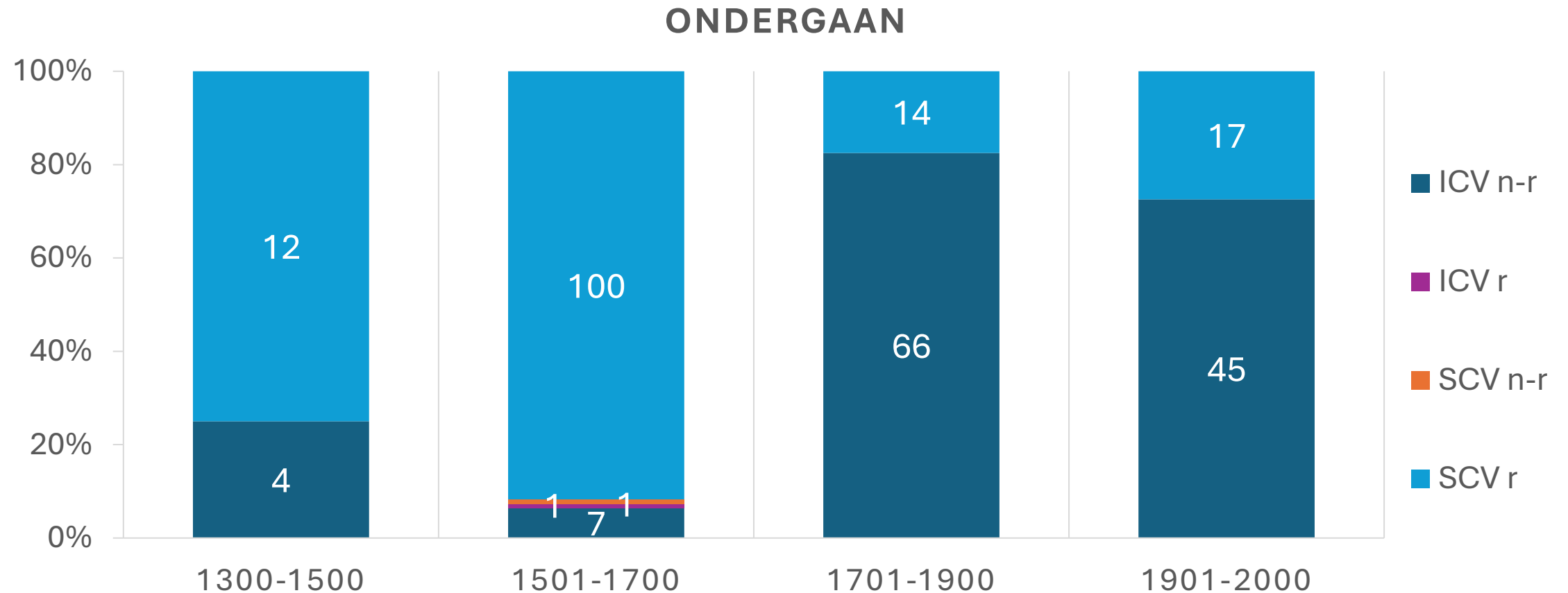
Results

Hypothesis testing and hypothesis generation

1) Hypothesis testing

Ondergaan ('to go down, undergo')

- Difference resultative/non-resultative

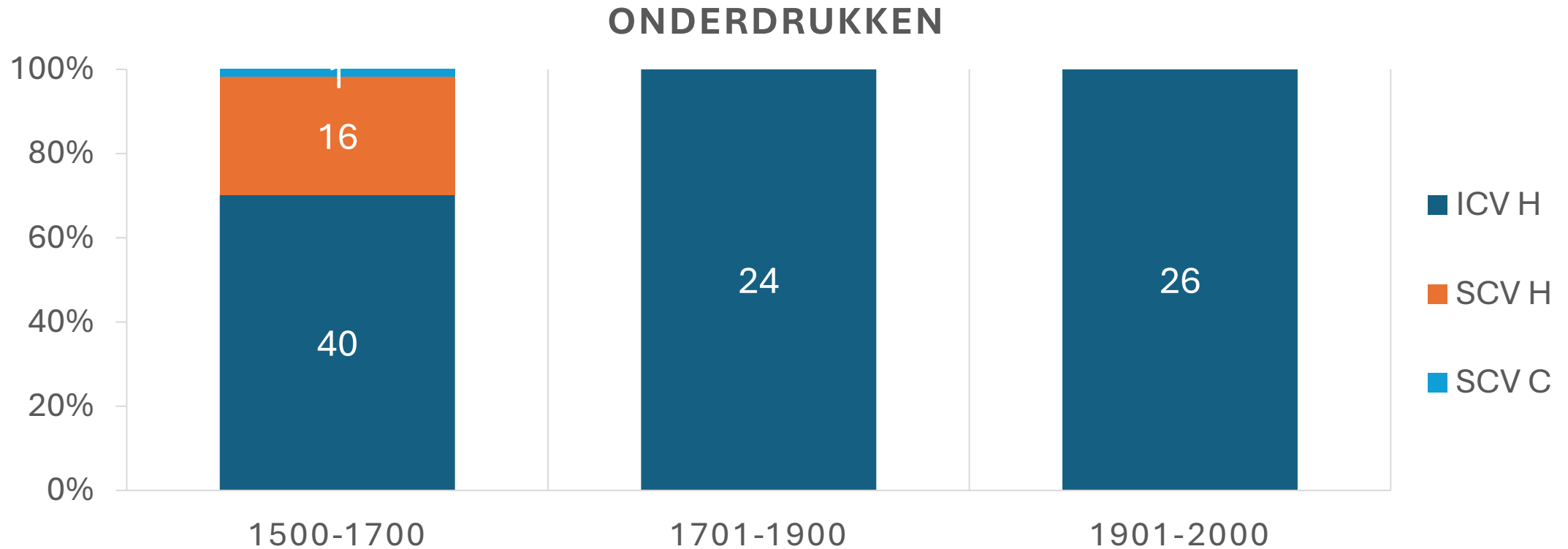


Exception

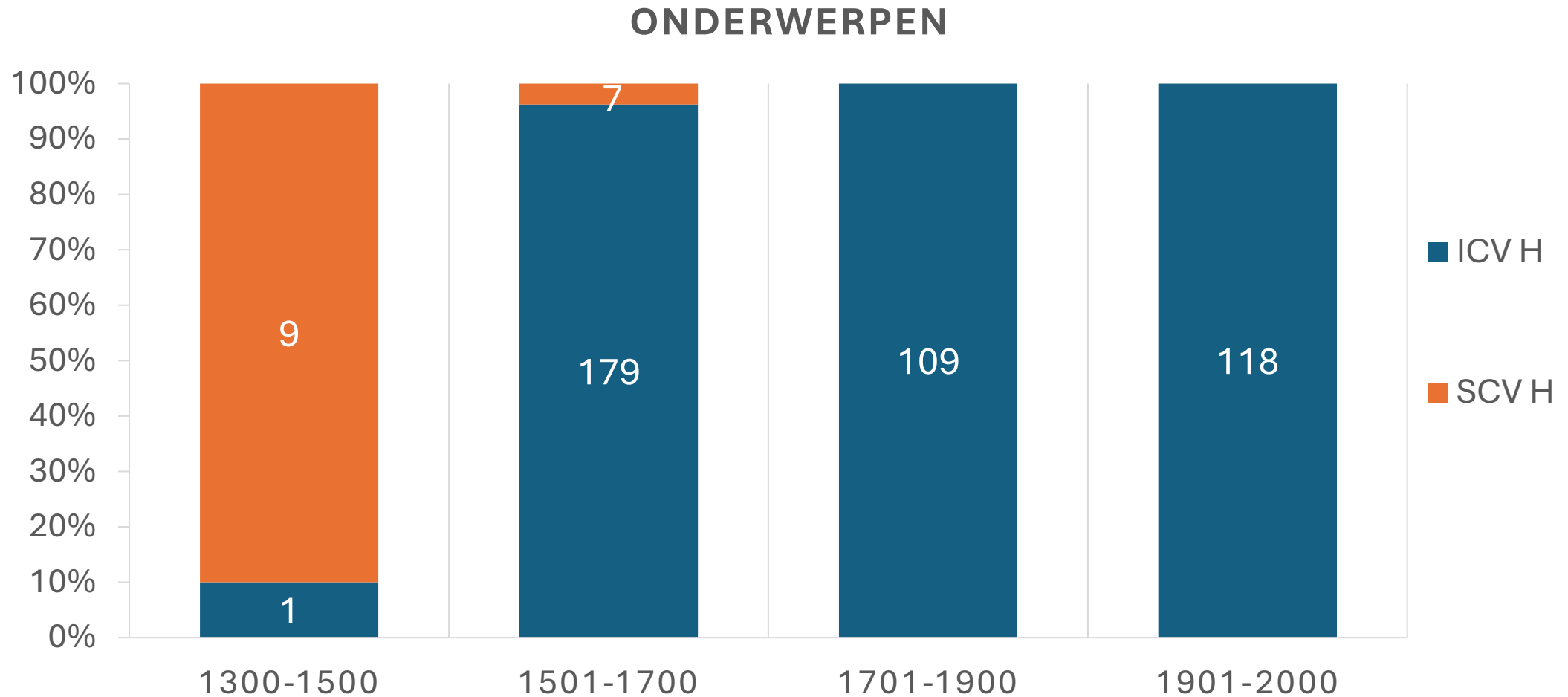
- Inseparable and resultative
 - "maer doe de Sonne **onder-gaen** was" (EDGeS corpus, nl_1648_Lutherse-allverses.txt)
 - ! Past participle marker 'g(h)e-' : not consistently used in Middle Dutch data
 - Possibly a remnant of this

Onderdrukken ('to suppress')

- Separable before inseparable
- Semantic change before formal change



Onderwerpen ('to subject')



Overhandigen ('to hand over'), *overtuigen* ('to convince') and *overreden* ('to persuade')

- Always inseparable
- ! No semantic change (vs. our hypothesis)
 - *Overhandigen* and *overreden*: always compositional
 - *Overtuigen*: always holistic

2) Hypothesis generation

Why resultative and inseparable?

Overtuigen ('to convince')

- Originally not resultative
- *tuigen = related to Middle Dutch « tugen »
- 'to prove by an incriminating statement that someone is guilty'
 - Attested in our data
- Probably developed according to non-resultative elements
- But:
 - Non-resultative SCVs of this subtype did not evolve into ICVs (cf. Blom)

Overreden ('to persuade') and *overhandigen* ('to hand over')

- *Overreden* ('to persuade')
 - loanword from the German *überreden*
 - Also inseparable
 - But: why?
- *Overhandigen* ('to hand over')
 - Development of the verb itself (Van Loey 1979): always stress on 'hand'

Handreiken (14th C) → *verhandreiken* (15th C) → *overhandreiken*
/behanden/behandigen (16th C) → *overhanden/overhandigen* (18th C)

Overhandigen ('to hand over'), *overtuigen* ('to convince') and *overreden* ('to persuade')

- “Transfer verbs” (cf. Dewell 1996)
 - Stress on abstract transition rather than on result
 - Between resultatives and path particles
 - Resultative, because transition/change affects direct object
 - Path particle, because transition rather than change of state
 - -> new (sub-)category needed?
- Analogy?
 - Most compound verbs with *over-* are inseparable

Onderwerpen and onderdrukken

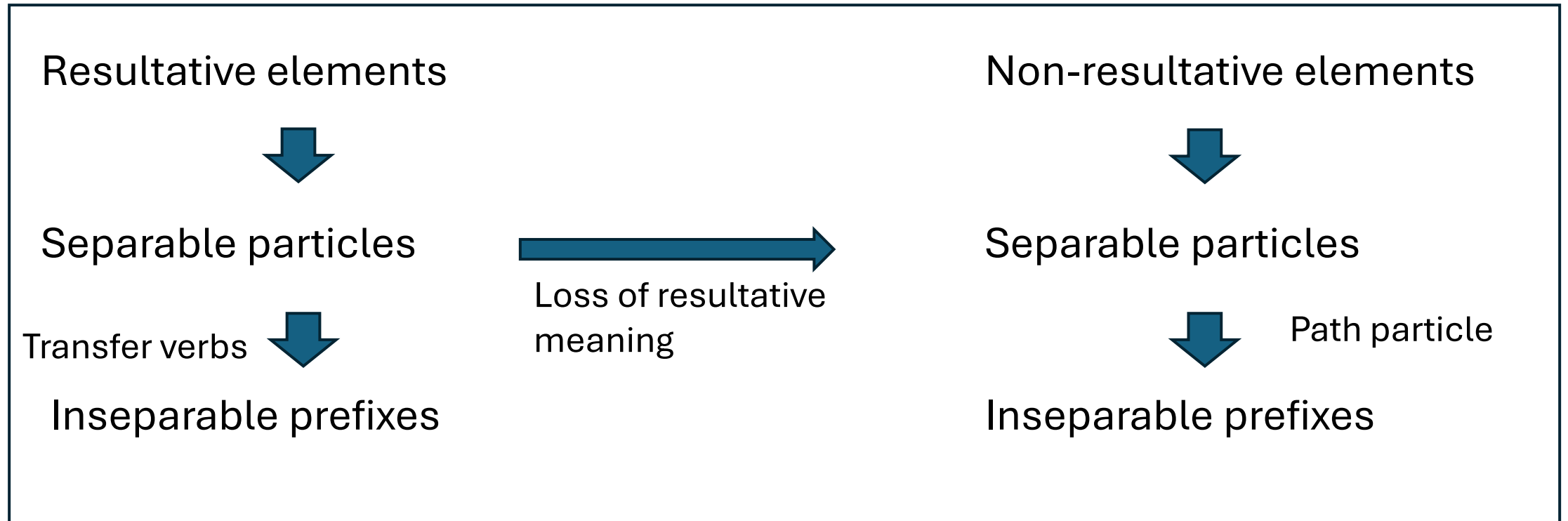
- -> Loss of compositional meaning?

Conclusions

Conclusion

- Testing of hypotheses
 - ☑ Resultative versus non-resultative (cf. = Blom, ≠ Booij)
 - ☑ Separable before inseparable (cf. = Booij, Blom)
 - But: not all verbs ! (Van der Horst?)
 - ☑ Semantic change -> formal change (cf. = Booij, Blom)
- Generation of hypotheses
 - Subtypes of resultatives possible?
 - “Transfer verbs” (*overreden, overhandigen, overtuigen*)
 - Holistic resultatives (*onderdrukken, onderwerpen*)
 - Still difficult to explain...

Grammaticalisation path



Thank you!

More information?

ann-sophie.vrielynck@uliege.be

References

- Blom, C. (2004). On the Diachrony of Complex Predicates in Dutch: Predicative and Nonpredicative Preverbs. *Journal of Germanic Linguistics* 16(1), 1-75.
- Blom, C. (2005). *Complex Predicates in Dutch: Synchrony and Diachrony*. Netherlands Graduate School of Linguistics.
- Blom, C. & G. Booij. (2003). The diachrony of complex predicates in Dutch: A case study in grammaticalization. *Acta Linguistica Hungarica* 50, 61-91.
- Booij, G. (2002). *The Morphology of Dutch*. Oxford: Oxford University Press.
- Bouma, G., Coussé, E., Dijkstra, T., & N. van der Sijs (2020). The EDGeS Diachronic Bible Corpus. *Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020)*, 5232-5239.
- Dewell, R. (1996). The Separability of German über-: A Cognitive Approach. *The construal of space in language and thought*, 109-133.
- De Vries, M. & A. te Winkel. (1851-1998). *Woordenboek der Nederlandsche taal*. Geraadpleegd op 26 februari 2024. <http://gtb.inl.nl>
- van der Horst, J. (2008). Vaste werkwoordelijke verbindingen in het Nederlands en het Afrikaans. *Atn&a*, 29-43.
- Van Loey, A. (1979). Overhandigen. *Verlagen en mededelingen van de Koninklijke Academie voor Nederlandse taal- en letterkunde (nieuwe reeks), jaargang 1979*, 85-104. Geraadpleegd op 26 februari 2024. https://www.dbnl.org/tekst/_ver016197901_01/_ver016197901_01_0009.php?q=overgehandigd#hl1
- Van Olmen, D. (2019). A diachronic corpus study of prenominal zo'n 'so a' in Dutch: Pathways, analogy and (inter)subjectification. *Functions of Language* 26(2), 217-248.