Chapter 5

Internal constituent variability and semantic transparency in N Prep N constructions in Romance languages

Inga Hennecke

University of Tübingen

Constructions of the type N Prep N represent one of the most controversial issues in Romance word formation. In particular, their lexical status and their degree of productivity are still crucial points of discussion. Hence, it remains unclear whether these constructions fall within the category of morphological word formation or of syntax. Furthermore, the possibilities for internal prepositional variation remain uncertain. This article takes a constructionist approach within the framework of construction morphology in order to describe the internal constituent variability and transparency of the prepositional element in N Prep N constructions in Spanish, Portuguese, and French, as in Sp. *juego de niños*, *juego para niños* ('kid's game') or in Sp. *cabaña de árbol* and *cabaña en árbol* ('tree house'). A qualitative analysis of large-scale corpus data from the TenTen corpus family indicates that Romance N Prep N constructions may undergo internal prepositional variation. The analysis focuses on the semantic relations of the internal nominal constituents and the semantic transparency of the constructions in the three Romance languages under investigation. The results indicate that semantic relations and semantic transparency play a role in the internal constituent variability of the prepositional element.

1 Introduction

Compounds of the type N Prep N, such as Sp. *bicicleta de montaña* 'mountain bike', Fr. *salle de bain* ('bath room'), or Pt. *história em quadrinhos* ('comic strip'), are generally considered to be the most problematic aspect of research on compounding and word formation in Romance languages. This is because these constructions represent nominal lexical units that clearly approach free syntactic

structures (de Bustos Gisbert 1986). Compounds of the type N Prep N have been treated very differently in research on compounding and have also been labeled with many different terms, such as syntagmatic compounds (Buenafuentes de la Mata 2010), syntactic compounds (Rio-Torto & Ribeiro 2009), improper compounds (Kornfeld 2009), phrasal lexemes (Masini & Thornton 2007), frozen multiword units (Guevara 2012), lexicalized syntactic constructions (Villoing 2012), lexicalized phrases (Fradin 2009), and syntactic words (Di Sciullo & Williams 1987). Generally, compounding is a mechanism whereby two lexical units are combined. Compounds of the type N Prep N are characterized as lexical units that consist of (at least) two lexical elements that are not orthographically combined. As a result, compounds of the type N Prep N, such as Sp. *traje de baño* ('bathing suit'), do not differ on a formal level from syntactic phrases of the type N Prep N, such as Sp. *libro para niños* ('book for children') (de Bustos Gisbert 1986: 69).

The most problematic issue in current research on compounding of the type N Prep N is the question of the delimitation of syntactic and lexical structures in Romance languages. As the treatment of these constructions is based largely on the theoretical background of the individual author, there is no general agreement on whether or not N Prep N constructions should be included in the class of compounds. Related to this issue is the question of whether these constructions emerge by means of productive word formation processes or are merely "fossilized" or lexicalized syntactic structures. These two crucial issues will be discussed and analyzed in this study, with a focus on one particular case of internal constituent variability, the alternation of the internal preposition in N Prep N constructions. A large-scale corpus analysis of this alternation in French, Spanish, and Portuguese supports the adoption of a constructionist approach within a framework of construction morphology. Such an approach allows the internal constituent variation of N Prep N constructions to be represented without recourse to traditional notions of lexicon and syntax.

2 Definition and classification of syntagmatic compounds

As mentioned above, constructions of the type N Prep N are often excluded from descriptions of Romance compounding. Typically, they are classified together with other compound-like constructions lacking an orthographical union, as in the examples in Table 1.

According to Masini, these examples are separated orthographically, show no strong degree of idiomaticity, and appear quite frequently in each of the four

Table 1: Phrasal lexemes in Romance languages (Masini 2009: 257)

Language	Types	Phrasal lexems	Lit.	Glosses
French	$[ADJ\ N]_N$	*premier violon*	first violin	'first violin'
Italian	$[N\ da\ N]_N$	*camera da letto*	room from bed	'bedroom'
Portuguese	$[N\ de\ N]_N$	*cadeira de rodas*	chair of wheels	'wheelchair'
Spanish	$[N\ ADJ]_N$	*luna nueva*	moon new	'new moon'

languages. The question nevertheless remains whether these constructions form part of the class of compounds.

According to Guevara (2012), Spanish syntagmatic compounds, such as *fin de semana* ('weekend') or *sabelotodo* ('know-it-all'), should be excluded from the class of Spanish compounds, as these units are clearly syntactic units that contain "certain effects of lexicalization and atomicity in their distribution" (Guevara 2012: 180). In the same way, Villoing (2012) excludes French constructions such as *fil de fer* ('iron wire') and *brosse à dents* ('tooth brush') from her description of French compounds, as they are "lexicalized syntactic constructions that behave like lexical units" (Villoing 2012: 35). The approach taken by Guevara and Villoing indicates, on the one hand, that constructions of the type N Prep N are often considered as syntactic units that lie outside of the core of word formation processes. For this reason, they are regularly neglected in research papers on Romance word formation. On the other hand, this approach shows that N Prep N constructions are frequently interpreted as lexicalized syntactic constructions and, more precisely, as syntactic constructions that have somehow attained a high degree of fixedness. If this is the case, they should also be excluded as belonging to the class of Romance-language compounds, as lexicalization cannot be considered a morphological word formation process.

There is an opposing perspective according to which the constructions mentioned in Table 1 constitute a productive type of word formation and clearly follow productive morphosyntactic rules. According to Rainer, constructions of the type N Prep N are "very productive lexical patterns, which normally continue to obey the rules of [...] syntax (for example, agreement rules), but may occasionally also deviate from them" (Rainer 2016: 2724). This perspective is not new and was already adopted by Benveniste (1974) in his work on French compounds of the type *robe de chambre* ('robe') and *plat à barbe* ('shaving bowl'), for which he claims indefinite productivity (Benveniste 1974: 172). In the course of the present paper, I will provide new empirical evidence in favor of this perspective using

large-scale corpus data. The analysis will show that N Prep N constructions in Romance languages are highly frequent and productive and that their internal variability follows clear morphological rules that can be mapped using construction morphology.

In order to distinguish N Prep N constructions from other phrase-like constructions, their characteristics must be clearly delineated. According to Buenafuentes de la Mata (2010), a syntagmatic compound may be defined as a lexical element that has been created by the fixation of a syntagm, which keeps its sentential structure, and therefore shows neither orthographic nor accentual union (Buenafuentes de la Mata 2010: 21ff.). De Bustos Gisbert (1986) states that Spanish N Prep N compounds differ from syntactic units on the syntactic level in two respects. First, they have a fixed word order, for example, *ojo de buey* ('porthole') cannot be reordered as *buey ojo de. Second, there is generally no unproblematic substitution of their constituents; for example *ojo de vaca ('eye of cow') (Val Àlvaro 1999: 4825). On a morphological level, he adds that N Prep N constructions show the same characteristics as other compounds in terms of gender and number agreement, the presence of composition markers, and the ability to undergo further derivation and to form collocations (de Bustos Gisbert 1986: 77). According to Masini, N Prep N constructions are of major interest, as they follow the syntactic rules of head modification of a nominal phrase by a prepositional phrase. This means that in Romance languages, N Prep N constructions are generally left-headed, and that inflectional processes are performed on the head of the construction (Masini 2009: 257). Val Àlvaro (1999: 4827) adds a fundamental characteristic on the semantic level: the absence of compositional meaning that may lead to syntactic reinterpretation of the complex nouns. This means that syntagmatic compounds, in contrast to syntactic units, represent one single naming unit at the semantic level; that is, they refer to one specific conceptual representation, as in Fr. *sac à main* ('purse').

In this paper, I will focus on the syntactic criteria given by de Bustos Gisbert, specifically, on the impossibility of constituent substitution. This criterion does not appear to be suitable for purposes of differentiating syntactic and lexical elements, as the delimitation between syntactic and lexical N Prep N constructions remains a matter of controversy. Here, I will show that variation of the internal preposition can be best explained within a constructional framework. I will then argue with regard to the internal preposition that not only is the substitution of internal constituents possible in N Prep N constructions, but it is also a rule-governed process and depends largely on semantic factors, particularly the semantic relation of the nominal constituents.

When investigating the semantic relations of constituents of N Prep N constructions, it is crucial to consider the notions of semantic transparency and semantic opacity. In current research, the term semantic transparency refers to the degree to which the meaning of a complex construction can be derived from the meaning of its constituents (Zwitserlood 1994). For example, the French N Prep N construction *salle de bains* ('bathroom') is considered semantically transparent, whereas the Spanish construction *ojo de buey* ('porthole', lit. 'bull's eye') is considered semantically opaque. Bell and Schäfer view semantic transparency and semantic opacity as scalar notions, lying at either end of a continuum (see Bell & Schäfer (2016) for a detailed discussion on semantic transparency). Later in the present study, I will discuss whether the semantic transparency of an N Prep N construction determines the possibility of internal constituent variation.

3 Internal constituent variation in N Prep N constructions: The role of the preposition

Characteristic of N Prep N constructions, and a crucial factor in their delimitation, is their resistance to paradigmatic variation. In the context of the delimitation of nominal compounds and noun phrases of the type N Prep N in Portuguese, Rio-Torto & Ribeiro (2012: 9) state that the "(im)possibility of lexical insertion" is one of the most important tests of compoundhood. They go further, claiming that if internal changing is allowed, "we are no longer dealing with compounds ([N[PrepN]]$_N$) but with noun phrases ([N[PrepN]]$_{NP}$)" (ibid.). When speaking of internal change, Rio-Torto and Ribeiro refer principally to changes in determination, as in Pt. *fim de semana* ('weekend') and Pt. *fim da semana* ('end of this week'), and to changes effected through insertion of lexical material, as in *fim da última semana* ('end of last week'). As these examples suggest, internal constituent variation is generally seen as a crucial test of delimitation between compounds and syntactic structures. Similarly, Masini argues that, for lexical elements, "paradigmatic variation is blocked, since the words in the construction cannot be substituted by a near-synonym, which should not be a problem for normal phrases" (Masini 2009: 259). Masini defines paradigmatic blocking as the inability to replace a constituent of the construction by another paradigmatically fitting constituent. She also refers to cases of paradigmatic blocking of a nominal unit of a N Prep N construction, as the Italian examples *casa di cura* ('nursing home') and **abitazione di cura* ('*nursing domicile'). In this case, *casa di cura* is a fixed naming unit that loses its semantic meaning when there is paradigmatic variation of a nominal element. The following analysis will show

that paradigmatic blocking holds particularly true for N Prep N constructions with a stronger degree of semantic opacity and idiomaticity. More transparent N Prep N constructions allow productive and rule-governed internal constituent alternation, as the analysis will show by means of the prepositional constituent.

In the literature, all references to a delimitation test of constituent variability neglect the prepositional constituent in N Prep N constructions. The prepositional element is fundamental in N Prep N constructions, but its status is far from clear. In all the Romance languages under investigation here, the preposition *de* is the most frequently used prepositional constituent in N Prep N constructions. In the case of Spanish, Buenafuentes de la Mata (2010) cites various examples of N Prep N constructions with prepositions other than *de*, such as *leche en polvo* ('milk powder'), *cita a ciegas* ('blind date'), *caridad con uñas* ('self-serving favor'), *pozo sin fondo* ('bottomless pit'), and *caballo con arcos* ('pommel horse'). She adduces the appearance of prepositions other than *de* as evidence for the structural complexity of N Prep N constructions in Spanish. The same case can be made for the other languages under investigation in this paper (i.e. French and Portuguese), which show the same ability to form N Prep N constructions with other prepositions.

This paper concentrates on a specific set of (partially) synonymous prepositions in French, Spanish, and Portuguese, which are Fr. *de* ('of'), *à* ('to'), *en* ('in'), and *pour* 'for'; Sp. *de* ('of'), *a* ('to'), *en* ('in'), and *para* ('for') as well as Pt. *de* ('of'), *a* ('to'), *em* ('in') and *para* ('for'). These prepositions may all appear in N Prep N constructions and they may all undergo internal alternation and variation in the three languages under investigation. Consider examples (1–3) from the TenTen corpora:

(1)   a.   Sp. *fuente de horno – fuente para horno* 'casserole'

       b.   Pt. *água de lavagem – água para lavagem* 'wash water'

       c.   Fr. *livre d'enfant – livre pour enfants* 'children's book'

(2)   a.   Sp. *motores de gasolina – motores a gasolina* 'gas engine'

       b.   Fr. *jauge d'essence – jauge à essence* 'fuel gauge'

       c.   Pt. *fogão de lenha – fogão a lenha* 'wood stove'

(3)   a.   Fr. *chemise de coton – chemise en coton* 'cotton shirt'

       b.   Pt. *bracelete de aço – bracelete em aço* 'steel bracelet'

       c.   Sp. *ciclismo de pisto – ciclismo en pisto* 'track cycling'

Example (1) shows internal variation of the prepositional elements *de* and *pour/para*. While the constructions containing *de* are considered to have a lexical

status, the constructions with *pour/para* are generally considered to be syntactic constructions, as they pass certain of the classification tests mentioned above. In contrast to the construction with *de*, they allow substitution and insertion, as the two tests of compoundhood demonstrate: Sp. *fuentes de vidrio para horno* ('glass casseroles for the oven'), *fuentes profundas para horno* ('deep casseroles for the oven'), but **fuentes de vidrio de horno* and **fuentes profundas de horno*. Example (3b) demonstrates the internal alternation of the prepositions *de* and *en/em*. Here, alternation is possible without changing the semantic context of the whole construction or its degree of semantic transparency. In example (2), the prepositions *de* and *a/à* alternate in N Prep N constructions without changing the lexical status of the respective constructions. Nonetheless, these constructions differ in their frequency of usage, productivity, and fixedness, as well as in their degree of lexicalization and of idiomaticity. Especially in French, alternation of *de* and *à* may indicate a change in meaning, as in *verre de vin* ('glass of wine') and *verre à vin* ('wine glass'). In this case, the interpretation of both constructions as two distinct products of word formation is reasonable (this specific case will be discussed in detail in the course of the corpus analysis). In other cases, such as Pt. *fogão de lenha – fogão a lenha* ('wood stove'), no clear semantic difference is visible, as attested by native speakers of Brazilian and European Portuguese: internal variation is possible inside one construction (a more detailed discussion of the examples will follow in the upcoming section). As mentioned above, authors including Rio-Torto & Ribeiro (2009) interpret constructions of the type shown in examples (1–3) as syntactic units, on the grounds that they do not pass all the delimitation tests for compoundhood. The following theoretical discussion and empirical analysis will show that it is neither necessary nor possible to draw a clear distinction between syntactic constructions and lexical constructions; the possibility of alternating prepositional elements in N Prep N constructions depends largely on the semantic function of the N2, and the fixedness, semantic transparency and the idiomaticity of the whole construction.

Another problem in analyzing alternation of prepositional elements concerns the role of the prepositions. In Romance languages, the prepositions *de* and *à/a* in particular have often been considered as semantically "empty" units that do not contain meaning. This perspective has often been applied to the prepositional element in N Prep N constructions: for example, Bartning (1993: 164) states that prepositions in French N Prep N constructions do not code any specific meaning and that they function only as linking elements. Similarly, Bartning (1993) refers to Cadiot (1997) by describing these prepositional elements as "colorless prepositions" (Bartning 1993: 164). For the French prepositions *de* and *à*, Bosredon & Tamba (1991: 44) use the term of "opérateur de couplage" ('linking operator'). Ca-

diot (1997) sees prepositions in N Prep N constructions as elements that express the operation of a construction or the denomination of a subclass of N1; he associates the prepositional element with a "referential calibration" of N1. In contrast, Laumann (1998: 32) notes the importance of distinguishing between different types of meaning when investigating the function and meaning of prepositions in nominal compounds. He differentiates between system meaning (*Systembe-deutung* – the sum of the meaning patterns of the constituents in the N Prep N construction), word meaning (*Wortbedeutung* – the meaning of the construction on the level of word formation), and lexicon meaning (*Wortschatzbedeutung* - the meaning of the construction as a naming unit in the lexicon).

Other authors interpret the possibility of elision of the prepositional element, as Sp. *ducha de teléfono > ducha teléfono* ('detachable shower head') or Sp. *crédito de vivienda > crédito vivienda* ('home loan'), as evidence of the semantic emptiness of the preposition. However, the elision of the prepositional element is only possible in certain strongly lexicalized constructions. Therefore, a counterargument may be based on the same evidence, given that the elision of the prepositional element is not possible in most cases. In the present paper, I argue that the elision of prepositional elements is not proof of a lack of semantic content. The elision can be explained in terms of common processes of language change that may or may not take place in certain lexicalization processes within complex units. The alternation between prepositional elements, as exemplified above is a productive word-formation process that differs clearly from mere lexicalization processes. Therefore, the following qualitative corpus analysis considers the internal constituent alternation of the prepositional element from a comparative perspective and does not focus on the elision of this element. The analysis adopts a constructionist approach based on Goldberg (1995; 2006), with a special focus on construction morphology as introduced by Booij (2010; 2015).

4 N Prep N constructions in construction grammar and morphology

Since Goldberg's seminal work *Constructions* (1995), the constructional approach has had a strong impact on linguistic research. Constructions are considered as conventionalized form-meaning pairs that can be found at all levels of abstraction in language, are dynamically formed, and may be changed continuously. They are acquired via general processes of abstraction, generalization, and categorization. Goldberg (2006: 5) considers any linguistic unit to be a construction,

if "some aspect of its form or function is not strictly predictable from its component parts". Furthermore, units are considered to be stored as constructions if they can be fully predicted and if they are sufficiently frequent (ibid.).

In his theory of construction morphology, Booij (2015) applied the general notions and concepts of construction grammar to morphological units that have traditionally been regarded as morphological. The underlying assumption of the theory of construction morphology is that a construction may have characteristics that cannot be derived from their constituents (Booij 2015: 3). Booij cites the example of the reduplication of nouns in Spanish in order to express the notion 'real', as in *un café café* ('a real coffee'). He establishes the notion of conceptual schemas and subschemas, defined as schematic representations of morphological constructions. These schemas represent a correlation between form and meaning:

(4) (Booij 2015: 2)
     $<[[x]V_i \text{ er}]N_j \leftrightarrow [\text{Agent of SEM}_i]_j>$

This example indicates that a word with base x, in this case an English infinitive verb form, can transform into a noun with the meaning 'agent of the base word (SEM)' by adding the suffix *-er* (Booij 2015: 2). The variable x denotes the phonological content of the base word, i denotes the meaning of the base word, and j shows that the meaning of the complete construction depends on the form of the complete construction (ibid.). Masini (2009: 261) applies the theory of construction morphology to constructions of the type N Prep N in Italian, taking them to represent an abstract template that is stored in the mental lexicon. Masini further notes that this abstract template features a certain degree of productivity and is associated with a concrete naming function (ibid.). By means of a specific inheritance mechanism, based on instance inheritance links (Goldberg 1995), constructions that are more and more specific can be derived from the abstract template. This may be done by categorical specification (filling an unspecified slot with a specific category), as in N Prep N or N Prep V, by lexical specification (filling a slot with specific lexical material), as in N *de* N or by a completely lexical construction, such as It. *casa di cura* (Masini 2009: 261). Figure 1 demonstrates the application of this theory to French constructions of the type N Prep N.

This figure shows the inheritance hierarchy from the abstract template [N1 *de* N2]$_N$, which here is an intermediate construction of the abstract template [N1 Prep Y] and [N1 Prep N2]. From the level [N1 *de* N2]$_N$, it is possible to proceed to a second intermediate lexical level, which indicates the semantic function of N2,
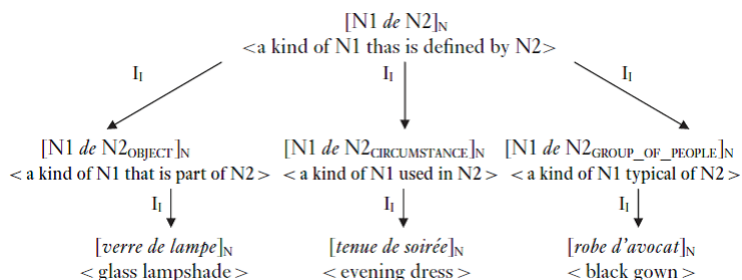
Inga Hennecke

Figure 1: Inheritance hierarchy for N Prep N templates in French (Masini 2009: 263)

and to conclude at a completely lexical level, which shows the lexical result with a concrete naming function. According to Masini (2009: 263), this model can also clarify and describe new occurrences of the N1 *de* N2 construction.

The following qualitative corpus analysis aims to apply the concept of construction morphology presented by Booij (2010; 2015) and exemplified by Masini (2009) to a cross-linguistic comparative analysis of large-scale corpus data for Spanish, French, and Portuguese N Prep N constructions. The focus of the analysis is on constituent variation of the internal prepositional element in N Prep N constructions in these three languages, and I will apply Masini's inheritance hierarchy template (Figure 1) to the internal variability of prepositional constituents. It is useful to include a further intermediate level prior to the first and second levels of the hierarchy for N Prep N templates mentioned above. This additional level contains the abstract template with the semantic function of N2, which in the following corpus analysis is shown to be a crucial factor in determining the possibility of internal prepositional variation. For the purposes of the present analysis, the inheritance hierarchy for N Prep N templates may be visualized as in Figure 2.

This figure shows the inheritance hierarchy adapted from Figure 1 by means of example (3a). As mentioned above, the added abstract intermediate levels are intended to reflect the possibility of that prepositional variability for certain N Prep N constructions and the dependence of this variability on the semantic function of the nominal constituents of the construction. The objectives of the following qualitative corpus analysis are to apply the inheritance hierarchy in Figure 2 to a large-scale corpus of natural speech data for Spanish, French and Portuguese and to compare the internal prepositional variability of N Prep N constructions in these three languages.
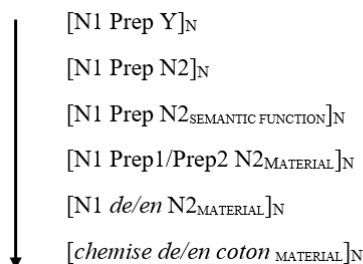
[N1 Prep Y]$_N$

[N1 Prep N2]$_N$

[N1 Prep N2$_{\text{SEMANTIC FUNCTION}}$]$_N$

[N1 Prep1/Prep2 N2$_{\text{MATERIAL}}$]$_N$

[N1 *de/en* N2$_{\text{MATERIAL}}$]$_N$

[*chemise de/en coton* $_{\text{MATERIAL}}$]$_N$

Figure 2: Inheritance hierarchy for internal variation in N Prep N templates in French (adapted from Masini 2009)

5 Qualitative corpus analysis

As mentioned above, the present corpus analysis is intended to investigate internal constituent alternation of the prepositional element in N Prep N constructions in Spanish, French and Portuguese. The focus is on the alternation between *de* and *à/a*, *de* and *en/em*, and *de* and *pour/para*. This study builds on a quantitative corpus survey on the internal alternation of the prepositional element in N Prep N constructions in Spanish, French, and Portuguese by means of large-scale corpus data (Hennecke & Baayen 2017). Hennecke & Baayen (2017: 144) showed that internal prepositional variation in the three languages under investigation is possible, but that these languages show different characteristics in terms of frequency and productivity of such alternations. The quantitative analysis of the three languages focused on frequency of types and token, productivity (i.e. probability of previously unobserved types), and population size (i.e. potential number of formations) (Hennecke & Baayen 2017: 139). The results show that Portuguese and, to a lesser extent, French, allows productive internal constituent variation of the prepositional element. In contrast, Spanish does not show productivity in internal variation, which is demonstrated by the absence of hapax legomena (ibid.). At the same time, Spanish has the greatest tendency to employ the preposition *de* in N Prep N constructions. In French, the prepositions *à* and *pour* are slightly more productive than in the other two languages. Moreover, French tends to avoid constructions using *avec*, whereas constructions with *com* are productive in Portuguese. The latter tendency may be explained by the fact that French prefers NA-constructions over constructions of the type N *avec* N.

The aim of the present corpus analysis is to investigate the results from the above-mentioned study from a qualitative perspective. In this qualitative survey, the internal prepositional variability will be investigated from a mostly semantic perspective, combined with a constructionist approach. Here, the focus will be on

which nominal semantic functions allow prepositional variability and whether the variability depends on the semantic transparency of the construction. To that end, this corpus analysis is based on the same dataset as in Hennecke & Baayen (2017), namely three web corpora from the TenTen corpus family from Sketchengine: the French corpus frTenTen12, the Spanish corpus esTenTen11 and the Portuguese corpus ptTenTen11. The TenTen corpora are large-scale web corpora with the counts displayed in Table 2.

Table 2: Corpus Information of the TenTen corpora for Spanish, French and Portuguese (https://the.sketchengine.co.uk)

	frTenTen12	esTenTen11	ptTenTen11
Tokens	11,444,973,582	10,994,616,207	4,626,584,246
Words	9,889,689,889	9,497,402,122	3,900,501,097
Sentences	456,065,104	407,205,587	190,221,913
Paragraphs	188,079,362	213,364,685	91,248,976
Documents	20,400,411	22,287,566	10,216,060

In order to perform a qualitative analysis of the data, all N Prep N constructions were extracted automatically from the corpora, keeping only those that appear with more than one internal prepositional element. The present analysis focuses exclusively on N Prep N constructions and therefore excludes constructions of the type N Prep Det N. The data were manually inspected by excluding grammaticalized constructions (for example Fr. *face à N*, Sp. *gracias a N* 'thanks to N'), binominal pairs (e.g. Fr. *temps en temps* 'time to time', Sp. *dia a dia* 'day to day'), and antonyms (Fr. *chien avec/sans laisse* 'dog with/without leash'). Table 3 demonstrates the underlying dataset for the qualitative analysis.

Table 3: Type and token counts for the underlying dataset with all pairs of nouns that are attested with at least two different internal prepositions

	Types	Tokens
French	1062	6991
Spanish	547	10219
Portuguese	6795	58932

This dataset shows important differences in type-token frequency between the three languages (Hennecke & Baayen 2017). Portuguese presents by far the

greatest number of different types and tokens of N Prep N constructions with more than one internal preposition. In contrast, Spanish has very few different types but a considerable number of tokens. This can be interpreted as a small number of different N Prep N constructions, but these few types appear quite often in the corpus data. The French data show a significantly higher number of different types than the Spanish data, but a lower number of different tokens. Here, more different types occur less often in the corpus data (for a detailed quantitative analysis of the data see Hennecke & Baayen 2017). In what follows, a qualitative analysis of selected pairs of internal prepositions is presented in order to investigate whether these differences also appear at a qualitative level, with a special focus on the semantic functions of the nominal constituents and the semantic transparency of the constructions. The specific semantic relations were established with regard to the current literature on the semantic relations of nominal constituents in nominal compounds (Gagné & Shoben 1997; Gagné & Spalding 2009; Girju et al. 2005). They were subsequently modified and adapted to the specific case of N Prep N constructions in the corpus data under investigation. It is not possible to list and discuss all occurrences of all types in the present paper; only selected examples will therefore be discussed and analyzed. Where necessary, references will be made to frequency of occurrence.

5.1 The preposition *de* in N Prep N constructions

In all three languages under investigation, the preposition *de* is most often used to combine two nominal expressions, as in Fr. *salle de bain* ('bathroom'), Sp. *botas de agua* ('rubber boots'), or Pt. *moinho de vento* ('wind mill'). Therefore, the preposition *de* appears in all pairs of internal prepositional variation analyzed in the following section. The three data sets also show internal variation for prepositions other than *de*, but these pairs are not the subject of the present analysis. As mentioned above, the preposition *de* has been much discussed; it has often been considered an "empty" or "colorless" preposition that lacks any kind of semantic content and that merely fulfills a linker role functions. This completely functional approach is not adopted in the present paper for reasons given above. In the present account, I follow a constructionist approach (see Masini 2009), in which the prepositional constituent in N Prep N constructions is an element of semantic consequence to the whole construction (Masini 2009: 262). Masini states the example of Italian N1 *di* N2 intermediate lexical constructions, which clearly differ semantically from N1 *a* N2 intermediate lexical constructions. With reference to Johnston and Busta (1996), she emphasizes that "the prepositions *da*,

di and *a* in Italian N+PREP+N expressions, under certain conditions and in combination with certain classes of nouns, are specialized for different kinds of modification" (Masini 2009: 262). In the further analysis, the statement from Masini will be refined, since in the present data, the intermediate lexical constructions N1 *de* N2 and other intermediate lexical constructions (e.g. *N1 para/pour* N2) may overlap semantically under certain conditions. These cases will be exemplified below in a cross-linguistic comparative analysis. In this analysis, the preposition is seen not as a semantically opaque constituent but as a constituent with a specific semantic value determined by the semantic functions of the nominal constituents.

The preposition *de* in French, Spanish, and Portuguese has been described as expressing various relations (Bartning 1993: 187). In binominal constructions, it expresses, for instance, a relation of possession (Sp. *el ordenador de Luis* 'Luis' computer', Fr. *la voiture de Jean* 'John's car'), characterization (Fr. *statut de valeur* 'status'), instrument (Fr. *coup de baton* 'blow'), material (Fr. *papier de soie* 'silk paper'), a part-whole relation (Sp. *puerta de casa* 'front door', Pt. *ponta do dedo* 'fingertip'), an affiliation (Fr. *fils de roi* 'king's son'), a content (Fr. *tasse de café* 'cup of coffee'), a defining characteristic (Sp. *hotel de lujo* 'luxury hotel') or a purpose (Pt. *vestido de noiva* 'wedding dress'). For more examples in French see Lang (1991: 291ff.).

5.2 Internal variation between *de* and *a/à*

Internal variation between the prepositional constituents *de* and *à* has been the subject of several articles and books on French prepositions and nominal syntagms (e.g. Anscombre 1990; Lang 1991; Bosredon & Tamba 1991; Cadiot 1997). However, it is interesting that this discussion has no equivalent in the literature on Spanish and Portuguese prepositions. This is because such internal variation does not take place in Spanish and only to a small extent in Portuguese. The sole example of internal variation of *de* and *a* in the Spanish corpus data is the following:

(5)   [N1 *de/a* N2$_{\text{TYPE/SPECIFICATION}}$]$_{\text{N}}$> *freno de/a disco*, 'disk brake'

Here, the construction containing *de* is far more frequent and the lexicalized form can be found in dictionaries. Still, the construction *freno a disco* also occurs regularly in the corpus data of the esTenTen corpus, with a frequency of 0.10 occurrences per million. However, the corpus data shows that the internal variation of *de* and *a* is neither frequent nor productive in Spanish, as only one example of

one type can be found in this large-scale internet corpus. In Portuguese, the pt-TenTen data shows at least two Intermediate lexical constructions with variation of *de* and *a* that present a certain productivity:

(6) [N1 *de/a* N2_{PURPOSE}]_N   [N1 *de/a* N2_{TYPE/SPECIFICATION}]_N
     forno de/a microondas, forno de/a   *lampião de/a gás, pilhas de/a com-*
     lenhas   *bustível*
     'microwave oven', 'wood stove'   'gas lantern', 'fuel cell'

The template [N1 *de/a* N2_{TYPE/SPECIFICATION}]_N, in particular, is frequently present in the corpus data and is expressed via different types, as in *motor de/a combustão* ('combustion motor') or *bomba de/a vácuo* ('vacuum pump'). It is striking that many of these types are technical terms. It is possible to perceive a semantic difference in both intermediate constructions, where the type N1 *a* N2 more clearly indicates the material part of the N2 constituent, the type N1 *de* N2 focuses semantically on complementing N1 and creating a construction that is a subtype of N1. However, the first sample surveys and questionnaires revealed that native speakers of European and Brazilian Portuguese do not perceive a difference in the semantic meaning patterns or, more precisely, in the semantics of the whole construction.

For the French data, a very different pattern appears in the analysis of internal variability of *de* and *à*:

(7) [N1 *de/à* N2_{PURPOSE}]_N   [N1 *de/à* N2_{TYPE/SPECIFICATION}]_N
     fil de/à pêche   *course à/d'obstacles*
     'fishing rod'   'obstacle course'

(8) [N1 *de/à* N2_{INGREDIENT}]_N   [N1 *de/à* N2_{CONTAINER}]_N
     crème au/de citron   *conteneur de/à déchets*
     'lemon creme'   'waste bin/bin with waste'

(9) [N1 *de/à* N2_{TRANSPORT}]_N
     course de/à vélo
     'biking trip'

The existing literature on *de-à* alternation in French emphasizes that there is a semantic difference between binominal constructions containing *de* and *à* and that this semantic difference affects not only the prepositional element itself but also the whole naming unit. This becomes very clear in more detailed analysis

of the examples from the template [N1 *de/à* N2$_{\text{CONTAINER}}$]$_N$. In all these examples, the intermediate lexical construction N1 *à* N2 designates the container itself, as in *flûte à champagne* ('champagne glass') or *corbeille à fruit* ('fruit bowl'). In contrast, the intermediate lexical construction N1 *de* N2 denotes the content of the container, as in *flûte de champagne* ('a glass of champagne') or *corbeille de fruit* ('a bowl of fruits'). In these cases, according to Cadiot (1997), *de* turns the interpretation of the construction toward the N2 and constructs a quantified image of the referent, whereas *à* turns the interpretation toward the N1 and permits a qualified image of the reference (Cadiot 1997: 44). That is, *de* carries an effect of quantification whereas *à* carries a semantic notion of qualification. For cases of the intermediate lexical construction [N1 *de/à* N2$_{\text{INGREDIENT}}$]$_N$, such as *salade d'écrevisses* and *salade aux écrevisses* ('crawfish salad'), Lang (1991) states that the preposition *à* connects N1 and N2, whereas the preposition *de* derives N1 from N2. That is to say that *à* describes an ingredient, whereas *de* describes a substance (Lang 1991: 283). In the same way, in the examples of [N1 *de/à* N2$_{\text{TYPE/SPECIFICATION}}$]$_N$ and [N1 *de/à* N2$_{\text{MEANS OF TRANSPORT}}$]$_N$, it can be seen that *à* points to the material object *vélo* or *obstacles*, whereas *de* more likely complements the N1, and hence the whole construction describes a subtype of N1. According to Cadiot (1997: 43), the semantic differences that occur through the variation of the prepositions *de* and *à* can be accounted for in terms of the more abstract categorization that is the opposition of intension and extension. On this view, *de* constructs an extensional reference directly, whereas *à* creates an extensional reference indirectly by passing over an intentional reference (Cadiot 1997: 62).

From a constructionist perspective, it can be stated that only in the template [N1 *de/à* N2$_{\text{CONTAINER}}$]$_N$ does the semantic value of the whole construction change, as in *conteneur de déchets* ('bin containing waste') and *conteneur à déchets* ('waste bin'). In this case only, we have two different naming units when *de* and *à* alternate. Therefore, only here is it appropriate to refer to two different constructions, [N1 *de* N2$_{\text{CONTAINER}}$]$_N$ and [N1 *à* N2$_{\text{CONTAINER}}$]$_N$, which lead to two different naming units at the lexical level. In all the other cases mentioned above, the variability of *de* and *à* does not lead to different semantic interpretations of the lexical outcome, but only to a difference in the semantic weight of certain meaning patterns in the interpretation. Therefore, in all other cases, the inheritance hierarchy from the previous section of this paper can be applied in order to capture the internal constituent variation.

To conclude this analysis, it can be stated that all constructions that allow internal constituent variation of the prepositional element are semantically transparent. The analysis shows that alternation of the internal prepositional constituent does not go along with the semantically more opaque constructions in

the languages under investigation, since normally in these cases, the semantic functions of the nominal constituents cannot always be clearly determined. In Spanish and Portuguese, the internal variation is only possible in very specific cases of semantic function of the nominal constituents. In French, on the other hand, the internal variation of *de* and *à* is more frequently used or observed, and appears to be governed by the semantic functions of the nominal constituents.

5.3 Internal variation between *de* and *em/en*

The variation between *de* and *en/em* in N Prep N constructions has received little attention in the literature. On a general level, Lang (1991: 411) states that in French, *en* between two nouns indicate the location of N1, as in *arc-en-ciel* ('rainbow') and *une ville en Italie* ('a city in Italy'), the characterization of N1, as in *ange en stuc* ('stucco angel'), a way of preparation of N1, as in *une salade en vinaigrette* ('a salad with dressing'), the material of N1, as in *robe en soie* ('silk dress'), the form in which N1 appears, as in *fleurs en bouquet* ('bouquet of flowers'), the condition in which N1 stands, as in *arbre en fleur* ('blooming tree'), or a field in which N1 operates, as in *expert en assurances* ('insurance expert'). According to Laumann (1998: 55), French *de* and *en* are not always interchangeable when N2 refers to the material of N1. On the basis of an analysis of French grammar and dictionary entries, Laumann states that *en* appears more regularly with a predicative supplement than *de*, gives more concrete information about the material, and is less strongly linked to the N1. However, the most important difference seems to be that *en* cannot appear in more opaque constructions with a (partially) idiomatic reading. Laumann (1998: 55) cites the examples of *homme de fer* ('iron man') and *yeux d'acier* ('steely eyes'), where it is not possible to substitute *en* for *de*. In the French data, most of these relations can also be seen in variations of *de* and *en*, as in the following examples:

(10) [N1 *de/en* N2$_{\text{MATERIAL}}$]$_N$ [N1 *de/en* N2$_{\text{FIELD}}$]$_N$
      chemise de/en coton   *étudiant de/en Sciences Po*
      'cotton shirt'   'student of politics'

(11) [N1 *de/en* N2$_{\text{LOCATION}}$]$_N$ [N1 *de/en* N2$_{\text{CONDITION}}$]$_N$
      course de/en montagne   *maison de/en vente*
      'mountain race'   'house for sale'

(12) [N1 *de/en* N2$_{\text{GROUP}}$]$_N$ [N1$_{\text{ACTION}}$ *de/en* N2$_{\text{MATERIAL}}$]$_N$
      dîner de/en famille   *dépenses d'/en énergie*
      'family dinner'   'energy expenditures'

In each of these cases, the variation between *de* and *en* does not trigger any strong meaning difference between the two construction types; that is, it is possible to talk about internal variability rather than about two different types of constructions referring to different naming units. Nonetheless, certain differences are visible, as Laumann (1998) pointed out. For instance, *en* is generally less closely linked to N1 and more often introduces a complement. Constructions with *en* also appear to have a lesser degree of fixedness and put the focus on the N2. The present analysis confirms Laumann's observation that the alternation of *de* and *en* is only possible in semantically transparent constructions that do not include any idiomatic meaning.

A very similar picture emerges from the analysis of the Portuguese data, as shown in the following examples:

(13) [N1 *de/em* N2$_{MATERIAL}$]$_N$ [N1 *de/em* N2$_{FIELD}$]$_N$
      bracelete de/em aço   *profissional de/em artes*
      'steel bracelet'   'art professional'

(14) [N1 *de/em* N2$_{LOCATION}$]$_N$ [N1 *de/em* N2$_{CONDITION}$]$_N$
      surf de/em ondas   *crianças de/em risco*
      'surf (on waves)'   'children at risk'

(15) [N1 *de/em* N2$_{GROUP}$]$_N$ [N1 *de/em* N2$_{MEDIUM}$]$_N$
      almoço de/em família   *comentário de/em áudio*
      'family lunch'   'audio commentary''

The Portuguese data show almost the same intermediate lexical constructions that function with a variation between *de* and *em*. The only difference is in the intermediate construction [N1 *de/em* N2$_{MEDIUM}$]$_N$, where N2 designates the medium via which N1 is transferred. In contrast, the French data offer the intermediate construction [N1$_{ACTION}$ *de/en* N2$_{MATERIAL}$]$_N$, which indicates a concrete action referring to a specific (raw) material. Nevertheless, from a quantitative perspective, the internal variation between *de* and *em/en* is by far more frequent in the Portuguese data.

From a quantitative perspective, the variation between *de* and *en* is quite rare in the Spanish data, but the qualitative analysis shows a more diverse picture:

(16) [N1 *de/en* N2$_{MATERIAL}$]$_N$ [N1 *de/en* N2$_{FIELD}$]$_N$
      construcción de/en madera   *grado de/en ingeniería*
      'wood construction'   'engineering degree'

(17)   [N1 *de/em* N2$_{\text{LOCATION}}$]$_N$ [N1 *de/en* N2$_{\text{CONDITION}}$]$_N$
       ciclismo de/en pista   *obras de/en construcción*
       'track cycling'   'construction site'

(18)   [N1 *de/en* N2$_{\text{MEDIUM}}$]$_N$
       entrevista de/en radio
       'radio interview'

The examples show that Spanish allows the same internal variation as Portuguese, except that the template [N1 *de/en* N2$_{\text{GROUP}}$]$_N$ was not present in the data. In Portuguese and French, there are no strong meaning differences between the two templates, and therefore they can be counted as variants rather than as two distinct forms. In Spanish, as in French and Portuguese, the same subtle differences in the degree of fixedness and focus of the constituents can be observed.

Overall, it is possible to state that the variation between *de* and *en/em* is possible in all three languages under investigation. The differences appear to exist at the quantitative level rather than in the specific semantic meaning patterns. In all three languages, different templates can demonstrate and explain the possible alternation between *de* and *en/em*. For most cases, these templates overlap in the three languages. Therefore, it is possible to apply the inheritance hierarchy mentioned in Section 4 to all of the examples.

5.4 Internal variation between *de* and *pour/para*

For French binominal compounds of the type N Prep N, Laumann (1998) states that the preposition *pour* occurs quite rarely. This may be explained by the fact that *pour* is less abstract than other prepositions, such as *de* or *à*: that is, *pour* indicates a very concrete meaning of purpose or determination, whereas *de* shows a less definite meaning pattern. Therefore, *de*, as a semantically more opaque constituent, offers a wider scope for application than *pour*, but in some cases both prepositions are interchangeable, as in the following examples:

(19)   [N1 *de/pour* N2$_{\text{USER}}$]$_N$ [N1 *de/pour* N2$_{\text{PURPOSE}}$]$_N$
       collier de/pour chien   *décoration de/pour mariage/table*
       'dog collar'   'marriage/table decoration'

(20)   [N1 *de/pour* N2$_{\text{USER(OBJECT)}}$]$_N$
       musique de/pour piano
       'piano music/music for piano'

The French data show that the variation between *de* and *pour* only is possible in cases where N2 designates a user (or a beneficiary), or where N2 specifies the purpose of N1. In all three templates given above, N2 serves to form a subtype of N1. However, the templates containing the preposition *de* point more clearly to the N1 and focus on the interpretation of the whole template as a subtype of N1. In templates containing the preposition *pour*, the preposition is clearly attached to the N2, and the semantic emphasis is on N2. Furthermore, the preposition *pour* clearly carries the interpretation 'for', whereas the constructions containing *de* leave room for ambiguous interpretation. While *musique pour piano* clearly designates music (a piece of music or composition) for piano, *musique de piano* may also refer to music played by a piano (and not necessarily composed for playing on a piano). In this sense, *pour* helps to resolve ambiguity and allows only the interpretation 'designed for'. For the Spanish data, the pattern is quite similar to the French data, as in the following examples:

(21) [N1 *de/para* N2$_{\text{USER}}$]$_{\text{N}}$ [N1 *de/para* N2$_{\text{PURPOSE}}$]$_{\text{N}}$
      club/ropa de/para niños   *alimentos de/para consumo*
      'children's club/clothes'   'consumer goods'

(22) [N1 *de/para* N2$_{\text{USER(OBJECT)}}$]$_{\text{N}}$
      juego de/para pc
      'PC game'

These cases show that the variation between *de* and *para* is possible only in contexts in which N2 semantically represents a user of (a person or an object) or a specific purpose for N1. These are the same templates that were found for the French data above. This result contradicts the findings from López (1970), who indicates that variation of *de* and *para* is also possible in contexts in which N1 designates a container, as in *cesto de/para basura* ('waste bin/bin for waste'). In her corpus data of Argentinian Spanish from Buenos Aires, Pacagnini (2003: 164) also finds constructions of the type *loción de/para limpieza* ('cleaning lotion') or *crema de/para hidración* ('hydration crème'), in which the preposition expresses the utility of an object. Furthermore, she describes examples of the type *lápiz de/para labios* ('lipstick') and *esmalte de/para uñas* ('nail polish'), in which N1 represents an instrument. From this, Pacagnini deduces a schema in which, on a continuum between morphology and syntax, *de* lies closer to the morphological pole, whereas *para* is closer to the syntactic pole. In this paper, I can confirm Pacagnini's hypothesis that N Prep N constructions in Spanish show a certain internal variation in respect of the prepositions *de* and *para*, which might therefore

be considered as lying at different points of a continuum between the morphological and the syntactic pole. In this case, it is evident that constructions with *para* are located closer to the syntactic pole than constructions with *de*. Pacagnini observes that 75 percent of the participants in her data used a determiner or a qualifying adjective with the preposition *para* in cases where N1 denotes an instrument, as in *loción de/para la limpieza* ('lotion for cleaning') or *esmalte para uñas sensibles* ('polish for sensitive nails') (Pacagnini 2003: 166). In the esTenTen corpus data, this type of variation between *de* and *para* does not occur at all. However, a closer look at the Portuguese data offers interesting findings:

(23) [N1 *de/para* N2$_{USER}$]$_N$ [N1 *de/para* N2$_{PURPOSE}$]$_N$
      brinquedos de/para crianças   *acessórios de/para decoração*
      'children's toys'   'accessories for decoration'

(24) [N1 *de/para* N2$_{USER(OBJECT)}$]$_N$ [N1 *de/para* N2$_{REASON}$]$_N$
      concerto de/para piano   *cirurgia de/para correção*
      'piano concert'   'reconstructive surgery'

(25) [N1 *de/para* N2$_{PERIOD}$]$_N$ [N1$_{INSTRUMENT}$ *de/para* N2]$_N$
      aluguel de/para férias   *produto de/para limpeza*
      'vacation rental'   'cleaning product'

(26) [N1 *de/para* N2$_{DETERMINATION}$]$_N$
      animais de/para abate
      'animals for slaughter'

The Portuguese data illustrate that variation of *de* and *para* is possible in a larger number of nominal semantic relations in Portuguese than in Spanish or French. On the one hand, Portuguese offers the same templates as French and Spanish: N2 as a user (object or person) and N2 as a specific purpose of N1. Portuguese also provides additional templates, including N2 as a specific time or period of time, and N2 designating a specific determination for N1 (which in most cases is a living being). One additional template, N1 being an instrument for N2, is of particular interest. Here, we find the Portuguese example *produto de/para limpeza* ('cleaning product'), which Pacagnini cited for Argentinian Spanish. This template appears to be productive in Portuguese, as shown in the additional examples *creme de/para mãos* ('hand cream') and *máscara de/para cílios*

('mascara for eyelashes'). Although our Spanish data contradict Pacagnini's findings for Spanish, the same template can be found in Portuguese. Further investigation of this phenomenon is necessary, particularly in light of the possibility that the Spanish used in Buenos Aires, where Pacagnini collected her data, may be influenced by Portuguese from Brazil. Initial informal speaker assessments of native Spanish speakers in Spain reveal that the template [N1$_{\text{INSTRUMENT}}$ *de/para* N2]$_{\text{N}}$ is not productive in Spain and that the template [N1$_{\text{INSTRUMENT}}$ *para* N2]$_{\text{N}}$ is considered incorrect.

The analysis of the variation between *de* and *para*, and *de* and *pour*, in Spanish, French, and Portuguese reveals that Portuguese has the largest number of templates at an intermediate lexical level for the variation of *de* and *pour/para*. The Spanish and French data overlap in their templates for the variation of *de* and *pour/para*, while the Spanish data from Buenos Aires (Pacagnini 2003) offer a slightly different picture. The analysis here supports the findings from the previous subsections on the semantic transparency of the constructions under investigation. The present analysis does not feature any (partially) opaque or (partially) idiomatic constructions. I mentioned at the beginning of this subsection that the prepositions *de* and *pour/para* vary in their semantic transparency; nevertheless, they undergo internal constituent variation in all three languages under investigation. While traditional accounts generally mention the different syntactic status of constructions containing *de* and *pour/para*, the constructionist approach introduced in Section 4 makes possible an unproblematic mapping of this internal constituent variation.

6 Conclusion

The present study of internal constituent variation in N Prep N constructions allows numerous conclusions to be drawn as to their nature in Romance languages as well as on the role and variability of the prepositional element. The discussion and analysis here have shown that it is not always possible or expedient to differentiate clearly between lexical and syntactic N Prep N constructions. In many cases, not even the numerous delimitation tests may lead to a clear distinction. Therefore, the present account has abandoned this strict, dichotomous distinction in favor of a more holistic approach. When considering internal constituent variability, the determining factor is not the lexical or syntactic status of the elements; instead, it is the nominal semantic relation expressed via the preposition. Here, it is not crucial to differentiate between the lexical status, e.g. *libro de niños* ('children's book') and the syntactic status, e.g. *libro para niños*

('children's book'). In order to conduct a fruitful qualitative comparative analysis of N Prep N constructions in Romance languages, it is necessary to adopt a theoretical account that does not focus on the lexicon-syntax distinction. In the present paper, construction morphology, a constructionist approach that expands the notion of construction to the word level, offers the appropriate tools for analysis. Following Masini (2009: 261), N Prep N constructions are analyzed as abstract templates, which are, to some degree, productive and associated with a naming function. For the present analysis, a constructionist inheritance hierarchy has been adapted to internal constituent variation in one construction (see Section 4). The latter analysis focused on the intermediate lexical level, that is, the alternation between [N1 Prep1 N2] and [N1 Prep2 N2], at which Prep1 and Prep2 designate alternative prepositions. This constructionist approach revealed the possible templates for prepositional variation in three different languages: Spanish, French, and Portuguese.

The analysis of three alternating pairs, specifically *de* and *à/a*, *de* and *en/em*, and *de* and *pour/para*, demonstrates important differences and common features between the languages. The quantitative aspect, which was not the primary focus of this paper, demonstrates the strong frequency and productivity of the different templates in Portuguese. This holds to a lesser extent in French and is even less in Spanish. This result is in line with the results from Hennecke & Baayen (2017). The qualitative analysis demonstrates that, in the underlying datasets, Portuguese offers the greatest number of different templates for internal prepositional variation, followed by French, and then Spanish. In this connection, it should be mentioned that Portuguese also offers the largest number of constructions (or types) for each template. This result confirms the impression from the quantitative study that Portuguese N Prep N templates are frequent in speech and are very productive. From a qualitative perspective, it is striking that most templates of internal prepositional variation exist across languages. In the case of the pair *de* and *en/em*, the templates that allow internal prepositional variation vary only slightly between the languages. For variation between *de* and *à/a*, the French data show the greatest tendency to internal variation. This is mainly because the preposition *à* is relatively productive and frequent in French, which is not the case for Spanish and Portuguese. In cases where French relies on the preposition *à*, Spanish and Portuguese mostly employ the preposition *de*, as in Fr. *verre à vin*, Sp. *copa de vino* and Pt. *copo de vinho* ('wine glass', in each case). Spanish does not offer any internal variation of *de* and *a*, whereas Portuguese shows certain tendencies in this direction. For the variation between *de* and *pour/para*, the French and Spanish data do not show any qualitative differences; that is, they overlap exactly in terms of which templates allow internal prepositional varia-

tion. Studies based on data from Argentinian Spanish indicated the existence of further templates; these were not found in the present data in Spanish, but many of them were present in the Portuguese data.

A very important finding from the qualitative analysis is that internal prepositional variation in the three languages is possible only for semantically transparent constructions. This can be explained by the fact that in opaque N Prep N constructions, the semantic relation between the nominal constituents often cannot be determined explicitly.

In conclusion, a constructionist approach to N Prep N constructions may solve certain problems in defining and delimitating these constructions in Romance languages. Furthermore, a constructionist approach allows an accurate investigation of the differences and common features of templates for internal prepositional variation in the three languages under investigation here. Future studies should investigate these templates in more detail, extending the approach to other types of internal variation.

References

Anscombre, Jean-Claude. 1990. Pourquoi un moulin à vent n'est pas un ventilateur. *Langue française* 86. 103–125.

Bartning, Inge. 1993. La préposition 'de' et les interprétations possibles des syntagmes nominaux complexes: Essai d'approche cognitive. In Anne-Marie Berthonneau & Pierre Cadiot (eds.), *Les prépositions: Méthodes d'analyse*, vol. 11 (Lexique), 163–191. Lille: Presse Universitaires de Lille.

Bell, Melanie J. & Martin Schäfer. 2016. Modelling semantic transparency. *Morphology* 26. 157–199.

Benveniste, Émile. 1974. Formes nouvelles de la composition nominale. In Émile Benveniste (ed.), *Problèmes de linguistique générale*, vol. 2 (Bibliothèque des sciences humaines), 163–176. Paris: Gallimard.

Booij, Geert. 2010. Compound construction: Schemas or analogy? A construction morphology perspective. In Sergio Scalise & Irene Vogel (eds.), *Cross-disciplinary issues in compounding* (Current issues in linguistic theory 311), 93–109. Amsterdam: John Benjamins.

Booij, Geert. 2015. Word-formation in construction grammar. In O. Peter Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word formation*, vol. 1, 188–202. Berlin & Boston: Mouton de Gruyter.

Bosredon, Bernard & Irène Tamba. 1991. Verre à pied, moule à gaufres: Préposition et noms composés de sous-classe. *Langue française* 91. 40–55.

Buenafuentes de la Mata, Cristina. 2010. *La composición sintagmática en español* (Instituto histororia de la lengua. Monografías 7). España: Cilengua.

Cadiot, Pierre. 1997. *Les prépositions abstraites en français* (Linguistique). Paris: Armand Colin/ Masson.

de Bustos Gisbert, Eugenio. 1986. *La composición nominal en español*. Salamanca: Universisad de Salamanca.

Di Sciullo, Anna Maria & Edwin Williams. 1987. *On the definition of word* (Linguistic Inquiry Monographs 14). Cambridge, MA: Springer.

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

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. & Edward J. Shoben. 1997. The influence of thematic relations on the comprehension of modifier-noun combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 23(1). 71–87.

Girju, Roxana, Dan Moldovan, Marta Tatu & Daniel Antohe. 2005. On the semantics of noun compounds. *Journal of Computer Speech and Language* 19(4). 479–496. Special Issue on Multiword Expressions.

Goldberg, Adele E. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago/London: The University of Chicago Press.

Goldberg, Adele E. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.

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

Hennecke, Inga & R. 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.

Kornfeld, Laura Malena. 2009. Romance: Spanish. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford handbook of compounding*, 436–453. Oxford: Oxford University Press.

Lang, Jürgen. 1991. *Die französische Präposition: Funktion und Bedeutung* (Beiträge zur Literatur-, Sprach- und Medienwissenschaft 95). Heidelberg: Carl Winter Universitätsverlag.

Laumann, Ferdinand. 1998. *Französische Nominalkomposita mit Präposition: Eine Untersuchung aufgrund der naturwissenschaftlich-technischen Fachsprache* (Studien zur Linguisitk 5). Münster: Lit.

López, Maria Luisa. 1970. *Problemas y métodos en el análisis de preposiciones*. Madrid: Gredos.

Masini, Francesca. 2009. Phrasal lexemes, compounds and phrases: A construcionist perspective. *Word Structure* 2(2). 254–271.

Masini, Francesca & Anna Thornton. 2007. Italian VeV lexical constructions. In Geert Booij, Angela Ralli & Sergio Scalise (eds.), *Morphology and dialectology*, 148–189.

Pacagnini, Ana Maria Judith. 2003. Compuestos sintagmáticos y alternancia preposicional. *Moenia* 9. 159–172.

Rainer, Franz. 2016. Italian. In O. Peter Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word formation*, vol. 4, 2712–2731. Berlin & Boston: Mouton de Gruyter.

Rio-Torto, Graça & Sílvia Ribeiro. 2009. Compounds in Portuguese. *Lingue e Linguaggio* 8(2). 271–291.

Rio-Torto, Graça & Sílvia Ribeiro. 2012. Portuguese compounds. *Probus. International Journal of Latin and Romance Linguistics* 24(1). 119–145.

Val Àlvaro, José Francisco. 1999. La composición. In Violeta Demonte & Ignacio Bosque (eds.), *Gramática descriptiva de la lengua espanola*, vol. 3, 4747–4842. Barcelona: Espasa Calpe.

Villoing, Florence. 2012. French compounds. *Probus. International Journal of Latin and Romance Linguistics* 24(1). 29–60.

Zwitserlood, Pienie. 1994. The role of semantic transparency in the processing and representation of Dutch compounds. *Language and Cognitive Processes* 9(3). 341–368.