

Benjamin Fagard
Lattice Laboratory (CNRS, ENS|PSL & Sorbonne Nouvelle)

José Pinto de Lima
Universidade de Lisboa

Complex prepositions in (Medieval) Portuguese: The question of language contact

Abstract

One important question in the current line of research concerning adpositions is how complex prepositions emerged in the diachrony of Portuguese. This is the question we address, focusing on the first centuries of available texts, and aiming to determine whether CPs have emerged independently from other European languages. Starting with definitions and lists of CPs in Modern Portuguese, and with a usage-based approach, we document the presence of several CPs as early as the 13th century. These findings do not exclude the importance of language contact, but are also compatible with an independent emergence of CPs in different European languages.

1. *Introduction*

Adpositions have been central in semantics in the last decades, with a huge body of literature, due in part to a strong focus on space and language which benefited from many studies on *simple* prepositions, mostly in European languages (from Brøndal 1950 to Vandeloise 1986; Zelinsky-Wibbelt 1993; or Tyler & Evans 2003) but also from other perspectives, especially typological (Svorou 1994; Hagège 2010) and diachronic (Di Meola 2000; Fagard 2010).

One important question which has emerged in this line of research is how *complex prepositions* emerged in European languages (Horst 2013, 182; see also Hoffmann 2005; Hüning 2014; and more recently Fagard et al. 2019; 2020). Indeed, one might wonder why these constructions are so frequent among European languages, and specifically whether language contact has had an impact

on their emergence, as seems to be the case at least in some of them (e.g., Breton or Albanese, see Rottet 2020; Bonnet and Fagard 2020).

In line with previous work by the authors, which has described both the emergence of complex adpositions in the diachrony of Portuguese (Lima 2014; 2019ab) and the influence of language contact, in the form of text translations, on this evolution (Fagard 2023), this paper aims to describe the emergence of CPs in the diachrony of Portuguese, focusing on the first centuries of available texts, in order to determine the importance of language-internal phenomena. The aim of our investigation is to search for clues indicating whether CPs have emerged independently from other European languages. Based on earlier research on other Romance languages (Cifuentes Honrubia 2003; Piunno and Ganfi 2017; Fagard 2020), we expected to find only few sequences in Medieval Portuguese, with a low relative frequency, and larger paradigms only in later stages (Classical & Contemporary Portuguese).

After Section 1, dedicated to the introduction, we discuss in Section 2 existing definitions and lists of CPs in European languages, and more specifically those existing for Modern Portuguese. Section 3 is dedicated to a presentation of our methodological approach. In Section 4, we present the general results of our corpus study, and go on in Section 5 to analyze in further detail three case studies, before concluding in Section 6.

2. *Complex Prepositions in Portuguese: Towards a Working Definition*

Grammars of Portuguese have provided varying definitions, labels and lists of Complex prepositions. In this section, we review these lists, which we then use as a starting point for our corpus search.

Complex prepositions are notoriously hard to define and identify: they are not a consensual category – or even a “helpful” one, according to Seppänen et al. (1994, 25). The definition adopted for this study relies on a definition of adpositions as relational expressions with 3 main functions: (i) introducing a nominal constituent; (ii) establishing a relation between this constituent and a (higher) constituent; (iii) helping determine the semantic nature of this relationship. This definition is valid of course for simple adpositions, as in (1) below in which *ẽ*, i.e., *em* ‘in’, introduces the constituent *sete maneiras* and links it syntactically and semantically to the verb *dividirse*.

(1) o peccado mortal açima dito se diuide ã sete maneiras.
'The mortal sin mentioned above is divided in seven ways' (Dom Diogo de Sousa, *Constituições do Bispado do Porto*, 1497, Corpus do Português)¹

Complex adpositions are defined in turn as functionally equivalent, but morphologically complex, typically having 3 formal components dedicated each to one of the 3 functions: this is illustrated in (2), where the semantic relationship is specified by the noun phrase *o pé* 'the foot' and the simple prepositions *a* 'at' and *de* 'of, from' fulfill functions (i) and (ii) above (for further details and a discussion of marginal cases and exceptions, see Stosic and Fagard 2019; Fagard et al. 2020).

(2) Lançou-se a dormir **ao pé do** muro, debayxo daquela ataude
'he threw himself down by the wall, below that coffin' (*Crónica de D. Fernando*, 1431-1443, Corpus do Português)

In Portuguese, this type of sequence has been identified mainly as a "prepositional phrase" (*locução prepositiva*, see e.g., Nunes 1930, 352; Cunha and Cintra 1984, 551-2) or as a "complex preposition" (*preposição composta*, see e.g., Cunha and Cintra 1984, 551-2). The inventories we have been able to identify in the literature range from a dozen sequences (Nunes 1930, 352) to several dozen (44 in Cunha and Cintra 1984, 551-2; 54 in Bechara 2002, 305; 62 in Raposo et al. 2013, 1503-4). Additionally, some authors introduce distinctions among complex prepositions; for instance, Raposo et al. (2013, 1503-4) distinguish prepositional phrases on the basis of the nature of the nucleus: they list 4 "noun-based" items, 17 "adverb-based" items and 41 "preposition-based" items (by default in Table 1).

As shown in previous works on complex adpositions in Romance languages (see e.g., Fagard et al. 2019; Fagard et al. 2020), most items follow one of two patterns, pattern A: [Preposition₁ + (Determiner +) Lexical nucleus + Preposition₂]; or pattern B: [Adverb + Preposition].² These patterns are illustrated in (3) (see also (2) above for an example with determiner) and (4), respectively.

(3) Esta foy a reuerença que fezerom os discipollos a xpisto que nõ o leixarõ caualgar em ouosso mas ante **em lugar de** sella lhe poserõ de baixo seus vistidos.

1 All examples are our own, and were selected from the corpus chosen for this study.

2 In Portuguese, sequences pertaining to Pattern B are often described as *advérbios relacionais*, i.e., "relational adverbs" (Raposo et al. 2013, 1500).

‘Here is the favor the disciples did for the Christ: they did not let him ride bareback but rather instead of a saddle they put their clothes beneath him’ (Gonçalo Garcia de Santa Maria, *Euangelhos e epistolas con suas exposições en romãçe*, 1497, Corpus do Português)

(4) E passou **aalem do** ryo.
‘and he went beyond the river’ (*Crónica Geral de Espanha*, 1344, Corpus do Português)

These patterns are common in Romance languages (Fagard et al. 2020), and even in other Indo-European languages, especially in Western Europe.

For Portuguese, we compiled the lists provided by the authors mentioned above, and obtained a total of 102 items which they consider to be Complex Adpositions. There is wide variation across lists, with partial agreement on only 41 sequences. Unsurprisingly, most sequences in this list follow pattern A (65 items, i.e., 63.7%) or B (22 items, i.e., 21.6%). Pattern A, though more productive in terms of types, may be expected to be less frequent overall (see Fagard, in press). Some sequences pertaining to Type A contain an adverb as lexical nucleus, while a few sequences could be classified as pertaining to Type A or B, depending on the analysis, e.g., *em cima de / encima de*. Indeed, this sequence could be analyzed as [Preposition *em* ‘in, on’ + nominal nucleus *cima* ‘top’ + Preposition *de* ‘of’] or as [Adverb *encima* ‘above’ + Preposition *de* ‘of’]. In such cases, the noun may be headed by a determiner (*na cima do ano* ‘at the end of the year’), as in (5), in which the substantive nature of the lexical nucleus is shown by the presence of a determiner.

(5) des ali rezou el, e non lle foi greu, / quanto lle mandara ela; e, com’ oý eu, / **na cima do ano** foy-o consigo levar.
‘from then on he prayed as much as she had commanded, and it did not seem hard to him; and as I heard, at the end of the year she took him to be with her’ (*Cantigas de Santa Maria* 1, 12th c., Corpus do Português)

An important specification in this definition is that the Complex Adposition must constitute a semantic unit, i.e., present at least some degree of semantic opacity. In some rare cases, there is direct evidence for this unity, e.g., when a CP in Portuguese is used to translate a simple preposition in Latin, as in (6).

(6) canis **circum** intestina. || cão **ao redor de** tripas
‘a dog around offal’ (Jerónimo Cardoso, *Dicionário de Latim-Português* 1, 1569-70, Corpus do Português)

For other sequences listed in Table 1, this might not – or not entirely – be the case. For instance, the sequence *ao conselho de* appears very early in the corpus; however, even later occurrences are not always prepositional, as illustrated in (7). Similarly, in (8), *em redor de* should not be analyzed as a CP, since *em redor* is an adverbial phrase modifying the verb *cobrir* ‘to cover’, while *de cores* ‘of colors’ is an argument of the same verb.

(7) E **mãdo ao Conçelho** de vihões que os enparë
 ‘And I order the Council of Vihães to protect them’ (*Documentos foraleiros da Chancelaria de D. Dinis-Bragança*, 16th c., Corpus do Português)

(8) Toda gloria da filha do rrey he dentro em faldas douro, **cubertas ë redor** de coores de muytas maneyras.
 ‘All the glory of the king’s daughter is within in golden borders, clothed all around with a wide variety of colors’ (*Orto do Esposo*, 15th c., Corpus do Português)

To take another example, *a braços com* might be better analyzed as part of the verb phrase *vir* (or *andar*) *a braços (com alguém)* ‘to come to blows (with someone)’: indeed, all five examples in the corpus (13th-17th c.) follow this pattern, as in (9).

(9) Terceira, soltura, segurança e atrevymento pera **vïir a braços** com qual quer homem, ainda que mais ryjo que elle seja.
 ‘3rd [advantage when it comes to waging war]: flexibility, self-confidence and boldness in coming to blows with any man, even if he should be more vigorous’ (*Livro da Ensinança de Bem Cavalgar Toda Sela*, 15th c., Corpus do Português)

Finally, some sequences are clearly grammaticalized, but seem to behave more like conjunctions than prepositions. The affinity between these two categories is indeed well-known (Pottier 1962: 195-197, Piot 1993). This is the case, for instance, of *a fim de* ‘in order to’, as illustrated in (10).

(10) Da outra parte pensou que era bem de enviar a eles huû messegeiro, da parte del-rei Tholomeu e da sua, **a fim de os pacificar** e trautar com eles paz (*Vida e feitos de Júlio Cesar*, 15th c., Corpus do Português)
 ‘On the other hand he thought it would be good to send them a messenger, in the name of king Tholomeu and his own, in order to allay them and negotiate peace with them’

Table 1 below lists the sequences we compiled from the literature.

TYPE A1 – WITH DETERMINER (PDXP)			
P1 = a, P2 = de (à/ao X de)			Other
à altura de, à conta de, à custa de, à direita de, à disposição de, à esquerda de, à frente de, à mercê de, à roda de, à volta de, ao lado de, ao longo de, ao pé de, ao redor de			na conta de, para a frente de, pela frente de
Type A2 – without determiner (PXP)			
P1 = a, P2 = de (a X de)	P1 = em, P2 = de (em X de)	P1 = por, P2 = de (por X de)	Other
a (meio) caminho de, a bel-prazer de, a conselho de, a despeito de, a fim de, a mais de, a meio de, a menos de, a par de, a partir de, a respeito de	em baixo de, em cima de, em favor de, em frente de, em lugar de, em prol de, em razão de, em redor de, em torno de, em troco de, em vez de	por baixo de, por causa de, por cima de, por debaixo de, por defronte de, por dentro de, por detrás de, por diante de, por perto de, por trás de	a braços com, a julgar por, a par com, a seguir a, com base em, de acordo com, de caras com, de cima de, de conformidade com, em frente a, para além de, para alguém de, para baixo de, para cima de, sem embargo de
TYPE B			
P2 = de (X de)	P2 = a (X a)	other	
abaixo de, acerca de, acima de, adiante de, além de, antes de, apesar de, após de, alguém de, atrás de, através de, cerca de, debaixo de, defronte de, dentro de, depois de, detrás de, diante de, fora de, junto de, longe de, mercê de, perto de, próximo de	até a, enquanto a, face a, frente a, graças a, junto a, quanto a	dentro em, para com, por entre	
Other			
de so, dentre, exceto, não obstante			

Table 1. Complex prepositions in Portuguese: list of potential sequences, per construction type; based on Nunes 1930; Cunha and Cintra 1984; Bechara 2002 and Raposo et al. 2013.

Though theoretically all simple prepositions may fill slots 1 and 3, these are typically occupied only by the most grammaticalized ones. Indeed, among the sequences we compiled, most are headed by *a*, *em* or *por*, and slot 2 is generally occupied by *de* or, less frequently, *a*, as shown in Table 1.

3. Methodological Approach

For this study, we adopted a usage-based approach. Using the list of sequences presented in Table 1 as a starting point, and including also several sequences identified as complex adpositions by Pinto de Lima (2014; 2019ab) and Fagard (2023), we controlled for the presence of CPs in a diachronic corpus of Portuguese, the *Corpus do Português* (Davies and Ferreira 2006), with a triangulation approach combining frequency counts (presented in Section 4) and qualitative analyses, esp. case studies (presented in Section 5).

The *Corpus do Português* is a very large corpus, with 1 billion words in total. We relied on the historical section of the corpus (from the 13th to the 20th century), which is much smaller, but of acceptable dimensions for a diachronic corpus, with 45 million words from nearly 57,000 texts. The interface enables the use of lemmas and of complex queries, with the possibility of sampling for large outputs, making it a very useful tool for diachronic studies.

For the quantitative part of our study, we tested all 102 sequences listed in Table 1. In order to check for the existence of other sequences, we also used open queries, testing for patterns A and B and their most frequent subpatterns described in Section 2: *à/ao X de*, *a/em X de*, *Adv a/de*.

The amount of noise can be limited by including in slots 1 and 2, i.e., as *Adposition₁* and *Adposition₂*, only those items which appear in the sequences listed in Table 1. Some sequences found as a result of these queries do not pertain to either pattern, but were found to be functionally equivalent, and were thus included in our analysis. One example of such a sequence is the pattern in which two adpositions are used in sequence: in such cases, two items which generally function as simple prepositions are found to behave jointly as a complex preposition with distinct semantics, as is the case in (11) for *para com*.

(11) Da grande caridade, e compayxão que o Infante tinha **para com** os pobres.
'Of the great charity and compassion that the Infant had for the poor' (*Crónica de D. Fernando*, 1431-1443, *Corpus do Português*)

We systematically analyzed the first occurrences to see when they were first used as complex adpositions, rather than as syntactically free sequences.

For the qualitative part of our study, we retrieved all occurrences of each sequence, and analyzed a random sample of 100 occurrences per century, for each sequence. The randomization was performed with the Excel software, af-

ter noise exclusion – noise being understood as resulting e.g., from cases of homonymy, as with *cabo* ‘head, end’ and ‘cape’, as in (12).

(12) Viagem de lisboa ate ao **cabo** de boa esperança
‘Trip from Lisbon to the Cape of Good Hope’ (*Virgeu de consolaçon*, 15th c., Corpus do Português)

Once retrieved, all occurrences were then analyzed in order to distinguish between referential uses (as in (2) or (5) above), i.e., those in which the lexical element in the sequence has a referent and may thus be for instance the antecedent of an anaphor, and non-referential uses, which display at least some degree of semantic bleaching. For instance, in example (3) above, the noun *lugar* is not preceded by a determiner, and its referential potential seems low. Indeed, in example (3) *em lugar de* could be glossed both *in the place of* (i.e., they put their robes where one usually attaches a saddle) and *instead of* (i.e., they did not attach a saddle but instead put their robes on the horse’s back). The existence of such ambivalent contexts, also called bridging contexts (Heine 2002), is typical of grammaticalization phenomena.

In referential occurrences, the sequence can generally be considered as non-adpositional. However, such occurrences typically pave the way for more grammaticalized uses of a sequence. This explains the frequent shift of constructions from the pattern [Preposition₁ + Determiner + Lexical nucleus + Preposition₂] to the pattern [Preposition₁ + Lexical nucleus + Preposition₂], as is the case for *ao cabo de* (14th c., in our corpus) and *a cabo de* (16th c., in our corpus).

We also aimed to describe the possible adpositional uses of each sequence, and distinguished broad semantic categories (spatial, temporal, abstract) in order to identify diachronic semantic shifts.

4. *Corpus Study – General Results*

Our quantitative analysis suggests that dozens of complex adpositions are already present in Medieval Portuguese. We were able to detect the presence of several CPs as early as the 13th c., for instance *em redor de* ‘behind’. Others seem to appear slightly later, e.g., in the 14th c., for instance *por parte de* ‘on behalf of’ or *em favor de* ‘in favor of’, centuries, as illustrated by examples (13-5).

(13) Senhor, se teverdes por ben, cavalguemos ambos soos e vaamos andar **en redor da** villa.
 ‘Lord, if you will, let the two of us ride alone **around** the town’ (*Crónica Geral de Espanha*, 1344, Corpus do Português)

(14) o ssenorio & directo do Couto de Bueu que uos Johan Martinz ia dito auyades **por parte de uosso padre & de uossa madre**
 ‘the land and right of land of Couto de Bueu which you, aforementioned Johan Martinz, have had from your father and your mother’ (*Textos Notariais. História do galego-português*, 1262-1300, Corpus do Português)

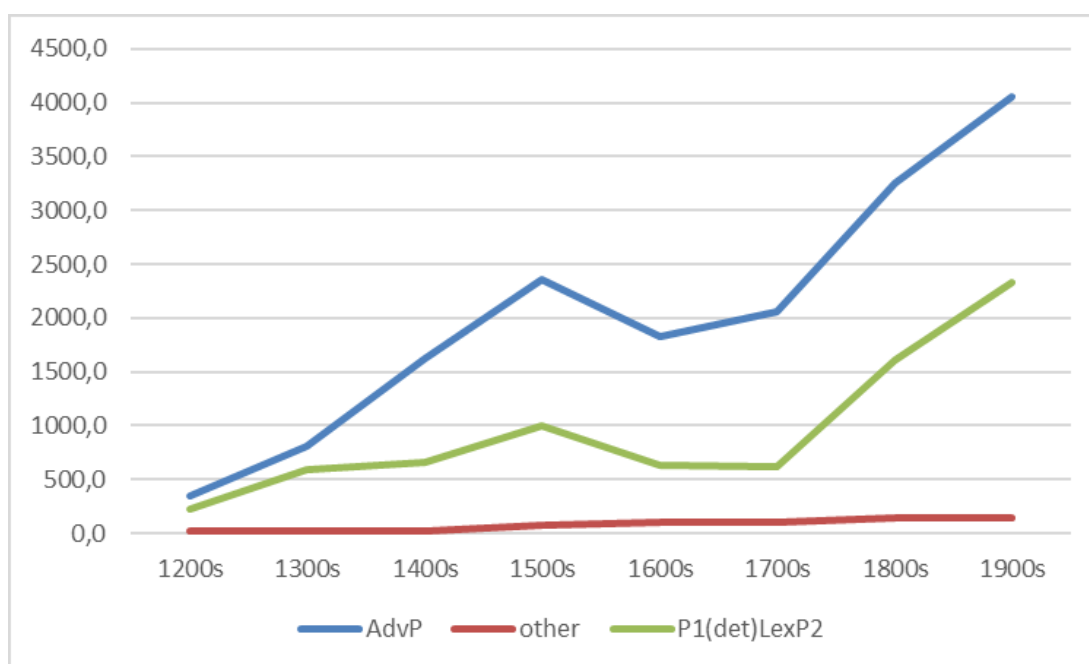
(15) bem se comprio esto em el Rei dom pedro. ca el mouido per sobeio queixume comtra seus Jrmaãos & outros do Reino por aazo da teemçom que tomaram **em fauor da** rrainha dona bramca & comtra os parentes de dona maria de padilha segumdo ouuistes.
 ‘This was well accomplished by King Dom Pedro, because he, moved by exaggerated grievance against his Brothers and others of the Kingdom on account of the dispute that they took in favor of the queen Dona Branca and against the parents of Dona Maria de Padilha, as you heard.’ (*Cronica de Dom Pedro*, Fernão Lopes, 15th c., Corpus do Português)

Our quantitative analysis suggests that the main pattern in terms of relative frequency is Pattern B, i.e., AdvP sequences. It is more frequent than all other patterns, throughout the diachrony of Portuguese (Table 2).

Relative Frequency, per type	1200s	1300s	1400s	1500s	1600s	1700s	1800s	1900s	Total
Type A1 (P1detLexP2)	29,0	103,3	181,5	329,2	222,1	197,3	772,5	777,0	596,6
Type A2 (P1LexP2)	190,6	489,3	483,0	664,7	415,2	420,6	840,0	1552,4	1054,7
Type B (AdvP)	350,3	803,8	1620,5	2356,9	1826,4	2054,2	3250,2	4061,5	3159,5
PP	14,5	17,5	24,0	69,0	79,2	72,5	62,2	38,7	49,7
Other	0	1,5	2,8	7,2	20,2	34,0	83,8	102,5	69,2
Total	584,4	1415,4	2311,8	3427,0	2563,3	2778,6	5008,8	6532,1	4929,7

Table 2. Relative frequency of specific patterns of Complex Prepositions in our corpus.

The frequency of this pattern, which appears clearly in Graph 1, is common to all Ibero-Romance languages, and distinguishes them from other Romance languages (Fagard 2010).



Graph 1. Relative frequency of Patterns A and B ([P1(det)LexP2] and [AdvP], respectively) *vs* other sequences, in our corpus.

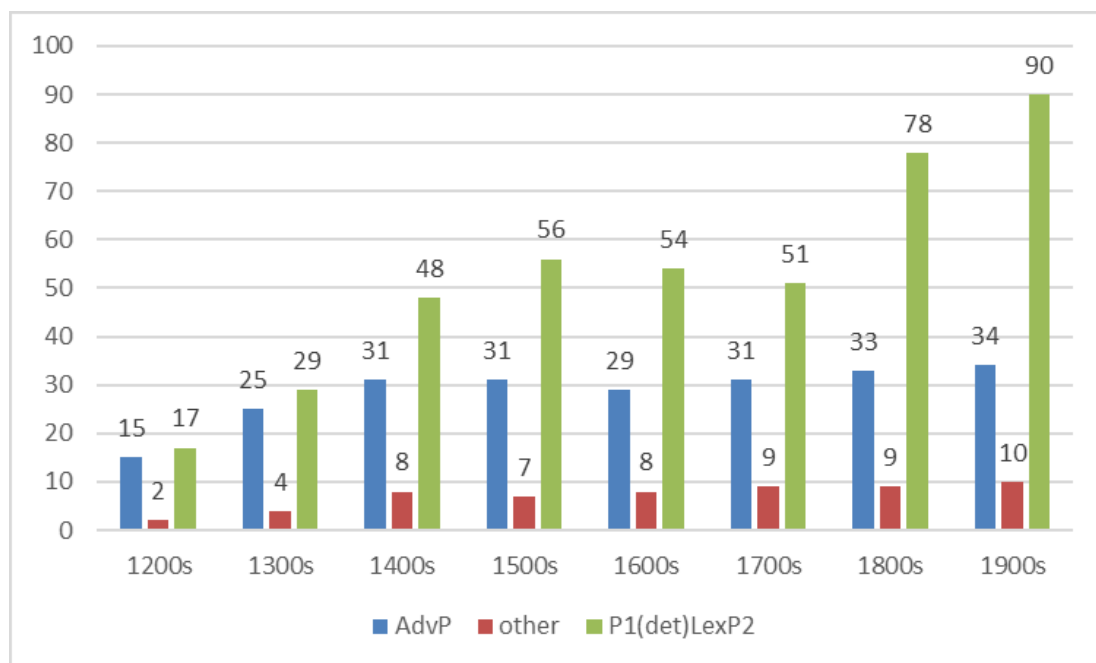
Another significant result is the emergence of new sequences throughout the Middle Ages, as observed in other languages, viz. French (Fagard 2020) or Italian (Piuino and Ganfi 2017).

As shown in Table 3, most new sequences first follow patterns A2 and B, then A1 and A2; by the end of the 17th century, over 68% of all CP-like sequences in the corpus are of type A (83/121).

Items per type	1200s	1300s	1400s	1500s	1600s	1700s	1800s	1900s	Total
Type A1 (P1detLexP2)	5	3	13	31	33	27	43	43	43
Type A2 (P1LexP2)	15	26	44	52	50	47	75	86	88
Type B (AdvP)	15	25	31	31	29	31	33	34	34
PP	2	3	4	4	4	4	3	4	4
Other	0	1	4	3	4	5	6	6	6
Total	37	58	96	121	120	114	160	173	175

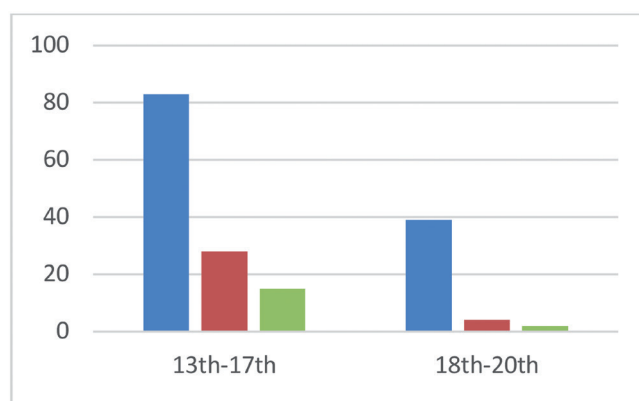
Table 3: CP-like sequences in the corpus, per type & century (relative frequencies per million words; Corpus do Português).

The cumulative frequency of Pattern A is similar to that of Pattern B at first, and gradually becomes twice as important (Graph 2).



Graph 2. Patterns A and B vs other types of CP-like sequences, per type & century (relative frequencies per million words; Corpus do Português).

The relative importance of pattern A tends to increase with time, as shown in Graph 3. This could indicate that the pattern itself is undergoing a process of grammaticalization (or constructionalization, i.e., the formation of entrenched form-meaning pairings; see Goldberg 1995, and for their use in a diachronic perspective Traugott and Trousdale 2013), and eventually started functioning as a fixed construction. This, in turn, explains the possibility of using the pattern to create new sequences as it seems *ex nihilo*, i.e., without a previous phase of grammaticalization, as shown by Hoffmann for English (2005). We have not yet identified such cases in the diachrony of Portuguese, but it would explain the uptake in frequency of the pattern; this remains an issue for future research.



Graph 3. Relative weight of patterns A and B vs other types of CP-like sequences in the diachrony of Portuguese (in % of all sequences identified as CP-like; Corpus do Português).

Minor sources of CPs which we have not included in Table 4 include sequences based on present participles, which may be considered as undergoing a grammaticalization process but not fully grammaticalized, e.g., *não obstante* ‘notwithstanding’. Such sequences may be shown to be functionally equivalent to prepositions or other CPs, e.g., *a despeito de* in the case of *não obstante*.

1st occurrence	AdvP	P ₁ detLexP ₂	P ₁ LexP ₂	other
1200s	<i>acerca de, acima de, adiante de, além de, dentro de, dentro em, depois de, diante de, fora de, fruto de, graças à, longe de, mercê de, quando de, quanto a</i>	<i>da parte de, na conta de, ao pé de, no meio de, na altura de</i>	<i>em nome de, em cima de, de cima da, em razão de, por força de, em lugar de, por cima de, por causa de, em redor de, com coita de, em prol de, em forma de, a mais de, em vez de, por parte de, em defesa de</i>	<i>de entre, de só</i>
1300s	<i>antes de, apesar de, após de, aquém de, até a, cerca de, detrás de, junto a, perto de, por entre</i>	<i>ao cabo de</i>	<i>em torno de, em busca de, em termos de, a par com, a par de, para cima de, em hora de, em virtude de, em pós de, por meio de, sem embargo de, sob pena de, por mercê de</i>	<i>para com, face a</i>

1400s	<i>abaixo de, através de, abaixo de, enquanto a, junto de, próximo de</i>	<i>ao longo de, ao caminho de, no caso de, na hora de, ao redor de, ao fim de, à hora de</i>	<i>para além de, a fim de, por diante de, por baixo de, a despeito de, de acordo com, para alguém da, em baixo de, a braços com, por abaixo de, em altura de, a propósito de, por detrás de, por dentro de, de frente de, em face de, em caso de, em favor de, por via de</i>	<i>dentre, devido a, não obstante, no que toca a</i>
1500s		<i>à conta de, na matéria de, à semelhança de, ao contrário de, ao encontro de, à disposição de, à volta de, ao cargo de, à margem de, ao lado de, à luz de, ao meio de, às custas de, à altura de, à mercê de, ao abrigo de, ao favor de, na sede de</i>	<i>a cargo de, a seguir a, a cabo de, em situação de, para baixo de, em frente de, em relação a, por conta de, em troca de, para lá de, em matéria de</i>	<i>a não ser</i>
1600s		<i>na base de, à beira de, à roda de</i>	<i>sem deixar de, em conformidade com, por perto de, a julgar por, em consequência de, em detrimento de, a título de</i>	
1700s	<i>relativamente a, frente a</i>	<i>ao nível de</i>	<i>a respeito de, a través de, por trás de</i>	<i>no que respeita a</i>
1800s		<i>à espera de, à frente de, à procura de, à excepção de, à direita de, à esquerda de, pela frente de, para a frente de</i>	<i>a favor de, a partir de, em vias de, a caminho de, de conformidade com, em meados de, de cara com, por defronte de, em função de, por volta de, com excepção de, em direcção a</i>	
1900s	<i>aquando de</i>		<i>a menos de, com base em, em fase de, a nível de, com vista a, em base de, em regime de, a conselho de</i>	

Table 4. CP-like sequences in the corpus, per type & century (Corpus do Português).

Once in place, in line with what could be expected in a grammaticalization or constructionalization perspective, these patterns seem to attract or license the emergence of similar sequences which may seem to fulfill only partially the required features. This could explain the existence of sequences in which the nature of the lexical nucleus is less than obvious. This is the case, for instance, of *mais* ‘more’ in *a mais de* ‘besides’.

5. Case Studies

In order to better understand the dynamics of this phenomenon, we focus here on three of these sequences, chosen because they illustrate different types of Complex Adpositions in the diachrony of Portuguese: *em redor de* ‘around, about’ is formed on a relational noun (also called *internal localization noun*, see Aurnague et al. 2000),³ which is typical of Type A CPs (see e.g., Stosic & Fagard 2019); *por parte de* ‘on the part of, on behalf of’ presents a case of CP based on a more abstract noun, which could explain the very different pattern of semantic shift it displays; finally, *ao longo de* seems to present a hybrid pattern, since it is based on a noun which is not relational, but does have a spatial meaning.

For each case study, we analyze the grammaticalization process underlying the emergence of the Complex Adposition, from a free syntactic sequence with a fully referential lexeme to a Complex Adposition.

5.1 *em redor de*

This sequence is quite representative, in that like many others it presents a high degree of internal variation, viz. there are several variants for P_I, but also for the lexical base itself, with no clear meaning shift, as can be seen in (16). In this paragraph, indeed, two distinct sequences – *em redor de* and *per d’arredor de* – are used to describe the same situation, in the same spatial layout (see above our example (13), which immediately precedes this excerpt).

(16) E cavalgarom ambos soos **per d’arredor da** villa
‘And the two of them rode alone **around** the town’ (*Crónica Geral de Espanha*, 1344, Corpus do Português)

3 Relational nouns sometimes grammaticalize into adpositions (Heine 1989; DeLancey 1997).

In Medieval Portuguese, this sequence remains mostly spatial, but a semantic shift eventually takes place, with a new use-type emerging which is quite present in the 20th c. (17-8) but seems to appear quite late: in this case, *ao redor de* introduces a quantity, and takes on a meaning of approximation.

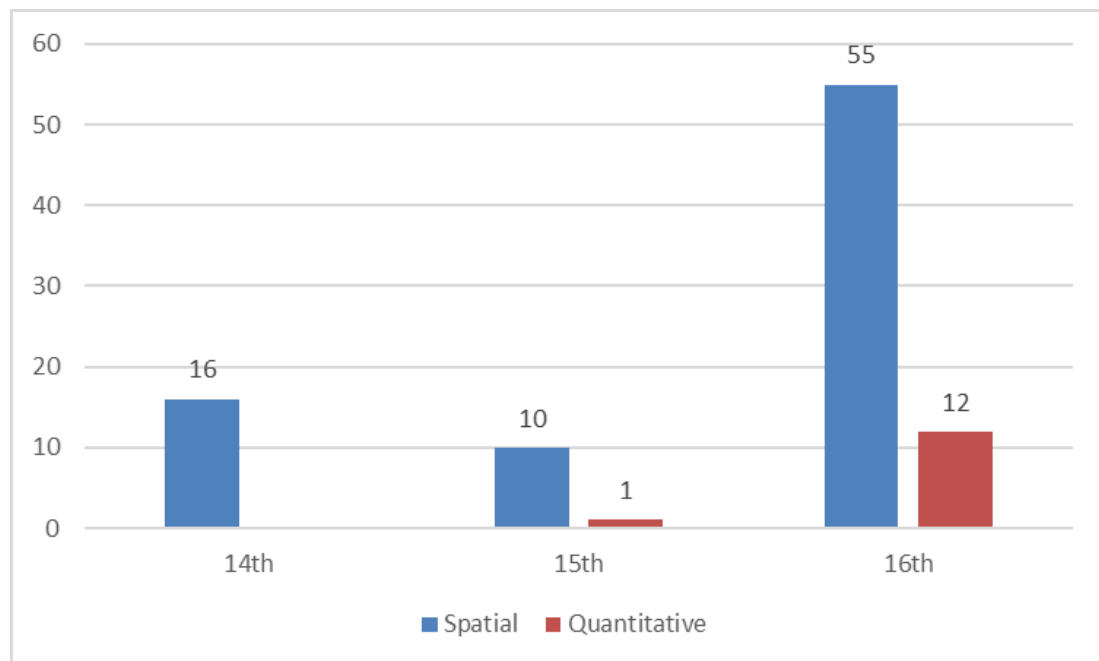
(17) ele tem o direito de durante seis meses levantar até vinte mil cruzeiros pagando um juro **ao redor de** dois por cento

‘he has the right to raise up to 20,000 Cruzeiros during six months, paying an interest rate of around two percent’ (Brazil, 20th c., Corpus do Português)

(18) O senador Jonas Pinheiro (PFL-MT) disse que a taxa foi de 16% quando a inflação estava em torno de 40% e, no ano passado, foi de 12% para uma inflação **ao redor de** 20%.

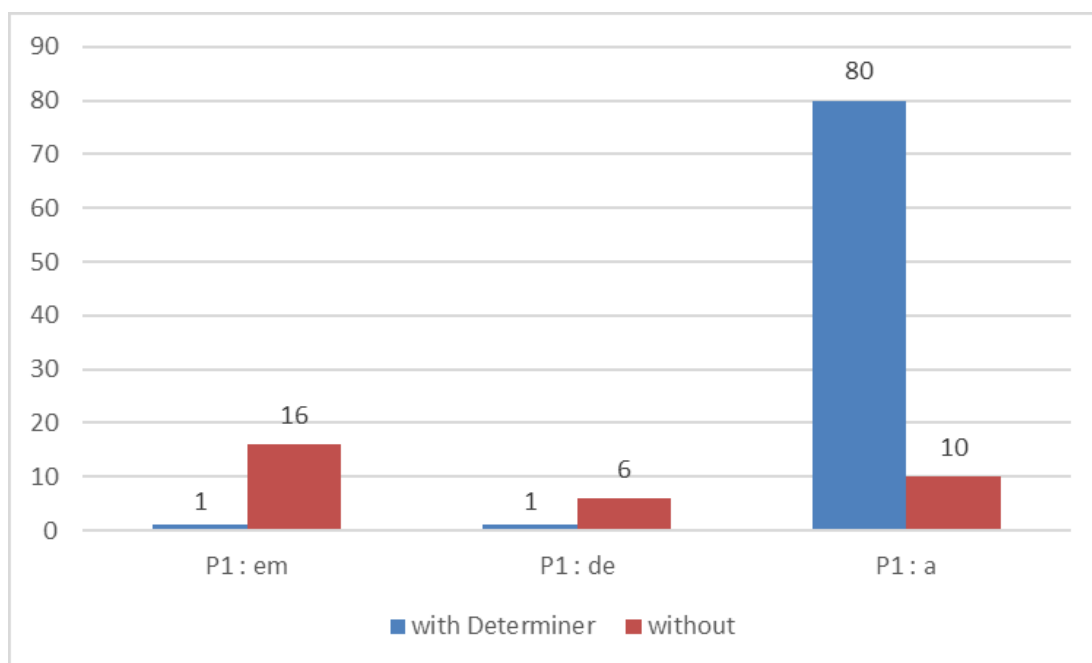
‘the senator Jonas Pinheiro (PFL-MT) said the rate was 16% when inflation was around 40% and last year at 12% with an inflation around 20%’ (*Governo antecipa normas para custeio da safra agrícola*, 16 de abril de 1997, Corpus do Português)

The emergence of quantitative uses seems gradual, as shown in Graph 4.



Graph 4: Spatial vs Quantitative uses of complex adpositions having *redor* as their lexical nucleus (Corpus do Português, 13th-15th c.)

Though the spatial use is found for various sequences, as illustrated in (19-20), the approximative use is found only for the sequence *ao redor de* (21), which is the most frequent variant (Graph 5).



Graph 5: Sequence types for complex adpositions having *redor* as their lexical nucleus (Corpus do Português, 13th-16th c.)

(19) Em este logar quero eu aver mynha sepultura. Mas, com todo esto, os Barboros, de que ja dissemos, nõ quedavã de dar guerra e correr toda a terra **em redor de** Cordova, de guisa que destroirom e estragarom Eceja e Carmona e toda a ribeira d'Alquyvir.

'In this place I wish to have my grave. But with all this the barbarians, as we have already said, did not cease to wage war and run all over the land around Cordoba, and as a result they destroyed Ecija and Carmona and all the river of Guadalquivir' (*Crónica Geral de Espanha de 1344*, 14th c., Corpus do Português)

(20) E o conde foisse entom pera sua terra com todolos outros cristããos que hy eram com elle. El rei, depois desto, mandou fazer a carcova **em redor da** cidade quando [sic] sayrom aa batalha que ouverom cõ Çolleima, assy como ja avemos dito.

'And the count then went to his land with all other christians who were there with him. After that, the king gave orders to build a ditch around the city when they left for the battle

that they fought against Suleiman, as we said earlier' (*Crónica Geral de Espanha de 1344*, 14th c., Corpus do Português)

(21) nunca deixou de aver dos nossos **ao redor de** - 200 - soldados feridos
'there were never less than about 200 of our soldiers wounded' (*Década oitava*, Diogo de Couto, 16th c., Corpus do Português)

In later centuries, another use-type arises, viz. temporal uses, as in (22).

(22) **ao redor do** meio-dia, apareceu o barbeiro da Lamosa, com um pifão, o desalmado, que até lhe caía a baba ao falar.
'Around midday appeared the barber of Lamosa, completely drunk, the wretched man, drooling as he talked' (Manuel Ribeiro, *A Planície Heróica*, 1927, Corpus do Português)

It should be noted that this shift from a spatial to a temporal use is common for spatial adpositions and adverbs (Haspelmath 1997; Heine and Kuteva 2002), including in Romance languages (Fagard 2010). The same goes for the semantic shift from spatial proximity to the more abstract notion of approximation, which is also found in other Portuguese adpositions, e.g., *cerca de* 'near, about'.

5.2 *por parte de*

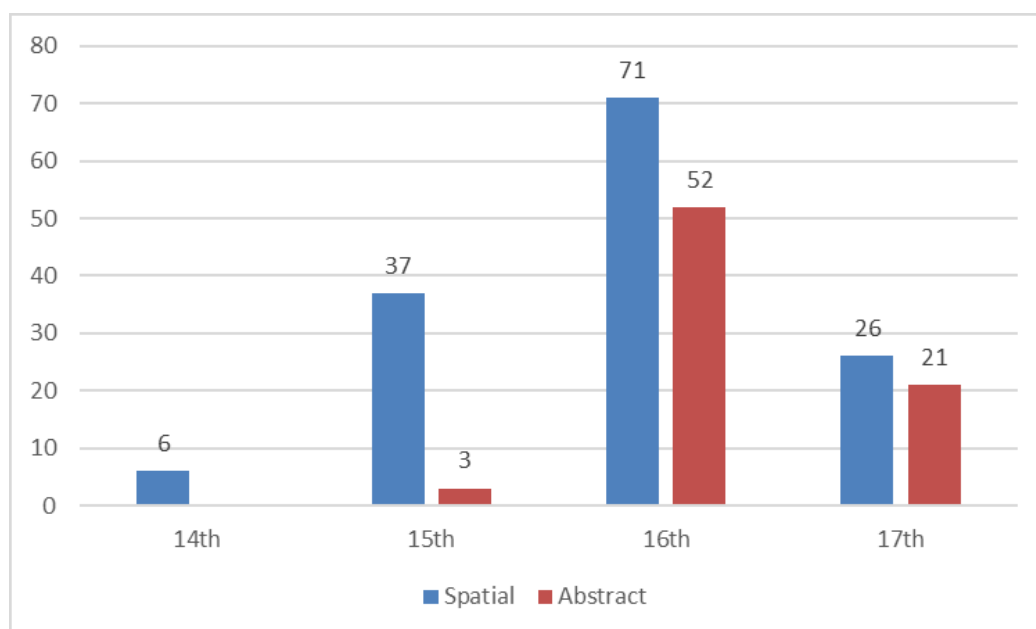
The sequence *por parte de* appears in the corpus from the 13th c. onward. In the first occurrences, the complement is always a human being, and the sequence typically introduces either the source of a command (e.g., complements of the verb *mandar* 'to order'), as in (23), or the origin of an inheritance, as in (24).

(23) E o dicto visitador mandou tomar a posse della **por parte da** hordem.
'and the above-mentioned visitor sent word to take possession thereof on account of the religious order' (*Tombos da Ordem de Cristo*, 1505-1509, Corpus do Português)

There is a gradual shift towards more abstract uses, as in (24); this shift is illustrated in Graph 6.

(24) teve sobre isso tantos contrastes e empedimento, **por parte do** induzimento dos mouros mercadores que ali resediam, peitando grossamente aos governadores da terra

‘he encountered many difficulties and hindrances, due to the instigation of the Moorish merchants residing there, fighting hard against the governors of the land’ (João de Barros, *Décadas da Asia* (*Década Terceira*, Livros I-X), 16th c., Corpus do Português)



Graph 6. Uses of *por parte de* in the corpus.

Once again, it should be noted that this shift from a spatial to a more abstract use is common for spatial adpositions and adverbs (Heine and Kuteva 2002), including in Romance languages (Fagard 2010).

5.3 *ao longo de*

The sequence *ao longo de* appears in the corpus from the 14th c. onward. In contrast to the two previous sequences, there seems to be a remarkable semantic stability over a long period, with exclusively spatial uses from the 14th to the 18th c., in our corpus. In these uses, *ao longo de* introduces a spatial complement describing the extent of a motion or of a static scene, as in (25).

(25) Jmdo do nos asy **ao longo da** costa por rrespeito [d]o vento que era pouco, com o vento da terra pera o orar e a viraram pera terra, de dia com a calma lançavamos anquoras.

‘As we went thus along the coast on account of the gentle wind, with the wind coming from the land in order to pray and turn it towards the land, during the day as all was calm we

cast the anchors.’ (*Diário da viagem de Vasco da Gama (O descobrimento das Índias)*, 1498, Corpus do Português)

However, in the 20th c., temporal uses appear, as in (26).

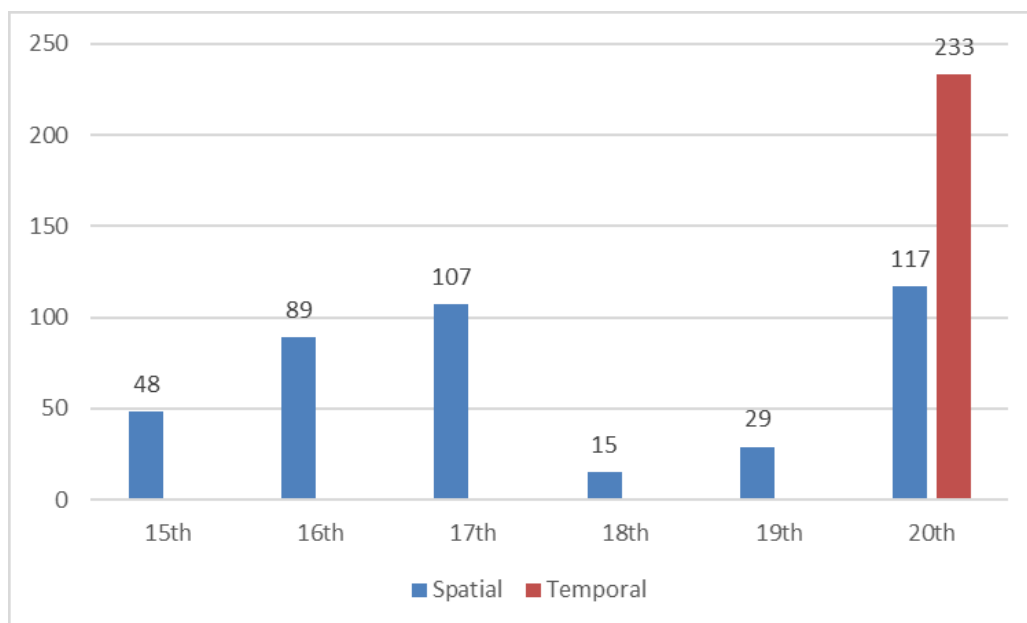
(26) A loucura é um problema que enfrentamos **ao longo de** toda a vida.
‘Madness is a problem we have to face through all our life’ (Ivan Angelo, 12 julho 1997, Corpus do Português)

The semantic shift may be illustrated by cases in which the relation can be seen as both spatial and temporal, as in (27), which constitutes another instance of bridging context.

(27) a obra Footstone, constituída por cinco montes de pedras de várias regiões, executada **ao longo de** uma caminhada de 126 milhas que o levou da costa irlandesa à linha costeira do mar do Norte.

‘the creation Footstone, made from five mounts of rocks from various regions, executed during a hike of 126 miles which led him from the coast of Ireland to the coast of the North Sea.’ (Richard Long, 20th c., Corpus do Português)

This semantic extension, which takes place very late in the diachrony of Portuguese, is quite abrupt, as illustrated in Graph 7.



Graph 7. Uses of *ao longo de* in the corpus.

As mentioned above, this shift from a spatial to a temporal use is common for spatial adpositions and adverbs, and found for instance, in Romance, for Italian *lungo* ‘along’ or French *au long de* ‘along’.

6. Conclusion

In this paper, we show that CPs appear early on in Portuguese: as early as the 13th c., with a constant renewal of the category, as new items emerge.

These new items typically follow one of two main patterns of CPs. One pattern is emblematic of Ibero-Romance; it involves an adverb followed by the preposition *a* ‘at’ or *de* ‘of’. This pattern is more frequent but less productive overall than the others.

The other pattern is more typical of CPs in Romance, and includes a lexical root linked to the rest of the sentence by two simple adpositions, generally *em* ‘in, on’, *a* ‘at’, *para* ‘to’ or *por* ‘by, through’ and *de* ‘of’.

Though the possibility of influence via language contact is not excluded, and can even be shown to have operated in the case of complex prepositions, our results show that these patterns are highly productive throughout the diachrony of Portuguese, which suggests that language-internal phenomena are at play.

Bibliography

- Aurnague, Michel, Boulanouar, Kader, Nespoulous, Jean-Luc, Borill, Andrée and Borillo, Mario. 2000. "Spatial Semantics: The Processing of Internal Localization Nouns." *Cahiers de Psychologie Cognitive – Current Psychology of Cognition* 19, no.1, 69-110.
- Bechara, Evanildo. 2002. *Moderna gramática portuguesa*. Rio de Janeiro: Editora Nova Fronteira.
- Bonnet, Guillaume and Fagard, Benjamin. 2020. "Complex Prepositions in Albanian: a first assessment." In *Complex Adpositions in European Languages*, edited by Benjamin Fagard, José Pinto de Lima, Elena Smirnova and Dejan Stosic, 265-98. Berlin and New York: De Gruyter.
- Brøndal, Viggo. 1950. *Théorie des prépositions. Introduction à une sémantique rationnelle*. Copenhagen: Munksgaard.
- Cifuentes Honrubia, José Luis. 2003. *Locuciones prepositivas – sobre la gramaticalización preposicional en español*. Alicante: Universidad de Alicante.
- Cunha, Celso and Cintra, Lindley. 1984. *Nova gramática do Português contemporâneo*. Lisbon: João Sá da Costa.
- Davies, Mark and Ferreira, Michael J. 2006. "Corpus do Português." <https://www.corpusdoportugues.org/> (accessed: 01/11/2023).
- DeLancey, Scott. 1997. "Grammaticalization and the Gradience of Categories: Relator Nouns and Postpositions in Tibetan and Burmese." In *Essays on Language Function and Language Type*, edited by Joan Bybee, John Haiman and Sandra A. Thompson, 51-69. Amsterdam and Philadelphia: John Benjamins.
- Di Meola, Claudio. 2000. *Die Grammatikalisierung deutscher Präpositionen*. Tübingen: Stauffenburg Verlag.
- Fagard, Benjamin. 2010. *Espace et grammaticalisation – L'évolution sémantique des prépositions dans les langues romanes*. Sarrebruck: Editions Universitaires Européennes.
- Fagard, Benjamin. 2020. "Prépositions et locutions prépositives." In *Grande Grammaire Historique du français*, edited by Christiane Marchello-Nizia, Ber-

nard Combettes, Sophie Prévost and Tobias Scheer, 856-86. Berlin and New York: De Gruyter.

Buzaglo Paiva Raposo, Eduardo, Bacelar do Nascimento, Maria Fernanda, Coelho da Mota, Maria Antónia, Segura, Luísa and Mendes, Amália. 2013. *Gramática do Português*, 2 vols. Lisbon: Fundação Calouste Gulbenkian.

Fagard, Benjamin. 2023. “Les prépositions complexes dans les langues romanes: évolution parallèle et contact linguistique.” In *Les prépositions complexes en français : théories, descriptions, applications*, edited by Dejan Stosic, Myriam Bras, Chiara Minoccheri and Océane Abrard, 137-56. Paris: L’Harmattan.

Fagard, Benjamin, Lima, José Pinto de and Stosic, Dejan. 2019. *Les prépositions complexes dans les langues romanes*. Special Issue of *Revue Romane: Revue Romane* 54, no. 1.

Fagard, Benjamin, Lima, José Pinto de, Smirnova, Elena and Stosic, Dejan. 2020. *Complex Adpositions in European Languages. A Micro-Typological Approach to Complex Nominal Relators*. Berlin and New York: De Gruyter.

Fagard, Benjamin, Stosic, Dejan and Lima, José Pinto de. 2020. “Complex adpositions in Romance Languages.” In *Complex Adpositions in European Languages*, edited by Benjamin Fagard, José Pinto de Lima, Elena Smirnova and Dejan Stosic, 33-64. Berlin, New York: De Gruyter.

Goldberg, Adele E. 1995. *A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

Hagège, Claude. 2010. *Adpositions*. Oxford: Oxford University Press.

Haspelmath, Martin. 1997. *From Space to Time – Temporal Adverbials in the World’s Languages*. München and Newcastle: Lincom Europa.

Heine, Bernd and Kuteva, Tania. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.

Heine, Bernd. 2002. “On the Role of Context in Grammaticalization.” In *New Reflections on Grammaticalization*, edited by Ilse Wischer and Gabriele Diewald, 83-101. Amsterdam and Philadelphia: John Benjamins.

Heine, Bernd. 1989. “Adpositions in African languages.” *Linguistique Africaine* 2, 77-127.

- Hoffmann, Sebastian. 2005. *Grammaticalization and English Complex Prepositions. A corpus-based analysis*. London and New York: Routledge.
- Horst, Joost van der. 2013. *Taal op drift: lange-termijnontwikkelingen in taal en samenleving*. Amsterdam: Meulenhoff.
- Hüning, Matthias. 2014. "Over Complexe Preposities en Convergentie." In *Paatroon en argument. Een dubbelfestbundel bij het emeritaat van William Van Belle en Joop van der Horst*, edited by Freek Van de Velde, Hans Smessaert, Frank Van Eynde and Sara Verbrugge, 433-45. Leuven: Universitaire Pers Leuven.
- Lima, José Pinto de. 2014. *Studies on Grammaticalization and Lexicalization*. Munich: Lincom Europa.
- Lima, José Pinto de. 2019a. "On Grammaticalized Complex Prepositions in Portuguese: Deployment, Shift, Redundancy, Complementation." *Revue Romane* 54, no. 1, 126-40.
- Lima, José Pinto de. 2019b. "Preposições complexas em Português." In *Estudos linguísticos e filológicos oferecidos a Ivo Castro*, edited by Ernestina Carrilho, Ana Martins, Sandra Pereira and João P. Silvestre, 1181-211. Lisbon: Centro de Linguística da Universidade de Lisboa.
- Nunes, José Joaquim. 1930. *Gramática histórica portuguesa*. Lisbon: Clássica.
- Piot, Mireille. 1993. "Les connecteurs du français." *Linguisticae Investigationes* 17, no. 1, 141-60.
- Piunno, Valentina and Ganfi, Vittorio. 2017. "Preposizioni complesse in italiano antico e contemporaneo. Grammaticalizzazione, schematismo e produttività." *Archivio Glottologico Italiano* 102, no. 2, 184-204.
- Pottier, Bernard. 1962. *Systématique des éléments de relation. Étude de morpho-syntaxe structurale romane*. Paris: Klincksieck.
- Rottet, Kevin J. 2020. "Complex Prepositions in Breton." In *Complex Adpositions in European Languages*, edited by Benjamin Fagard, José Pinto de Lima, Elena Smirnova and Dejan Stosic 195-232. Berlin and New York: De Gruyter.
- Seppänen, Aimo, Rhonwen Bowen and Trotta, Joe. 1994. "On the So-Called Complex Prepositions." *Studia Anglica Posnaniensia* 29, 3-29.

Stosic, Dejan and Fagard, Benjamin. 2019. “Les prépositions complexes en français. Pour une méthode d’identification multicritère.” *Revue Romane* 54, no. 1, 8-38.

Traugott, Elizabeth C. and Trousdale, Graeme. 2013. *Constructionalization and Constructional Changes*. Oxford: Oxford University Press.

Benjamin Fagard is a linguist at the CNRS (France), currently Visiting Scholar at Indiana University (Bloomington, Indiana). The main focus of his research is on semantic change in the course of grammaticalization processes, studied in diachrony and with a contrastive approach, mainly on French and other Romance languages. His publications include papers on grammaticalization phenomena (simple and complex adpositions, conjunctions, discourse markers), on motion in a typological perspective, and on the dynamics of language evolution.

José Pinto de Lima is assistant professor at the University of Lisbon. His main fields of interest are linguistic pragmatics, grammaticalization and language change, particularly concerning European Portuguese and research on complex prepositions.