Syntactic alternations of Dutch psych verbs
A corpus study

Dirk Pijpops & Dirk Speelman

RU Quantitative Lexicology and Variational Linguistics

SLE, Napels, 01.09.2016

Psych verbs' problem

Problem for linking theories of argument realization:

(e.g. Bresnan & Kanerva 1989, Dowty 1991, Baker 1997, Van Valin 2004,…)

If argument realization is predictable by semantics, then why do verbs that express a very similar mental state, exhibit different argument realizations?

(Croft 1993; Kitis 2009; Verhoeven 2010; Grafmiller 2013)

Psych verbs' problem

John likes Elizabeth ↔ *Elizabeth pleases John*

John fears Elizabeth ↔ *Elizabeth frightens John*

Psych verbs' problem

Experiencer-subject		Stimulus-subject
John likes Elizabeth	↔	*Elizabeth pleases John*
John fears Elizabeth	↔	*Elizabeth frightens John*

- Experiencer: experiences the mental state
- Stimulus: causes the mental state

Psych verbs' problem

Reflexive construction

Transitive construction

John ergert zich aan Elizabeth ↔ *Elizabeth ergert John*

'Elizabeth annoys John'

Psych verbs' problem

amuseren ('to amuse')
bedroeven ('to sadden')
benieuwen ('to make curious')
berouwen ('to rue')
ergeren ('to annoy')
frustreren ('to frustrate')
generen ('to embarrass')
interesseren ('to interest')
irriteren ('to irritate')
ontroeren ('to emotionally move')
opwinden ('to arouse')
plezieren ('to make happy')
spijten ('to regret')
storen ('to disturb')
verbazen ('to amaze')
verbijsteren ('to amaze')
verblijden ('to gladden')
verdrieten ('to grieve')
vergenoegen ('to content')
verheugen ('to rejoice')
vermaken ('to entertain')
verontwaardigen ('to indignify')
vervelen ('to bore')
verwonderen ('to surprise')
…

Psych verbs' problem

amuseren ('to amuse')
bedroeven ('to sadden')
benieuwen ('to make curious')
berouwen ('to rue')
ergeren ('to annoy')
frustreren ('to frustrate')
generen ('to embarrass')
interesseren ('to interest')
irriteren ('to irritate')
ontroeren ('to emotionally move')
opwinden ('to arouse')
plezieren ('to make happy')
spijten ('to regret')
storen ('to disturb')
verbazen ('to amaze')
verbijsteren ('to amaze')
verblijden ('to gladden')
verdrieten ('to grieve')
vergenoegen ('to content')
verheugen ('to rejoice')
vermaken ('to entertain')
verontwaardigen ('to indignify')
vervelen ('to bore')
verwonderen ('to surprise')
…

⇨ *ergeren* ('to annoy')
interesseren ('to interest')
storen ('to disturb')
verbazen ('to amaze')

Hypotheses

- Agentivity hypothesis

Since the agentivity of the participants in a mental state – stimulus and experiencer – can vary, variation arises as to which one is appointed subject position. The most agentive of the two participants in the mental state to be expressed, is then realized as the subject.

Hypotheses

- Agentivity hypothesis

Since the agentivity of the participants in a mental state – stimulus and experiencer – can vary, variation arises as to which one is appointed subject position. The most agentive of the two participants in the mental state to be expressed, is then realized as the subject.

- Umbrella term: Hopper & Thompson (1980), Grimshaw (1990), Dowty (1991), Langacker (1991), Croft (1993), Zaenen (1993), Pesetsky (1995), Vanhoe (2002),...

Agentivity hypotheses

- Type level

Verbs that express a mental state in which the experiencer is more agentive, will be more compatible with experiencer-subject argument constructions.

Lexical construction

interesseren
'to take interest'
↳ EXPERIENCER > stimulus

Lexical construction

interesseren
'to take interest'
↳ EXPERIENCER > stimulus

Argument construction

NP V reflex PP

↳ compatible with EXPERIENCER > stimulus

Lexical construction

interesseren
'to take interest'
↳ EXPERIENCER > stimulus

Lexical construction

verbazen
'to amaze'
↳ STIMULUS > experiencer

Argument construction

NP V reflex PP

↳ compatible with EXPERIENCER > stimulus

Lexical construction

interesseren
'to take interest'
↳ EXPERIENCER > stimulus

Argument construction

NP V reflex PP

↳ compatible with EXPERIENCER > stimulus

Lexical construction

verbazen
'to amaze'
↳ STIMULUS > experiencer

Argument construction

NP V NP

↳ compatible with 'STIMULUS > experiencer'

Agentivity hypotheses

- Type level
  - ↳ The verb determines the argument construction
  - ↳ An argument construction may be more compatible with a certain lexical meaning of a verb
  - ↳ The argument constructions do not necessarily carry meaning, independently from verbs
  - ↳ Difference in agentivity = difference in meaning between verbs

Agentivity hypotheses

- Token level

The speaker will put the participant which he or she wants to represent as more agentive, in subject position.

Lexical construction

interesseren
‘to take interest’

Lexical construction

interesseren
'to take interest'

Argument construction

'EXPERIENCER > stimulus'

Lexical construction

interesseren
'to take interest'

Argument construction

NP V reflex PP

'EXPERIENCER > stimulus'

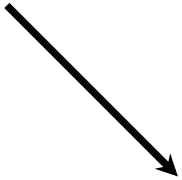
Utterance

John interesseert zich voor Elizabeth
'John takes an interest in Elizabeth.'

↳ EXPERIENCER > stimulus

Lexical construction

interesseren
'to take interest'



Argument construction

NP V reflex PP

'EXPERIENCER > stimulus'

Argument construction

NP V NP

'STIMULUS > experiencer'

Utterance

John interesseert zich voor Elizabeth
'John takes an interest in Elizabeth.'

↳ EXPERIENCER > stimulus

Lexical construction

interesseren
'to take interest'

Argument construction

NP V reflex PP

'EXPERIENCER > stimulus'

Argument construction

NP V NP

'STIMULUS > experiencer'

Utterance

John interesseert zich voor Elizabeth
'John takes an interest in Elizabeth.'

↳ EXPERIENCER > stimulus

Utterance

Elizabeth interesseert John
'Elizabeth interests John.'

↳ STIMULUS > experiencer

Agentivity hypotheses

- Token level
  - ↳ The speaker determines the argument construction
  - ↳ Argument constructions do carry meaning independently from verbs
  - ↳ Difference in agentivity = difference in meaning between the occurrences of a single verb

Agentivity hypotheses

- Type level

Verbs that express a mental state in which the experiencer is more agentive, will be more compatible with experiencer-subject argument constructions.

- Token level

The speaker will put the participant which he or she wants to represent as more agentive, in subject position.

Agentivity hypotheses

- Type level

Verbs that express a mental state in which the experiencer is more agentive, will be more compatible with experiencer-subject argument constructions.

- Token level

The speaker will put the participant which he or she wants to represent as more agentive, in subject position.

⇨ Intertwined, but logically independent

Agentivity hypotheses

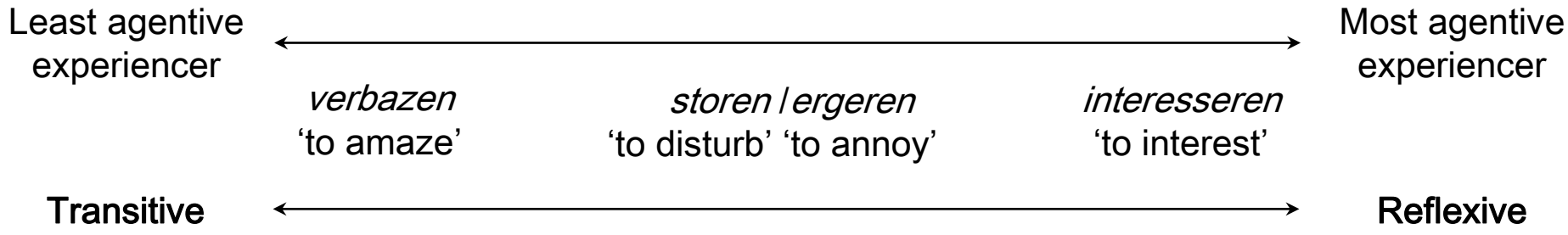
- Type level: operationalization taken over from Van de Velde (2004)

Experiencer	Volition	Control	Responsibility
ergeren (‘to annoy’)			
interesseren (‘to interest’)			
storen (‘to disturb’)			
verbazen (‘to amaze’)			

Agentivity hypotheses

- Type level: operationalization taken over from Van de Velde (2004)

Experiencer	Volition	Control	Responsibility
ergeren (‘to annoy’)			
interesseren (‘to interest’)			
storen (‘to disturb’)			
verbazen (‘to amaze’)			



Agentivity hypotheses

- Token level: operationalization taken over from Levin & Grafmiller (2012)

⇨ Code for animacy: more animate stimuli will generally be more agentive, therefore, language users will be more inclined to put animate stimuli in subject position.

Agentivity hypotheses

Category	Corpus examples
animate	*We* ('we'), *Tony Herreman, de grote Israëlische delegatie* ('the large Israeli delegation')
concrete	*het wrak* ('the wreckage'), *auto's* ('cars'), *de hoge kostprijs van het apparaat* ('the high cost of the machine')
event	*het breken met de regel* ('the breaking with the rule'), *ned.-belgië* ('the Netherlands vs. Belgium'), *het gegroet* ('the greeting')
abstract	*de landelijke politiek* ('national politics'), *de inefficiëntie* ('the inefficiency'), *die hypocrisie* ('that hypocrisy')
proposition	*dat de afgelaste wedstrijden zo laat worden ingehaald* ('that the cancelled games are rescheduled so late'), *dat ik daar moeilijkheden mee krijg* ('that that will get me into trouble')
inanimate residual category	*iets* ('something'), *dit soort dingen* ('this kind of stuff'), *wat* ('what')

Etymology hypothesis

- The argument construction of psych verbs is not determined by their psychological meaning, but rather by their (ties with a former) physical meaning.

(Klein & Kutcher 2002)

Etymology hypothesis

Storen: ‘to destroy’. Still present concerning connections & bird nests

Ergeren: ‘to damage’. No longer present.

Verbazen: ‘to make someone speak or act senselessly’. No longer present.

Interesseren: Never had a physical meaning in the history of Dutch. Latin ‘to be inbetween’

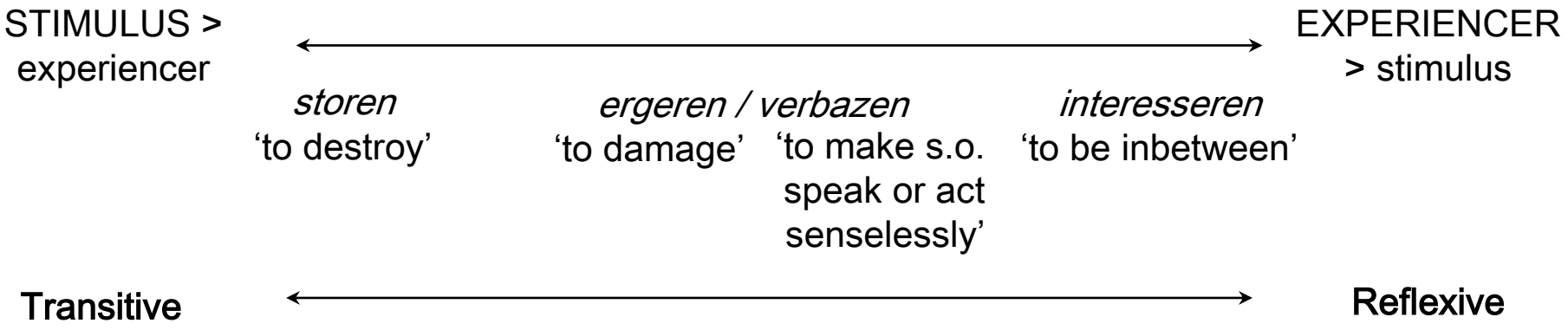
Etymology hypothesis

Storen: 'to destroy'. Still present concerning connections & bird nests

Ergeren: 'to damage'. No longer present.

Verbazen: 'to make someone speak or act senselessly'. No longer present.

Interesseren: Never had a physical meaning in the history of Dutch. Latin 'to be inbetween'



Topicality hypothesis

- Informationally heavier constituents are more likely to take up object position, lighter constituents to take up subject position.

(Bresnan et al. 2007; Colleman 2009; Levin & Grafmiller 2012;…)

	Transitive	⟷			**Reflexive**
Stimulus-Topicality	1st person	2nd person	3rd person-pronoun	Definite noun	Indefinite noun
Experiencer-Topicality	Indefinite noun	Definite noun	3rd person-pronoun	2nd person	1st person

Hypotheses

- Agentivity Hypothesis
  - Type level — *Verb*
  - Token level — *Stimulus-Animacy*
- Etymology Hypothesis — *Verb*
- Topicality Hypothesis — *Stimulus-Topicality*, *Experiencer-Topicality*

Corpora

- Representative cross-cut of spoken & written Dutch
  - Corpus of Spoken Dutch 10 million words (Oostdijk 2002)
  - Synchronic ConDiv corpus 42 million words (Grondelaers et al. 2000)

Dataset

Excluded:

- All instances of *interesseren* in which the meaning was ‘to motivate’ rather than ‘to interest’.

  Op termijn hoopt GroenLinks een aantal van deze vrouwen te interesseren voor raadswerk.
  ‘In the long run, GroenLinks hopes to motivate a number of these women for council work.’

- All instances of *storen* (‘to disturb’) with a clear physical meaning
- All participles: *gestoord* (‘crazy’)
- All instances in which the experiencer or stimulus were not expressed (cf. Levin & Grafmiller 2012: 25)
- All instances in which a free relative clause functioned as the experiencer

⇨ 2785 instances

Variables

Response variable

- Variant: *transitive, reflexive*

Explanatory variables

- Hypothesis-driven variables
  - Verb: *ergeren, interesseren, storen, verbazen*
  - Stimulus-Animacy: *animate, inanimate, concrete, event, abstract*
  - Stimulus-Topicality: *first person, second person, third person-pronoun, definite noun, indefinite noun*
  - Experiencer-Topicality: *first person, second person, third person-pronoun, definite noun, indefinite noun*
- Nuisance variables
  - Stimulus-Number: *singular, plural*
  - Experiencer-Number: *singular, plural*
  - Negation: *with, without*
  - Finiteness: *finite, infinitive*
  - Tense: *present, past, future, conditional*
  - Country: *Belgium, the Netherlands*
  - Register: *chat, informal speech, formal speech, e-mail, mass newspaper, quality newspaper*

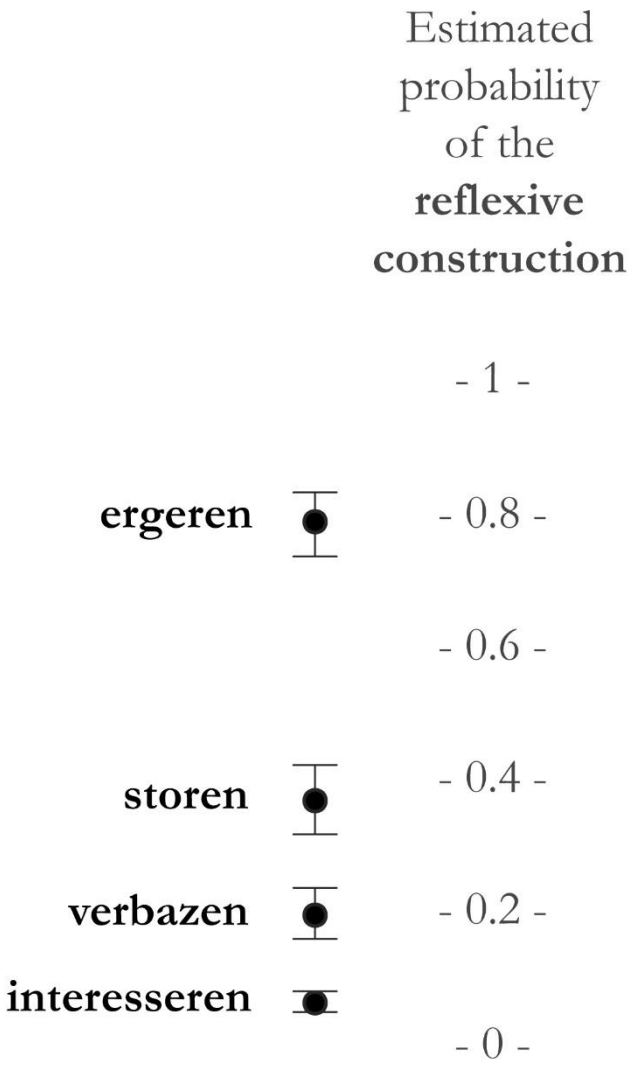
Analysis

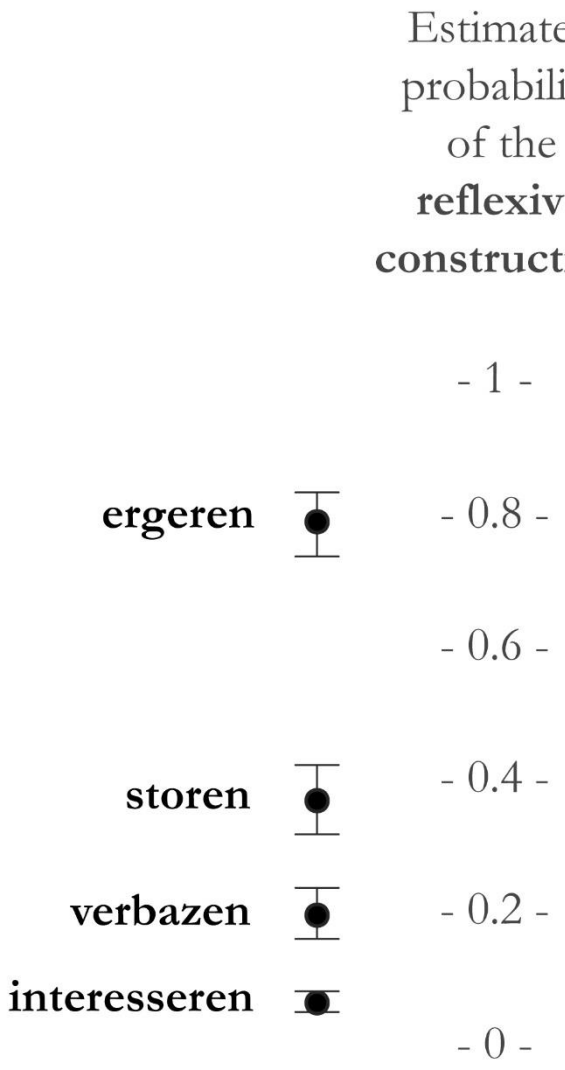
- Logistic regression
  - Dummy coding
  - Bidirectional stepwise variable selection procedure (Venables & Ripley 2002: 173-177)
  - Each predictor significantly improves the model
  - No more parameters than the number of occurrences of the least frequent response level divided by 20
  - Residual deviance not much higher than the degrees of freedom
  - Hosmer-Lemeshow-Cessie goodness of fit test did not return a significant p-value
  - Variance Inflation Factors < 4

(Baayen 2008; R Core Team 2013; Speelman 2014)

Explanatory variable	Level	Estimate	Confidence interval 2.5%	97.5%	P-value
	intercept	-3.78	-4.27	-3.30	< 0.0001
Verb	*interesseren*	Reference level			
	ergeren	4.00	3.60	4.41	< 0.0001
	storen	2.14	1.80	2.48	< 0.0001
	Verbazen	1.26	0.90	1.64	< 0.0001
Stimulus-Topicality	*third person pronoun*	Reference level			
	first person	0.42	-0.49	1.32	0.3600
	second person	-0.05	-1.44	1.10	0.9342
	definite noun	1.80	1.45	2.16	< 0.0001
	Indefinite noun	1.97	1.51	2.45	< 0.0001
	subordinate clause	0.59	0.18	1.00	0.0048
Experiencer-Topicality	*first person*	Reference level			
	second person	0.69	0.31	1.07	0.0003
	third person pronoun	1.40	1.08	1.71	< 0.0001
	definite noun	1.38	1.02	1.75	< 0.0001
	indefinite noun	1.34	0.83	1.86	< 0.0001
Country	*Belgium*	Reference level			
	the Netherlands	0.88	0.64	1.13	< 0.0001
Stimulus-Animacy	*proposition*	Reference level			
	animate	-1.37	-1.91	-0.84	< 0.0001
	inanimate	0.34	-0.11	0.79	0.1404
	concrete	-0.08	-0.54	0.37	0.7163
	event	-0.63	-1.17	-0.10	0.0198
	abstract	0.94	0.44	1.44	0.0002
Tense	*present*	Reference level			
	past	0.20	-0.11	0.51	0.1967
	future	-0.82	-1.48	-0.21	0.0112
	conditional	-1.60	-2.42	-0.89	< 0.0001
Negation	*without*	Reference level			
	with	-0.61	-0.89	-0.35	< 0.0001
Stimulus-Number	*singular*	Reference level			
	plural	0.68	0.32	1.04	0.0002
Register	*e-mail*	Reference level			
	chat	-0.22	-0.64	0.18	0.2810
	informal speech	0.02	-0.42	0.45	0.9335
	formal speech	-0.29	-0.67	0.09	0.1331
	mass newspaper	-0.69	-1.03	-0.36	< 0.0001
	quality newspaper	-0.35	-0.78	0.07	0.1057

AIC	2071
C-index	0.904
Total number of occ.	2785
Transitive occ.	1911
Reflexive occ.	874


Estimated probability of the **reflexive construction**
- 1 -
- 0.8 -
- 0.6 -
- 0.4 -
- 0.2 -
- 0 -
ergeren
storen
verbazen
interesseren



Type level agentivity hypothesis

interesseren
'to interest'

ergeren / storen
'to annoy' 'to disturb'

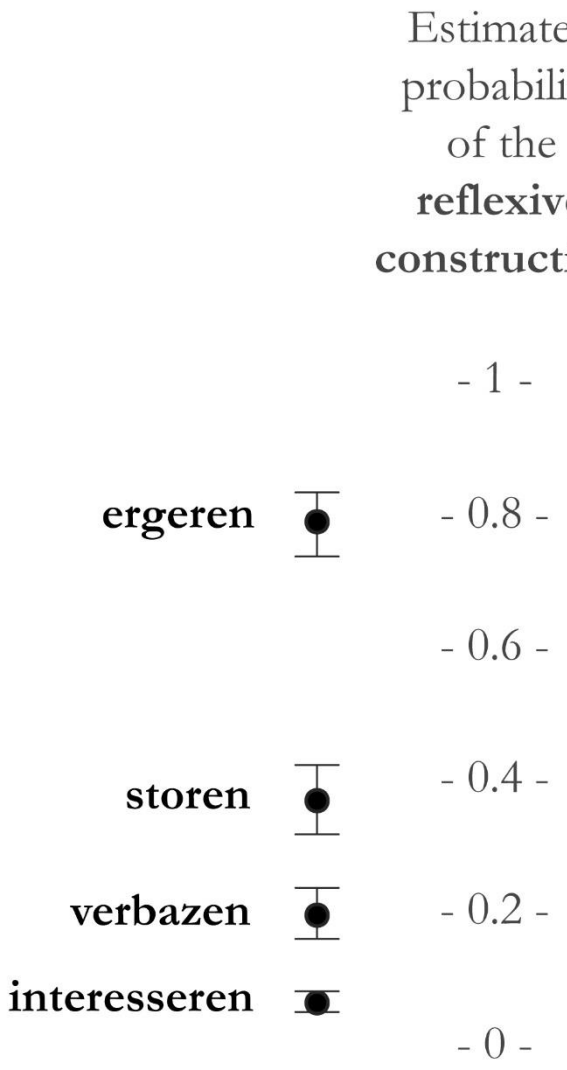
verbazen
'to amaze'

Etymology hypothesis

interesseren
'to interest'

ergeren / verbazen
'to annoy' 'to amaze'

storen
'to disturb'



Type level agentivity hypothesis

interesseren
'to interest'

ergeren / storen
'to annoy' 'to disturb'

verbazen
'to amaze'

Etymology hypothesis

interesseren
'to interest'

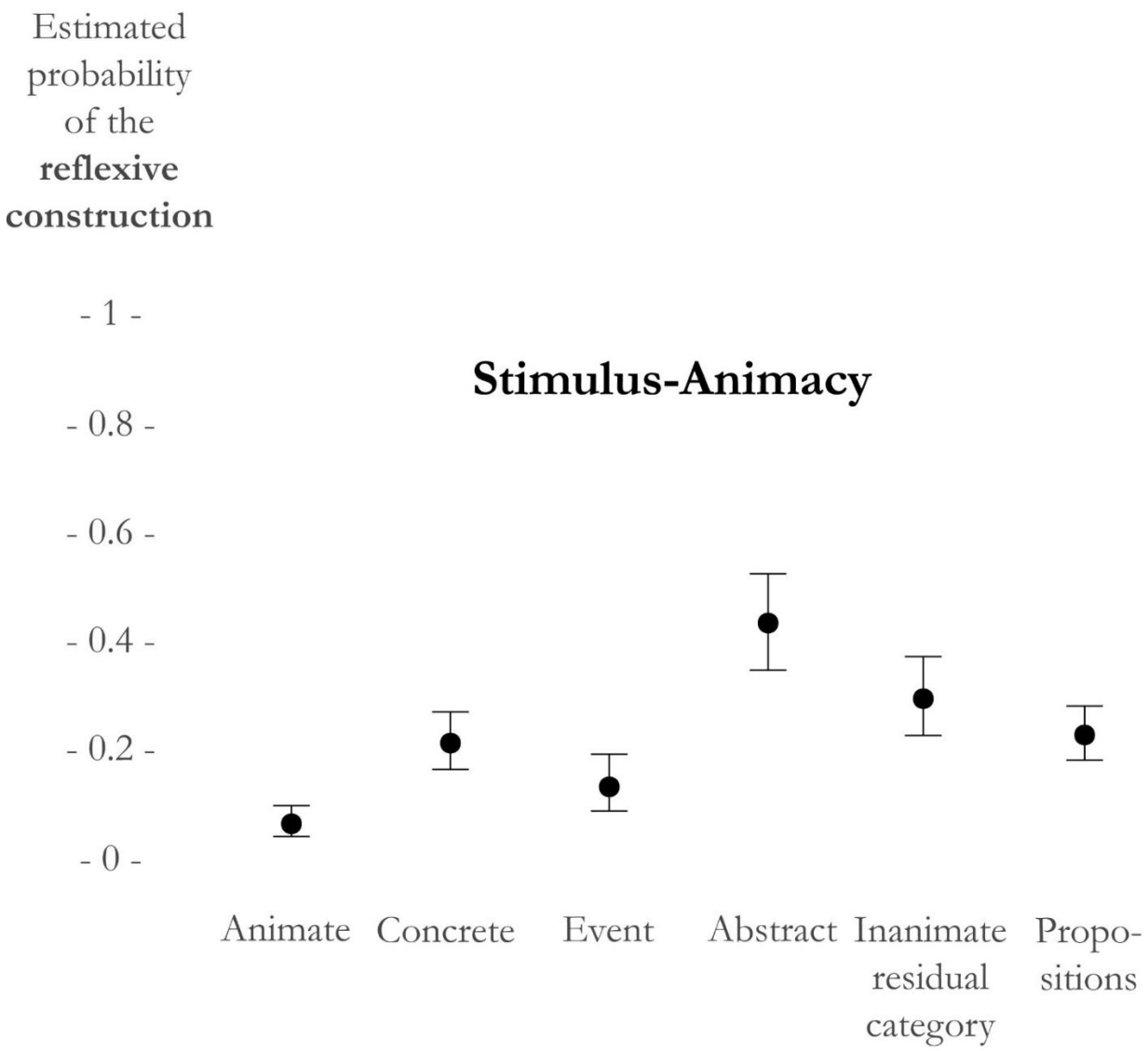
ergeren / verbazen
'to annoy' 'to amaze'

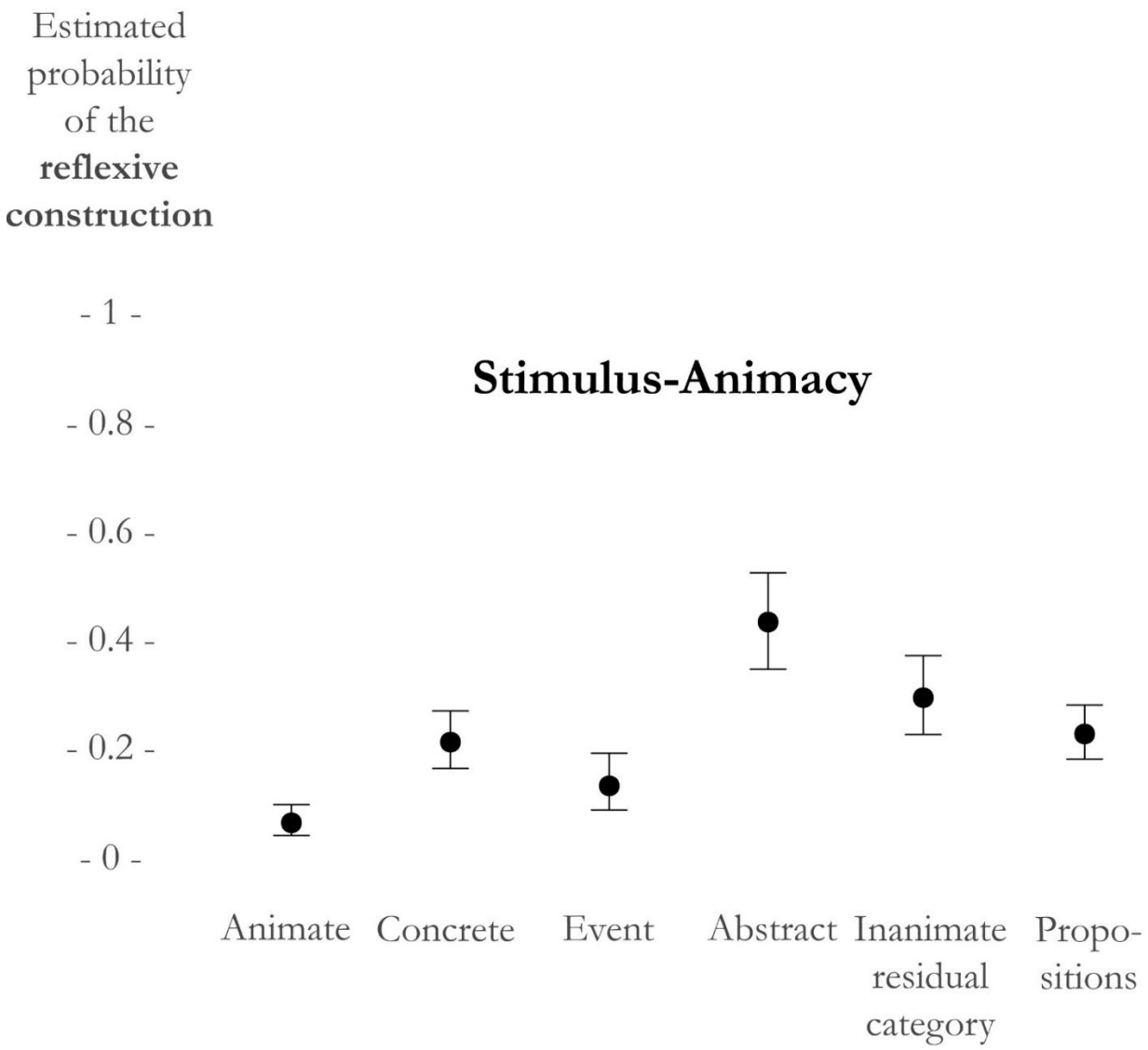
storen
'to disturb'

Not confirmed

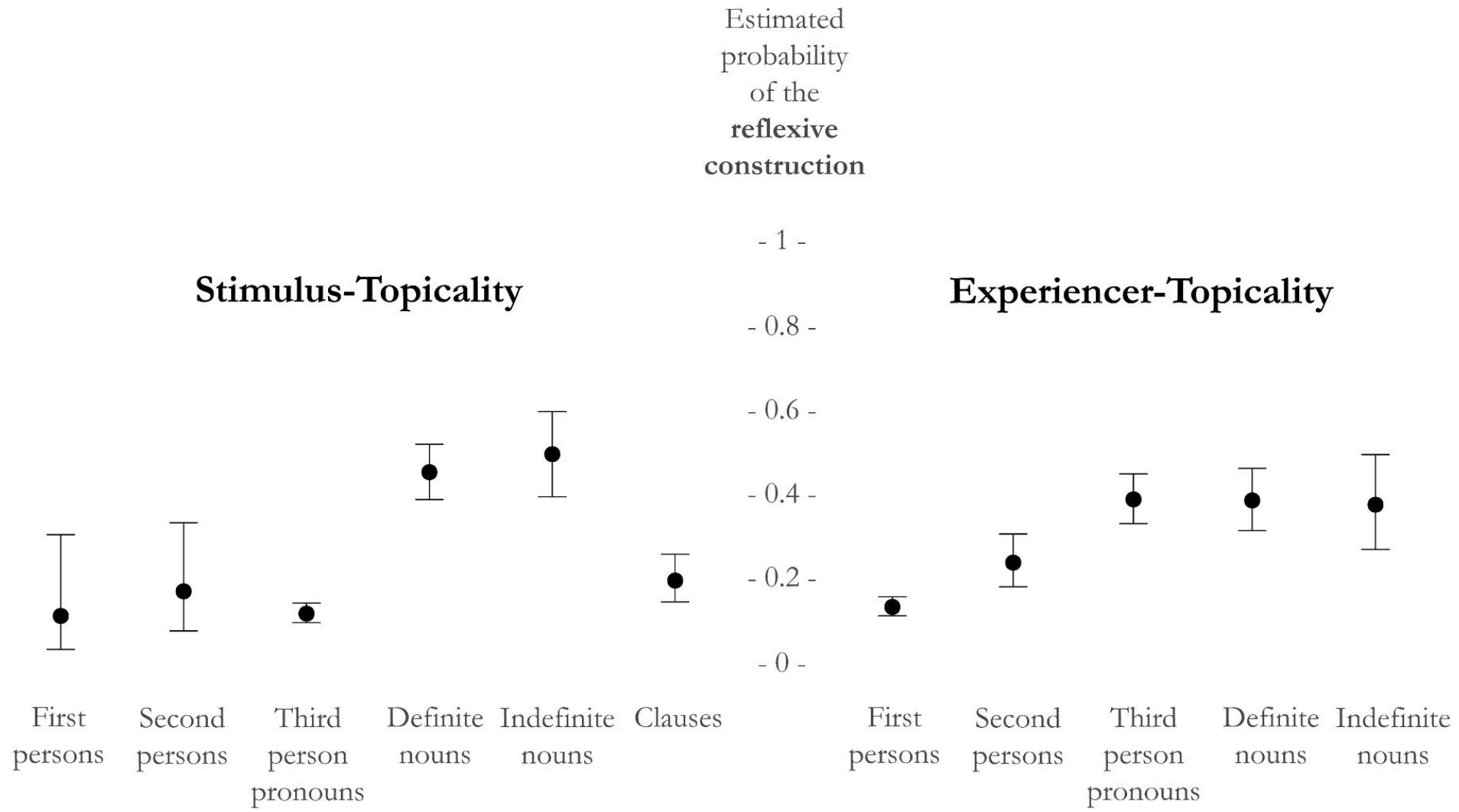
		Type level agentivity hypothesis	Etymology hypothesis
ergeren			
storen		*interesseren* ‘to interest’	*interesseren* ‘to interest’
verbazen		*ergeren / storen* ‘to annoy’ ‘to disturb’	*ergeren / verbazen* ‘to annoy’ ‘to amaze’
interesseren		*verbazen* ‘to amaze’	*storen* ‘to disturb’

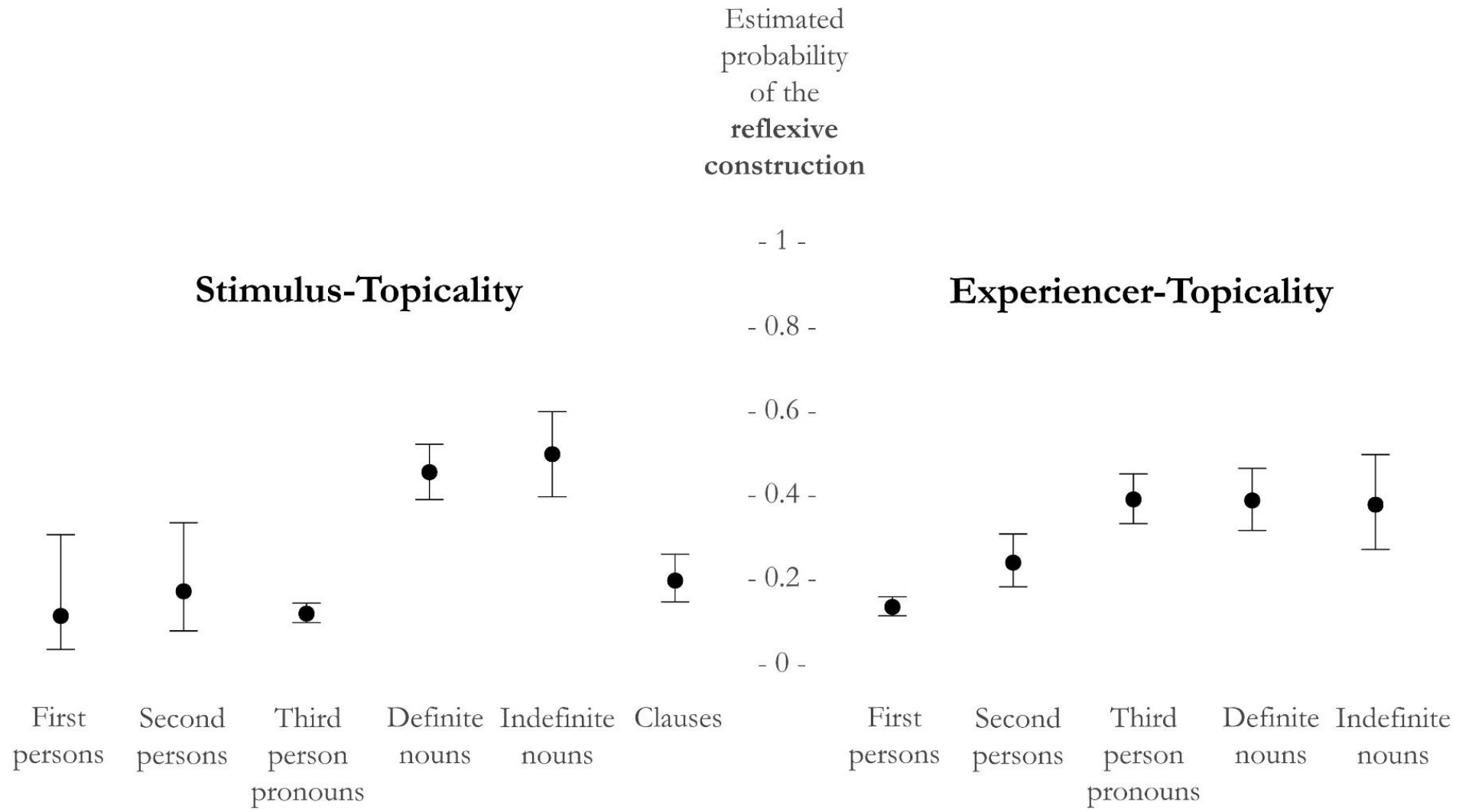
Not confirmed


Estimated probability of the **reflexive construction**
- 1 -
- 0.8 -
- 0.6 -
- 0.4 -
- 0.2 -
- 0 -
Stimulus-Animacy
Animate
Concrete
Event
Abstract
Inanimate residual category
Propo-sitions

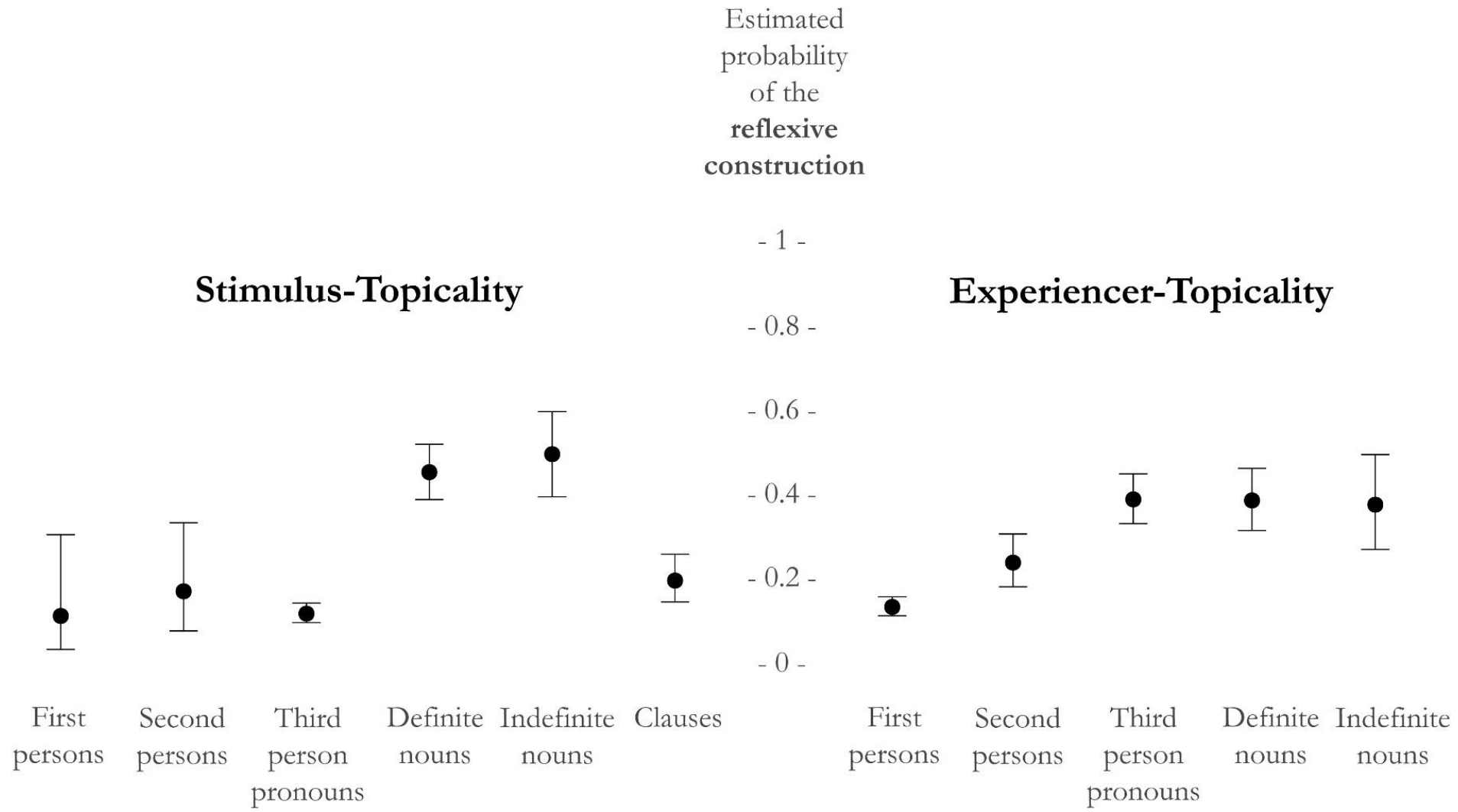

Estimated probability of the **reflexive construction**
- 1 -
- 0.8 -
- 0.6 -
- 0.4 -
- 0.2 -
- 0 -
Stimulus-Animacy
Animate
Concrete
Event
Abstract
Inanimate residual category
Propo-sitions

Confirmed


Estimated probability of the **reflexive construction**
- 1 -
- 0.8 -
- 0.6 -
- 0.4 -
- 0.2 -
- 0 -
Stimulus-Topicality
Experiencer-Topicality
First persons
Second persons
Third person pronouns
Definite nouns
Indefinite nouns
Clauses
First persons
Second persons
Third person pronouns
Definite nouns
Indefinite nouns


Estimated probability of the **reflexive construction**
- 1 -
- 0.8 -
- 0.6 -
- 0.4 -
- 0.2 -
- 0 -
Stimulus-Topicality
Experiencer-Topicality
First persons
Second persons
Third person pronouns
Definite nouns
Indefinite nouns
Clauses
First persons
Second persons
Third person pronouns
Definite nouns
Indefinite nouns

Confirmed



Confirmed

Not confirmed

Results

- Agentivity Hypotheses
  - Type level — not confirmed
  - Token level — confirmed
- Etymology Hypothesis — not confirmed
- Topicality Hypothesis — revised version confirmed

Conclusions

- Failure to confirm the type level agentivity hypothesis: findings cannot be used in support of linking theories
- Confirmed the token level agentivity hypothesis: findings do indicate that argument constructions contribute to the compositional meaning of the utterance, separately from the verb
- Important to distinguish between the type and token level for theoretical and methodological reasons
- Findings dovetail with Levin & Grafmiller (2012) and Grafmiller (2013: 167-202)

References

Baayen, Rolf Harald. 2008. *Analyzing linguistic data: a practical introduction to statistics using R*. Cambridge University Press.

Bresnan, Joan. 2007. Is syntactic knowledge probabilistic? Experiments with the English dative alternation. In Sam Featherston & Wolfgang Sternefeld (eds.), *Roots: Linguistics in Search of Its Evidential Base*, 77–96. Berlin: Mouton de Gruyter.

Colleman, Timothy. 2009. Verb disposition in argument structure alternations: a corpus study of the dative alternation in Dutch. *Language Sciences* 31(5). 593–611.

Croft, William. 1993. Case marking and the semantics of mental verbs. In James Pustejovski (ed.), *Semantics and the Lexicon*, 55–72. Dordrecht: Kluwer.

Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67(3). 547–619. Oostdijk, Nelleke, Wim Goedertier, Frank Van Eynde, Louis Boves, Jean-Pierre Martens, Michael Moortgat and Harald Baayen. 2002. Experiences from the Spoken Dutch corpus project. 2002.

Grafmiller, Jason. 2013. The Semantics of Syntactic Choice: An Analysis of English Emotion Verbs. Stanford University.

Grimshaw, Jane. 1990. *Argument structure*. Cambridge: MIT press.

Grondelaers, Stefan, Katrien Deygers, Hilde Van Aken, Vicky Van den Heede and Dirk Speelman. 2000. Het CONDIV-corpus geschreven Nederlands. *Nederlandse Taalkunde* 5(4). 356–363.

Hopper, Paul and Sandra Annear Thompson. 1980. Transitivity in Grammar and Discourse. *Language* 56(2). 251–299.

Klein, Katarina and Silvia Kutscher. 2005. *Lexical Economy and Case Selection of Psych-Verbs in German. Working papers of the SFB 282, Theory of the Lexicon [Arbeitspapiere des SFB 282 Theories des Lexicons] Nr. 122*.

Lakoff, George. 1977. Linguistic Gestalts. *Papers Regional Meeting Chicago Linguistics Society*, 236–287.

Langacker, Ronald Wayne. 1995. *Foundations of cognitive grammar*. Stanford: Stanford University Press.

Levin, Beth and Jason Grafmiller. 2012. Do you always fear what frightens you? In Tracy Holloway King & Valeria de Paiva (eds.), *From Quirky Case to Representing Space: Papers in Honor of Annie Zaenen*, 21–32. Stanford: CSLI Publications.

Oostdijk, Nelleke, Wim Goedertier, Frank Van Eynde, Louis Boves, Jean-Pierre Martens, Michael Moortgat and Harald Baayen. 2002. Experiences from the Spoken Dutch corpus project. 2002.

Pesetsky, David. 1995. *Zero Syntax*. Cambridge: MIT Press.

Plevoets, Koen. 2008. Tussen spreek- en standaardtaal. Een corpusgebaseerd onderzoek naar de situationele, regionale en sociale verspreiding van enkele morfosyntactische verschijnselen uit het gesproken Belgisch-Nederlands. Dissertation University of Leuven.

R Core Team. 2013. *R: A language and environment for statistical computing. R Foundation for Statistical Computing*. Vienna. http://www.r-project.org/.

Speelman, Dirk. 2014. Logistic regression: A confirmatory technique for comparisons in corpus linguistics. In Dylan Glynn & Justyna A. Robinson (eds.), *Corpus Methods for Semantics: Quantitative studies in polysemy and synonymy*, 487–533. (Human Cognitive Processing [HCP]). Amsterdam: John Benjamins.

Speelman, Dirk and Dirk Geeraerts. 2009. Causes for causatives: the case of Dutch “doen” and “laten.” In Ted Sanders & Eve Sweetser (eds.), *Causal Categories in Discourse and Cognition*, 173–204. Berlin: Mouton de Gruyter.

Vanhoe, Henk. 2002. Aspects of the syntax of psychological verbs in Spanish: a lexical functional analysis. In Miriam Butt & Tracy H. King (eds.), *Proceedings of the LFG02 Conference*, 373–389. Stanford: CSLI Publications.

Velde, Freek Van de. 2004. De Middelnederlandse onpersoonlijke constructie en haar grammaticale concurrenten: semantische motivering van de argumentstructuur. *Nederlandse Taalkunde* 9(1). 48–76.

Velde, Freek Van de. 2014. Degeneracy: the maintenance of constructional networks. In Ronny Boogaart, Timothy Colleman & Gijsbert Rutten (eds.), *Extending the Scope of Construction Grammar*, vol. 1, 141–179. Berlin: Mouton de Gruyter.

Verhoeven, Elisabeth. 2010. Agentivity and stativity in experiencer verbs: Implications for a typology of verb classes. *Linguistic Typology* 14(2-3).

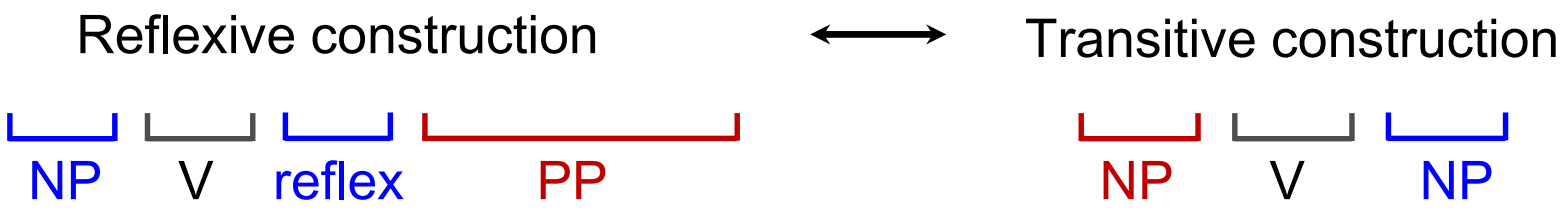
Thanks!

for further information:

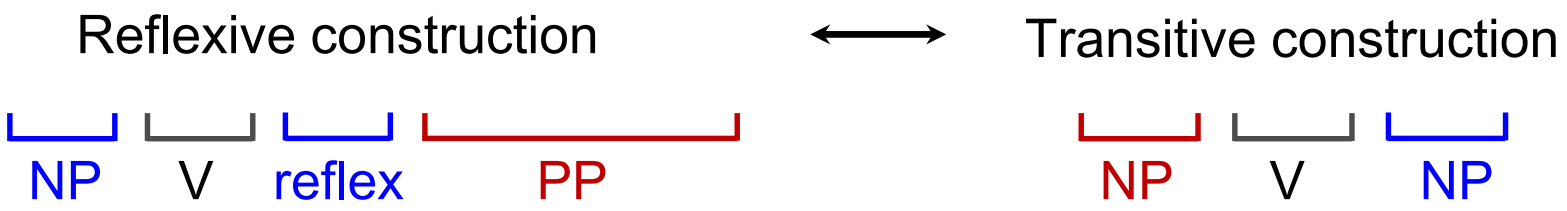
dirk.pijpops@kuleuven.be
dirk.speelman@kuleuven.be

Dirk Pijpops & Dirk Speelman. ‘Alternating argument constructions of Dutch psychological verbs. A theory-driven corpus investigation’. Accepted pending minor revisions by Folia Linguistica

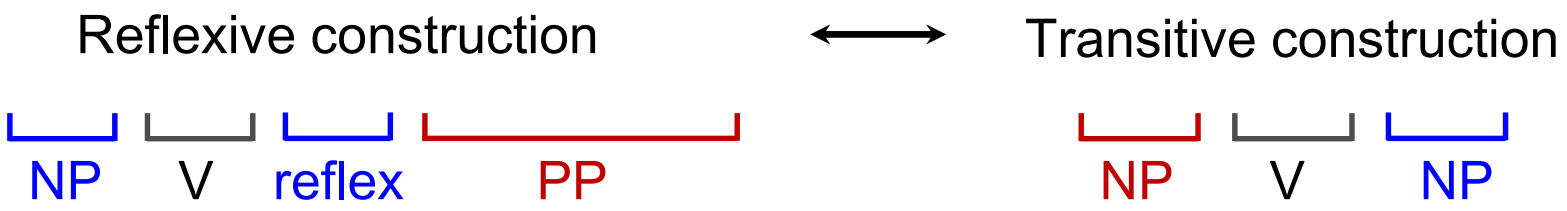
- Heavy stimulus



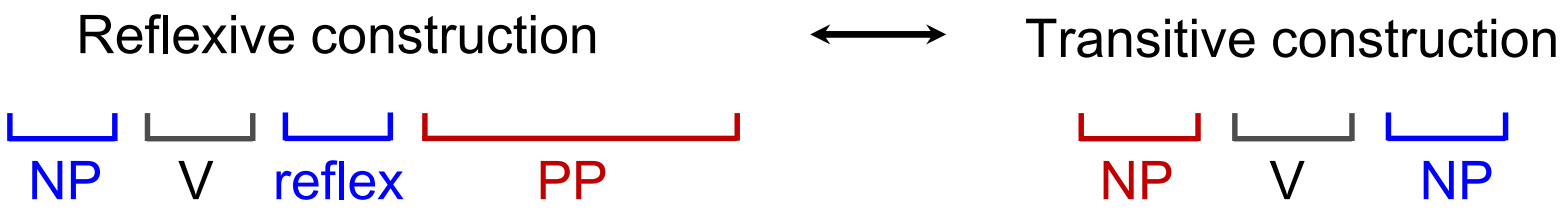
- Heavy stimulus
  - Reflexive: prepositional object
  - Transitive: subject



- Heavy stimulus
  - **Reflexive: prepositional object**
  - Transitive: subject

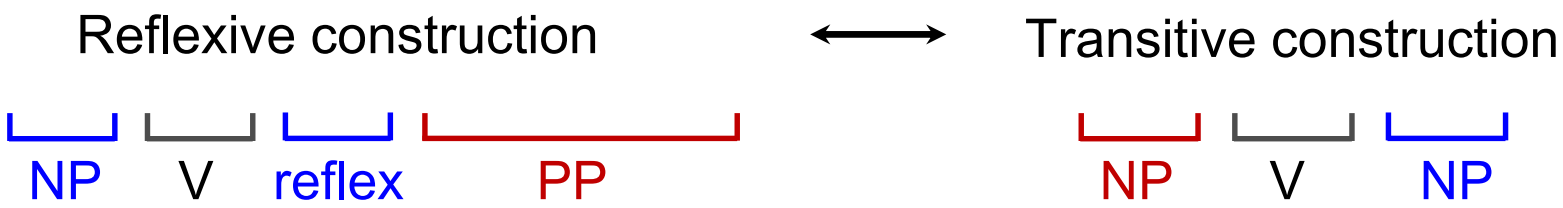


- Heavy stimulus
  - **Reflexive: prepositional object**
  - Transitive: subject
- Heavy experiencer
  - Reflexive: subject
  - Transitive: direct object



- Heavy stimulus
  - **Reflexive: prepositional object**
  - Transitive: subject
- Heavy experiencer
  - Reflexive: subject
  - Transitive: direct object

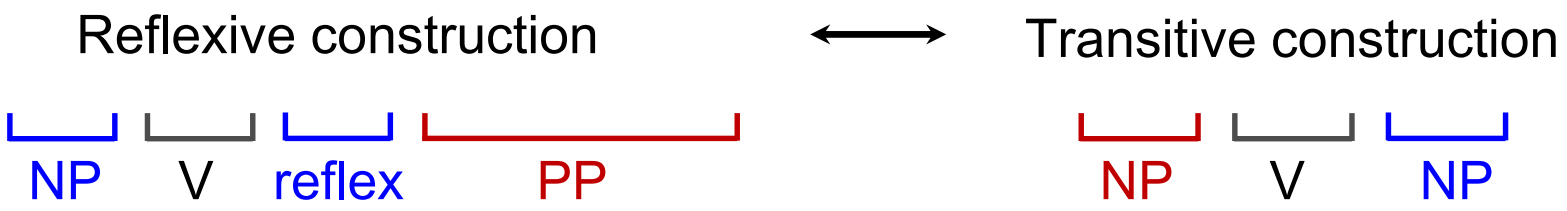
⇨ Reflexive: reflexive pronoun and preposition can enlarge the distance between stimulus and experiencer



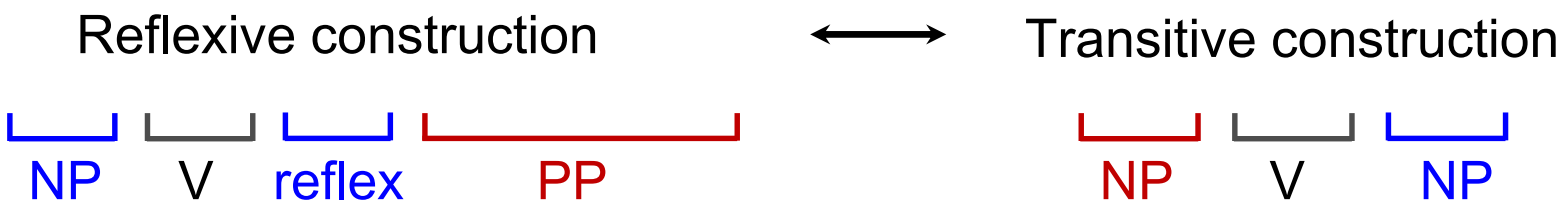
- Heavy stimulus
  - **Reflexive: prepositional object**
  - Transitive: subject
- Heavy experiencer
  - Reflexive: subject
  - Transitive: direct object

⇨ Reflexive: reflexive pronoun and preposition can enlarge the distance between stimulus and experiencer

⇨ Transitive: stimulus and experiencer need to placed close to one another



- Heavy stimulus
  - **Reflexive: prepositional object**
  - Transitive: subject
- Heavy experiencer
  - **Reflexive: subject**
  - Transitive: direct object

⇨ Reflexive: reflexive pronoun and preposition can enlarge the distance between stimulus and experiencer

⇨ Transitive: stimulus and experiencer need to placed close to one another