Understanding PP placement in written Dutch

A corpus-based multifactorial investigation
of the principal syntactic, semantic and discursive determinants

Annelore Willems & Gert De Sutter

Research unit EQTIS, Department of Translation, Interpreting and Communication
Ghent University, Belgium
{Annelore.Willems}; {Gert.DeSutter}@UGent.be

Abstract— **The present paper aims at refining current knowledge about the driving forces in Dutch constituent ordering of non-predicate prepositional phrases (PPs) as well as re-evaluating the common assumption in traditional Dutch reference grammars that the middle field position is the standard slot. Building on journalistic data in the Dutch Parallel Corpus, it is first shown that non-predicate PPs are significantly more often placed in postfield position (the structural position after the final verb cluster) than in middle field position (the position before the final verb cluster), which indicates that the postfield position rather than the middle field position should be considered the standard slot for PPs in written Dutch. Second, a binary logistic regression model was fitted with PP placement as a function of several syntactic, semantic and discursive predictor variables. This model is able to describe, explain and predict more than 80% of the variation in the data set, leading to a much better understanding of the mechanisms underlying PP placement in written journalistic Dutch. On the basis of our findings, we furthermore propose a refined and more fine-grained version of the theoretical framework in which PP placement in Dutch is traditionally described and understood.**

Keywords—syntactic variation; PP placement; Dutch; multivariate corpus analysis; logistic regression.

I. INTRODUCTION

An interesting case of syntactic variation in Dutch is the placement of non-predicate prepositional phrases (henceforth: PPs). In Dutch subordinate clauses, for instance, the PP can be placed either before the final verb cluster (the so-called *middle field position*; example 1) or after the final verb cluster (the *postfield position*; example 2):

(1) … [dat]₁ₛₜ ₚₒₗₑ de trainer **door een laptop naast het veld** middle field [vervangen wordt]final verb cluster
… [that]₁ₛₜ ₚₒₗₑ the coach **by a laptop next to the field** middle field [replaced will be]final verb cluster

(2) … [dat]₁ₛₜ ₚₒₗₑ de trainer middle field [vervangen wordt]final verb cluster **door een laptop naast het veld** postfield [dpc-cam-002333-nl]
… [that]₁ₛₜ ₚₒₗₑ the coach middle field [replaced will be]final verb cluster **by a laptop next to the field** postfield
'that the coach will be replaced by a laptop next to the field.'

In most reference grammars of Dutch as well as scientific literature, it is assumed that the middle field is the standard slot for all non-verbal clausal material, including PPs (especially in [1], the standard grammar of Dutch, but also in generative-linguistic literature (e.g., [2, 3, 4 and 5]) and in a few corpus studies (e.g., [6, 7, 8 and 9])). Direct empirical evidence for this assumption is sparse, but unequivocal: there is diachronic evidence that middle field position increasingly became the more dominant position for clausal material ever since Middle Dutch [10, 11, 12, 13, 14 and 15], there is contrastive-linguistic evidence which shows that middle field position is the more normal position in the structurally similar German language [16], and there are a few corpus studies that have revealed that middle field position is the most frequent position for PPs in spoken Dutch (cf. [6 and 9]). Although these studies seem to provide strong proof for the *standard middle field* assumption, no evidence is available for written (journalistic) Dutch yet.

It is furthermore remarkable how little is known about the driving forces behind the assumed shifting of PPs from the middle field to the postfield, especially given its central status in Dutch syntactic literature. This is in stark contrast with other variable syntactic structures in Dutch, which have been often studied in well-designed and statistically advanced corpus studies, like presence or absence of Dutch *er* ('there') in presentative sentences [17], verb order in subordinate clauses [18 and 19], direct object scrambling [20] and the dative alternation [21]. A case in point is the standard grammar of Dutch (see [1]), which states that it is unclear what governs PP placement. Nevertheless, it points to two explanatory mechanisms, viz. the complexity of the middle field on the one hand and information distribution on the other hand. The former mechanism is assumed to affect PP placement in that the length of the middle field is positively correlated with a PP shift to the postfield position. It is argued that shifting PPs to the postfield position helps to avoid an overstrained middle field, which would cause an increased parsing and processing effort (cf. [22] and [23]; following [24]), as language users have to wait considerably longer for the semantically important verbal constituent in the second pole (see e.g., [1, 25, 26]). The effect of syntactic complexity has been empirically verified in [6 and 9].

In contrast, the effect of information distribution on PP placement is often mentioned, but has not yet been studied empirically, as a consequence of which its status is unclear: in [1], it is suggested that PPs can either be extraposed because of high informative value or because of low informative value.

II. Data

Data for this study were extracted from the journalistic component of the Dutch Parallel Corpus, a 10-million-word, parallel corpus of Dutch, English and French [27]. The reason for only focussing on the journalistic component is purely practical: journalistic data is easy to collect and is qualitative data (produced by proficient language users who operate in a well-known context). Given the aims of this study, we obviously selected only the Dutch data, thus ignoring the parallel English and French texts. As the Dutch part of the corpus contains translated and non-translated Dutch texts as well as Belgian and Netherlandic Dutch, we first verified whether this had an effect on PP placement, which was not the case (resp. χ^2 = 1.07, df = 1, p = 0.3; χ^2 = 0.3, df = 1, p = 0.6). Following Delaere [28], we excluded semi-journalistic texts, such as magazines from financial institutions (since the context in which these magazines are produced as well as the objective is quite different from prototypical journalistic texts). Finally, we checked whether there was a difference (in PP placement) between news reporting articles on the one hand and comment articles on the other, which again was not the case (χ^2 = 2.83, df = 1, p = 0.09).

From this part of the Dutch Parallel Corpus, we extracted all subordinate clauses with one or more PPs in the middle or postfield. The reason to extract subordinate clauses only is that subordinate clauses always contain two verbal poles, which is a condition sine qua non for this study, as we need a clear demarcation between the middle and postfield; in main clauses the final verb cluster can remain empty, as a consequence of which it cannot be decided whether a given PP is located in the middle field or in the postfield. In order to reduce the amount of data, we only selected subordinate clauses with the grammatical (semantically empty) conjunction *dat* ('that'). This yielded an initial data set of 5,234 sentences, which was then manually filtered according to a strict set of criteria. The main criteria were: (i) the PP is located in a subordinate clause introduced by *dat* and (ii) the position of the PP is variable. As a consequence, predicate PPs in the middle field were deleted, as they are not subject to variation. Consider example (3) where the PP *in staat* ('able') is a nominal predicate, and can therefore only be placed in the middle field:

(3)   … [dat]$_{\text{1st pole}}$ het   in staat   is om spionagesatellieten [uit te schakelen]$_{\text{final verb cluster}}$
… [that]$_{\text{1st pole}}$ it   able is to espionage satellites [disconnect]$_{\text{final verb cluster}}$.
'that it is able to disconnect espionage satellites.'

Additionally, all identical clauses (produced by the same author/company) were removed from the dataset, otherwise one single data point would be counted twice. We also removed all PPs that are syntactically dependent on other PPs, NPs, APs or AdvPs (see example 4, where the underlined PP *van God* ('of God') is dependent on the NP *het woord* ('the word').

(4)   … [dat]$_{\text{1st pole}}$ de aartsengel Gabriël het woord van God [heeft verkondigd]$_{\text{final verb cluster}}$.

… [that]$_{\text{1st pole}}$ the archangel Gabriël the word of God [has preached]$_{\text{final verb cluster}}$.
'that the archangel Gabriël has preached the word of God.'

The reason for not selecting dependent PPs is that we wanted to rule out the potential effect of this factor (but this type of data will obviously be added to the data set in future investigations).

Finally, clauses with a PP which is located before the subject, as in example (5), were eliminated for a methodological reason (as we will explain below, we operationalized the weight of the middle field as the number of words between the subject and the finite verb (in the final verb cluster)). In that respect, PPs that occur before the subject, have to be considered preposed, and hence irrelevant for this study.

(5)   … [dat]$_{\text{1st pole}}$ bij elk bod boven de 1,2 miljoen euro   [zijn huidige   werkgever]$_{\text{subject}}$ zijn ontslag [[moet]$_{\text{finite verb}}$ aanvaarden]$_{\text{final verb cluster}}$.
… [that]$_{\text{1st pole}}$ with   each offer above 1,2 million euros [his current employer]$_{\text{subject}}$ his   resignation [[must]$_{\text{finite verb}}$ accept]$_{\text{final verb cluster}}$.
'that his current employer must accept his resignation with each offer above 1.2 million euros.'

After manually checking all data, we obtained a data set of 1,718 relevant clauses.

III. RESULTS (1): MIDDLE FIELD AS STANDARD SLOT

In this section, we want to answer the first research question, viz. is the middle field position the standard slot for PPs in written Dutch? If the traditional assumption of the middle field position as the standard topological field and the postfield as an expansion tank holds, one can reasonably expect that PPs are placed relatively more frequently in the middle field. However, Figure 1 shows that, in contrast to spoken discourse (40% postfield position; cf. [6] and [9]), the postfield position is the more frequent position for PPs in journalistic written discourse (60% postfield position); the difference is statistically significant (χ^2 = 52.28, df = 1, p < .0001). More specific analyses have pointed out that this distributional difference is not influenced by the predictor variables mentioned below.
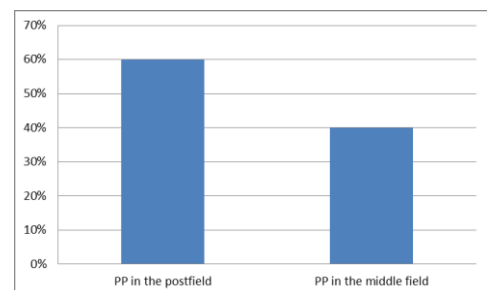

Figure 1: General distribution of PP placement in journalistic Dutch

The theoretical implication of this result is that the structuralist idea of the middle field as the only standard slot should be abandoned in favor of a more usage-based

theoretical model that considers middle field and postfield as equivalent positions; depending on the mode of communication, language users tend to favor the middle field or the postfield as standard position for PPs. The frequency difference between the spoken and written mode also raises the interesting question as to why speakers arrange PPs in a clause differently than writers. Although we cannot provide a definitive answer to such a process-related question on the basis of this corpus study only, it seems plausible to relate this difference in PP placement to the different production circumstances of journalistic texts vs. spontaneous discourse (cf. [29] and [30]). In spontaneous dialogic speech, speakers must encode and produce their utterances under heavy time constraints. As a consequence, and following the incremental approach of language production, speakers start speaking without completely planning the syntactic structure of the entire sentence [29, 30, 31, 32]. As the verbal brace construction is an inherent aspect of Dutch syntax, speakers must decide during speaking whether or not to postpone certain constituents after the final verb cluster. One can easily imagine that speakers wait as long as they themselves can remember the final verb cluster (speaker-oriented clause design), as a consequence of which they postpone fewer PPs than writers, who have the time to find the optimal constituent distribution for their readers (audience-oriented clause design). From that perspective, (professional) writers do have the time to find the optimal distribution of clausal material over the middle field and postfield.

IV. RESULTS (2): MECHANISMS GOVERNING PP PLACEMENT

In this section, we try to find out which factors determine PP placement in written Dutch. In order to achieve this, we fitted a logistic regression model[A] with PP position (middle field vs. postfield) as binary response variable and 12 predictor variables. 6 of these predictors are related to a more basic linguistic concept, viz. *syntactic complexity*: length and complexity of the PP itself, length and complexity of the middle field, length and complexity of the postfield[B]. Length was measured in terms of 'number of words' (this operationalization correlates highly with the operationalization in terms of 'number of syllables'), complexity was measured in terms of number and depth of different embedded structures. Three types of embedded structures were taken into account: subordinate clauses (6), attributes that modify a phrase (7) and appositions (8).

(6) … [dat]$_{1st pole}$ u [omringd wordt]$_{2nd pole}$ door mensen **die uw plannen kunnen uitvoeren en financieren**
… [that]$_{1st pole}$ you [surrounded are]$_{2nd pole}$ by people **who your plans can execute and finance**
'that you will be surrounded by people who can execute and finance your plans.'

(7) … [dat]$_{1st pole}$ administraties en consultants bij het ontwerp **van nieuwe programma's** [gebruikmaken]$_{2nd pole}$ van de informatie **over de opzet van bepaalde types programma's in andere landen**.
… [that]$_{1st pole}$ managements and consultants by the design **of new programs** [make use]$_{2nd pole}$ of the information **about the set-up of different kinds of types of programs in other countries**.

'that managements and consultants make use of the information about the set-up of different kinds of types of programs in other countries by the design of new programs.'

(8) … [dat]$_{1st pole}$ de Mexicaanse president, **Lazaro Cardenas**, asiel [gaf]$_{2nd pole}$ aan Trotski en zijn vrouw, **Natalia**.
… [that]1st pole the Mexican president, **Lazaro Cardenas**, asylum [gave]$_{2nd pole}$ to Trotski and his wife, **Natalia**.
'that the Mexican president, Lazaro Cardenas, gives asylum to Trotski and his wife, Natalia.'

For each embedded structure we found inside the PP, the middle field or the postfield, we increased the value for complexity with one unit. Following Szmrecsányi [33], however, we decided to double the value in case of subordinate clauses, as "subordinators are the most straightforward indicators of increased embeddedness – and thus of high complexity" [33: 1034].

Next to these syntactic complexity variables, we also included 3 discourse-related predictor variables. More particularly, we investigated the effect of definite vs. indefinite nouns (vs. non-nouns) (cf. [6 and 12]), but also of discursive importance (i.e., the number of times the referent in the PP is mentioned in the preceding and following context – based on the idea of 'persistence' [17, 34]) and the effect of discourse accessibility. This was measured by classifying all data according to 10 accessibility types (based on the Given/New-Taxonomy developed by [35]), in which new (inaccessible) information and old (completely accessible) information, such as pronouns, are at the ends of the hierarchy (see Figure 2).

Figure 2: Accessibility hierarchy

Finally, three additional predictors were added: the grammatical function and the semantic class of the PPs as well as the strength of association between a specific PP and a specific verb. For grammatical function, we distinguished between PP as an object, an adverbial adjunct and a complement. For semantic class, we used the semantic classification in [1], which again results in three (semantic) categories. In order to measure the strength of association between the verb and the PP, we used collostructional analysis ([36]), which showed that PPs that occur with copulative verbs and verbs that are part of an idiom have a clear preference for the postfield position. Based on these results, we incorporated the verb type as a final predictor variable (particle verbs, copulative verbs, main verbs and verbs in idioms).

The results of the logistic regression analysis are presented in Table 1 (we only present the predictors that have a significant effect on PP placement).

TABLE 1.

	Odds ratio	p-value
Length of the midfield	1.077	0.012 *
Length of the postfield	0.406	2.27e-07 ***
Length of the PP	1.361	< 2e-16 ***
Discourse accessibility	1.174	5.44e-08 ***
Particle verbs	1.627	0.004 **
Verbs in idioms	2.142	0.000 ***
Copulative verbs	2.435	4.57e-05 ***
Semantic category 1: e.g. location,…	0.330	1.22e-08 ***
Semantic category 2: time	0.102	5.91e-16 ***

First, it should be noted that the overall quality of the model is very good, with a c-index of 0.82 – also, the model does not suffer from multicollinearity[C].

Second, the results show that this type of syntactic variation too (cf. the introduction for related research focussing on other types of variation in Dutch) shows up to be multifactorial in nature.

Third, the model shows that both syntactic complexity and discourse-related aspects play a unique, decisive role in PP placement, something which was suggested in previous literature, but never empirically tested. More particularly, the multivariate analysis points out that the length of the PP and the length of the middle field are positively correlated with postfield position, which is in line with what was previously found. More remarkably, the length of the postfield, which is a new predictor, also influences the word order. In fact, it works as an inhibitory variable: the heavier the postfield, the fewer PPs in postfield position. As for the discourse-related variables, the results show that less accessible PPs are placed significantly more often in the postfield, which is in line with the general literature on the Given-before-New principle (see among others [1 and 37]): the given (and thus accessible) information is placed before the new (less accessible) information.

As for the semantic class of the PP, our results show that locative and temporal PPs prefer middle field position. Finally, verb type affects PP placement significantly in that postfield position is preferred in sentences with a semantically 'empty' verb, such as copulative verbs and verbs in idioms. Although it is not completely clear how to interpret this result, one could imagine that these verbs can be processed more quickly (cf. [38] and [39]), as a result of which they are preferably placed before a less accessible clausal element (the PP).

[A] We also performed a mixed model with a random factor 'verbs', the fixed factors from table 1 and the random slopes. The predictive power of that model was 0.901.
[B] When statistically exploring the operationalization of length in terms of words and syllables, both operationalizations yielded high correlation scores (r = 0.93-0.98). For that reason, we will restrict the discussion in the remainder of this text to one of the length operationalizations, viz. in terms of words, which is the standard procedure in most corpus studies.
[C] Following [17], we treat vif ≥ 4 as the threshold for multicollinearity.

V. CONCLUSION

The results of this study are relevant and innovative for Dutch language studies, and especially for Dutch syntax. First, the general distribution presented in this paper refines the traditional idea of the middle field as the standard position for PPs. Instead of the middle field as standard slot and the postfield as an expansion tank, we argue for a bidirectional model, in which both the middle field and postfield positions are equivalent; journalists then place PPs either in the middle field or postfield depending on the level of occupation by other constituents, the size of the PP itself, the discourse accessibility of the PP, and the semantic status of the main verb and the PP itself. Secondly, the multifactorial analysis adds further proof to the existence of a so-called probabilistic grammar, in which word order is not determined in a clear-cut all-or-nothing manner, but on the basis of fine-grained linguistic and contextual constraints, which language users seem to internalize through exposure and use (cf. [40]).

REFERENCES

[1] **ANS** – Haeseryn, W., Romijn, K., Geerts, G., de Rooij, J. and van den Toorn, M.C. *Algemene Nederlandse Spraakkunst*. Groningen/Deurne: Martinus Nijhoff – Wolters Plantyn, 1997.

[2] Koster, J. "Het werkwoord als spiegelcentrum". *Spektator*, 3, 601-618, 1974.

[3] Koster, J. *Locality Principles in Syntax*. Dordrecht: Foris, 1978.

[4] Hoekstra, T. *Transitivity: grammatical relations in Government-Binding theory*. Dordrecht: Foris, 1984.

[5] Zwart, J.W. *PP Extraposition from NP in Dutch*. Ms. U. of Groningen, 1990.

[6] Jansen, F. "Hoe krijgt een spreker zijn woorden op een rijtje? Taalgebruiksaspekten van de 'PP over V' konstruktie". In: J. G. Kooij (Ed.), *Aspekten van de woordvolgorde in het Nederlands*. Leiden: Vakgroep Nederlandse Taal- and Letterkunde, 70-104, 1978.

[7] Jansen, F. "On tracing conditioning factors of movement rules: extraposition of PP in spoken Dutch". In: M. Van de Velde and W. Vandeweghe (Ed.), *Sprachstruktur, Individuum und Gesellschaft*. Tübingen: Niemeyer, 83-93, 1979.

[8] Jansen, F. "Tangconstructies en geen reacties daarop". *Taal en Tongval*, 3, 135-151, 1990.

[9] Braecke, C. "Uit de tang' of [+ or -] prominent?". *Taal en Tongval*, 3, 125-134, 1990.

[10] De Meersman, A. *Woordvolgorde in 14de-eeuws Brabants*. Gent : Koninklijke academie voor Nederlandse taal- en letterkunde, 1980.

[11] De Schutter, G. "In en uit de tang in de Middelnederlandse bijzin". *Leuvense Bijdragen*, 77, 385-401, 1988.

[12] De Schutter, G. "Extrapositie in het Middelnederlands. Proza- en rijmteksten als informatiebron voor syntactisch onderzoek". *Taal en Tongval*, 47, 195-204, 1995.

[13] De Schutter, G. "De zinsbouw in het vroege Middelnederlands. Een case study". *Taal en Tongval*, 55, 41-67, 2003.

[14] Burridge, K. *Syntactic change in Germanic. Aspects of language change in Germanic with particular reference to Middle Dutch*. Amsterdam: John Benjamins pubishing company, 1993.

[15] Coussé, E. **"**Focus, complexiteit en extrapositie. Over de veranderende woordvolgorde in het Nederlands". *Neerlandistiek.nl,* 9 (4), 1-32, 2009.

[16] Van de Velde, M. "De Ausklammerung in het Duits en het Nederlands". *Studia Germanica Gandensia*, 14, 119-142, 1973.

[17] Grondelaers, S.A. and Speelman, D. "A variationist account of constituent ordering in presentative sentences in Belgian Dutch". *Corpus Linguistics and Linguistic Theory*, 3, 161-193, 2007.

[18] De Sutter, G. "Naar een corpusgebaseerde, cognitief-functionele verklaring van de woordvolgordevariatie in tweeledige werkwoordelijke eindgroepen". *Nederlandse Taalkunde*, 12, 302-330, 2007.

[19] Coussé, E., Arfs, M. and De Sutter, G. "Variabele werkwoordsvolgorde in de Nederlandse werkwoordelijke eindgroep. Een taalgebruiksgebaseerd perspectief op de synchronie en diachronie van de zgn. rode en groene woordvolgorde". In: G. Rawoens (Ed.), *Taal aan den lijve. Het gebruik van corpora in taalkundig onderzoek en taalonderwijs*. Ghent: Academia Press, 29-47, 2008.

[20] Van Bergen, G. and de Swart, P. „Scrambling in Spoken Dutch. Definiteness versus weight as determinants of word order variation". *Corpus Linguistics and Linguistic Theory*, 6 (2), 267-295, 2010.

[21] Colleman, T. "Beyond the dative alternation: the semantics of the Dutch *aan*-dative". In: D. Glynn and K. Fischer (Eds.), *Quantitative Methods in Cognitive Semantics: Corpus-driven Approaches*. Berlin/New York: Mouton de Gruyter, 271-303, 2010.

[22] Gibson, E. "The dependency locality theory: A distance-based theory of linguistic complexity". In: Y. Miyashita, A. Marantz, and W. O'Neil (Eds.*), Image, language, brain*. Cambridge: MA: MIT Press, 95-126, 2000.

[23] Wasow, T. *Postverbal behavior*. Stanford, CA: CSLI Publications, 2002.

[24] Behaghel, O. „Beziehungen zwischen Umfang und Reihenfolge von Satzgliedern". *Indogermanische Forschungen*, 25, 110-142, 1909.

[25] Van Haeringen, C.B. "'Tangconstructies' en reacties daarop". *De Nieuwe Taalgids*, 40, 1-7, 1949.

[26] Zwart, J.W. *The syntax of Dutch*. Cambridge: Cambridge University Press, 2011.

[27] Macken, L., Declercq, O. and Paulussen, H. "Dutch Parallel Corpus: a Balanced Copyright-Cleared Parallel Corpus". *META*, 56 (2), 374-390, 2011.

[28] Delaere, I. "Do translations walk the line? Visually exploring translated and non-translated texts in search of norm conformity." Phd, 2015.

[29] Ferreira, F. and Swets, B. "How Incremental Is Language Production? Evidence from the Production of Utterances Requiring the Computation of Arithmetic Sums". *Journal of Memory and Language*, 46, 57-84, 2002.

[30] Allum, P. H. and Wheeldon, L.R. "Planning Scope in Spoken Sentence Production: The Role of Grammatical Units". *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(4), 791-810, 2007.

[31] Levelt, W. J. M. 1989. *Speaking: From intention to articulation*. Cambridge, MA: MIT Press, 1989.

[32] Wheeldon, L. and Lahiri, A. 1997. "Prosodic units in speech production". *Journal of Memory and Language*, 37, 356-381, 1997.

[33] Szmrecsányi, B. "On operationalizing syntactic complexity". In: G. Purnelle, C. Fairon and A. Dister (Eds.), *Le poids des mots. Proceedings of the 7th International Conference on Textual Data Statistical Analysis*. Louvain-la-Neuve, March 10-12, 2004, Vol. 2. Louvain-la-Neuve: Presses universitaires de Louvain, 1032-1039, 2004.

[34] Givón, T. *Topic Continuity in Discourse: A Quantitative Cross Language Study*. Amsterdam: John Benjamins, 1983.

[35] Grondelaers, S., Speelman, D., Drieghe, D., Brysbaert, M. & Geeraerts, D. "Introducing a new entity into discourse: Comprehension and production evidence for the status of Dutch er "there" as a higher-level expectancy monitor". *Acta Psychologica* 130, 153-160, 2009.

[36] Gries, S. Th. & Stefanowitsch, A." Extending collostructional analysis: a corpus-based perspective on 'alternations'". *International Journal of Corpus Linguistics* 9(1). 97-129, 2004.

[37] Gundel, J.K., Hedberg, N. and Zacharski, R. "Cognitive status and the form of referring expressions". *Language*, 69, 274-307, 1993.

[38] Segalowitz, S.J. and Lane, K.C. ‚Lexical access of function versus content words'. *Brain and Language*, 75(3): 376-389, 2000.

[39] Tremblay, A., Derwing, B., Libben, G. and Westbury, C. „Processing advantages of lexical bundles: Evidence from self-paced reading and sentence recall tasks". *Language Learning*, 61(2): 569-613, 2011.

[40] Bybee, J. and Beckner, C. "Usage-based Theory". In: B. Heine and H. Narrog (eds.), *The Oxford Handbook of Linguistic Analysis*. Oxford: Oxford University Press, 826-855, 2011.