

At what level of schematicity should we investigate argument structure variation?

Dirk Pijpops

RU Lilith, University of Liège

1. Dirk De Wilde **verlangt (naar)** wat meer creativiteit.
'Dirk De Wilde desires some more creativity.'
2. Automobilisten **zoeken** altijd **(naar)** de goedkoopste oplossing.
'Car drivers always look for the cheapest solution.'
3. Samen **zoeken** zij **(naar)** een oplossing.
'Together, they are looking for a solution.'
4. Hij **zoekt (naar)** de juiste woorden.
'He is looking for the right words.'
5. 'What's in a name' dachten we en **peilden (naar)** jullie mening.
"What's in a name", we thought, and gauged your opinion.'

1. Dirk De Wilde **verlangt (naar)** wat meer creativiteit.

'Dirk De Wilde desires some more creativity.'

2. Automobilisten **zoeken** altijd **(naar)** de goedkoopste oplossing.

'Car drivers always look for the cheapest solution.'

3. Samen **zoeken** zij **(naar)** een oplossing.

'Together, they are looking for a solution.'

4. Hij **zoekt (naar)** de juiste woorden.

'He is looking for the right words.'

5. 'What's in a name' dachten we en **peilden (naar)** jullie mening.

'“What's in a name”, we thought, and gauged your opinion.'

Transitive construction vs. *Naar*-construction

Transitive vs. Prepositional
verlang-cxn *verlang-naar*-cxn

Transitive vs. Prepositional
zoek-cxn *zoek-naar*-cxn

Transitive vs. Prepositional
peil-cxn *peil-naar*-cxn

...

At what level of schematicity should we investigate argument structure variation?

Lexical approach

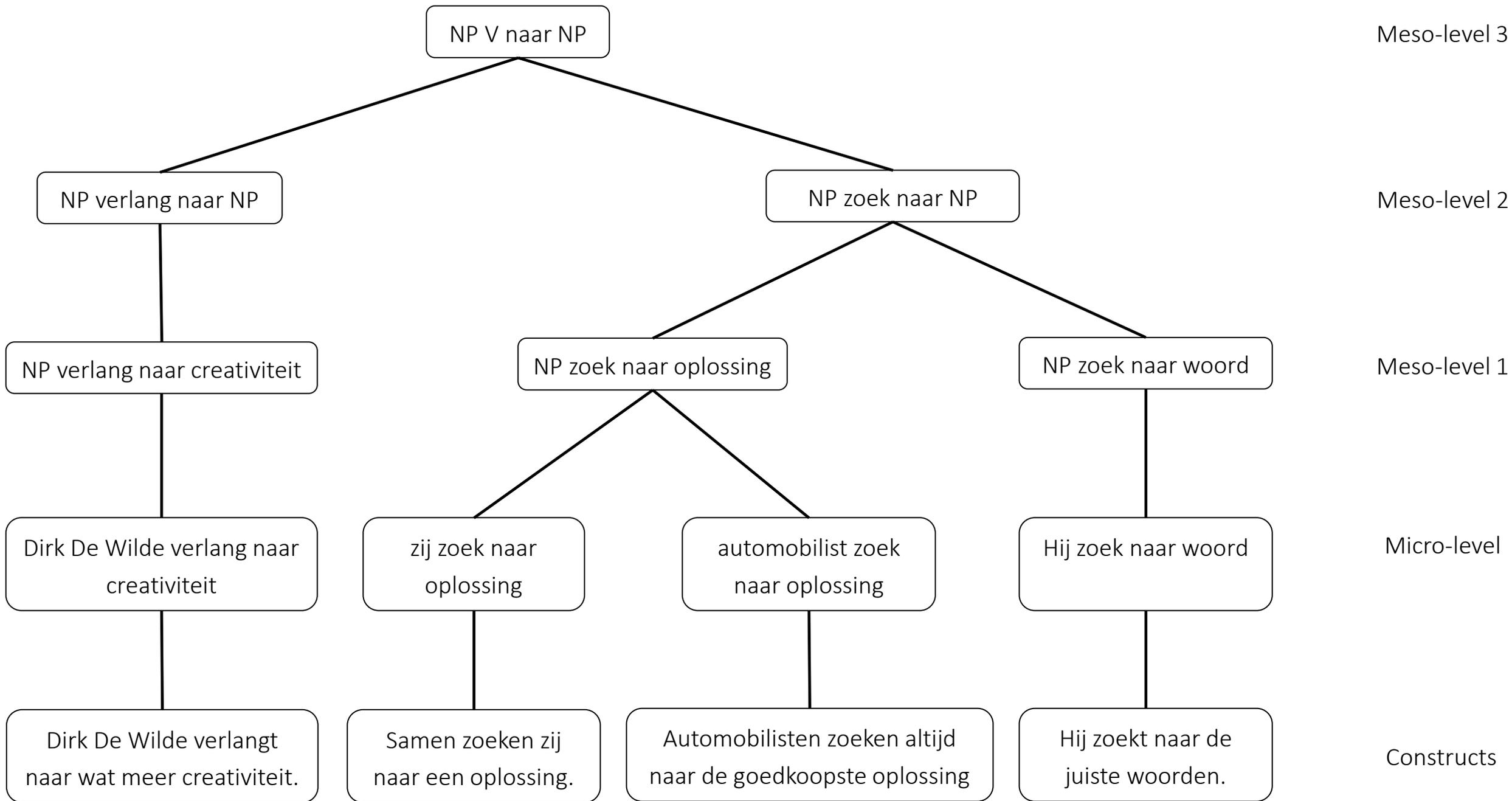
vs.

Phrasal approach

Hybrid approach

At what level of schematicity should we investigate argument structure variation?

Which are the possible levels of schematicity?



Formulating a hypothesis at Meso-level 3

Formulating a hypothesis at Meso-level 3

- Transitive construction vs. *naar*-construction

NP V NP

vs.

NP V naar NP

- Lexical Origin Hypothesis: schematic constructions obtain their meaning from their most prototypical lexical slot fillers

Transitive construction

Collexeme	Collostr. strenght
<i>hebben</i> 'have'	4,101,778
<i>krijgen</i> 'get'	2,602,924
<i>doen</i> 'do'	2,410,269
<i>zien</i> 'see'	1,297,011
<i>nemen</i> 'take'	1,190,952

Naar-construction

Collexeme	Collostr. strenght
<i>gaan</i> 'go'	923,830
<i>kijken</i> 'look'	532,119
<i>trekken</i> 'pull'	272,823
<i>verwijzen</i> 'refer'	258,146
<i>komen</i> 'come'	251,058

- Principle of Semantic Coherence: a lexical filler, e.g. a verb, will more readily combine with a construction that is semantically coherent, i.e. that has a similar or a compatible meaning

Transitive construction

Collexeme	Collostr. strenght
<i>hebben</i> 'have'	4,101,778
<i>krijgen</i> 'get'	2,602,924
<i>doen</i> 'do'	2,410,269
<i>zien</i> 'see'	1,297,011
<i>nemen</i> 'take'	1,190,952

Naar-construction

Collexeme	Collostr. strenght
<i>gaan</i> 'go'	923,830
<i>kijken</i> 'look'	532,119
<i>trekken</i> 'pull'	272,823
<i>verwijzen</i> 'refer'	258,146
<i>komen</i> 'come'	251,058

For each alternating $verb_i$:

VERBAL SEMANTIC COHERENCE TO THE NAAR-CXN =

$$10 \left(\frac{\sum_{n=1}^5 sim_{cm}(\overrightarrow{\text{naar-cxn collexeme}_n}, \overrightarrow{verb_i})}{5} - \frac{\sum_{n=1}^5 sim_{cm}(\overrightarrow{\text{trans-cxn collexeme}_n}, \overrightarrow{verb_i})}{5} \right)$$

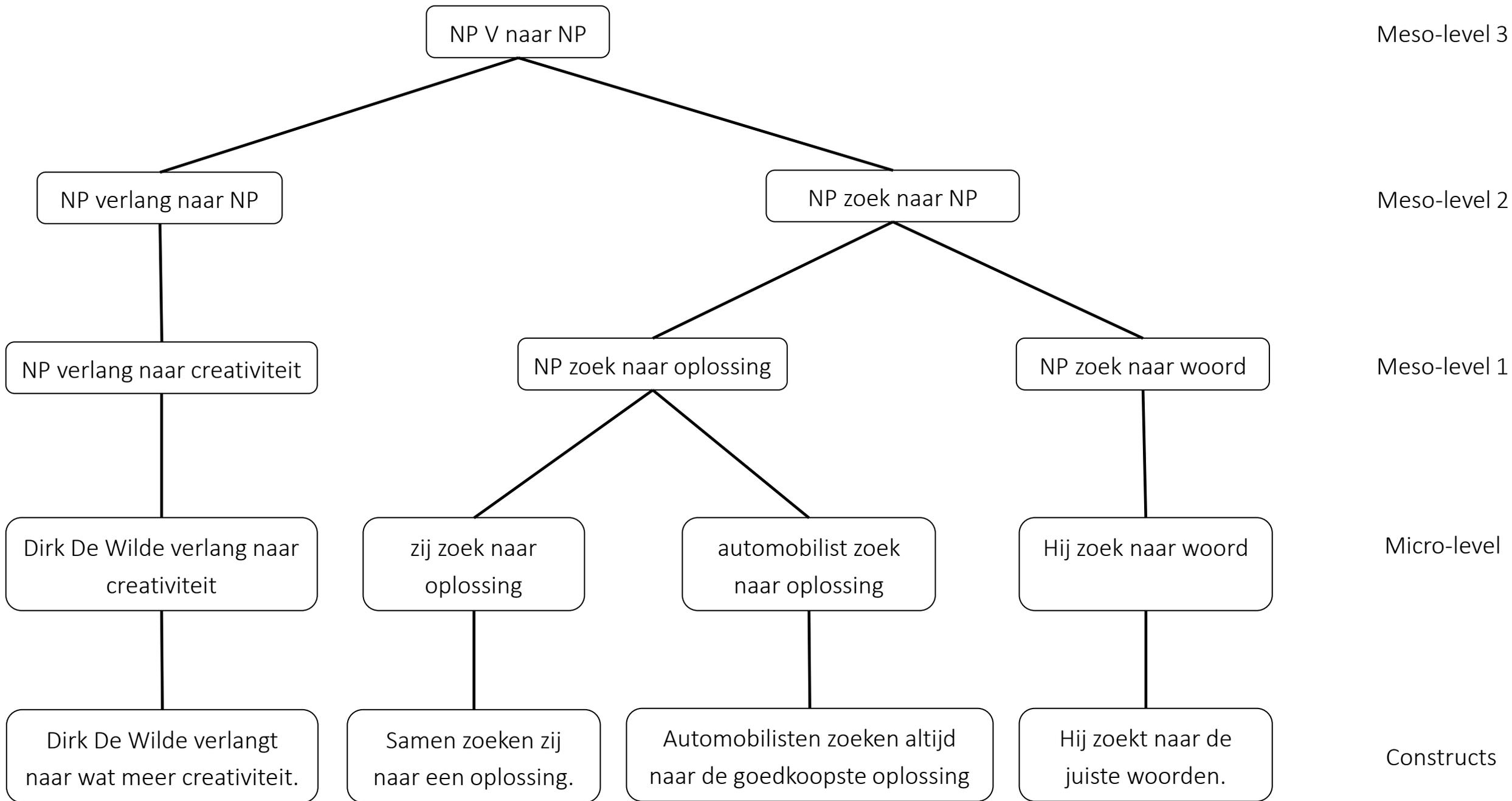
Transitive construction

Collexeme	Collostr. strenght
<i>hebben</i> 'have'	4,101,778
<i>krijgen</i> 'get'	2,602,924
<i>doen</i> 'do'	2,410,269
<i>zien</i> 'see'	1,297,011
<i>nemen</i> 'take'	1,190,952

Naar-construction

Collexeme	Collostr. strenght
<i>gaan</i> 'go'	923,830
<i>kijken</i> 'look'	532,119
<i>trekken</i> 'pull'	272,823
<i>verwijzen</i> 'refer'	258,146
<i>komen</i> 'come'	251,058

Hypothesis: The more similar an alternating verb's meaning is to the meaning of the *naar*-construction, and the less similar it is to the meaning of the transitive construction, i.e. the higher its value for VERBAL SEMANTIC COHERENCE TO THE NAAR-CONSTRUCTION, the stronger its predilection for occurring in the *naar*-construction.



Formulating hypotheses at Meso-level 2

Formulating a hypothesis at Meso-level 2

- Transitive *verlang*-construction vs. *verlang-naar*-construction

NP verlang NP

vs.

NP verlang naar NP

- Lexical Origin Hypothesis: schematic constructions obtain their meaning from their most prototypical lexical slot fillers

Transitive *verlang*-construction

Collexeme	Collostr. strenght
<i>bewijs</i> 'proof'	179
<i>tegenprestatie</i> 'countereffort'	154
<i>bijdrage</i> 'contribution'	131
<i>offer</i> 'sacrifice'	129
<i>garantie</i> 'guaranty'	127

Verlang-naar-construction

Collexeme	Collostr. strenght
<i>kind</i> 'child'	262
<i>leven</i> 'life'	255
<i>rust</i> 'rest'	234
<i>dood</i> 'death'	213
<i>huis</i> 'house'	181

- Principle of Semantic Coherence: a lexical filler, e.g. a theme, will more readily combine with a construction that is semantically coherent, i.e. that has a similar or a compatible meaning

Transitive *verlang*-construction

Collexeme	Collostr. strenght
<i>bewijs</i> 'proof'	179
<i>tegenprestatie</i> 'countereffort'	154
<i>bijdrage</i> 'contribution'	131
<i>offer</i> 'sacrifice'	129
<i>garantie</i> 'guaranty'	127

Verlang-naar-construction

Collexeme	Collostr. strenght
<i>kind</i> 'child'	262
<i>leven</i> 'life'	255
<i>rust</i> 'rest'	234
<i>dood</i> 'death'	213
<i>huis</i> 'house'	181

For each theme lemma_i:

SEMANTIC COHERENCE TO THE *VERLANG-NAAR-CXN* =

$$10 \left(\frac{\sum_{n=1}^5 \text{sim}_{cm}(\overrightarrow{\text{verlang-naar-cxn collexeme}_n}, \overrightarrow{\text{theme lemma}_i})}{5} - \frac{\sum_{n=1}^5 \text{sim}_{cm}(\overrightarrow{\text{trans-verlang-cxn collexeme}_n}, \overrightarrow{\text{theme lemma}_i})}{5} \right)$$

Transitive *verlang*-construction

Collexeme	Collostr. strenght
<i>bewijs</i> 'proof'	179
<i>tegenprestatie</i> 'countereffort'	154
<i>bijdrage</i> 'contribution'	131
<i>offer</i> 'sacrifice'	129
<i>garantie</i> 'guaranty'	127

Verlang-naar-construction

Collexeme	Collostr. strenght
<i>kind</i> 'child'	262
<i>leven</i> 'life'	255
<i>rust</i> 'rest'	234
<i>dood</i> 'death'	213
<i>huis</i> 'house'	181

Hypothesis: The higher a theme lemma's value for SEMANTIC COHERENCE TO THE *VERLANG-NAAR*-CONSTRUCTION, the stronger its predilection for the *verlang-naar*-construction.

Formulating a hypothesis at Meso-level 2

- Transitive *zoek*-construction vs. *zoek-naar*-construction

NP zoek NP

vs.

NP zoek naar NP

- Lexical Origin Hypothesis: schematic constructions obtain their meaning from their most prototypical lexical slot fillers

Transitive zoek-construction

Collexeme	Collostr. strenght
<i>oplossing</i> 'solution'	24,675
<i>toevlucht</i> 'refuge'	14,695
<i>contact</i> 'contact'	14,416
<i>heil</i> 'salvation'	13,871
<i>toenadering</i> 'rapprochement'	7,944

Zoek-naar-construction

Collexeme	Collostr. strenght
<i>oplossing</i> 'solution'	22,564
<i>manier</i> 'manner'	3,738
<i>alternatief</i> 'alternative'	3,677
<i>overlevende</i> 'survivor'	2,887
<i>oorzaak</i> 'cause'	2,372

Transitive zoek-construction

Zoek-naar-construction

Collexeme	Collostr. strenght
<i>oplossing</i> 'solution'	24,675
<i>toevlucht</i> 'refuge'	14,695
<i>contact</i> 'contact'	14,416
<i>heil</i> 'salvation'	13,871
<i>toenadering</i> 'rapprochement'	7,944

Collexeme	Collostr. strenght
<i>oplossing</i> 'solution'	22,564
<i>manier</i> 'manner'	3,738
<i>alternatief</i> 'alternative'	3,677
<i>overlevende</i> 'survivor'	2,887
<i>oorzaak</i> 'cause'	2,372

For each theme lemma:

SEMANTIC COHERENCE TO THE ZOEK-NAAR-CXN =

$$10 \left(\frac{\sum_{n=1}^5 \text{sim}_{cm}(\overrightarrow{\text{zoek-naar-cxn collexeme}_n}, \overrightarrow{\text{theme lemma}_i})}{5} - \frac{\sum_{n=1}^5 \text{sim}_{cm}(\overrightarrow{\text{trans-zoek-cxn collexeme}_n}, \overrightarrow{\text{theme lemma}_i})}{5} \right)$$

Transitive *zoek*-construction

Collexeme	Collostr. strenght
<i>oplossing</i> 'solution'	24,675
<i>toevlucht</i> 'refuge'	14,695
<i>contact</i> 'contact'	14,416
<i>heil</i> 'salvation'	13,871
<i>toenadering</i> 'rapprochement'	7,944

Zoek-naar-construction

Collexeme	Collostr. strenght
<i>oplossing</i> 'solution'	22,564
<i>manier</i> 'manner'	3,738
<i>alternatief</i> 'alternative'	3,677
<i>overlevende</i> 'survivor'	2,887
<i>oorzaak</i> 'cause'	2,372

Hypothesis: The higher a theme lemma's value for SEMANTIC COHERENCE TO THE *ZOEK-NAAR*-CONSTRUCTION, the stronger its predilection for the *zoek-naar*-construction.

Overview

Hypothesis at Meso-level 3:

- The higher an alternating verb's value for VERBAL SEMANTIC COHERENCE TO THE *NAAR*-CONSTRUCTION, the stronger its predilection for the *naar*-construction.

Hypotheses at Meso-level 2:

- The higher a theme lemma's value for SEMANTIC COHERENCE TO THE *VERLANG-NAAR*-CONSTRUCTION, the stronger its predilection for the *verlang-naar*-construction.
- The higher a theme lemma's value for SEMANTIC COHERENCE TO THE *ZOEK-NAAR*-CONSTRUCTION, the stronger its predilection for the *zoek-naar*-construction.

Testing the hypothesis at Meso-level 3

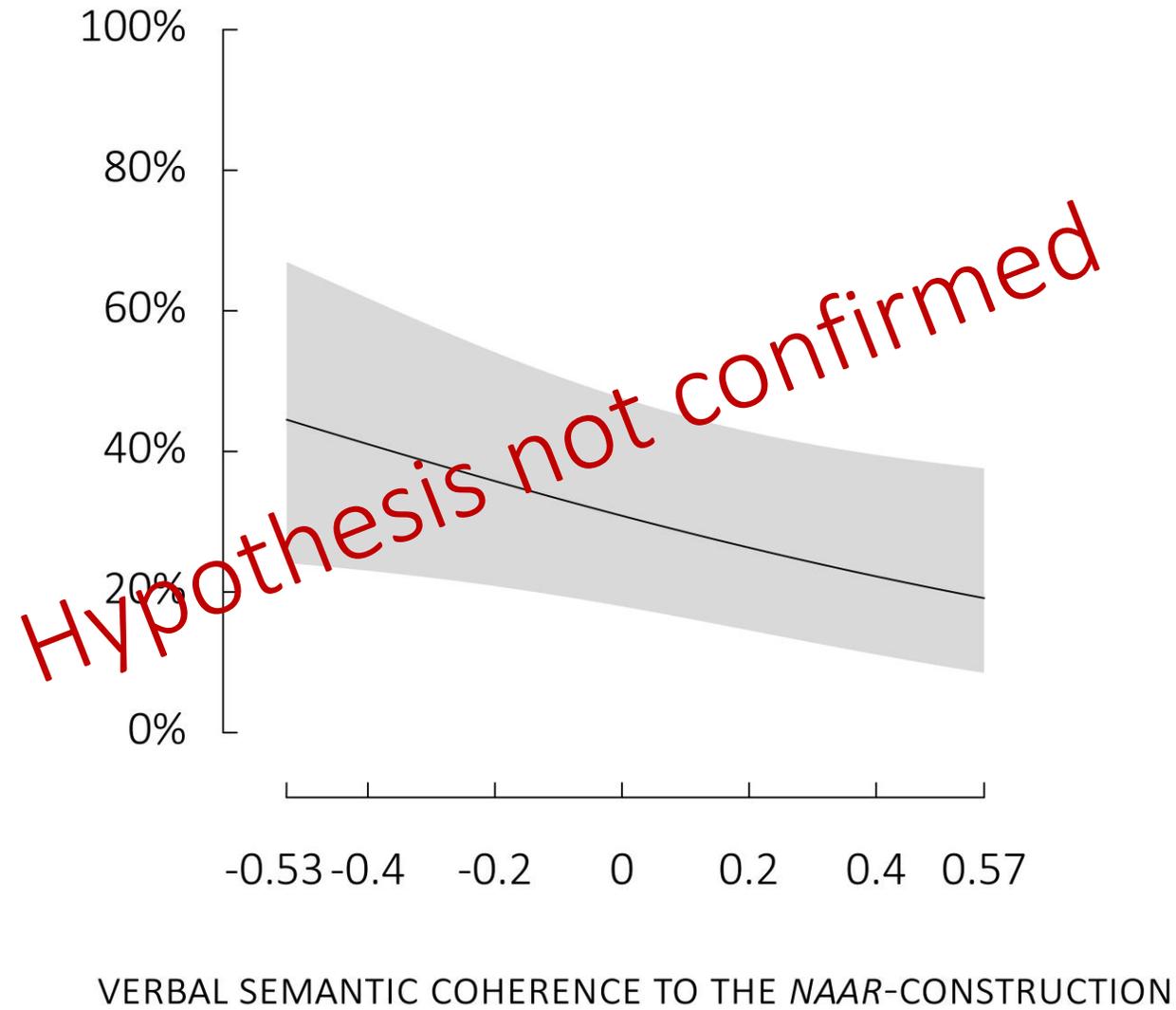
Data

- Extracted all instances of the 13 alternating verbs *bellen* 'ring', *graaien* 'grasp', *grabbelen* 'scramble', *grijpen* 'grab', *happen* 'snap', *jagen* 'hunt', *opbellen* 'ring up', *peilen* 'gauge', *schoppen* 'kick', *telefoneren* 'phone', *verlangen* 'desire', *vissen* 'fish' and *zoeken* 'search'
- From the Sonar-corpus (except the chat material, Tweets, text messages and discussion lists)
- Only those for which the country of origin is known, and the theme is expressed and not extraposed
- Subjected to manual checking
- Left with 76,138 transitive observations and 17,530 prepositional observations

Analysis

- All data in one dataset
- Mixed logistic regression model: choice of variant
 - VERBAL SEMANTIC COHERENCE TO THE *NAAR*-CONSTRUCTION
 - COUNTRY
 - THEME COMPLEXITY
 - VERB-THEME ORDER
 - Interaction between THEME COMPLEXITY and VERB-THEME ORDER
- Random effects:
 - CORPUS COMPONENT
 - VERB

Estimated
probability
of the
prepositional
variant

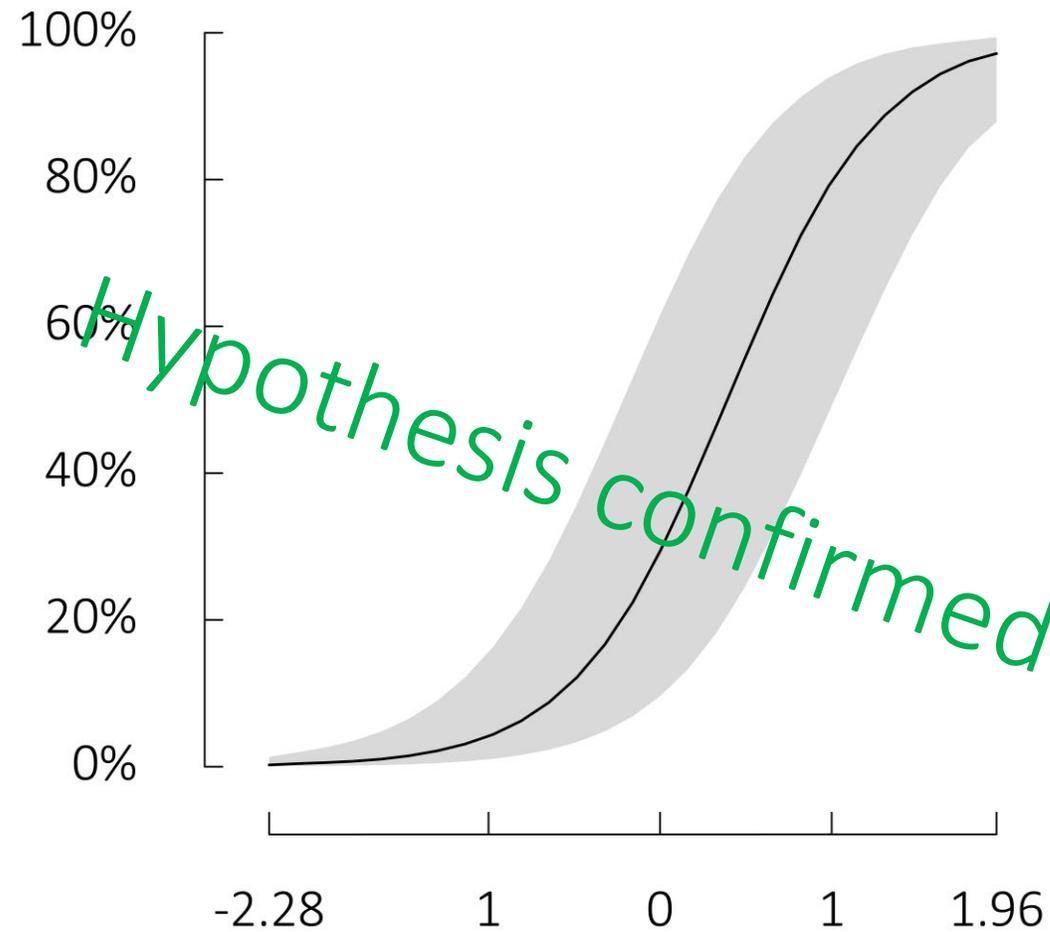


Testing the hypotheses at Meso-level 2

Verlangen 'desire'

- Limit the dataset to only the instances of *verlangen* 'desire', excluding the top 5 collexemes and their diminutives: 725 transitive observations vs. 914 prepositional observations
- Mixed logistic regression model: choice of variant
 - SEMANTIC COHERENCE TO THE *VERLANG-NAAR*-CONSTRUCTION
 - OBJECTAL COHERENCE TO THE *NAAR*-CONSTRUCTION
 - COUNTRY
 - THEME COMPLEXITY
 - VERB-THEME ORDER
 - Interaction between THEME COMPLEXITY and VERB-THEME ORDER
- Random effects:
 - CORPUS COMPONENT
 - THEME LEMMA

Estimated
probability
of the
prepositional
variant

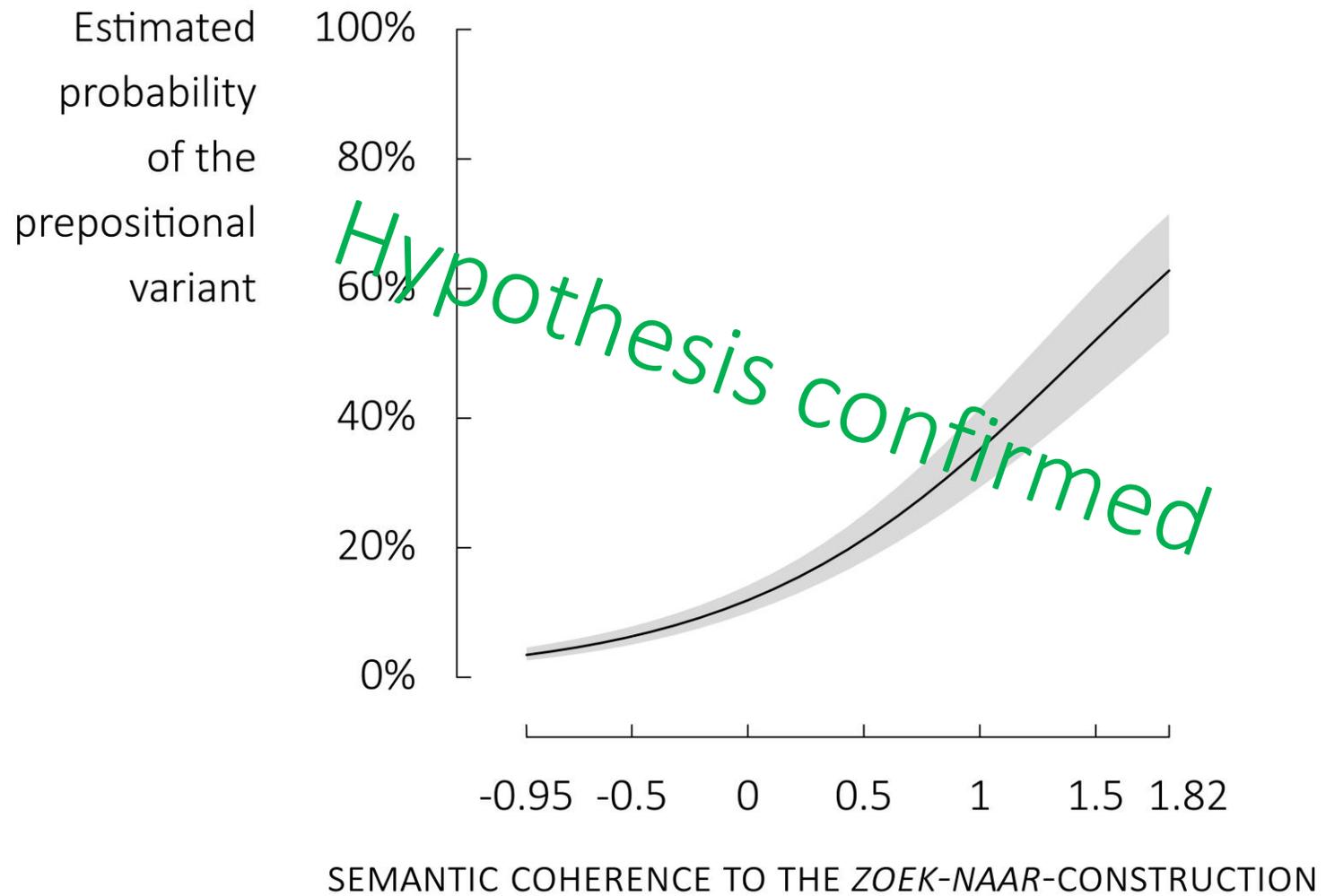


SEMANTIC COHERENCE TO THE VERLANG-NAAR-CONSTRUCTION

$p < 0.0001$

Zoeken 'search'

- Limit the dataset to only the instances of *zoeken* 'search', excluding the top 5 collexemes and their diminutives : 40,662 transitive observations vs. 8,843 prepositional observations
- Mixed logistic regression model: choice of variant
 - SEMANTIC COHERENCE TO THE *ZOEK-NAAR*-CONSTRUCTION
 - OBJECTAL COHERENCE TO THE *NAAR*-CONSTRUCTION
 - COUNTRY
 - THEME COMPLEXITY
 - VERB-THEME ORDER
 - Interaction between THEME COMPLEXITY and VERB-THEME ORDER
- Random effects:
 - CORPUS COMPONENT
 - THEME LEMMA



p < 0.0001

Overview

Hypothesis at Meso-level 3:

- The higher an alternating verb's value for VERBAL SEMANTIC COHERENCE TO THE *NAAR*-CONSTRUCTION, the stronger its predilection for the *naar*-construction. **Not confirmed**

Hypotheses at Meso-level 2:

- The higher a theme lemma's value for SEMANTIC COHERENCE TO THE *VERLANG-NAAR*-CONSTRUCTION, the stronger its predilection for the *verlang-naar*-construction. **Confirmed**
- The higher a theme lemma's value for SEMANTIC COHERENCE TO THE *ZOEK-NAAR*-CONSTRUCTION, the stronger its predilection for the *zoek-naar*-construction. **Confirmed**

Conclusions

Conclusions

- *Verlangen* 'desire' and *zoeken* 'search' seem to use the alternation to express their own, idiosyncratic semantic construals, viz. 'desire as demand' vs. 'desire as longing for' and 'search as seek to make/get' vs. 'search as literally looking for'
- Further findings:
 - *Zoeken* 'search' partially functions at an even lower level of schematicity, viz. Meso-level 1
 - *Peilen* 'gauge' also functions at Meso-level 2, expressing 'gauge as directly assess' vs 'gauge by asking'
 - Telephonic verbs, motoric verbs and venatic verbs appear to function at a higher level of schematicity

Conclusions

- Non-semantic motivating factors
 - Lectal differences: low level
 - Processing-related factors: very stable, high level

Conclusions

- At what level of schematicity should we investigate argument structure variation?

Argument structure variation may function at various levels of schematicity, depending on the case study at hand

→ Methodological procedure:

1. Define a number of possible levels of schematicity
2. Formulate hypotheses at various levels
3. Systematically put them to the test

Want to know more?

Pijpops, Dirk. 2019. *How, why and where does argument structure vary? A usage-based investigation into the Dutch transitive-prepositional alternation*. PhD-thesis University of Leuven.

Pijpops Dirk, Dirk Speelman, Stefan Grondelaers & Freek Van de Velde. Manuscript. *Incorporating the multi-level nature of the construction into hypothesis testing*.

dirk.pijpops@uliege.be

