Argument alternations of the Dutch psych verbs

A corpus investigation

Dirk Pijpops
Research Foundation – Flanders
Quantitative Lexicology and Variational Linguistics
University of Leuven
Leuven, Belgium
dirk.pijpops@kuleuven.be

Dirk Speelman
Quantitative Lexicology & Variational Linguistics
University of Leuven
Leuven, Belgium
dirk.speelman@kuleuven.be

***Abstract*—This paper presents a corpus study of the alternation between the reflexive and transitive argument constructions of the Dutch psych verbs *ergeren* ('to annoy'), *interesseren* ('to interest'), *storen* ('to disturb') and *verbazen* ('to amaze'), as in *Jij ergert je aan mij* vs. *Ik erger jou* (both 'I annoy you'). Logistic regression analysis revealed that the choice of the language user was driven by – in order of decreasing importance – the choice of verb, the morphological form of the stimulus, the animacy of the stimulus, the morphological form of the experiencer, and a number of nuisance variables. However, verbs whose lexical meaning entailed a more agentive experiencer did not more often realize this experiencer in subject position than other verbs, nor could the preference of the verbs be predicted by looking at their etymology.**

Keywords—psych verbs; Dutch; argument realization; logistic regression; agentivity

I. Introduction

In Dutch, a number of psychological verbs may realize their experiencer in both subject (1) and object (2) position, in respectively a reflexive and transitive argument construction.[1] As opposed to similar alternations in English, such as *fear-frighten* and *like-please* [1], the same verb is used in both constructions. In other words, the only difference between both variants lies in the argument constructions. As such, any observed difference in meaning or usage can be directly attributed to these argument constructions, which makes this alternation a particularly interesting case study on the mechanisms behind argument realization.

(1) Reflexive construction:

Daar erger ik me groen en geel aan. (CGN)
There annoy I myself green and yellow to
'That greatly annoys me.'

(2) Transitive construction:

Dit […] ergerde de Romeinen mateloos. (ConDiv)
This […] annoyed the Romans excessively
'This […] excessively annoyed the Romans.'

We seek to answer the following research question, using a corpus-based alternation study, as in e.g. [2]–[5].

i. What makes the language user opt for the reflexive respectively transitive argument construction of the Dutch psychological verbs?

The alternation will be investigated for four verbs, namely *ergeren* ('to annoy'), *interesseren* ('to interest'), *storen* ('to disturb') and *verbazen* ('to amaze').

II. Hypotheses

A. Agentivity hypothesis

Different theoretical frameworks make rather similar predictions regarding the argument realization of the psychological verbs. One such prediction is here called the agentivity hypothesis, which is, be it in varying forms, put forward in o.a. [6]–[13]. It may be summarized as follows.

> For mental states or events, it is not always clear which of the participants, i.e. the stimulus or the experiencer, is more agentive. This causes variation in argument realization. The participant that is most agentive is assigned subject position.

This hypothesis can be seen as operating on two levels, which are not often strictly discriminated. These are the type level, i.e. the level of the verb, and the token level, i.e. the level of the utterance. At the type level, the agentivity hypothesis states that mental events or states which entail a more agentive experiencer, will be more likely to realize this experiencer in subject position. In other words, verbs whose lexical meaning

The research in this paper has been conducted thanks to funding of the Research Foundation Flanders. The paper is a summary an article that is currently under review.

[1] In this paper, we will use the terms *experiencer* and *stimulus* as practical designators for respectively the participant experiencing the mental state and the participant that is the cause or object of the mental state. This is done to remain in keeping with previous research [8], [10], [13]. However, we do not mean to ascribe these terms special theoretical status as thematic roles or similar concepts.

attributes a more agentive role to the experiencer, will be more compatible with experiencer-subject constructions.

The operationalization of the agentivity hypothesis at the type level is taken over from [14, pp. 53–55] and embodied by the variable *Verb*. Reference [14] uses introspection to attribute three of the four agentivity features of [15], namely *volition*, *control* and *responsibility*, to the experiencers of the verbs under scrutiny.[2] In Table 1, the same is done to the experiencers of the four verbs of the present study.

Doing this leads us to consider *interesseren* ('to interest') to entail the most agentive experiencer, while the experiencers of *ergeren* ('to annoy') and *storen* ('to disturb') are tied for agentivy features. Finally, the experiencer of *verbazen* ('to amaze') is considered least agentive. Preference for the transitive construction is therefore expected to rise from *interesseren* to either *ergeren* or *storen* and finally to *verbazen.*

As the attribution of the agentivity features relies on introspection, it can be easily criticized. However, even if one disagrees with the exact attribution in Table 1, one might still agree that generally, taking interest in something requires more active commitment than being amazed by something. Important for the operationalization is only the ranking of the verbs according to the agentivity of their experiencer, not the exact attribution of agentivity features itself.

Meanwhile, at the token level, the agentivity hypothesis predicts that given a particular utterance, the language user will put the currently most agentive participant in subject position. For instance, suppose that the language user is annoyed, and he wants to stress that this is because the stimulus is actively trying to annoy him. He would then express this active annoyance by putting the stimulus in subject position. As opposed to the type level agentivity hypothesis, the meaning facet of agentivity is thus not part of the meaning of the verb, but is rather added separately to the utterance by the argument construction.

The operationalization of the agentivity hypothesis at the token level is taken over from [1], who measure the agentivity of the stimulus through its animacy and concreteness. Reference [1] assumes that animate stimuli are usually more agentive than inanimate stimuli, and concrete objects are more agentive than abstract entities. The operationalization, embodied by the variable *Stimulus-Animacy*, thus predicts that utterances with animate stimuli will prefer the transitive construction, while inanimate concrete objects, and even more so, abstract entities, will prefer the reflexive construction.[3]

TABLE I. ATTRIBUTION OF AGENTIVITY FEATURES

Experiencer	Volition	Control	Responsibility
ergeren ('to annoy')	+	-	+/-
interesseren ('to interest')	+	+	+
storen ('to disturb')	+	-	+/-
verbazen ('to amaze')	-	-	-

B. Etymology hypothesis

The etymology hypothesis is inspired on the work of [13], who posit that it's not the psychological meaning of the psychological verbs which determines their argument construction, but rather their (ties with a former) physical meaning. Etymological inquiry revealed that each of our four verbs, except *interesseren* ('interesseren'), once held a physical meaning. These physical meanings, 'to damage' for *ergeren*, 'to destroy' for *storen*, and 'to make someone act senselessly' for *verbazen*, all entailed more agentive stimuli [16]–[22]. *Storen* ('to disturb') has the strongest bond with its physical meaning, as it is still present in contemporary Dutch, e.g. concerning connections and bird nests [22]. As such, we expect that *storen* most strongly favors the transitive construction, followed by either *ergeren* ('to annoy') or *verbazen* ('to amaze'), while *interesseren* ('to interest') most often appears in the reflexive construction.[4]

Note that even someone who does not accept the claims of [13], may still find the etymology hypothesis to be intuitively appealing. It only assumes that, as the meaning of a verb changes from physical to psychological, its argument construction does not (instantly) change with it.

C. Topicality hypothesis

The topicality hypothesis presents the influence of information structure. It is operationalized through the variables *Stimulus-* and *Experiencer-Topicality*. These variables present a scale ranging from the first and second persons, to the third person pronouns, the definite nouns and the indefinite nouns. It is expected that preference for object position rises as we go to the end of this scale. As this hypothesis is confirmed in nearly all corpus studies on argument alternations (o.a. [1]–[3]), we do not expect any surprises here. However, the hypothesis might present itself as an alternative to the agentivity hypothesis. It may be interesting to see whether the alternation is more strongly determined by a semantic variable such as *Stimulus-Animacy* or a syntactic one like *Stimulus-Topicality.*

III. DATA

The data were extracted from the Corpus of Spoken Dutch (CGN) [23] and the ConDiv corpus [24]. These corpora were chosen because together, they represent a representative cross-

[2] The fourth feature, source, was not used, as it status was considered problematic [14, p. 54].

[3] In principle the same could be done for the experiencer, but the experiencers in our dataset were almost exclusively animate, rendering such a variable useless. Except for the categories *animate, concrete, event* and *abstract*, used in [1], we also employed a category *inanimate.* Stimuli were assigned to this category if they were clearly inanimate, but also too vague to fall into one of the other inanimate categories: e.g. *iets* ('something') or *dit soort dingen* ('this kind of stuff').

[4] This order differs from the one predicted by the type level agentivity hypothesis.

cut of spoken and written Dutch from around the turn of the millenium.[5]

The alternation under scrutiny is not an idiosyncrasy of a few Dutch verbs, but a notable property of many of them.[6] From this set, the verbs *ergeren* ('to annoy'), *interesseren* ('to interest'), *storen* ('to disturb') and *verbazen* ('to amaze') were selected, because they yielded ample occurrences of both variants in both corpora.

The corpora were searched for all instances of these four verbs. Next, all these instances were manually checked and the following occurrences of the verbs had to be excluded from the dataset. First, all instances of *interesseren* were removed in which the meaning was 'to motivate' rather than 'to interest', as in (3).

(3) *Op termijn hoopt GroenLinks een aantal van deze*
on term hopes GroenLinks a number of these
vrouwen te interesseren voor raadswerk. (ConDiv)
women to interest for council_work
'In the long run, Groenlinks hopes to motivate a number of these women for council work.'

Second, all instances of *storen* ('to disturb') in which a clear physical meaning was present, as they only allow for a transitive construction, and are not instances of a psych verb to begin with. Third, all participles, because a large part of them were adjectives, which often had specialized meanings, such as *gestoord* ('crazy'). These were judged to yield too few usable instances to warrant labor-intensive manual annotation. Lastly, all instances were kept out of the analysis in which the experiencer or stimulus were not expressed, following [1, p. 25], or in which a proposition filled one of these roles. This is because, except in a number of exceptional contexts, the subject is obligatorily expressed in Dutch [25, p. 1131], [26, p. 566]. As such, when the experiencer or stimulus is not expressed, the argument construction that assigns subject position to this participant is not possible. The propositions were not taken up in the analysis because it is unclear which position they hold on the employed animacy scale [1, pp. 25–27].

The resulting dataset still included 1810 occurrences. This dataset was then manually and automatically enriched with the following variables.

A. Response variable

- Variant: *transitive, reflexive*

B. Hypothesis-driven variables

- Verb: *ergeren, interesseren, storen, verbazen*
- Stimulus-Animacy: *animate, inanimate, concrete, event, abstract*
- Stimulus-Topicality: *1st person, 2nd person, 3rd person-pronoun, definite noun, indefinite noun*
- Experiencer-Topicality: *1st person, 2nd person, 3rd person-pronoun, definite noun, indefinite noun*

C. 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*
- Medium: *written, spoken*

IV. Analysis

A logistic regression model was composed using a bidirectional stepwise variable selection procedure run in R [27].[7] This model is presented in Table 2; the effect plots of the hypothesis-driven variables can be found in Fig. 1.

The variable *Verb* does not confirm the agentivity hypothesis at the type-level. Neither does *Verb* behave as predicted by the etymology hypothesis.[8] Conversely, the variable *Stimulus-Animacy* does more or less confirm the animacy hypothesis at

[5] The Corpus of Spoken Dutch contains 10 million words of transcribed speech from a wide range of informal and formal registers. From the ConDiv corpus, we have not made use of the small diachronic component, nor of the legal material from the Bulletins of Acts, Orders and Decrees. These components returned too few occurrences of the verbs to be useful. The remaining material from ConDiv that was used, encompasses chat logs, e-mails and articles from mass and quality newspapers, in total numbering around 42 million words.

[6] Instances of the following psych verbs can be readily found on the internet in both argument constructions: *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 baffle'), *verblijden* ('to gladden'), *verdrieten* ('to grieve'), *vergenoegen* ('to content'), *verheugen* ('to rejoice'), *vermaken* ('to entertain'), *verontwaardigen* ('to indignify'), *vervelen* ('to bore') and *verwonderen* ('to surprise'). This is not intended as an exhaustive list.

[7] The ordinal variables were implemented using polynomial contrasts and the categorical variables through dummy coding. All variables as well as all two-way interactions were fed into the selection procedure. The model was then checked for the following criteria [37, p. 221], [38]. Each predictor which did not significantly improve the model, was dropped. No more parameters were allowed than the number of occurrences of the least frequent variant divided by 20. As the automatic variable selection procedure orders the predictors from most to least important, all predictors above this threshold were removed from the model. The residual deviance was not much higher than the degrees of freedom, and the VIF's were smaller than 4. Still, the HLC-test returned an almost significant p-value, indicating that important predictors may still be missing from the model. The success level of the response variable is the reflexive construction.

[8] The observed ranking of verb from strongest preference for the reflexive to the transitive construction, i.e. *ergeren* ('to annoy') – *storen* ('to disturb') – *verbazen* ('to amaze') – *interesseren* ('to interest'), does not change when raw

TABLE II. REGRESSION MODEL

AIC:	**1365.3**	**Total number of occurrences:**	**1810**
C-index:	**0.907**	**Transitive occurrences:**	**1169**
		Reflexive occurrences:	**641**

Predictors	**Levels and polynomials**	**Estimates**	**Confid. intervals**		**P-values**
			2.5%	**97.5%**	
	Intercept	1.20	0.62	1.77	< 0.0001
Verb	*ergeren*		Reference level		
	interesseren	-4.14	-4.64	-3.67	< 0.0001
	storen	-1.93	-2.38	-1.49	< 0.0001
	verbazen	-3.38	-3.89	-2.89	< 0.0001
Stimulus-Topicality	*linear*	1.96	1.26	2.69	< 0.0001
	quadratic	0.60	0.02	1.20	0.0434
	cubical	-0.87	-1.75	-0.12	0.0335
	^4	-0.43	-1.02	0.26	0.1770
Stimulus-Animacy	*linear*	0.97	0.61	1.34	< 0.0001
	quadratic	0.10	-0.22	0.42	0.5532
	cubical	1.14	0.71	1.58	< 0.0001
	^4	0.04	-0.32	0.40	0.8226
Experiencer-Topicality	*linear*	1.05	0.58	1.52	< 0.0001
	quadratic	-0.65	-1.03	-0.27	0.0008
	cubical	-0.19	-0.58	0.18	0.3160
	^4	0.26	-0.08	0.60	0.1301
Country	*Belgium*		Reference level		
	The Netherlands	0.83	0.54	1.12	< 0.0001
Stimulus-Number	*singular*		Reference level		
	plural	0.62	0.27	0.98	0.0006
Experiencer-Number	*singular*		Reference level		
	plural	-0.52	-0.92	-0.13	0.0100
Negation	*with*		Reference level		
	without	0.46	0.14	0.79	0.0052
Medium	*written*		Reference level		
	spoken	0.37	0.01	0.72	0.0416

the token level. Although the levels *concrete* and *event* don't quite behave as expected, we still find that animate stimuli have a preference for the transitive argument construction as compared to the four inanimate categories and especially as compared to the abstract stimuli.

The topicality hypothesis has been confirmed by *Stimulus-Topicality*, but *Experiencer-Topicality* behaves exactly opposite to what was predicted. In hindsight however, such behavior may not be as aberrant as it appears on first sight. For the stimulus, the choice between the transitive and reflexive construction entails a choice between subject and prepositional object position. The prepositional object is typically associated with post-field position in Dutch [28] and is thus suited for heavy informational weight. The experiencer, however, has a choice between subject and direct object. Both of these are associated with light informational weight. However, the reflexive pronoun and the preposition of the reflexive construction provide a way to enlarge the distance between a heavy experiencer and a heavy stimulus. From this, it follows that when confronted with a heavy experiencer and heavy stimulus, the reflexive construction will be preferred. When both are light, the transitive construction is better suited.

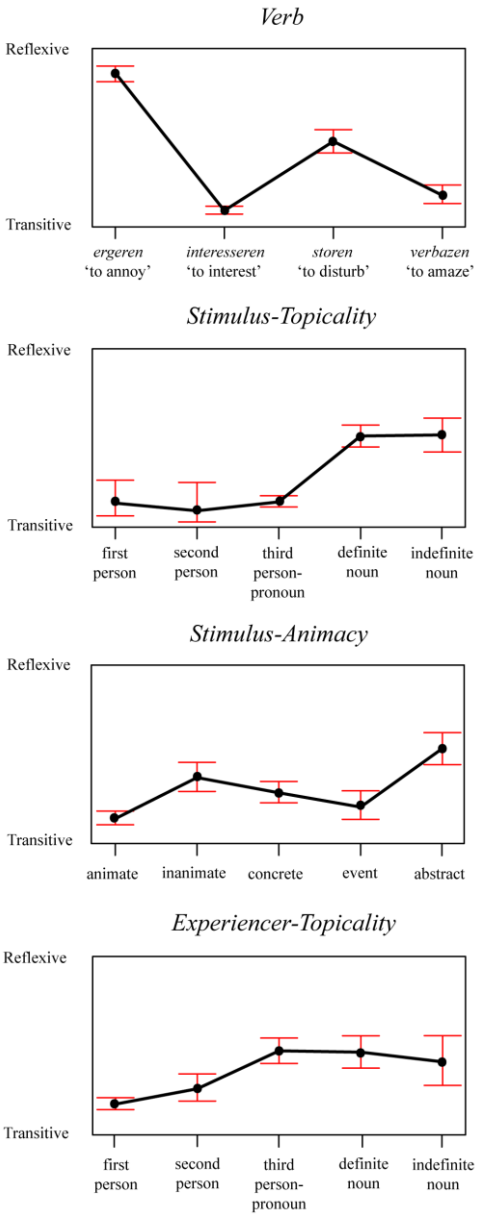

Fig. 1. Effect plots of the hypothesis-driven variables

For *Stimulus-Topicality*, we find the critical transition to be from the pronouns to the nouns. For *Experiencer-Topicality*, the difference between the persons is more essential. The reason is that nearly 95% of the stimuli are third persons, while more than 80% of the experiencers are pronouns. As such, employing the alternation to accommodate for informational differences between pronoun and noun is more expedient for stimuli, while accommodating differences between the persons greater benefits the experiencers.

numbers are calculated, nor when the occurrences with propositions as experiencer or stimulus are added to the dataset.

V. Conclusions

To end with, we shortly summarize the relevance of this study for theories of argument realization. First, the study has shown that inter- and intralingual generalizations such as the agentivity and topicality hypothesis definitely seem possible (cf. [1], [2], [10], [29]).

Second, our failure to confirm the type level agentivity hypothesis means that this study cannot be used in support of theories which aim to deduce the argument realization of a verb from its lexical meaning, such as Baker's UTAH [30], [31], the Lexical Mapping Theory [32], [33], Langacker's flow of energy [9], Van Valin's Default Macrorole Assignment Principles [34] etc. However, we must underline that the results certainly do not contest the existence of such mechanisms. There are several reasons for this. Most importantly, the present study examined no more than four verbs from a single language. These four may very well happen to present the exception, rather than the rule. Also, the operationalization of the type level argentivity hypothesis was only determined in a manual fashion and might have missed out on other important lexical semantic features. Still, this study does show that caution may be in order when applying the agentivity hypothesis too rigidly at the type level.

Finally, by confirming the token level agentivity hypothesis, we have shown that a semantic property such as agentivity may have an important effect on argument realization at the token level. That is, argument constructions do seem to add meaning to utterances, separately from the meaning of the verb [29], [35], [36]. However, the influence of such semantic properties can easily be superseded by the morphological form of the participants.

Acknowledgment

We would like to thank two anonymous reviewers for their helpful comments on an earlier version of this paper.

References

[1] B. Levin and J. Grafmiller, "Do you always fear what frightens you?," in *From Quirky Case to Representing Space: Papers in Honor of Annie Zaenen*, T. H. King and V. de Paiva, Eds. Stanford: CSLI Publications, 2012, pp. 21–32.

[2] J. Bresnan, A. Cueni, T. Nikitina, and R. H. Baayen, "Predicting the dative alternation," in *Cognitive Foundations of Interpretation*, G. Bouma, I. Krämer, and J. Zwarts, Eds. Amsterdam: Royal Netherlands Academy of Science, 2007, pp. 69–94.

[3] T. Colleman, "Verb disposition in argument structure alternations: a corpus study of the dative alternation in Dutch," *Lang. Sci.*, vol. 31, no. 5, pp. 593–611, Sep. 2009.

[4] D. Speelman and D. Geeraerts, "Causes for causatives: the case of Dutch 'doen' and 'laten,'" in *Causal Categories in Discourse and Cognition*, T. Sanders and E. Sweetser, Eds. Berlin: Mouton de Gruyter, 2009, pp. 173–204.

[5] D. Pijpops and F. Van de Velde, "A multivariate analysis of the partitive genitive in Dutch. Bringing quantitative data into a theoretical discussion," *Corpus Linguist. Linguist. Theory*, p. Published online, ahead of print, 2014.

[6] P. Hopper and S. A. Thompson, "Transitivity in Grammar and Discourse," *Language (Baltim).*, vol. 56, no. 2, pp. 251–299, 1980.

[7] J. Grimshaw, *Argument structure*. Cambridge: MIT press, 1990.

[8] D. Dowty, "Thematic proto-roles and argument selection," *Language (Baltim).*, vol. 67, no. 3, pp. 547–619, 1991.

[9] R. W. Langacker, *Foundations of cognitive grammar*. Stanford: Stanford University Press, 1995.

[10] W. Croft, "Case marking and the semantics of mental verbs," in *Semantics and the Lexicon*, J. Pustejovski, Ed. Dordrecht: Kluwer, 1993, pp. 55–72.

[11] D. Pesetsky, *Zero Syntax*. Cambridge: MIT Press, 1995.

[12] H. Vanhoe, "Aspects of the syntax of psychological verbs in Spanish: a lexical functional analysis," in *Proceedings of the LFG02 Conference*, 2002, pp. 373–389.

[13] K. Klein and S. Kutscher, "Lexical Economy and Case Selection of Psych-Verbs in German," 2005.

[14] F. Van de Velde, "De Middelnederlandse onpersoonlijke constructie en haar grammaticale concurrenten: semantische motivering van de argumentstructuur," *Ned. Taalkd.*, vol. 9, no. 1, pp. 48–76, 2004.

[15] G. Lakoff, "Linguistic Gestalts," in *Papers Regional Meeting Chicago Linguistics Society*, 1977, pp. 236–287.

[16] W. Pijnenburg and T. Schoonheim, *Oudnederlands Woordenboek*. Leiden: Instituut voor Nederlandse lexicografie, 2009.

[17] W. Pijnenburg, *Vroegmiddelnederlands woordenboek: woordenboek van het Nederlands van de dertiende eeuw in hoofdzaak op basis van het Corpus-Gysseling*. Leiden: Instituut voor Nederlandse Lexicologie, 2001.

[18] E. Verwijs and J. Verdam, *Middelnederlandsch woordenboek*. Zedelgem: Zedelgem Flandria Nostra, 1991.

[19] A. Lasch, C. Borchling, G. Cordes, and D. Möhn, *Mittelniederdeutsches Handwörterbuch*. Neumünster: Wachholtz, 1956.

[20] M. de Vries and L. A. te Winkel, *Woordenboek der Nederlandsche taal*. 's-Gravenhage: Nijhoff, 1998.

[21] M. Philippa, F. Debrabandere, A. Quak, T. Schoonheim, and N. van der Sijs, *Etymologisch woordenboek van het Nederlands*. Amsterdam: Amsterdam University Press, 2009.

[22] T. den Boon and D. Geeraerts, *Van Dale Groot woordenboek van de Nederlandse taal*, 14th ed. Antwerpen/Utrecht: Van Dale Lexicography, 2005.

[23] L. van Eerten, "Over het Corpus Gesproken Nederlands," *Ned. Taalkd.*, vol. 12, no. 3, pp. 194–215, 2007.

[24] S. Grondelaers, K. Deygers, H. Van Aken, V. Van den Heede, and D. Speelman, "Het CONDIV-corpus geschreven Nederlands," *Ned. Taalkd.*, vol. 5, no. 4, pp. 356–363, 2000.

[25] W. Haeseryn, K. Romijn, G. Geerts, J. de Rooij, and M. van den Toorn, *Algemene Nederlandse Spraakkunst*. Groningen: Nijhoff, 1997.

[26] J. van der Horst, *Geschiedenis van de Nederlandse syntaxis*. Leuven: Universitaire Pers Leuven, 2008.

[27] R Core Team, *R: A language and environment for statistical computing. R Foundation for Statistical Computing*. Vienna, 2013.

[28] A. Willems and G. De Sutter, "Distance-to-V, length and verb disposition effects on PP placement in Belgian Dutch. A corpus-based multifactorial investigation," in *Proceedings of Quantitative Investigations in Theoretical Linguistics 5*, 2013, pp. 103–105.

[29] T. Colleman and B. De Clerck, "'Caused motion'? The semantics of the English to-dative and the Dutch aan-dative," *Cogn. Linguist.*, vol. 20, no. 1, pp. 5–42, 2009.

[30] M. Baker, *Incorporation : a theory of grammatical function changing*. Chicago: University of Chicago press, 1988.

[31] M. Baker, "Thematic Roles and Syntactic Structure," in *Elements of Grammar: Handbook in Generative Syntax*, L. Haegeman, Ed. Dordrecht: Kluwer, 1997, pp. 73–137.

[32] J. Bresnan and J. Kanerva, "Locative Inversion in Chichewa: A Case Study of Factorization in Grammar," *Linguist. Inq.*, vol. 20, no. 1, pp. 1–50, 1989.

[33] J. Bresnan and J. M. Kanerva, "The Thematic Hierarchy and Locative Inversion in UG. A Reply to Paul Schachters Comments,"

in *Syntax and Semantics 26: Syntax and the Lexicon*, T. Stowell and E. Werhli, Eds. San Diego: Academic Press, 1992, pp. 111–135.

[34] R. J. Van Valin, "Semantic Macroroles in Role and Reference Grammar," in *Semantische Rollen*, R. Kailuweit and M. Hummel, Eds. Tübingen: Gunter Narr, 2004, pp. 62–82.

[35] A. E. Goldberg, *Constructions: a construction grammar approach to argument structure*. Chicago: University of Chicago press, 1995.

[36] A. E. Goldberg, "Argument Structure Constructions versus Lexical Rules or Derivational Verb Templates," *Mind Lang.*, vol. 28, no. 4, pp. 435–465, 2013.

[37] R. H. Baayen, *Analyzing linguistic data: a practical introduction to statistics using R*. Cambridge University Press, 2008.

[38] D. Speelman, "Logistic regression: A confirmatory technique for comparisons in corpus linguistics," in *Corpus Methods for Semantics: Quantitative studies in polysemy and synonymy*, D. Glynn and J. A. Robinson, Eds. Amsterdam: John Benjamins, 2014, pp. 487–533.