Romance N Prep N constructions in visual word recognition:

An eye-tracking study of French, Spanish and Portuguese

Abstract

N Prep N constructions such as Sp. *bicicleta de montaña* 'mountain bike' are very productive and frequent in Romance languages. They commonly have been classified as *syntagmatic compounds* that show no orthographic union and exhibit an internal structure that resembles free syntactic structures, such as Sp. *libro para niños* 'book for children'. There is no consensus on how to best distinguish lexical from syntactic N Prep N constructions. The present paper presents an explorative eye-tracking study on N Prep N constructions, varying both lexical type (lexical vs. syntactic) and preposition across three languages, French, Spanish and Portuguese. The task of the eye-tracking study was a reading aloud paradigm of the constructions in sentence context. Constructions were fixated on less when more frequent, independent of lexical status. There was also modest evidence that a higher construction frequency afforded shorter total fixation durations, but only for lower deciles of the response distribution. The (construction-initial) head noun also received fewer fixations as construction frequency increased, and also when the head noun was more frequent. The second fixation durations on the head noun also revealed an effect of lexical status, with syntactic constructions receiving shorter fixations at the 5th and 7th deciles. The probability of a fixation on the preposition decreased with preposition frequency, but first fixations on the preposition increased with preposition frequency. The prepositions of Portuguese, the language with the richest inventory of prepositions, received more fixations than the prepositions of French and Spanish. The observed pattern of results is consistent with models of lexical processing in which reading is guided by knowledge of both higher-level constructions and knowledge of key constituents such as the head noun and the preposition.

Introduction

N Prep N constructions in Romance languages are a very productive and frequent pattern in nominal word formation, as in Sp. *puerta de casa* 'front door' or Pt. *roupão de banho* 'bathrobe'. While N Prep N constructions may form lexicalized compounds, they also may maintain their free syntactic status, such as in Pt. *café com leite* 'coffee with milk' or in Sp. *molde para dulces* 'tin for cake'. The distinction between lexicalized and free syntactic forms has been subject to many discussions. It is for that reason that N Prep N constructions and other types of so-called *syntagmatic compounds* have for a long time been considered the most problematic issue in research on Romance compounding (e.g. Moyna 2011: 38). As Buenafuentes de la Mata (2007) states, syntagmatic composition is one of the most productive but also of the most problematic issues in Spanish (Buenafuentes de la Mata 2007: 83).

In general, the discussion about N Prep N constructions in Romance languages revolves around two crucial issues. The first issue lies at the interface of morphology and lexicology. It considers the question, whether N Prep N constructions may be regarded as fully lexicalized or even fossilized expressions or whether they underlie productive patterns and processes of word formation. The second issue lies at the interface of morphology and syntax and considers the morphosyntactic status of the constructions. There is no consensus in the literature as to whether N Prep N constructions are a part of the class of compounds or whether they count as free syntactic word combinations. In order to address these issues from an empirical perspective, the present eye-tracking study examines the processing of N Prep N constructions in three Romance languages: Spanish, French and Portuguese.

Defining and determining the common characteristics of syntagmatic compounds plays a crucial role in the definition and delimitation of these multiword units. Various scholars proposed delimitation and classification tests in order to determine the lexical or syntactic status of a respective N Prep N construction. Still, there are as yet no clear criteria for classifying N Prep N compounds as such or for delimiting them from other free syntactic structures. More recent approaches even suggested that both types of N Prep N constructions may lie on a continuum between syntax and lexicon (e.g. Rio-Torto & Ribeiro 2012).

Research on the processing of multi-word units has been at the center of psycholinguistic research since its beginning. The processing of multi-word units, especially compounds, played an important part in the development of models of morphological processing. Research on processing of compounds and other multi-word units until today has mainly focused on the English language. However, many languages may resort to other compounding types that have not yet been investigated in detail. Romance languages prefer compounding of the type N Prep N in order to express two connected nominal concepts, such as in Fr. *salle de bain* 'bathroom'. While other types of compounding, such as N N or N A, also exist in Romance languages, the construction N Prep N is particularly frequent and productive (see Hennecke & Baayen 2017). This type of compounding, which has also been called syntagmatic compounding for its syntagmatic and sentence-like structure, has not yet been investigated from a psycholinguistic and experimental perspective. In the present study, we consider the differences in the frequency distributions of the construction N Prep N, the semantic relation of the constituents, the lexical or syntactic status of the whole construction and in addition the possible variation in the use of the prepositional constituent. The results of the study will be discussed in the light of current discussions on the processing and modeling of compounds and multi-word units. Of special interest to us is whether lexicalized compounds and free syntactic constructions show processing differences in visual word recognition. This raises the question of the consequences of there being multiple words that have to be read for visual word recognition, since individual constituents of a construction are likely to be processed before the whole construction has been scanned. With regards to this research question, frequency effects may provide further information on the processing of lexicalized compounds and free syntactic constructions. Furthermore, the semantic relations between the nominal constituents will be examined in detail, following up on earlier work on semantic relations in English compounds (e.g. Gagné & Shoben 1997; Gagné & Spalding 2009). It is noteworthy that the Romance construction under investigation makes the semantic relation explicit to some extent, by means of the prepositional element. At the same time, a given preposition may subserve the expression of different semantic relations. It therefore is an open question whether relational effects are actually detectable when reading Romance N Prep N constructions, over and above effects of the preposition.

The present paper presents a comparative eye-tracking study, exploring a cluster of relative research questions in an exploratory way. First, it addresses the question of whether N Prep N constructions in French, Spanish and Portuguese show processing differences that are related to their frequency distribution, their syntactic or lexical status, their semantic relational structure or the type of internal preposition. Second, it examines whether the differences in productivity of the internal variation in the three different languages, observed by Hennecke and Baayen (2017), also gives rise to processing differences between the languages in an eye-tracking study.

As an overall aim, the present study seeks to clarify whether syntactic and lexical N Prep N constructions are treated differently in visual word recognition and which role different factors play in the processing of these items.

State of the Art in the Literature

The status of compounds of the type N Prep N is a hotly debated topic in Romance word formation. In the past, syntagmatic compounds have also been designated as phrasal or syntactic compounds, as lexicalized phrases or syntactic words (see e.g. Bisetto & Scalise 2005; Rio-Torto & Ribeiro 2009; Villoing 2012; Fraudin 2009; Kornfeld 2009). While the terminology and definition of syntagmatic compounds is far from clear, even their classification is disputed and they generally form a relatively heterogeneous group. Rio-Torto and Ribeiro (2012) include syntagmatic compounds (in their terminology phrasal compounds) of the type $[N[PrepN]]_N$ as in *caminho de ferro* (way of iron) 'railway', $[[N[PrepV]_N$ as in *ferro de engomar* (iron of starch) 'flat iron', $[VconjV]_N$ *come e dorme* (eat and sleep) 'eat and sleep', $[NA]_N$ as in *era cristã* (era Christian) 'Christian era' as well as $[AN]_N$ *belas-artes* (beautiful arts) 'fine arts' in their classification of Portuguese (Rio-Torto & Ribeiro 2012: 7). According to them, syntagmatic compounds lie on a continuum between morpho-lexical and syntactic structures (ibd.). Generally, in research on Romance compounding, it has been agreed that syntagmatic compounds do not show any graphical union and therefore share the same overall structure as syntactic phrases. Still, Buenafuentes de la Mata notes that syntagmatic compounds are not mere lexical compounds without graphic union, but more complex, as they present very specific characteristics (Buenafuentes de la Mata 2010: 24). One of the crucial distinctive features of syntagmatic compounds is their lack of compositional meaning; this is why they can be syntactically reinterpreted as complex nouns (Val Àlvaro 1999: 4827). Nevertheless, syntagmatic compounds share certain characteristics with syntactic phrases, as they follow the usual patterns of head modification in nominal and prepositional phrases (Masini 2009: 257). According to Masini (2009), syntagmatic compounds share the same overall structure with syntactic phrases, but differ only in the compositionality of their meaning. Still, many studies tend to exclude syntagmatic compounds from the general class of morphological compounds as they are supposedly not formed by a morphological process of word formation (e.g. Villoing 2012: 35), would be mere 'frozen' and lexicalized syntactic structures (e.g. Guevara 2012: 179) and do not follow productive word formation processes (ibd.). Still, Guevara acknowledges that these constructions show "certain effects of lexicalization and atomicity in their distribution" (Guevara 2012: 180). We agree with Rainer (2016), who points out that syntagmatic compounds of the type N Prep N are a very productive lexical pattern in Romance languages, which continues to mostly follow the rules of syntax (Rainer 2016: 2624, see Hennecke & Baayen 2017 for a comparative quantitative corpus analysis).

The present study investigates how Romance N Prep N constructions are perceived and processed in visual word recognition. Psycholinguistic research deals, amongst other things, with the question of whether a complex multi-word unit, such as an N Prep N construction, is first perceived as a whole whether each constituent is perceived separately before processing the whole construction. In order to answer this question, the role of the frequency of the whole construction and its constituents in Romance N Prep N constructions is potentially informative.

Frequency effects of whole words (i.e. compounds) and frequency effects tied to constituents have played a central role from the beginning of research on morphological processing. There is currently no consensus on the exact role of constituent frequency nor as to how the whole word and its constituent frequency interact (see Baayen, Kuperman & Bertram 2009 and Schmidtke et al. 2016 for a detailed discussion). Early studies, such as Taft and Forster (1976) posited that the constituent effect precedes the compound effect. In their studies, constituent frequency effects are interpreted as a sign for early morphological decomposition, while whole word frequency effects indicate subsequent recombination. More recent studies, such as Rastle, Davis and New (2004) support the view that constituents of complex words are perceived earlier than the whole complex word and state that obligatory decomposition occurs in an early stage of processing (see also Pylkkänen et al. 2004; Solomyak & Marantz 2010). In contrast, the results from Schmidtke et al. (2016) indicate that the whole word effect precedes the constituent effect and that the whole construction is perceived before its single constituents are (see also Giraudo & Grainger 2001; Feldman et al. 2009). Giraudo & Grainger (2001) explain morphological constituent effects by means of post-lexical access processes. In their interpretation constituent effects arise later than whole word effects due to initial access being to the meaning of the whole word. Still, there is a general consensus that multi-word expressions and compounds that are considered to be lexical trigger both constituent and whole word effects. Baayen, Kuperman & Bertram (2009: 17) support the view that the different frequency effects interact in a complex dynamic system. Following Kuperman et al. (2008), the whole word effect may also reflect the probability of co-occurrence of the constituents as well as their combinatorial semantics (Baayen, Kuperman & Bertram 2009: 17). While frequency undoubtedly plays a very important role in the acquisition and processing of morphological information, it has been shown that other factors, such as probability and productivity measures, should also be taken in account (see Kuperman et al. 2008). For the present study, three potential outcomes are possible: 1. the present study may find no frequency effects for N Prep N constructions in Romance languages, 2. the compound effect precedes the constituent effect (e.g. Schmidtke et al. 2016), 3. the constituent effect precedes the compound effect (e.g. Taft & Forster 1976).

Furthermore, we are interested in whether frequency effects occur for lexical compounds only or whether they also appear in syntactic constructions. In previous research, it has been shown that frequency may not only occur for lexical constructions but also influences syntactic processing (e.g. Cutler 1983; Gahl & Garnsey 2004; Ellis 2002 for an overview). Syntactic frequency effects have been detected and unquestionably interact with the frequency of the individual constituents in a syntactic construction. In this context, the question of the differences between syntactic (compositional) and lexical (non-compositional) processing has been approached in different models. Traditional models of morphological processing can generally be divided into two approaches: dual-system theories (e.g. Chomsky 1965, Pinker 1991, Pinker & Ullman 2002, Ullman 2004) and single-system theories (e.g. Bates and MacWhinney 1989; Rummelhart & McClelland 1986, Frost et al. 1997) (for a discussion of both approaches see Snider & Arnon 2012). Most traditional dual-system theories assume that compositional and non-compositional multi-word units are processed differently in the course of speech processing. Traditional dual-system theories postulate the existence of a mental lexicon – a store of lexical forms and information about their semantic and syntactic properties

– as well as a grammar component that computes forms that are not stored as such in the mental lexicon. According to this approach, forms that are stored in the lexicon and forms that are computed by grammar are learned differently and are even governed by different neural substrates (Ullman 2001). In contrast, single-system theories do not assume a distinction between lexically stored and grammatically computed forms. More precisely, grammatical and lexical items are learned and processed in the same way and underlie the same cognitive mechanisms. Single-system models may further be divided into models taking a connectionist approach and models taking an exemplar approach. While in exemplar models, all linguistic items are linked to clusters of memory traces (e.g. Pierrehumbert 2001), connectionist models posit that a single network, or set of coupled networks, underly the processing of all forms (e.g. Rumelhart & McClelland 1986) (Snider & Arnon 2012: 128). According to both approaches, all linguistic units, regardless of their length and semantic complexity, are subserved by one single cognitive mechanism. in these approaches, there is no distinction between 'stored' (i.e. 'lexical') and 'computed' (i.e. 'grammatical') forms. In yet other computational approaches, there are no internal lexical representations whatsover, but only high-dimensionality modality-specific input and output states. (Baayen, Chuang, Shafaei-Bajestan & Blevins 2019). As this line of research has not addressed compounding, we do not consider it in the remainder of this study.

As the present paper focuses on the distinction of lexical and grammatical (subsequently referred to as 'syntactic') forms in N Prep N constructions, it contributes to the discussion on dual-system and single-system access in morphological processing. Lexical N Prep N constructions, such as Fr. *lit d'enfant* 'crib', may be considered as non-compositional and 'stored' units, whereas syntactic N Prep N constructions, such as Fr. *film pour enfants* 'children movie', can be regarded as compositional and 'computed' units. However, both types of constructions share the same surface form. Therefore, dual-system theories should posit a processing difference between both forms, while single-system theories would predict no processing difference. Snider and Arnon (2012) mention that it is hard to distinguish between compositional and non-compositional forms and therefore question the usefulness of the distinction between 'stored' and 'computed' forms (Snider & Arnon 2012: 130).
Hence, for the present eye-tracking study, we could expect the following results. First, frequency effects may occur for lexical items only. If syntactic constructions are exempt from frequency effects, there may be an interaction between lexicality and frequency of the constructions. Second, frequency effects may occur for lexical and syntactic items without any significant difference in processing. In this case, it might be possible that there is a qualitative difference in processing on the constituent level. That is to say, even if we do not find a processing difference on construction level, differences may occur in processing of the different constituents.

As mentioned before, it remains controversial how to properly differentiate between lexical compounds and syntactic phrases. In line with this discussion, several studies have sought to establish clear tests and criteria in order to distinguish between clearly lexical and clearly syntactical N Prep N constructions (e.g. Bouvier 2000 for French, Val Àlvaro 1999 for Spanish, and Rio-Torto & Ribeiro 2012 for Portuguese). One common test is to check the interchangeability of the constituents, which should not be possible for lexical units, e.g. Sp.

ojo de buey (eye of bull) 'pothole', **ojo de vaca* (eye of cow) (Val Àlvaro 1999: 4825). A second test is to insert new constituents, as for example determiners or complements, as in Sp. *caja fuerte* (box strong) 'safe' and * *caja muy fuerte* (box very strong) (ibd.). Except for the interchangeability test and the insertion test, studies rely on tests of pronominalization and head ellipsis. It is thus not possible to pronominalize lexical compounds, as in Sp. **lo fuerte de la caja* (the strong of the box) or to omit their head, as in Sp. **caja fuerte y grande* (box strong and big). Another important test, which relies on the semantics of the unit, is the test of compositionality. According to Val Àlvaro (1999), the absence of a compositional meaning is the most crucial distinctive feature for lexical N Prep N constructions (Val Àlvaro 1999: 4827). Hence, the construction forms a lexical unit if it presents one single naming unit, that is to say if it refers to one single concept, such as Fr. *salle de bain* (room of bath) 'bathroom' (see Štekauer 2005). For the present study, this semantic criterion also was considered very important in determining the lexical status of the constructions in the eye-tracking task. Despite these tests, it is not always clear how to distinguish between lexical and syntactic units. As mentioned earlier, delimitation tests rely on the fact that constituents in lexical syntagmatic compounds are not interchangeable and constitute a non-compositional unit. Still, reality is far more complex than that, as Buenafuentes de la Mata (2009) already mentions. She points out that N Prep N constructions may indeed show a certain form of paradigmatic behavior, given that they may form paradigmatic networks, such as in Sp. *oso pardo* (bear brown) 'brown bear', *oso panda* (bear panda) 'panda bear' and *oso polar* (bear polar) 'polar bear' (Buenafuentes de la Mata 2009: 14). In the same way, constructions of the type N Prep N may form these kinds of networks, such as Fr. *salle de bain* (room of bath) 'bathroom', *salle de séjour* (room of stay) 'living room', *salle de classe* (room of class) 'class room' etc. These paradigmatic networks can be described as schemas from a constructional perspective (see e.g. Masini 2009, Booij 2010 and Hennecke 2019 for a detailed discussion and application of constructional approaches). Constructional schemas express morphological patterns and specify how to form new complex words by considering the form-meaning relationship between words (Booij 2018).

In this framework, the internal variation of the prepositional element in N Prep N constructions is highly relevant. All three Romance languages under investigation exhibit internal variation of the preposition in certain cases, such as in:

French	**Spanish**	**Portuguese**
chemise de/en coton (shirt of cotton) 'cotton shirt	*construcción de/en madera* (construction of wood) 'wood construction'	*escada de/em caracol* (stair of/in snail) 'spiral stair'
course à/d'obstacles (course with/of obstacles) 'obstacle course'	*motor de/a gasolina* (motor with/of gas) 'gas engine'	*forno de/a micro-ondas* (oven with microwave) 'microwave oven'
collier de/pour chien (collar for dog) 'dog collar'	*ropa de/para niños* (clothes for children) 'children's clothing'	*balança de/para pesagem* (scale for weighing) 'scale'

This kind of variation appears in all of the three languages but is by far the most productive in Portuguese and the least productive in Spanish (see Hennecke & Baayen 2017). While the prepositional variation occurs most commonly with the preposition *de* as well as the prepositions *en/em*, *à/a*, *pour/para*, *avec/con/com*, internal prepositional variation is possible with a large variety of prepositions, such as in Fr. *medicaments sur/sous ordonnance* (medicines under/over prescription) 'prescription drugs', Pt. *apetite ao/pelo risco* (appetite for risk) 'risk appetite' or Pt. *hóquei em/sobre patins* (hockey in/on skates) 'roller hockey'. Again, Portuguese shows the greatest variation, Spanish and French are much more restricted in what they allow. The existence of internal prepositional variation as such already challenges the above-named delimitation test of interchangeability and the non-compositionality of the constituents of a compound-like lexical unit. However, the paradigmatic pairs also constitute a problem for the delimitation of lexical and syntactic units. In many cases, it cannot be determined unequivocally whether a specific construction should be classified as a lexical or a syntactic unit, as for example in Fr. *jouet pour enfants* 'children's toy' or in Sp. *persona con discapacidad* 'handicapped person'. While both constructions are quite lexicalized and also constitute one single naming unit, they still allow some of the delimitation tests, such as insertion of lexical material, for instance in Fr. *jouet éducatif pour enfants* (toy educational for children) 'educational children's toy' or as in Sp. *persona con gran discapacidad física* (person with big disability physical) 'person with strong physical disability'. As stated by Rio-Torto & Ribeiro (2010) the insertion of lexical material may also transfer a lexical unit to a syntactic status and provoke a shift from a generic to a more referential reading, as in Pt. *fim de semana* (end of week) 'weekend' and Pt. *fim da semana* (end of the week) 'end of the week'. Hence, for the present study, internal variation in N Prep N compounds in Romance languages is of particular interest. This question is closely linked to the above-named question of the lexical and syntactic status of the N Prep N constructions. Since in many cases the variation of the preposition goes hand in hand with a shift in the lexicality of the expression and may influence the possible results of the present study.

However, as mentioned before, the possible internal variations and the paradigmatic prepositional pairs vary in each language and appear to be linked to the semantics of the construction. Hennecke (2019) investigated the semantic relational structures of a set of N Prep N constructions with paradigmatic variation from large-scale corpus data. Results show that internal variation indeed largely depends on the conceptual relation between the two nominal constituents in N Prep N constructions. The role of the relational structure of compounds has also been investigated from a psycholinguistic perspective. It has been shown in various studies, mostly focusing on English, that the relational structure between head and modifier affects compound processing (e.g. Gagné & Shoben 1997; Gagné & Spalding 2009; for a detailed discussion see Spalding & Gagné 2011 and Schmidtke et al. 2016). Results show that relation-priming effects mostly occur with the modifier (e.g. Gagné & Shoben 1997; Maguire & Cater 2004). Several studies (e.g. Gagné & Shoben 1997; Gagné & Spalding 2014) also investigated a relational competition effect in that increased competition of relational structures leads to increased processing costs. The CARIN and RICE theories, which were developed from these findings (Gagné & Shoben 1997; Spalding et al. 2010), are based on a potential competition of relational structures during the semantic composition of compounds and compound-like phrases. However, Gagné and Shoben (1997) point out that the type of the semantic relation

may play a less prominent role in processing than other variables, such as availability of the relation and competition (Gagné & Shoben 1997: 72-73). Gagné and Spalding (2014) as well as Spalding and Gagné (2014) already mention the influence of relational entropy and point out that semantic relation may have an effect on processing both in opaque and transparent compounds. These studies were complemented by the study of Schmidtke et al. (2016) that promotes entropy as a measure of relative competition in order to analyze the origin of the competition effect. Entropy measures include probability and estimation measures on the semantic relations in a specific compound (see Schmidtke et al. 2016: 558). They find clear evidence that increased entropy of relational competition leads to increased processing costs, that is to say slower reaction times. They conclude that only relevant meanings associated with a single compound affect language processing, and that the effect observed only reflects a typical neighborhood competition effect. The present eye-tracking study aims to add to these findings on the semantic relation of English compounds by investigating the semantic relation of the nominal constituents in N Prep N constructions in Romance languages. Here, we investigate if the above-mentioned results apply to Romance languages as well in that the relational structure between head and modifier also affects compound processing in Romance N Prep N constructions. In N Prep N constructions, the semantic relation between the head and the modifier is overtly expressed to a considerable extent by the prepositional element. That is, the semantic relation is fully implicit in English but much more explicit in Romance languages. This raises the question of whether our study can find any effect that goes further than the preposition. The present study only addresses the type of semantic relation, as already investigated in the qualitative corpus study from Hennecke (2019), and not primarily on the relational competition.

In summary, it can be stated that N Prep N constructions remain a very controversial topic in Romance linguistics. In line with current psycholinguistic and experimental investigations and theoretical developments, it seems appropriate to investigate N Prep N constructions with experimental methods in order to clarify their status and their role in language processing. As many recent and relevant studies on the processing of compounds and multi-word units rely on eye-tracking methods, the present explorative study also uses this experimental method. Data from eye-tracking studies can provide important information about the online processing of N Prep N constructions during the reading process.

Participants

64 native speakers participated in this experiment. 22 native speakers from France and the Maghreb – 15 female and 7 male – participated in the French task. The Spanish task was completed by 22 Spanish native speakers of European and American Spanish – 12 female and 10 male. 20 native speakers of Brazilian Portuguese – 14 female and 6 male – took part in the Portuguese task. All participants were aged between 18 and 50 years and acquired their native language under the age of 4. All participants were uninformed about the aims of the task and participated in exchange for payment. All participants reported normal vision. Before the task, all participants signed a consent form that informed them about the procedure and the risks of the task. Furthermore, all participants completed a short self-assessment questionnaire about

their language biography in the language of the latter task[1]. The questionnaire aimed at collecting general information about the participants, such as their age, sex and nationality. Furthermore, the questionnaire contained four short questions about their language acquisition and language competence. The objective of the questionnaire was to gather information on the three groups of participants from the three different tasks in order to test for potential differences between the subject groups.

Materials

All N Prep N constructions with internal prepositional variability where extracted for each of the three languages from the TenTen Corpora: FrTenTen for French, SpTenTen for Spanish and PtTenTen for Portuguese[2]. This initial dataset contained 6991 French tokens (1062 types), 10,219 Spanish tokens (547 types) as well as 58,932 Portuguese tokens (6795 types) (see Hennecke & Baayen 2017 for a quantitative evaluation~~ of the data~~). 200 pairs of N Prep N constructions were extracted from this dataset for each language, a total of 400 stimuli for each language. Furthermore, 200 filler stimuli were extracted from the TenTen Corpora. The 200 pairs of target stimuli N Prep N constructions all varied with respect to the internal prepositional constituent but shared the same first noun (N1) and second noun (N2), as for example in the pair *jouet d'enfant* and *jouet pour enfant* 'children's toy'. All constructions were left-headed, as is common for compounds in Romance languages, and fairly compositional. That is to say, the whole dataset did not contain any idioms. The extracted pairs of N Prep N constructions are designed to reflect the frequencies with which these corresponding distributions occur in our initial dataset and by extension in natural language. Therefore, most pairs included an N Prep N construction with the preposition *de*, which is by far the most frequent preposition in N Prep N constructions in Romance languages and therefore equally frequent in pairs with prepositional variability. Table 1 shows the distribution of prepositional elements in the sets of experimental stimuli for the three languages under investigation:

Spanish		**French**		**Portuguese**	
a	21	*à*	65	*a*	28
bajo	1	-		-	
con	29	*avec*	3	*com*	41
contra	1	-		*contra*	2
de	187	*de*	188	*de*	184
en	73	*en*	87	*em*	50
entre	1	*entre*	5	*entre*	4

1 The questionnaires for each language were based on the LEAP-Q questionnaires from Blumenfeld & Kaushanskaya (2007). Available online at: https://www.bilingualism.northwestern.edu/leapq/

2 https://www.sketchengine.eu

para	52	*par*	8	*para*	60
por	29	*pour*	40	*por*	26
sobre	6	*sur*	4	*sobre*	5
Total	**400**		**400**		**400**

Table 1. Distribution of the prepositions in the experimental stimuli pairs.

For each stimulus, we extracted the frequency of the construction as a whole, the frequencies of each constituent, as well as the frequencies of each pair of constituents from the corpora. Furthermore, all experimental stimuli were coded for their lexical status (lexical or syntactic) on the basis of various delimitation tests (see e.g. Bouvier 2000), with a particular focus on the semantic criterion mentioned above. In cases of doubt, all items that were not clearly identifiable as lexical units were coded as syntactic constructions. Furthermore, all items were coded for the type of internal prepositional variation as well as for the conceptual relation of the nominal constituents. The conceptual relations were created on the basis of relevant literature on semantic relations of nominal constituents in compounds (e.g. Shoben 1991, Gagné & Shoben 1997, Gagné & Spalding 2009, Girju et al. 2005). For the present analysis, 11 conceptual relations were selected and the experimental stimuli were classified according to these relations (for a detailed analysis of the conceptual relations of the dataset see Hennecke 2019). The following table presents these conceptual relations with a paraphrase and an example. The paraphrase was created following the relational categories in Shoben (1991). For ease of exposition, all examples in table 2 were taken from the French dataset.

Conceptual Relation	**Paraphrase**	**Example**
Action	N2 performs N1	*danse de couple* 'couple danse'
Cause	N1 is used for N2	*filtre de poussière* 'dust collector'
Container	N1 is a container for N2	*coffre à outils* 'toolbox'
Feature	N2 is a special feature of N1	*cheval à bascule* 'rocking horse'
		escalier à vis 'spiral stair'
Location	N1 is located in N2	*route de montagne* 'mountain road'
Material	N1 is made out of N2	*cravatte en soie* 'silk tie'
Part	N2 is a part of N1	*collier de perles* 'pearl necklace'

Purpose	N1 is made for N2	*boucle d'oreille* 'earring'
Target_Object	N1 is used by N2 (object)	*housses de chaise* 'chair cover'
Target_Person	N1 is used by N2 (person)	*casque de pompier* 'firefighters helmet'
Topic	N1 is about N2	*cours d'histoire* 'history class'

Table 2. Conceptual relations of the nominal constituents in N Prep N constructions.

Analogously to the experimental stimuli, a set of trigram filler constructions was created. The filler stimuli contained more and less lexicalized and grammaticalized constructions, for instance verbal constructions, such as Sp. *tener en cuenta* 'to consider', prepositional constructions, such as Sp. *en virtud de* 'according to', lexicalized constructions, such as Sp. *antes de ayer* 'day before yesterday', as well as free established constructions, such as Sp. *no pasa nada* 'no problem'.

Datasets for the French, Spanish and Portuguese task were created and coded in the same way.

Design

For each of the three languages, the 400 target N Prep N constructions were divided into two sets of 200 constructions each. Stimuli containing the same prepositional constituent and the same coded semantic relation were equally distributed among the two sets. Two stimuli belonging to the same stimulus pair did not appear in the same set. For each target stimulus, an Interest Area (IA) was determined for each of the three constituents of the construction, in order to enable a subsequent analysis of the constituents' fixation duration and count.

Each participant received one set of 200 target stimuli as well as 200 filler constructions. All experimental stimuli in each set were pseudo-randomized using Experiment Builder[3] and appeared in different orders for each participant. A short practice trial including 10 filler constructions was presented to all participants at the beginning to the task. Hence, each participant received a total of 410 stimuli.

Procedure

All three experiments were conducted using an EyeLink 1000 tracking device[4]. The recording of the eye movements was performed on the right eye only. Participants were seated in a soundproof booth in front of the computer screen and the microphone. Calibration of the pupil

3 https://www.sr-research.com/experiment-builder/

4 https://www.sr-research.com

took place at the beginning of the experiment, with recalibration after the practice trials, after each block of 25 stimuli, and after a short break in the middle of the task, i.e. after 200 trials.

A trial started with the presentation of a fixation cross at the left center of the screen. Once a participant had properly fixated the cross, the stimulus was presented in the center of the screen in white letters on black background. The first letter of the stimulus was located at the same position as the fixation cross in order to ensure, as far as possible, that all participants started reading from the same point. The participants read every stimulus out loud and their speech was recorded. To exit a trial, participants were asked to fixate on a white circle at the right center of the screen after reading the stimulus. The fixation of the white circle triggered the initiation of the next trial.

Predictor variables

For the analysis, different measures of fixation duration and fixation counts were applied. We examined the first fixation duration (*Ffixdur*) and the second fixation duration (*Sfixdur*) of each constituent of the construction as well as the gaze duration of the whole construction (*Totfixdur*) as response variables. Furthermore, fixation counts of each of the constituents (*FcountN1, FcountPrep, FcountN2*) as well as for the fixation count whole construction (*Totfcount*) were subjected to statistical analysis.

All measures were obtained from the Interest Area Report, which was based on the Interest Areas defined for each constituent in the construction. Data points with fixations that lay outside of the assigned Interest Areas were discarded.

Seven frequency variables were included: the trigram frequency of the whole trigram construction (*Wholefreq*), the frequency of each of the three constituents (*N1freq*, *Prepfreq*, *N2freq*) as well as the bigram frequencies of two pairs of constituents (*N1Prepfreq*, *PrepN2freq*, *N1N2freq*). Morphological family sizes of the left and the right constituent were calculated from the frequency counts of *N1Prepfreq* and *PrepN2freq*. All frequency measures were obtained from the TenTen Corpora.

In addition, four factorial predictors were included. The semantic relation of the two nominal constituents (*Semrel*) was coded as a multi-level factor; the different levels are listed in table 2. We modeled this predictor with random intercepts. We also coded the type of preposition (*Preptype*), e.g. *de*, as well as the type of prepositional variation in the respective pair of target stimuli (*Varprep*), e.g. *de-para* as random intercepts, as both variables show many levels. This predictor was also included as a random-effect factor. As a last language-internal variable, we coded the lexical status of the construction, (*Lexstat*), using treatment coding with 'lexical' as reference level and 'syntactic' as contrast.

Furthermore, word length was included as a variable in all analyses. This included two length measures, both counted as number of letters, the length of each individual word and the construction as a whole.

For comparison across languages, we added a factor including the respective language (*Language,* reference level Spanish). Random intercepts were included for both the participant and construction. In what follows, we report models from which predictors that did not contribute noticeably to the model fit were removed.

Results

Statistical analyses of the data were performed using the `R` software for statistical computing (R Core Team 2017). Statistical analyses were only performed for the data on the target stimuli, using generalized additive models as implemented in the `mgcv` package (Wood 2017) as well as quantile regression models as implemented in the `qgam` package (Fasiolo 2017). Speech disfluencies in reading aloud were noted during the experiment and trials including disfluencies in the relevant part of the sentence (N Prep N) where excluded from analysis. Analyses were performed on all first, second and total fixation (gaze) durations of each constituent and the whole construction as well as for the fixation counts of each constituent and the whole construction. Poisson generalized additive models were used for the analysis of the fixation counts. Quantile regression analyses were performed on the fixation duration data as the fixation durations were characterized by a distribution that resisted transformation into a Gaussian distribution. Quantile regression also enabled us to clarify the role of the predictor variables for each quantile of the distribution. For reasons of clarity, general information about fixation points and fixations on the prepositional element is listed separately for each language. Subsequently, the results of the comparative analysis are provided in the section *statistical evaluation*.

Fixation counts

For the Spanish task, a total of 36,278 fixations were recorded. For this dataset, there were 7,287 trials in which the preposition was skipped in reading (20%). The preposition *de* had the highest number of omissions, followed by the preposition *en* and the preposition *para*. Most omissions were recorded for the prepositional pair *de* and *en* and for the prepositional pair *de* and *para.* The percentages of non-fixated prepositions are displayed in table 3.

For the French task, a total of 35,232 fixations were recorded. A total of 7,722 prepositions were not fixated (22%). The preposition *de* was most often skipped, followed by the preposition *en* and *à*. Non-fixations occurred most frequently for the prepositional pairs *de* and *en, de* and *à*, as well as *de* and *pour*.

For the Portuguese task, a total of 32,759 fixations were recorded. In the Portuguese data, the preposition was not fixated in a total of 6,464 trials (19%). The preposition *de* was most frequently skipped, followed by *em*, *a* and *para*. Regarding the prepositional pairs of Portuguese prepositions, the pair *de* and *para were* most frequently skipped, followed by the pairs *de* and *em, de* and *com*, and *de* and *a*.

Spanish	*a*	*bajo*	*con*	*contra*	*de*	*en*	*entre*	*para*	*por*	*sobre*
Omitted fixations	5%	0.1%	5.9%	0.2%	45%	15.6%	0.2%	10.1%	6.1%	1.1%

Portuguese	*a*	-	*com*	*contra*	*de*	*em*	*entre*	*para*	*por*	*sobre*
Omitted Fixations	8.7%		10.4%	0.4%	51.5%	13.9%	0.7%	12.9%	5.9%	1.1%

French	*à*	-	*avec*	-	*de*	*en*	*entre*	*par*	*pour*	*sur*
Omitted Fixations	17%		0.5%		49.3%	21%	1%	1.9%	7.7%	1.2%

Table 3. Frequency of skipped prepositions by language and by preposition type (in percentages).

Table 3 clarifies that all three languages show quite similar tendencies in skipping certain prepositions in reading. For all of the three languages, the most frequent prepositions are skipped most often: The preposition *de* is most commonly skipped in reading, followed by *en/em* and *para* in Spanish and Portuguese, and *en* and *à* in French.

We used the generalized additive mixed model (GAMM) with a Poisson link to model the counts of fixations on the N1, the preposition, and the construction as a whole. In the analyses that follow, predictors and interactions of predictors that were not well supported were removed from the models. Model summaries concern these trimmed models. All analyses, including the analyses of the fixation durations, are part of an explorative inquiry into the structure of the experimental data. We therefore used a Bonferroni-adjusted alpha-level 0.05/10 = 0.005 for deciding whether there might be a signal in the noise.

Table 4 presents the model fitted to the fixation counts for the construction. The (logarithm of the) fixation count varied with the frequency of the construction, as shown in the lower left panel of Figure 1. A thin plate regression spline smooth was required for capturing a modest non-linearity in the frequency effect, which was stronger, with a more negative slope, for the higher construction frequencies. This model also included random intercepts for subjects and items, both of which were well supported statistically.

A. parametric coefficients	Estimate	Std. Error	z-value	Pr(>|z|)
intercept	0.7	0.03	20.2	< 0.0001

B. smooth terms:	edf	Ref.df	Chi.sq	p-value
smooth construction frequency	2.2	2.5	22.4	< 0.0001
random intercepts subject	62.3	63	7018.6	< 0.0001
random intercepts item	539.1	1191	1010.4	< 0.0001

Table 4. Summary of the GLMM with Poisson link fitted to the fixation counts for the construction as a whole.

Table 5 reports the Poisson GAMM fitted to the fixation counts on the N1. An effect of lexical status was just significant at the traditional 0.05 significance level, but falls short of expectations given our Bonferroni-adjusted confidence interval. We therefore refrain from interpreting this effect. Construction frequency received much stronger support, its non-linear effect is shown in the upper left panel of Figure 1. Apparently, this construction frequency effect comes into play only for the higher-frequency constructions. Counts varied consistently across subjects and across items, as indicated by the solid support for by-subject and by-item random intercepts.

A. parametric coefficients	Estimate	Std. Error	t-value	Pr(>|t|)
intercept	3.3	0.1	24.5	< 0.0001
Lexical status: syntactic	0.1	0.04	1.9	0.048

B. smooth terms:	edf	Ref.df	F	p-value
smooth construction freq	2.9	3.2	5.1	0.00151
random intercepts subject	62.01	63.0	108.4	<0.0001
random intercepts item	750.3	11191.0	1.8	<0.0001

Table 5. Model summary for the Poisson GAMM fitted to the fixation counts for N1.

The model for the fixation count on the preposition is presented in Table 6. Construction frequency emerged as linear with a negative slope, but uncertainty about the estimated coefficient is too large to allow any conclusions to be drawn. French constructions elicited fewer fixation counts compared to Spanish constructions. Portuguese constructions, by contrast, elicited slightly, but not significantly, more fixations as compared to Spanish. A non-linear effect of preposition frequency received good support, its partial effect is presented in the upper left panel of Figure 1. The gradient of this effect is reduced for lower preposition frequencies, and tapers off slightly for the highest preposition frequencies. Furthermore, random intercepts for the factor preposition type (specifying which specific preposition was present in the construction) were well supported.

A. parametric coefficients	Estimate	Std.Error	t-value	Pr(>|t|)
intercept	0.1	0.06	2.5	0.0107
language: French	-0.3	0.03	-9.8	<0.0001
language: Portuguese	0.02	0.04	0.6	0.5
construction frequency	-0.01	0.004	-1.8	0.06

B. smooth terms	edf	Ref.df	Chi.sq	p-value
smooth preposition frequency	7.1	7.8	109.8	<0.0001
random intercepts for preposition type	14.3	19.0	145.0	<0.0001

Table 6. Summary of the Poisson GLMM for the fixation counts on the preposition.

The two other predictors that we were interested in, the conceptual relation of the nominal constituents (*Semrel*), and the type of prepositional variation (*Varprep*) did not reveal any significant effects at all.

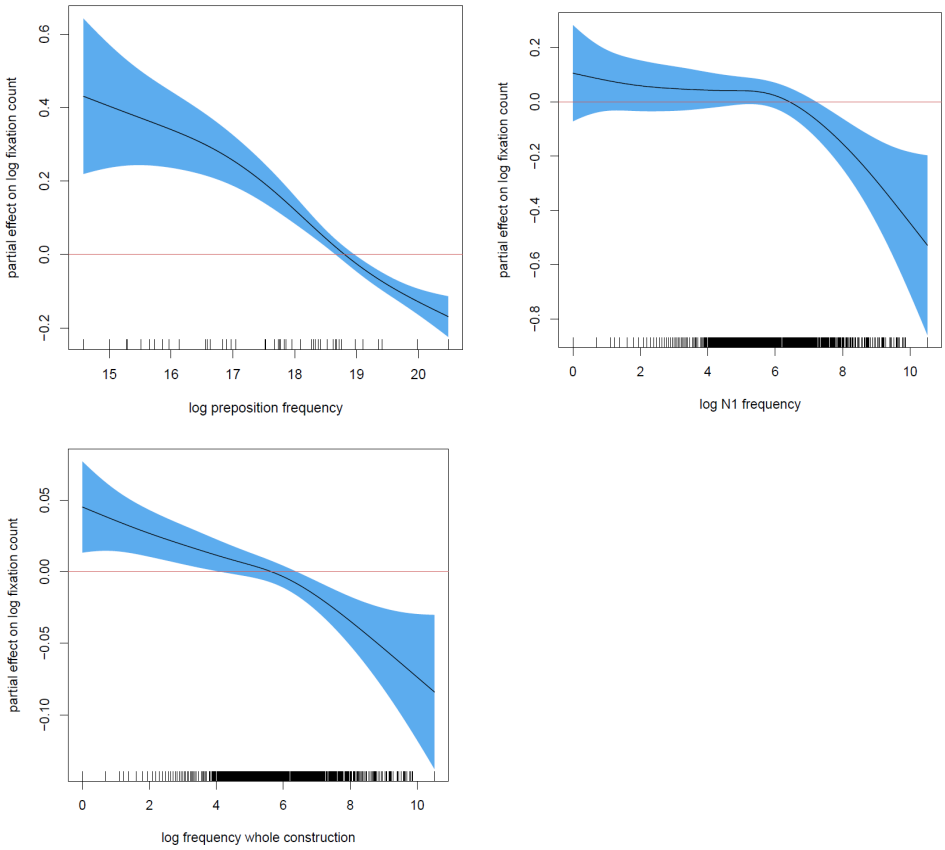

Figure 1. Partial effects of the preposition frequency on the fixation counts of the preposition (upper left), of the N1 frequency on the fixation counts of the N1 (upper right), and on the fixation counts of the whole construction (lower left).

Analysis of fixation durations

We used quantile regression analyses for the fixation durations, using the `qgam` package in `R` (Fasioli 2017). Quantile regression analysis was chosen as a statistical tool of analysis as the fixation durations were characterized by a distribution that, even after a Box-Cox transformation, was not even close to Gaussian. Furthermore, quantile regression analysis offers the advantage of providing insight into whether investigate the effects of predictors are constant or rather variable across the quantiles of the distribution. The five quantiles that we inspected ranged from 0.1 to 0.9 in steps of 0.2. In our analyses, we included random intercepts for subjects, but not for items. With both grouping factors included, the model becomes overspecified. We therefore only included random intercepts for subjects, as subjects are almost always a greater source of variance over which, without subject-specific ancillary variables, we have no control.

We first consider fixation durations on the N1. The QGAMM did not detect any effect of language, no frequency effects, and only a hint that syntactic constructions might have shorter fixation durations at the 5-th and 7-th deciles, as shown in the left panel of Figure 2. Unfortunately, given our Bonferroni-adjusted confidence level (0.005), it is questionable

whether these effects are truly present. The second fixation durations on the N1, by contrast, revealed clear effects of lexical status at the 5-th and 7-th deciles, with syntactic constructions having shorter second fixation durations as compared to lexical constructions. At the last decile (0.9), constructions in French and Spanish were fixated on some 50 ms longer compared to constructions in Portuguese (p =0.009 and 0.018 respectively). Although these differences do not reach our criteria for significance, it is noteworthy that there may be a systematic increase in second fixation durations across the deciles (e.g., for Portuguese, the differences with Spanish start at 6.6, and then change across subsequent deciles from 10.0, 14.7, and 22.0, to 47.3).

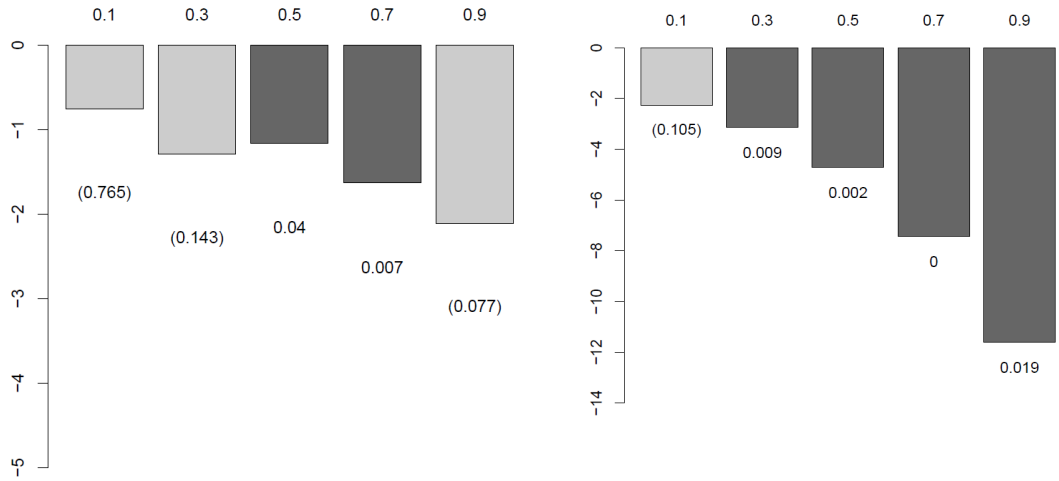

Figure 2. Partial effect of construction type at selected deciles of the distributions of first fixation durations (left) and second fixation durations (right) on the N1. Bars in black highlight where fixation durations to syntactic constructions are significantly shorter (for alpha = 0.05) than those to lexical constructions. As the present analyses are explanatory, it is prudent to use a more stringent significance level (alpha = 0.005). Under this criterion, significant effects are present only at the 0.5 and 0.7 deciles for second fixation durations.

At the deciles 0.1, 0.3, 0.5, and 0.7, we also observed a well-supported effect of N1 frequency on second fixation durations on the N1, with a negative slope that increased slightly across deciles (Figure 3). At the last decile (0.9), the absolute magnitude of the slope was greatest, but its confidence interval is huge, indicating that the N1 frequency effect is submerged in noise in the longest second fixation durations.

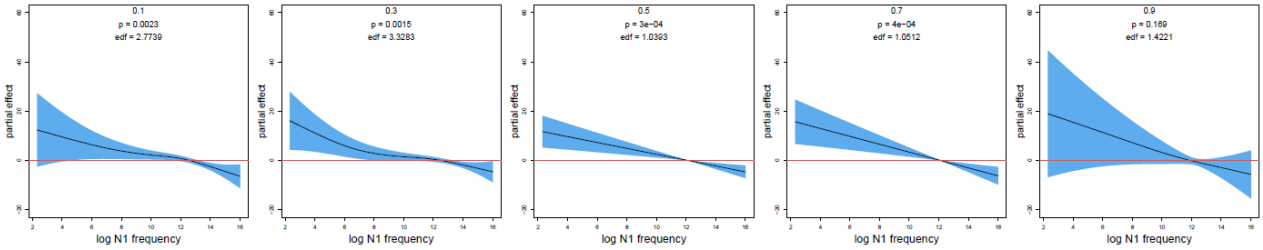

Figure 3. Partial effects in a QGAMM of N1 frequency effect on the second fixation duration of the N1.

A quantile GAM analysis of first fixation durations on the preposition did not reveal robust consistent effects across the distribution for language nor for construction type. Random intercepts for subject were consistently well supported (all p <0.0001). Furthermore, preposition frequency revealed a nonlinear S-shaped effect with an upward swing for the middle frequency range (see Figure 4). This effect almost met our criterion for significance at the 0.7 decile.

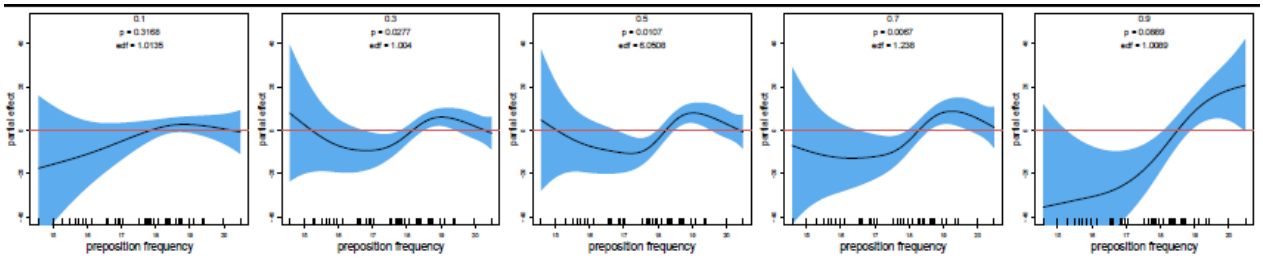

Figure 4. Partial effects in a QGAMM of preposition frequency on first fixation durations on the preposition.

Construction frequency had an effect with opposite sign, that was close to significance according to our Bonferroni-corrected alpha level at the 7th decile (p = 0.0067, see Figure 5). At this decile, the effect of construction frequency was linear with negative slope. Whereas a greater frequency of the preposition tended to induce longer fixation durations, a greater frequency of the construction afforded shorter fixation durations.

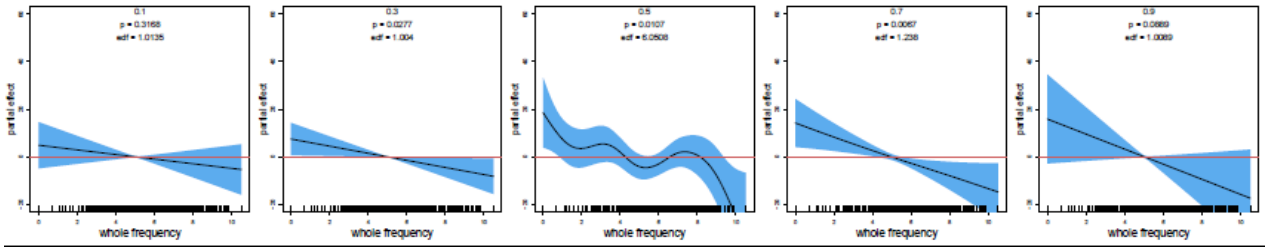

Figure 5. Partial effects in a QGAMM of construction frequency on first fixation durations on the preposition.

Fixation durations on the N2 did not reveal any effects of our lexical predictors.

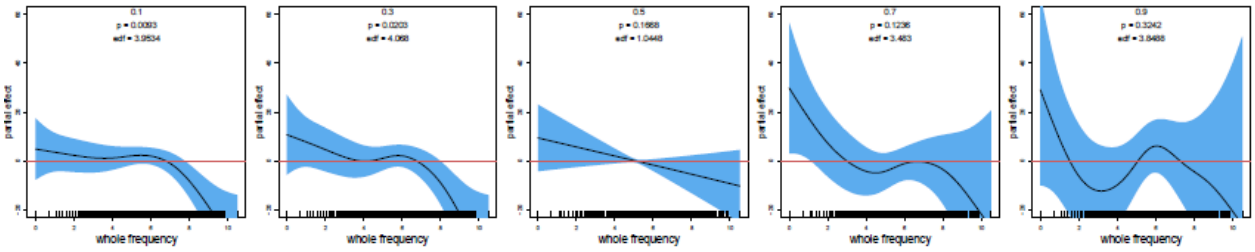

Figure 6. Partial effects in a QGAMM of construction frequency on total fixation durations on the whole construction.

Our final analysis addressed the total fixation duration on the construction as a whole. This analysis revealed a possible effect of construction frequency at the first two deciles (0.1 and 0.3, see Figure 6), with shorter total fixation durations for the highest frequency constructions. However, given our Bonferroni-adjusted alpha level, these effects are only suggestive of a possible trend.

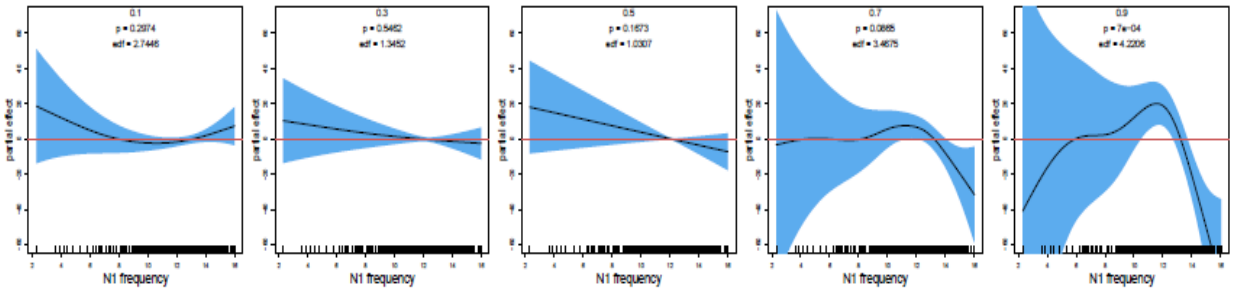

Figure 7. Partial effect in a QGAMM of N1 frequency on deciles of total fixation durations on the whole construction.

N1 frequency emerged with a nonlinear effect at the later deciles (0.7, 0.9), reaching full significance only at our last decile (Figure 7). Here, we find an almost linear downward trend for the higher N1 frequencies. If the head noun is truly frequent, the construction is read more quickly (see Figure 4). None of our other lexical predictors revealed trustworthy effects. However, random intercepts for subjects were well supported. Interestingly, although an effect of lexical status was present in the second fixation on N1, the head noun, lexical status was not reflected in the total fixation durations on the construction. We will return to this finding in the general discussion.

It is noteworthy that construction frequency never interacted with lexical status. Although we tested for effects of the semantic relation between the constituents, this predictor was never supported by our models. The preposition used in the construction (*Preptype*) also was not predictive for any of the fixation durations. Further, the predictor variables of constituent length and word length did not show any significant effect in our models.

General discussion

Tables 7 and 8 summarize the effects of our predictors on fixation counts and fixation durations. The analysis of the fixation counts (Table 6) revealed that prepositions in the French constructions were skipped most often as compared to Portuguese and Spanish. The low skipping rates for Spanish and Portuguese are probably due to French having the smallest

number of different prepositions that can appear in the construction. In all of the three languages, the preposition *de* is skipped most frequently, followed by *en/em* and *para* in Spanish and Portuguese as well as by *en* and *à* in French. The preposition *de*, which is skipped most frequently in all languages, also is the most frequent yet the most polysemous preposition in N Prep N constructions in these languages. It is noteworthy that the preposition that may express the broadest range of semantic relations is skipped most often. Here, frequency may have a larger impact than the range of semantic relations. In Spanish and Portuguese, *en/em* and *para* are also very frequently used in this construction, whereas French forms N Prep N constructions frequently with the preposition *à* (see Hennecke 2019). The statistical analysis (Table 6) provided strong support for an increasing rate of skipping as preposition frequency is increased. Fixation durations on the preposition, however, did not reveal well-supported effects. This fact may also be due to the mostly monosyllabic nature of the prepositional elements. What is clear is that the prepositional element in N Prep N constructions has an effect on the processing of the constructions, and this effect presents itself equally in syntactic and lexical constructions. This argues against the position that prepositions in N Prep N constructions are mere linking elements that are irrelevant and hence would not be fixated in reading.

Predictors / Constituent	N1	Preposition	Construction
Whole construction frequency	fixation count decreases with increasing frequency		fixation count decreases with increasing frequency
N1 frequency			
Preposition frequency		fixation count decreases with increasing frequency	
N2 frequency			
Semantic relation			
Preposition type		random intercepts	
Preposition variation			
Lexical status			
Language		French fewer fixations	

Table 7. Overview of the predictors and their effects on the fixation counts.

Predictors/ Constituents	N1	Prep	Construction
Whole construction frequency		(shorter first fixation duration with increasing frequency, at 0.7 decile)	shorter total fixation duration at early deciles for increasing frequency
N1 frequency	shorter second fixation duration for greater frequency except for 0.9 decile		shorter total fixation duration at late deciles for increasing frequency
Preposition frequency		(longer first fixation durations with increasing frequency at 0.7 decile)	
N2 frequency			
Semantic relation			
Preposition type			
Preposition variation			
Lexical status	shorter durations for syntactic constructions		
Language			

Table 8. Overview of predictors and their effects on the fixation durations.

The other constituent for which we observed clear effects is the first noun. A greater frequency of the noun itself afforded shorter second fixation durations. At the same time, the number of fixations on the first noun decreased with construction frequency. Thus, unlike the preposition, the processing of which appears to be dominated by its own properties (frequency, identity, and language), the processing of the head noun is modulated not only by the head noun's own frequency, but also by the frequency of the construction it is used in.

Total fixation counts and total fixation durations likewise revealed both construction-level and constituent-level effects. The total number of fixations as well as total fixation duration decreased with increasing construction frequency. The head noun showed an effect of its own frequency on total fixation duration at late deciles of its distribution.

An effect of lexical status emerged in the second fixation durations of the first constituent (N1), with significantly shorter fixation durations for syntactic constructions. This effect goes hand in hand with an effect of N1 frequency on the second fixation durations of the N1. It is striking that the constituent effect and the lexical status of the construction only appear for the head constituent (N1). The modifier constituent N2 and the preposition do not exhibit an effect of lexical status. A possible interpretation of the effect of lexical status is that in syntactic constructions the N1 does not carry as much information about the meaning of the construction as is the case for lexical constructions. Since lexical constructions are fixed by definition, they do not suffer from the discriminative uncertainty that comes with syntactic constructions allowing for additional words being inserted before or after the preposition. It is conceivable that participants are quickly fixating further into the construction in order to narrow down the possible interpretations of what they are reading, thus quickly reducing their semantic uncertainty. If correct, this would imply that in lexical constructions, the head N1 would then

already carry more specific information about the semantics of the construction, allowing the reader to dwell longer on the N1 while parafoveally processing the preposition and possibly the second noun.

Interestingly, lexical and syntactic constructions do not differ statistically with respect to the number of fixations and total fixation durations that they require. This suggests that the total costs of reading either type of construction are very similar, but that the point at which costs are predominantly incurred differs, earlier for lexical constructions, and later for syntactic constructions.

Recent research using survival analysis to analyze reaction times is indicating that in lexical processing, whole-word effects show up early in the distribution of reaction times, whereas constituent effects emerge later (Schmidtke et al., 2017). The present data provide partial support for the hypothesis that higher-order units take priority in time over lower-order constituents: Construction frequency is predictive for early deciles of the distribution of total fixation times, whereas N1 frequency is predictive for later deciles of the distribution of total fixation times. However, the effect of N1 frequency on N1 second fixation duration is present already at the early deciles, even though its effect shows up for total fixation durations only at late deciles. Interpretation of this effect is complicated by the finding that the number of fixations on N1 diminishes for increasing construction frequency. In other words, where we are observing an effect of N1 frequency for the second fixation duration on N1, this is likely to be due to the frequency of the construction being low. From this constellation of results, we tentatively draw the conclusion that reading Romance prepositional constructions is similar to reading complex words in English in that whole-form effects appear to precede in time effects related to constituents, provided that the whole form is of sufficient frequency.

How is the effect of construction frequency to be interpreted theoretically? Baayen, Dijkstra, & Schreuder (1997) argued for whole-word access representations for complex words, whereas Baayen, Kuperman and Bertram (2009) proposed to interpret holistic effects as reflecting joint probabilities of constituents. Tremblay & Baayen (2010) speculated that these effects might reflect the experience gained with interpreting constituents. Arnon & Snider (2010) suggest that word sequences may have their own representations (see also Tremblay et al., 2011). Although speakers have intuitions about how frequently multi-word units are used (Shaoul, Westbury, & Baayen, 2009), just as they have intuitions about how often words are used (Gernsbacher, 1984), postulating units above the word level has the disadvantage that the numbers of such units is extremely large (Baayen, Hendrix, & Ramscar, 2013). Possibly, frequency effects above the word level are an emergent property of discrimination learning (Baayen, Milin, et al., 2011). Whatever the proper interpretation of frequency effects above the word level may be, the existence of these frequency effects is consistent with usage-based approaches to language (Bybee, 2010).

Whereas properties of the head noun are predictive for both the processing of the head noun itself and for the processing of the construction, properties of the second noun (the modifier) were never predictive. This result is consistent with the hypothesis that the head noun plays a more important role than the modifier noun (see for instance Juhasz et al. 2003, Libben 2003 for English, Jarema et al. 1999 for French and Marelli et al. 2009 for Italian). However, since the N1 is also the first constituent read by the participants, headedness and order are

confounded, and further experimental evidence is required to further support this hypothesis (but see Goral, Libben, Baayen, 2019, for cross-linguistic evidence that headedness and constituent order have distinct effects).

The present study did not reveal any effects of semantic relation. The semantic relation between head and modifier did not predict any of the experimental measures which we used in this study to gauge lexical processing. This result indicates that the different types of semantic relations do not trigger differences in semantic processing. It might be argued that the absence of an effect may be due to the fact that not primarily the type of the relation causes a significant processing difference, but rather other factors, such as the availability and the competition of the relation (see Gagné, Shoben 1997: 72-73).

Furthermore, it is possible that the type of the task and the structure of the construction influence the processing of the semantic relation. It may be argued that a reading task does not require lexical access and that results would be different for lexical decision tasks. For Romance languages, it is more like that part of the relation is already spelled out in the compound. Tthe prepositional element in Romance N Prep N constructions already makes explicit, to some extent, the semantic relation between N1 and N2. The effect of preposition type (i.e., which specific preposition is used) seen in the fixation counts on the preposition may therefore already indicate a relational competition effect in disguise.

Conclusion

This explorative and comparative eye-tracking study aimed at investigating several specific aspects of the processing of N Prep N constructions in French, Spanish and Portuguese. It focused on processing differences that are related to their frequency distribution, their syntactic or lexical status, their semantic relational structure or the type of internal preposition and prepositional variation.

The research question that motivated this study was whether lexical and syntactic prepositional constructions in Romance have a different stauts, as this topic is discussed controversially in research on Romance word formation. The only evidence that emerged in our experiment for a qualitative difference concerned the fixation time on the head noun, which was shorter for syntactic constructions. The total fixation time on the construction as a whole, as well as the total number of fixations, were statistically indistinguishable for the two types of construction. Although there is no difference in the overall difficulty of reading lexical versus syntactic constructions, the head noun appears to be the locus of extended processing for lexical constructions. We think that lexical constructions, due to the absence of syntactic variation following the head noun, may allow for more parafoveal processing of the preposition and possibly the modifier noun as well. Although both lexical and syntactic constructions are onomasiological units, and although for both constructions we observed construction frequency effects, frequency effects were also well attested for the head noun and the preposition. This suggests that the prepositional constructions of Romance show the same integration of knowledge of lower-level and higher-level units that is also observed for English (Tremblay et al., 2011) and that also characterizes idiom processing (Geeraert & Baayen, 2017).

The present study makes a contribution to the discussion on N Prep N constructions in Romance languages. The results suggest some evidence for a difference in processing: However, the evidence is subtle and in need of replication. Still, future investigations are needed to back the results of this first explorative study. It is of particular interest in further studies to explore the role of the head constituent and the semantic role of the prepositional element in lexical N Prep N constructions.

References

Arnon, Inbal & Neal Snider. 2010. More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language* 62. 67-82.

Bates, E., & MacWhinney, B. 1989. Functionalism and the competition model. *The crosslinguistic study of sentence processing*, *3*, 73-112.

Baayen, R. H., Chuang, Y. Y., Shafaei-Bajestan, E., & Blevins, J. P. 2019. The discriminative lexicon: A unified computational model for the lexicon and lexical processing in comprehension and production grounded not in (de) composition but in linear discriminative learning. *Complexity*. https://www.hindawi.com/journals/complexity/2019/4895891/

Baayen, R.H., Dijkstra, T., and Schreuder, R. 1997. Singulars and plurals in Dutch: Evidence for a parallel dual route model. *Journal of Memory and Language*, 37, 94-117.

Baayen, R. H., Kuperman, V., & Bertram, R. 2010. Frequency effects in compound processing. *Compounding, Amsterdam/Philadelphia: Benjamins*, 257-270.

Baayen, R. H., Hendrix, P., & Ramscar, M. 2013. Sidestepping the combinatorial explosion: An explanation on n-gram frequency effects based on naive discriminative learning. *Language and Speech* 56, 329-347.

Baayen, R. H., Milin, P., & Ramscar, M. 2016. Frequency in lexical processing. *Aphasiology*, *30*(11), 1174-1220.

Baayen, R. H., Milin, P., Đurđević, D. F., Hendrix, P., & Marelli, M. 2011. An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological review*, *118*(3), 438.

Bisetto, A., & Scalise, S. 2005. The classification of compounds. *Lingue e linguaggio*, *4*(2), 319-0.

Booij, Gert. 2010. Compound constructions: Schemas or analogy? A construction morphologist perspective. In Sergio Scalise & Irene Vogel (eds.), *Cross-disciplinary issues in compounding* (Current Issues in Linguistic Theory 311), 93-109. Amsterdam / Philadelphia: John Benjamins.

Booij, Gert. 2018. The construction of words: Introduction and Overview. In Gert Booij (ed.), *The construction of words: Advances in construction morphology* (Studies in Morphology 4), 3-18. Cham: Springer.

Buenafuentes de la Mata, Cristina. 2009. La formación de palabras compuestas: del latín al español. In *Diachronic linguistics* (pp. 213-238). Documenta Universitaria.

Buenafuentes de la Mata, Cristina. 2010. *La composición sintagmática en español.* San Millán de la Cogolla: Cilengua.

Bybee, J. 2010. *Language, usage and cognition*. Cambridge University Press.

Chomsky, N. 1965. *Aspects of the theory of syntax*. MIT Press.

Cutler, A .1983. Lexical complexity and sentence processing. In Flores d'Arcais & Jarvella (eds). *The Process of Language Understanding.* Chichester: Wiley.

Ellis, N. C. 2002. Frequency Effects in Language Processing. In Studies in Second Language Acquisition 24 (2). 143-188.

Fasiolo, M., Goude, Y., Nedellec, R., & Wood, S. N. 2017. Fast calibrated additive quantile regression. URL: https://arxiv.org/abs/1707.03307.

Feldman, L. B., O'Connor, P. A., & del Prado Martín, F. M. 2009. Early morphological processing is morphosemantic and not simply morpho-orthographic: A violation of form-then-meaning accounts of word recognition. *Psychonomic bulletin & review*, *16*(4), 684-691.

Fradin, Bernhard. 2009. Romance: French. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford handbook of compounding*, 417-435. Oxford: University Press.

Frost, Ram, Kenneth Forster & Avital Deutsch. 1997. What can we learn from the morphology of Hebrew? A masked-priming investigation of morphological representation. *Journal of Experimental Psychology: Learning, Memory & Cognition* 23(4). 829-856.

Gagné, Christina L. & Edward J. Shoben. 1997. Influence of thematic relations on the comprehension of modifier-noun combinations. *Journal of Experimental Psychology: Learning, Memory & Cognition* 23(1). 71-87.

Gagné, Christina L. & Thomas L. Spalding. 2009. Constituent integration during the processing of compound words: Does it involve the use of relational structures? *Journal of Memory and Language* 60(1). 20-35.

Gagné, Christina L. & Thomas L. Spalding. 2013. Conceptual composition: The role of relational competition in the comprehension of modifier-noun phrases and noun-noun compounds. In Brian H. Ross (ed.), *The psychology of learning and motivation* (Volume 59), 97-130. Amsterdam: Academic Press.

Gagné, Christina L. & Thomas L. Spalding. 2014. Conceptual composition: The role of relational competition in the comprehension of modifier-noun phrases and noun-noun compounds. *Psychology of Learning and Motivation* 59. 97-130.

Gagné, Christina L. & Thomas L. Spalding. 2014. Relation diversity and ease of processing for opaque and transparent English compounds. In Rainer, F.; Dressler, W.; Gardini, F. and Luschützky, H.C. (Eds.). *Morphology and Meaning: Current Issues in Linguistic Theory*. Amsterdam: John Benjamins, 153-162.

Gahl, S., Garnsey, S.M. 2004. Knowledge of grammar, knowledge of usage: syntactic probabilities affect pronunciation variation. In Language 80(4). 748-774.

Geeraert, K., Newman, J., and Baayen, R. H. (2017). Idiom variation: Experimental data and a blueprint of a computational model. In Christiansen, M., and Arnon, I. (Eds.) *More than Words: The Role of Multiword Sequences in Language Learning and Use*. Special issue of *Topics in Cognitive Science*, 9, 653-669.

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of experimental psychology: General*, *113*(2), 256.

Giraudo, Hélène S. & Jonathan Grainger. 2001. Priming complex words: Evidence for supralexical representation of morphology. *Psychonomic Bulletin & Review* 8(1). 127-131.

Girju, R., Moldovan, D., Tatu, M., & Antohe, D. 2005. On the semantics of noun compounds. *Computer speech & language*, *19*(4), 479-496.

Guevara, Emiliano R. 2012. Spanish compounds. *Probus. International Journal of Latin and Romance Linguistics* 24(1). 175-195.

Hennecke, Inga. 2019. Internal constituent variability and semantic transparency in N Prep N constructions in Romance languages. In Sabine Schulte im Walde & Eva Smolka (eds.), The role of constituents in multiword expressions: An interdisciplinary, cross-lingual perspective, 131–156. Berlin: Language Science Press

Hennecke, Inga & Harald Baayen. 2017. A Quantitative Survey of N Prep N Constructions in Romance Languages and Prepositional Variability. *Quaderns de Filologia: Estudis Lingüístics* 22. 129-146.

Jarema, G., Busson, C., Nikolova, R., Tsapkini, K., & Libben, G. 1999. Processing compounds: A cross-linguistic study. *Brain and Language*, *68*(1-2), 362-369.

Juhasz, B. J., Starr, M. S., Inhoff, A. W., & Placke, L. 2003. The effects of morphology on the processing of compound words: Evidence from naming, lexical decisions and eye fixations. *British Journal of Psychology*, *94*(2), 223-244.

Kornfeld, L. M. 2009. IE, romance: Spanish. In *The Oxford handbook of compounding*.

Kuperman, V., Bertram, R. & Baayen, R.H. (2008) Morphological dynamics in compound processing, *Language and Cognitive Processes*, 23:7-8, 1089-1132

Kuperman, V., Schreuder, R., Bertram, R., & Baayen, R. H. 2009. Reading polymorphemic Dutch compounds: Toward a multiple route model of lexical processing. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(3), 876.

Lau, E., Rozanova, K., Phillips, C. 2007. Syntactic prediction and lexical frequency effects in sentence processing. University of Maryland Working Papers.

Libben, G., Gibson, M., Yoon, Y. B., & Sandra, D. 2003. Compound fracture: The role of semantic transparency and morphological headedness. *Brain and language*, *84*(1), 50-64.

Libben, G., Goral, M., & Baayen, R. H. 2018. What does constituent priming mean in the investigation of compound processing? *The Mental Lexicon*, *13*(2), 269-284

Maguire, Phil & Arthur Cater. 2004. Is conceptual combination influenced by word order? *42nd Annual Meeting of the Association for Computational Linguistics*. http://www.cs.nuim.ie/~pmaguire/publications/French2004.pdf (19 Nov, 2019).

Marelli, Marco, Davide Crepaldi & Claudio Luzzatti. 2009. Head position and the mental representation of nominal compounds. A constituent priming study in Italian. In: *The Mental Lexicon* 4(3). 430-454.

Masini, F. 2009. Phrasal lexemes, compounds and phrases: A constructionist perspective. *Word structure*, *2*(2), 254-271.

Pierrehumbert, J. B. 2001. Exemplar dynamics: Word frequency. *Frequency and the emergence of linguistic structure*, *45*, 137.

Pinker, S. 1991. Rules of language. *Science*, *253*(5019), 530-535.

Pinker, S., & Ullman, M. T. 2002. The past and future of the past tense. *Trends in cognitive sciences*, *6*(11), 456-463.

Pylkkänen, L., Feintuch, S., Hopkins, E., & Marantz, A. 2004. Neural correlates of the effects of morphological family frequency and family size: an MEG study. *Cognition*, *91*(3), B35-B45.

Rainer, Franz (2016): Italian. In: *O. Peter Müller, Ingeborg Ohnheiser, Susan Olsen und Franz Rainer (Hg.): Word Formation. An International Handbook of the Languages of Europe, Bd. 4.* Berlin, Boston: Mouton de Gruyter, S. 2712–2731.

R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

Rastle, K., Davis, M. H., & New, B. 2004. The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin & Review*, *11*(6), 1090-1098.

Rio-Torto, G., & Ribeiro, S. 2009. Compounds in portuguese. *Lingue e linguaggio*, *8*(2), 271-292.

Rio-Torto, G., & Ribeiro, S. 2010. Unidades pluriverbais - Processamento e ensino. In: As interfaces da gramática, 151-165, Araraquara, SP: UNESP Editora

Rio-Torto, G., & Ribeiro, S. 2012. Portuguese compounds. *Probus*, *24*(1), 119-145.

Rummelhart, D. E., McClelland, J. L., & PDP Research Group. 1986. Parallel distributed processing.

Shaoul, C., Westbury, C. F., & Baayen, H. R. 2013. The subjective frequency of word n-grams. *Psihologija*, *46*(4), 497-537.

Schmidtke, D., Kuperman, V., Gagné, C. L., & Spalding, T. L. 2016. Competition between conceptual relations affects compound recognition: the role of entropy. *Psychonomic bulletin & review*, *23*(2), 556-570.

Schmidtke, D., Matsuki, K., & Kuperman, V. 2017. Surviving blind decomposition: A distributional analysis of the time-course of complex word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(11), 1793-1820.

Shoben, E. J. 1991. Predicating and nonpredicating combinations. In P.J. Schwanenflugal(ed.), *The psychology of word meanings,* 117-135.Hillsdale, NJ: Erlbaum.

Snider, Neal & Inbal Arnon. 2012. A unified lexicon and grammar? Compositional and non-compositional phrases in the lexicon. In Dagmar Divjak & Stefan Th. Gries (eds.), *Frequency effects in language representation*, 127-164. Berlin & Boston: De Gruyter Mouton.

Solomyak, O., & Marantz, A. 2010. Evidence for early morphological decomposition in visual word recognition. *Journal of Cognitive Neuroscience*, *22*(9), 2042-2057.

Štekauer, P. 2005. “Onomasiological Approach to Word-formation.” In: Pavel Štekauer and Rochelle Lieber (eds.), Handbook of Word-formation. Dordrecht: Springer.

Spalding, T. L., Gagné, C. L., Mullaly, A., & Ji, H. 2010. Relation-based interpretation of noun-noun phrases: A new theoretical approach. *New impulses in word-formation*, 283-315.

Spalding, Thomas L. & Christina L. Gagné. 2011. Relation priming in established compounds: Facilitation? *Memory & Cognition* 39(8). 1472-1486.

Spalding, Thomas L. & Gagné, Christina L. 2014. Relational diversity affects ease of processing even for opaque English compounds. The Mental Lexicon, 9, 48-66.

Taft, M., & Forster, K. I. 1976. Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of verbal learning and verbal behavior*, 15(6), 607-620.

Tremblay, A., and Baayen, R. H. (2010). Holistic processing of regular four-word sequences: A behavioral and ERP study of the effects of structure, frequency, and probability on immediate free recall. In D. Wood, *Perspectives on formulaic language: Acquisition and communication*. London: The Continuum International Publishing Group, 151-173.

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

Ullman, M. T. 2001. The declarative/procedural model of lexicon and grammar. *Journal of psycholinguistic research*, *30*(1), 37-69.

Ullman, M. T. 2004. Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, *92*(1-2), 231-270.

Val Àlvaro, J.F. 1999. La composición. In I. Bosque & V. Demonte (Eds.), Gramática descriptiva de la lengua española. Madrid: Espasa, 757-4841.

Villoing, F. 2012. French compounds. *Probus*, *24*(1), 29-60.

Wood. 2017. Wood S (2017). Generalized Additive Models: An Introduction with R, 2nd edition. Chapman and Hall/CRC.

Zwitserlood, Pienie. 2018. Processing and representation of morphological complexity in native language comprehension and production. In Geert Booij (ed.), *The construction of words: Advances in construction morphology* (Studies in Morphology 4), 583-602. Cham: Springer.