

Benjamin Fagard, Dejan Stosic & José Pinto de Lima

Complex adpositions in Romance: Emergence and variation

Abstract

In this paper, we describe the systems of complex adpositions in modern Romance languages, and discuss their status and definition. We do so with a corpus-based analysis focusing on three Romance languages, French, Portuguese and Romanian, and show that some features of complex adpositions are common to most if not all varieties, while some seem to be specificities of one or more of these languages. One case in point is the way such constructions can be tested, for instance with possessives – a test which does not seem to affect French and Portuguese (Catalan, etc.) CAs in the same way: the clear contrast found in French between non-lexicalized constructions and CAs on account of their intolerance of the possessive is not to be found in Portuguese. Finally, we formulate a hypothesis as to the development of CAs in various Romance languages, suggesting it may not have been parallel.

Keywords

Romance languages, corpus studies, contrastive linguistics, grammaticalization, lexicalization

Benjamin Fagard

CNRS, ENS & Paris Sorbonne Nouvelle; PSL

Lattice laboratory, Ecole Normale Supérieure, 1 rue Maurice Arnoux, 92120 Montrouge, France

benjamin.fagard@ens.psl.eu

Dejan Stosic

University of Toulouse Jean Jaurès & Laboratory CLLE (CNRS & UT2J)

5, allées Antonio Machado, 31058 Toulouse Cedex 9, France

dejan.stosic@univ-tlse2.fr

José Pinto de Lima

Universidade de Lisboa, Faculdade de Letras

Rua Professor Simões Raposo, nº 20, 3-B, 1600-662 LISBOA, Portugal

pintodelima@outlook.pt

1 Introduction

This article aims at describing formal and functional properties of complex adpositions in Romance languages. The paper is organized as follows: we provide an overview of Romance languages (sections 2.1 and 2.2) as well as of the methodology employed in our study (Section 2.3). We propose functional and formal definitions of simple and complex adpositions (Section 3.1); we briefly sketch an overview of complex adpositions in Romance languages (Section 3.2), and describe other possible uses of these constructions (3.3). We go on to describe the various patterns of complex adpositions, focusing first on the main patterns (4.1), then on existing subtypes (4.2), before addressing issues of frequency and productivity (4.3). We then provide a diachronic account of the emergence of complex adpositions in French (5.1-5.3). We go on to examine variation in Romance languages, showing both commonalities (6.1) and divergence (6.2).

2 Language description

2.1 Romance languages – sociolinguistic and geographical features

Romance languages are a subgroup of the vast family of Indo-European languages. They originate from Latin, but have been for centuries under strong influence from other languages, both Indo-European and non-Indo-European, mainly Germanic, Arabic, Slavic and Celtic. They are mostly spoken in Europe (where all Romance languages are attested), in the Americas (mainly French, Spanish and Portuguese), and in Africa (mainly French and Portuguese), but there are also smaller, isolated communities elsewhere. The total number of mother-tongue speakers is around 700 million, among which around 200 million in Europe (by *Ethnologue*¹ counts, which we compiled). There are also Creoles based on Romance languages, with a total of around 11 million speakers in various parts of the world, from the Caribbean to the Indian Ocean; most speakers use a French-based Creole (around 10 million), followed by Portuguese-based (700,000 speakers) and Spanish-based Creoles (300,000 speakers). However, from a typological point of view, Creoles in general, and even Romance-based ones, are rather homogeneous, and very different from European Romance, and it does not seem justified to address them jointly: we believe Romance-based Creoles deserve a separate, dedicated study. For questions of simplicity, we shall henceforth use the term “Romance” to refer exclusively to European Romance languages, to the exclusion of Romance-based Creoles, and focus on European varieties.

Table 1: Number of speakers for the main Romance languages (in millions)

Variety	Number of speakers in the country of origin: all users (<i>L1 only</i>)	Total number of speakers: all users (<i>L1 only</i>)
Romansch	0.04	0.04
Occitan	0.5 ²	2

¹ In its web version (www.ethnologue.com, accessed November 19th, 2018). Note that some numbers seem low; for French, for instance, according to the ‘Organisation internationale de la francophonie’, the actual number of native speakers in the world is 115 million, with 274 million speakers in total (Wolff (ed), 2014). The numbers in Table 1 should therefore be taken with caution; however, they do provide a rough idea of the number of speakers and the relative (numerical) importance of each Romance language.

² According to Sibille (2010), the number of speakers could be estimated at 700,000 based on data gathered in 1999. The actual number of speakers of Occitan is actually very difficult to evaluate, on account of the complex socio-linguistic situation – Occitan is a mosaic of dialects with no or little standard form, spoken mainly in

Galician	2.3	2.3
Catalan	8.8 (3.7)	9.2 (4.1)
Romanian	18.7	23.4
Italian	59	67.8 (64.8)
French	63.2 (60.2)	285 (76.8)
Portuguese	9.9	236.5 (222.7)
Spanish	46.4 (42.7)	513 (442.4)

2.2 Romance languages – major typological characteristics

Romance languages are by and large inflectional as far as verbs are concerned, with a heavy Tense-Aspect-Modality system, but not so much for nouns and adjectives, which display variable gender and number marking and no case marking, with the exception of Romanian. Romance languages generally have preposed definite and indefinite articles. They have a large set of prepositions and conjunctions, which sometimes partly overlap, as in ‘before (prepositional)’ vs ‘before (conjunction)’: French *avant de / avant que*, Italian *prima di / prima che*, Portuguese *antes de / antes que*, etc. (see Section 3.3). French differs from “Central Romance” on a number of points besides word order and pronouns; it displays for instance fully grammaticalized analytic future and past tenses, as well as two distinct paradigms for demonstratives (demonstrative pronouns vs. demonstrative determiners). Romanian, on the other hand, presents a set of distinctive Balkan features. For the nominal sphere, the most notable features are the postposed definite article (e.g. *lup.ul* ‘wolf.DET.NOM/ACC’) and a residual set of nominal case markers with genitive-dative fusion (e.g. *lup.ul.e* ‘wolf.DET.VOC’, *lup.ul.ui* ‘wolf.DET.GEN/DAT’). This specificity of French on the one hand (see e.g., Koch 2002) and Romanian on the other can also be observed in the case of complementizers: while in ‘Central’ Romance there is a two-way distinction (*if* vs *that* complementizers), with a clear opposition between indicative and subjunctive complementation, Romanian is set apart by its three-way distinction and French by the higher degree of grammaticalization of the subjunctive (Fagard et al. 2016).

2.3 Methodology

In the following sections, we investigate formal and functional aspects of Romance complex adpositions. In order to investigate their uses, we combined different methodologies.

In a first phase, we based our hypotheses on the existing literature (for instance Borillo 2000, 2002, Cifuentes Honrubia 2003, Hoffmann 2005, Leeman 2007, Fagard and De Mulder 2007, Kurzon and Adler 2008, Lima 2014, 2019a, 2019b, Stosic to appear), and on traditional philological instruments, i.e. dictionaries and grammars (see the Reference section)³. In a second phase, we relied on corpora to check the validity of our results. The use of corpora for determining how speakers make use of language is now quite well established, as one can see from the flourishing subfield of corpus-based grammars (e.g., for English: Biber et al. 1999; for French: Abeillé, Godard and Delaveau in press, and in diachrony Marchello-Nizia et al. in prep. etc.). It is based on the idea that linguistic systems are best viewed as language(s) in use,

France, Italy and Spain. This estimate goes well beyond Bernissan’s (2012) estimate, which seems very low compared to previous estimates (generally around 2 million speakers, see e.g. the Euromosaic report in 1996).

³ Grammars do not systematically provide lists of complex adpositions, and even when they do, they are typically quite limited. However, there is a tendency in the last decade for grammars to provide better accounts of Complex Adpositions. For instance, for Portuguese, the grammar of Raposo et al. lists 50 items or so (in a 3-page description, Raposo et al. 2013, 1503-1506); the *Nueva gramática de la lengua española* (2009) is another good example, with a fair section on complex adpositions (2276-2283).

and that grammar is shaped by frequency effects (see e.g., Bybee and Hopper 2001, Bybee 2010) – in other words, it seems to us that it is more interesting to look at language use and find out from there what rules the speakers are following than to rely exclusively on introspection. The corpora we used are the following (links to the corresponding databases are given in the Reference section):

- French: Valibel, PFC, Clapi, Frantext, frTenTen;
- Italian: Badip, itTenTen;⁴
- Portuguese: Corpus do Português (Davies and Ferreira 2006), ptTenTen11, CRPC (*Corpus de Referência do Português Contemporâneo*), CETEM-Público;
- Romanian: CORV (Dascălu-Jinga 2002), Ruxândoiu, Romanian web corpus;
- Spanish: CREA, CORDE, Corpus del Español (Davies 2002), esTenTen.

We believe that these ‘cross-examinations’ guarantee the validity of our results.

3 General description of Romance complex adpositions

3.1 Nominal relation marking and its syntactic realizations in Romance

Before giving an overview of Romance complex adpositions, we briefly describe what we mean by nominal relation marking. We define it as the situation that arises when a noun phrase, or another constituent with similar distributional properties, is governed by a higher-ranking constituent, or by the sentence as a whole.

This situation has different syntactic realizations in Romance languages: case marking (only in Romanian) (1); simple adposition (2); complex adposition (3-7). Let us examine them in detail.

1. tatăl Anei (Romanian)
father-DET Ann-GEN
‘Ann’s father’

2. Le père d’Anne (French)
DET.M.SG father of.Ann
‘Ann’s father’

3. Il habite près de l’ étoile
he live.PRS.IND.3SG near to DET.F.SG star
‘He lives near the “place de l’étoile”’

4. o João sentou-se ao pé da rapariga (Portuguese)
DET John sit.PST.IND.3SG at.DET.M.SG foot of.DET.F.SG girl
‘John sat down by the girl’

5. Il vient au-devant de nous (French)
he come.PRS.IND.3SG at.DET.M.SG-front to us

⁴ For Italian, see also Piunno and Ganfi (2019).

6. May nu intenționează să demisioneze, **în pofida** rezultatului (Romanian)
 May NEG intend.PRS.IND.3SG COMP resign.PRS.SBJ.3SG, in spite result.DET.DEF.M.GEN.SG
 conservatorilor în alegerile pe care le-a convocat
 conservative.GEN.PL in election.GEN.PL for which it-AUX.PRS.IND.3SG summon.PST.PTCP
 ‘May does not intend to resign despite the results of the conservatives in the elections she summoned’
7. **În pofida** a ceea ce cred susținătorii unui (Romanian)⁶
 in spite at that COMP.REL believe.PRS.IND.3PL supporter.NOM.PL one.GEN.SG
 Grexit, devalorizarea monetară este doar o iluzie
 Grexit, devaluation.DET.DEF.F.NOM.SG monetary.F.SG be.PRS.IND.3SG only DET.INDF.F.NOM.SG illu
 ‘In spite of what the supporters of Grexit believe, monetary devaluation is only an illusion’

In (1), the genitive case marks the relation between the proper noun *Ana* “Ann” and the NP *tatăl* “the father”. The same relation is marked in (2) by the simple adposition *de* “of”, realized as *d’*. In these two examples, the genitive case and the adposition *de* can be said to have the same function. The same can be said of English *of*, German *von*, with an alternation in both languages between case marking and prepositional marking of the genitive.

Similarly, whereas e.g. Modern French and Portuguese use the preposition *de* to introduce the complement of various types of complex adpositions (3-5), Romanian may use *de* but it alternates with case marking and the use of *a* (6-7). This complex alternation is also found in Old French, with three constructions used to express the genitive: case marking and the adpositions *à* and *de* (e.g. *la mort le roi* ‘the king’s death (lit. *the death the king*, with an oblique case marking on *king*)’ vs *la mort du/au roi* ‘the king’s death (lit. *the death of.the/at.the king*)’).

In (8), there are two distinct adpositions, each marking a different semantic (and syntactic) relation. The adposition *de* introduces the argument of the verb *vient* “comes” and marks it as one type of argument, as shown by example (9).

8. Il [vient **de** [chez sa mère]] (French)
 he come.PRS.IND.3SG from at his.F.SG mother
 ‘He comes from his mother’s place’
9. Il vient **à** Paris (French)
 he come.PRS.IND.3SG at Paris
 ‘He comes to Paris’

In contrast, the adposition *chez* establishes a type of spatial relation between the general motion event (*venir de* ‘comes from’) and the following NP *sa mère* ‘his mother’. The two adpositions are thus independent from each other, and can vary independently, as shown in (10-11).

⁵ <https://www.news.ro/externe/may-nu-intentioneaza-sa-demisioneze-in-pofida-rezultatului-conservatorilor-inalegerile-pe-care-le-a-convocat-1922403609002017061017026254>, accessed June 6th, 2017.

⁶ <http://politeia.org.ro/stiri-revista-presei/revista-presei-19-iunie-mitomanul-se-blindeaza-cu-ajutorul-lui-nitu-sial-lui-oprea/38748/>, accessed June 6th, 2017.

10. Il vient **chez** sa mère (French)
 he come.PRS.IND.3SG at his.F.SG mother
 ‘He comes to his mother’s place’
11. Il vient **de** Rome (French)
 he come.PRS.IND.3SG from Rome
 ‘He comes from Rome’

Conversely, Fr. *près de*, Pt. *ao pé de* and Fr. *au-devant de* in examples (4-6) seem to be distributional equivalents of simple adpositions, the main difference being that they are morphologically complex. The fact that there is a functional and more specifically a semantic equivalence between such constructions and simple adpositions has been well documented e.g. for French (Melis 2003: 112-114) and Portuguese (Lima 2014, 2019a, 2019b). However, corpus studies show that this equivalence is not complete; for instance, the polysemy of complex adpositions seems much more limited than that of simple adpositions, and they have few or no purely grammatical uses (Fagard and De Mulder 2007, Fagard 2009, 2012).

Finally, Rom. *în pofida* (+ genitive marking of the dependent noun) in example (6) shows similar distributional and semantic properties, as is shown by the existence of the variant *în pofida a* (7) which corresponds more closely to the complex adposition in (6); the difference is the use of case marking (genitive) or adposition (*a* ‘at’) to establish the relation with the NP.

3.2 Canonical and non-canonical complex adpositions

Following the definition given in the introductory chapter, we define a complex adposition as having the same functions and distributional properties as simple adpositions (a-c), but with different formal characteristics (d-f).

The **typical** functions of simple adpositions are as follows – they:

- a. introduce a constituent
- b. establish a relation between this constituent and a higher constituent, or the sentence as a whole
- c. contribute to determining the semantic nature of this relationship

Complex adpositions are functionally equivalent to simple ones, i.e. their typical functions also correspond to features (a-c), but, additionally:

- d. they are MORPHOLOGICALLY complex
- e. canonically, functions (a-c) are performed by different components:
 - i. a simple adposition (or case marker) introduces the constituent
 - ii. another simple adposition (or case marker) establishes the relation with the higher constituent
 - iii. a lexical nucleus determines the semantic nature of the relation

Adpositions fulfill three different functions, and complex adpositions are – by definition! – complex, i.e. made up of different lexemes or morphemes. Therefore, as illustrated in Table 2,

a *canonical* complex adposition would be one in which each one of the three functions of adpositions (a-c) is realized by a different component, while a non-canonical complex adposition would be one in which there is no such one-on-one mapping.

Table 2: Canonical vs non-canonical complex adpositions

		Function			meaning/gloss
		(b)	(c)	(a)	
'canonical'	(Portuguese)	<i>ao</i>	<i>pé</i>	<i>de</i>	'next to (lit. <i>at.the foot of</i>)'
	(French)	<i>à</i>	<i>côté</i>	<i>de</i>	'next to (lit. <i>at side of</i>)'
	(French)	<i>en</i>	<i>dessous</i>	<i>de</i>	'beneath (lit. <i>in underneath of</i>)'
	(Portuguese)	<i>por</i>	<i>debaixo</i>	<i>de</i>	'beneath (lit. <i>by underneath of</i>)'
	(Portuguese)	<i>ao</i>	<i>lado</i>	<i>de</i>	'next to (lit. <i>at.the side of</i>)'
	(Portuguese)	<i>em</i>	<i>cima</i>	<i>de</i>	'above, on (lit. <i>on top of</i>)'
	(Portuguese)	<i>em</i>	<i>vez</i>	<i>de</i>	'instead of (lit. <i>in turn of</i>)'
	(Portuguese)	<i>em</i>	<i>relação</i>	<i>a</i>	'in relation to, concerning (lit. <i>in relation at</i>)'
	(Portuguese)	<i>por</i>	<i>meio</i>	<i>de</i>	'by means of (lit. <i>through means of</i>)'
	(Portuguese)	<i>por</i>	<i>causa</i>	<i>de</i>	'because of (lit. <i>through cause of</i>)'
'non-canonical'	(Portuguese)	<i>graças</i>		<i>a</i>	'owing to (lit. <i>thanks to</i>)'
	(French)	<i>grâce</i>		<i>à</i>	
	(Portuguese)	<i>quanto</i>		<i>a</i>	'concerning (lit. <i>how.much at</i>)'
	(French)	<i>quant</i>		<i>à</i>	
	(Portuguese)	<i>no</i>	<i>que respeita</i>	<i>a</i>	'concerning (lit. <i>in.the what concerns at</i>)'
	(French)	<i>en</i>	<i>ce qui concerne</i>		'concerning (lit. <i>in that which concerns</i>)'
	(French)	<i>à</i>	<i>travers</i>		'through (lit. <i>at transverse</i>)'

Complex adpositions may thus be non-canonical because there is no specific item dedicated to the introduction of a complement (e.g. Fr. *en ce qui concerne, à travers*), or no item establishing the relation to a higher constituent (e.g. Pt. *graças a, quanto a*). It can also be on account of the complexity of one of the slots, as in Pt. *no que respeita a*.

3.3 Other uses of complex adpositions

Simple adpositions often have other uses than adpositional ones, for instance subordination (9) (cf. Section 2.2). While this is sometimes taken to be a natural extension of their uses, to the point that Pottier (1962: 195-197) considers that e.g. French *avant* should be analyzed as one morpheme with different uses, depending on the context: adverb, preposition, conjunction. In (12-14), he would thus not posit the existence of an adposition, a conjunction and an adverb with similar form and meaning (*depuis* ‘since’), but of only one morpheme *depuis* with three different uses.

12. il est changé **depuis** la guerre (Goncourt, *Journal*, 1896)
 it be.PRS.IND.3SG change.PST.PTCP since the war
 ‘it has changed since the war’

13. il est changé **depuis que** la guerre a commencé
 it be.PRS.IND.3SG change.PST.PTCP since COMP the war AUX.PRS.IND.3SG start.PST.PTCP
 ‘it has changed since the war started’

14. il est changé **depuis**
 it be.PRS.IND.3SG change.PST.PTCP since
 ‘it has changed since’

One limitation of this view is that it does not account for the fact that despite the frequent polyfunctionality of such morphemes, be it within a given language or across languages (cf. the other chapters of this volume, which show that uses of adpositions as adverbs or conjunctions are well represented across Europe), it is not universal. Some items have uses as adposition and adverb but not as conjunction, for instance Modern French *devant* ‘in front (of)’; others have all three uses but with differences in form, e.g. Italian *prima* and Portuguese *antes* ‘before’, which cannot be used as adpositions without a functional *di / de* ‘of’ (*prima / prima che / prima di, antes / antes que / antes de*). What is true, however, is that adverbs, conjunctions and adpositions do seem functionally similar. To put it simply, an adposition links a sentence with a noun phrase, a conjunction links a sentence with a subclause, and it could be said that an adverb just anaphorically incorporates the noun phrase: in (11), *depuis* can stand for ‘since the war’ if it is contextually salient.

Though it is not systematic, there seems to be a similar feature in complex adpositions: there are frequent formal correspondences between complex adpositions and complex conjunctions, for instance *au lieu de / au lieu que* (‘instead of’ / ‘except that’) in Medieval and Modern (rather non-standard) French (15).

15. au lieu que dans les écoles parisiennes et dans le primaire
 in.the stead that in the.PL school.PL Parisian.PL and in the primary
 et le secondaire c’était le hasard qui faisait les choses
 and the secondary it.be.PST.3SG the chance that do.PST.3SG the thing.PL
 [some people have received this officially...] ‘while in Parisian schools, and in primary and secondary schools, what did that was chance’ (CFPP2000_16-01_MARC-JAKOBSON_H_83_16E)

Beyond these uses, which remain functionally similar, simple adpositions may also grammaticalize further, acquiring a wide range of functions, including uses as determiner (16), complementizer (17-18) (see e.g. English *to* used as complementizer or infinitive marker, German *zu*, Dutch *te*, etc.), differential object marking (19) (DOM, Regional French, 16; see also Romanian *pe*, Spanish *a*, Regional Italian *a*, among other Romance varieties), etc.

16. donne-moi **de** l' eau
 give-me of DET.DEF.F.SG water
 'give me (some) water'
17. elle m' a dit **de** venir
 she me.ACC AUX.PRS.IND.3SG say.PST.PTCP.M.SG of come.INF
 'she told me to come'
18. Il vient **pour** manger
 he come.PRS.IND.3SG for eat.INF
 'he comes to eat'
19. [ils] veulent t'interroger **à** toi et moi hein (rires)
 (they) want.PRS.IND.3PL CL.ACC.2SG.question.INF DOM you and me huh (laughs)
 they want to question you and me both (laughs) (PFC, Midi-Pyrénées, Fagard and Mardale 2014)

In such contexts, these morphemes are no longer adpositions. It should be noted that complex adpositions do not seem to display this characteristic of simple adpositions, at least in Romance.

4 Complex adpositions in French

In this section, we describe in detail the system of complex adpositions found in one Romance language, namely French.

4.1 Main schema

In terms of types, the most productive schema of complex adpositions in French is made up of three to four elements, i.e. [PREP₁ (DET) N PREP₂]:

- 1) preposition₁ (typically *à*)
- 2) definite article
- 3) noun
- 4) preposition₂ (typically *de*)

This schema accounts for complex adpositions such as *au-dessus de* 'above', *au lieu de* 'instead of', *à l'instar de* 'like', etc. For each of these four elements, there is a wide range of variation: the first and fourth – adpositions – are most often *à* and *de*, but other adpositions can be found in their place, such as *en* 'in', *avec* 'with', etc., including sometimes in a given construction (e.g. *en-dessous de* vs *au-dessous de* 'beneath'); they can also be absent. The same applies for the determiner, which is rarely present. Finally, the third element is not necessarily a noun; it can be a verb, an adverb, etc. The next subsection, on subtypes of complex adpositions in French, provides more detail on the extent of this variation.

4.2 Subtypes

The most productive subtypes in Modern French are the main schema, [PREP₁ (DET) N PREP₂], and variants thereof; there are also, however, a number of marginal subtypes (see Table 3).

Table 3: Typical patterns of CAs in Modern French (adapted from Stosic, to appear, on the basis of grammars, dictionaries and existing literature on CAs; the list cannot be exhaustive, but is representative. Numbers indicate the number of *constructions* found for each pattern, not their frequency).

Pattern	Total	Examples
P ₁ DET N <i>de</i>	27	<i>à l'égard de</i> 'concerning', <i>à la fin de</i> 'at the end of', <i>au bord de</i> 'at the edge of', <i>sous l'égide de</i> 'under the aegis of', <i>sur les traces de</i> 'on the steps of', <i>au-dessous de</i> 'below'
P ₁ N <i>de</i>	122	<i>à base de</i> 'with, made from', <i>en dépit de</i> 'in spite of', <i>sous prétexte de</i> 'under the pretense of', <i>sur ordre de</i> 'by order of'
P ₁ N P ₂ (P ₂ ≠ <i>de</i>)	20	<i>par comparaison avec</i> 'in comparison to', <i>par rapport à</i> 'in relation to'
ADV P ₂	20	<i>auprès de</i> 'beside', <i>autour de</i> 'around'
P ₁ ADV P ₂	16	<i>à moins de</i> 'unless, short of', <i>en dehors de</i> 'outside'
P ₁ P ₂	13	<i>d'après</i> 'according to, based on', <i>par-devers</i> 'on'
P ₁ ADV	11	<i>à même</i> 'next to, directly on', <i>par-dessus</i> 'on top of'
N P ₂	11	<i>dos à</i> 'back to', <i>face à</i> 'in front of', <i>grâce à</i> 'thanks to'
P ₁ DET ADJ N <i>de</i>	10	<i>au beau milieu de</i> 'in the middle of', <i>au fin fond de</i> 'in the depths of'
Other patterns	63	<i>au plus bas de</i> 'at the lowest of', <i>pour ce qui regarde</i> 'concerning', <i>quant à</i> 'according to, concerning', <i>quitte à</i> 'at the risk of', <i>proche de</i> 'close to', <i>à égale distance de</i> 'at equal distance from', <i>pas loin de</i> 'not far from', <i>au vu et au su de</i> 'known to', <i>en passant par</i> 'by way of, including', <i>mis à part</i> 'excluding', <i>il y a</i> 'ago'

The initial preposition is most typically *à*, which appears roughly in half of the constructions, less frequently *en*, *sous*, *par*, *de*, *sur* and *dans*, rarely other adpositions (see Table 4a); the final preposition is almost always *de* (Table 4b).

Tables 4a & 4b: Patterns of CAs in Modern French: (a) initial and (b) final adposition (adapted from Stosic, to appear)

Prep ₁	Productivity (number of CAs)
<i>à</i> 'at'	240
<i>en</i> 'in'	79
<i>sous</i> 'under'	50
<i>par</i> 'by, via'	35
<i>de</i> 'of, from'	31
<i>sur</i> 'on'	19
<i>dans</i> 'in'	18
others	26

Prep ₂	Productivity (number of CAs)
<i>de</i> 'of, from'	476
<i>à</i> 'at'	28
<i>avec</i> 'with'	10
others	11

4.3 Number & frequency

The frequency of simple adpositions in Modern French is far from homogeneous. On the one hand, a few adpositions are very frequent: *à*, *de*, *en* are the most typical *functional* adpositions (maybe not the best term, see Lehmann 2019, Footnote 2), which are highly grammaticalized, and generally considered as a subtype of simple adpositions. Their frequency ranges from 15‰ (i.e. per thousand words) to 70‰ in Modern Corpora (Fagard, to appear). On the other hand, some fifty to a hundred other simple adpositions, identified as *lexical* adpositions, display much lower frequency levels, generally less than 1‰ (i.e. a thousand occurrences per million words), as shown in Table 5. There is also much variation in the frequency levels of complex adpositions; however, it can be shown that their frequency is much lower than that of functional adpositions, and even lower than some lexical adpositions (Table 5).

Table 5: Relative frequency of simple and complex adpositions in Modern French (GGHF corpus, 19th-20th c., roughly 5 million words)⁷.

Frequency range (%)	Simple adpositions	Complex adpositions
above 5	<i>de</i> ‘of, from’, <i>à</i> ‘at’, <i>dans</i> ‘in, into’, <i>en</i> ‘in’, <i>pour</i> ‘for’, <i>par</i> ‘by’	
between 1 and 5	<i>sur</i> ‘on’, <i>avec</i> ‘with’, <i>sans</i> ‘without’, <i>chez</i> ‘at’	
between 0.5 and 1	<i>sous</i> ‘underneath’, <i>après</i> ‘after’, <i>entre</i> ‘between’, <i>jusque</i> ‘until’, <i>contre</i> ‘against’, <i>vers</i> ‘toward’, <i>depuis</i> ‘since’, <i>devant</i> ‘in front of’	
between 0.1 and 0.5	<i>pendant</i> ‘during’, <i>avant</i> ‘before’, <i>dès</i> ‘from, since’, <i>malgré</i> ‘despite’, <i>parmi</i> ‘among’, <i>derrière</i> ‘behind’, <i>selon</i> ‘according to’, <i>par-dessus</i> ‘above, on top of’	<i>autour de</i> ‘around’, <i>d’après</i> ‘according to, based on’, <i>au fond de</i> ‘at the bottom of’, <i>auprès de</i> ‘beside’, <i>quant à</i> ‘according to, concerning’, <i>au-dessous de</i> ‘below’
between 0.01 and 0.1	<i>hors</i> ‘outside’, <i>sauf</i> ‘except’, <i>envers</i> ‘towards’, <i>suivant</i> ‘according to’, <i>durant</i> ‘for, during’, <i>environ</i> ‘about’, <i>excepté</i> ‘except, besides’, <i>moyennant</i> ‘by means of, in exchange for’	<i>il y a</i> ‘ago’, <i>au bord de</i> ‘at the edge of’, <i>à la fin de</i> ‘at the end of’, <i>par rapport à</i> ‘in relation to’, <i>grâce à</i> ‘thanks to’, <i>en dépit de</i> ‘in spite of’, <i>en dehors de</i> ‘outside’, <i>à l’égard de</i> ‘with respect to’, <i>face à</i> ‘in front of’, <i>dos à</i> ‘back to’, <i>proche de</i> ‘close to’, <i>à même</i> ‘next to, directly on’, <i>en passant par</i> ‘by way of, including’, <i>sous prétexte de</i> ‘under the pretense of’
less than 0.01	<i>outré</i> ‘beyond, besides’, <i>touchant</i> ‘concerning’, <i>vu</i> ‘seen, given, owing to’, <i>nonobstant</i> ‘despite’, <i>par-delà</i> ‘beyond’, <i>concernant</i> ‘concerning’, <i>hormis</i> ‘save, except’, <i>devers</i> ‘on’, <i>fors</i> ‘except’ (archaic)	<i>pas loin de</i> ‘not far from’, <i>quitte à</i> ‘even if it means, at the risk of’, <i>à moins de</i> ‘unless, short of’, <i>au beau milieu de</i> ‘in the middle of’, <i>sur les traces de</i> ‘on the steps of’, <i>mis à part</i> ‘excluding’, <i>à base de</i> ‘with, made from’, <i>au fin fond de</i> ‘in the depths of’, <i>sous l’égide de</i> ‘under the aegis of’, <i>au plus bas de</i> ‘at the lowest of’, <i>au vu et au su de</i> ‘known to’, <i>sur ordre de</i> ‘by order of’, <i>à égale distance de</i> ‘at equal distance from’, <i>par comparaison avec</i> ‘in comparison to’, <i>pour ce qui regarde</i> ‘concerning’, <i>pardevers</i> ‘on’

5 A diachronic account: The emergence of Complex Adpositions in French

In this section, we describe the emergence of CAs in the diachrony of French. Their emergence as a category has been very slow in French, from a few constructions with limited frequency in Old French to a very large paradigm including a few quite frequent constructions in Modern French.

⁷ For simple adpositions, we included all items with ‘PRP’ tags, i.e. those categorized as ‘prepositions’, but checked manually and excluded some marginal items such as Old French *ès* ‘in’ (lit. in.DET.PL), which appears twice in the corpus but only in quotations of a medieval text. For complex adpositions, we included those listed in Table 3.

5.1 From Latin to Old French

In Classical Latin, there seems to be no paradigm of complex adpositions. Only a few constructions with similar uses – i.e. complex nominal relators – can be found, based on nouns such as *causa* ‘cause, motive’, *gratia* ‘favor, esteem’, *modus* ‘measure, extent’. These nouns, in the Ablative case, can govern another noun in the Genitive (20-21): they then display distributional properties which are quite similar to those of CAs.

20. **exempli** **causā** (Latin)
 example.GEN.SG motive.ABL.SG
 ‘for (the sake of) (the) example’
21. **dei** / **tui** **gratiā**
 god.GEN.SG / your.GEN.SG favor.ABL.SG
 ‘thanks to god/to you’

In Old French (9th-13th centuries), similar constructions can be found, though their (i) frequency, (ii) productivity and, more importantly, (iii) patterns are different. The frequency is hard to compare, in part because the available corpora are not of the same type or size, for historical reasons, in part because the frequency of the Latin constructions listed above is extremely low. The main pattern found in Old French is that typical of CAs (Prep₁ Det N Prep₂), and is a Romance innovation. In terms of types, the productivity of this pattern (i.e. its type frequency)⁸ is also more important in Old French, though it remains limited: we found a total of 52 constructions which seem to be used as CAs in our corpus (Table 6).

Table 6: Main patterns of CAs in Old French (GGHF corpus).

Pattern	Constructions
<i>à</i> ‘at’ (+ DET) + N + P ₂	16
<i>en</i> ‘in’ (+ DET) + N + P ₂	27
<i>de</i> ‘of, from’ / <i>par</i> ‘by, via’ / <i>pour</i> ‘for’ (+ DET) + N + P ₂	9
Total	52

This pattern is illustrated in (22) with *por l’amor de* ‘thanks to’ (lit. ‘for the love of’), which also provides evidence of semantic bleaching in CAs. An alternate pattern can be found, much as CA in Modern Romanian: [PREP₁ (DET) N + CASE MARKING], with the object case on the governed noun, illustrated in (23) with [*en som [lo mur]_{NP}]_{PP}, lit. ‘[in/on top [the wall]_{NP}]_{PP}’.*

22. Tout **por** **l’** **amor** **de** cest afaire li varlés au
 all for the.M.SG love of this.M.SG matter the.M.NOM.SG valet at.the.M.SG
 cuer debonaire a fait le poulain jus abatre
 heart good-natured AUX.IND.PRS.3SG do.PST.PTCP.M.SG the.M.SG colt down put.INF
 ‘Because of this matter, the good-hearted valet had the colt put down’ (*Eracle*, 12th c., p. 58, v. 1883)
23. Chascuns qui l’ ot corant i vet,
 each.M.SG.NOM who.NOM PRO.SG.ACC hear.PRS.IND.3SG run.PRS.PTCP there go.PRS.IND.3SG,
 d’ **en** **som** **lo** **mur** lo vont veoir

⁸ I.e. the number of constructions which are based on the pattern, regardless of their frequency in language use.

of on top the.M.SG.ACC wall PRO.M.SG.ACC go.PRS.IND.3PL see.INF
 ‘Anyone who heard him ran up to him, they went to watch him from upon the wall’ (*Eneas*, 12th c., v. 7340-1)

Given the fact that there are only two cases in Old French (subject case and object case), the distinction between this second type and a true adpositional use ([PREP [NP]]_{PP}) is not obvious, and seems to rely mainly on the analysis of the central element as retaining nominal features or having undergone reanalysis. Thus, it seems that *de lez* ‘next to (lit. ‘of side’, ‘from side’, ‘from beside’)’ has undergone full reanalysis and can be considered a ‘simple’ adposition, given the fact that it can be reinforced by a prefix, yielding the surprising form *dedelez* ‘next to’ (lit. ‘of of beside’ or ‘from from beside’). The productivity of the paradigm of constructions such as *en som* is hard to evaluate, depending mostly on the analysis retained for individual constructions.

5.2 Middle and Classical French

In Middle and Classical French (roughly from mid 13th c. through 18th c.), the productivity and frequency of the main pattern of CAs steadily increase. The importance of this phenomenon in Middle French is already noted by Shears (1922: 19)⁹. There is some renewal: not all constructions which appeared in Old French survive, and many new constructions appear, including for instance *au regard de* ‘with respect to’ (lit. ‘at the eye gaze of’). However, quite a few remain and gain frequency, for instance *au long de* ‘along’ (lit. ‘at the length of’) or *au-dessus de* ‘above’ (lit. ‘at the upper-side of’). As suggested in the chapter on CAs in Germanic languages (Hüning et al., this volume), and in line with construction grammar approaches (Goldberg 1995, Traugott and Trousdale 2013), it may be considered that Complex Adpositions may undergo constructionalization to some extent, losing variability and acquiring internal coherence. For instance, in the GGHF corpus, the sequence *au bord de* ‘at the edge of’ seems to appear in the 13th century, with a very low relative frequency in Middle and Pre-Classical French (roughly 4 occ. per million words). In Classical (and Modern) French, along with other similar sequences, it becomes much more frequent (37 occ. per million words) and thus reaches a frequency which is akin to that of simple (lexical) prepositions. At that period, for this construction, the proportion of occurrences with modification drops from 14% (2 in 14 occ.) to less than 1% (3 in 355 occ.).

The result is that, by our counts, in Middle French some 20 CAs have a frequency of about or more than 10 occurrences per million words, with more than 60 CAs in total; in Classical French, more than 30 CAs have a frequency above 10 occ. per million, with more than 70 CAs in total. The cumulative (relative) frequency of CAs seems to have reached 200 to 300 occurrences per million (0.2-0.3‰) in Middle French, and more than 600 per million (0.6‰) in Classical French. The rise in frequency continues after that: the cumulative frequency of CAs listed in Table 5 (Section 4.3, Modern French – 19th and 20th centuries) is roughly 1.3‰. This is still much lower than functional or even lexical adpositions during the same period; however, it shows the gradual emergence of the category, from a few infrequent constructions to a productive, consistent and relatively frequent pattern.

⁹ He thus writes that the tendency to substitute prepositional phrases for simple prepositions is one of the most remarkable features of Middle French (“l’un des caractères les plus remarquables du moyen français est la tendance marquée à substituer aux prépositions simples des locutions prépositives”; our translation).

5.3 Modern French

In the later centuries, this pattern becomes much more productive, yielding hundreds of new constructions, some of which not only acquire a high relative frequency, but also tend to become lexicalized; when that happens, they tend to deviate from the typical (and original) pattern.

Table 7: Most productive patterns of CAs in Modern French

pattern	examples
Prep ₁ + N + Prep ₂	<i>à cause de</i> ‘because of’, <i>à côté de</i> ‘next to’, <i>à défaut de</i> ‘for want of’, <i>à fleur de</i> ‘on the surface of’, <i>à force de</i> ‘by dint of’, <i>à raison de</i> ‘at the rate of’, <i>en face de</i> ‘in front of’, <i>de peur de</i> ‘for fear of’, <i>par rapport à</i> ‘with regard to’, <i>par suite de</i> ‘as a result of’, <i>pour cause de</i> ‘on account of’, <i>sous couleur de</i> ‘under the guise of’, <i>en faveur de</i> ‘in favor of’, <i>en guise de</i> ‘by way of’, <i>en raison de</i> ‘because of’, <i>en dépit de</i> ‘despite’
Prep ₁ + Adv + Prep ₂	<i>en dessous de</i> ‘below’, <i>en-dehors de</i> ‘outside, apart from’, <i>en sus de</i> ‘above’, <i>en amont de</i> ‘uphill from’, <i>en arrière de</i> ‘behind’, <i>en aval de</i> ‘downhill from’, <i>en bas de</i> ‘below’, <i>en deçà de</i> ‘short of’, <i>en dedans de</i> ‘within’, <i>en plus de</i> ‘besides’
Prep ₁ + Det + N + Prep ₂	<i>à l’aide de</i> ‘by means of’, <i>à l’égard de</i> ‘with respect to’, <i>à l’endroit de</i> ‘concerning’, <i>à l’entour de</i> ‘around’, <i>à l’exception de</i> ‘except’, <i>à l’exclusion de</i> ‘except’, <i>à l’instar de</i> ‘like’, <i>à l’issue de</i> ‘after’, <i>à l’occasion de</i> ‘on the occasion of’, <i>à la faveur de</i> ‘thanks to’, <i>à la merci de</i> ‘at the mercy of’, <i>à l’insu de</i> ‘unbeknownst to’, <i>à l’intérieur de</i> ‘inside’, <i>au bord de</i> ‘on the edge of’, <i>au fond de</i> ‘at the bottom of’, <i>au lieu de</i> ‘instead of’, <i>au milieu de</i> ‘in the middle of’, <i>au moyen de</i> ‘thanks to’, <i>au pied de</i> ‘at the foot of’, <i>au prix de</i> ‘by means of’, <i>au travers de</i> ‘by means of’, <i>du côté de</i> ‘near’, <i>le long de</i> ‘along’, <i>de la part de</i> ‘on behalf of’
Prep ₁ + Det + Adv + Prep ₂	<i>à l’arrière de</i> ‘behind’, <i>à l’encontre de</i> ‘counter to’, <i>au dedans de</i> ‘within’, <i>au dehors de</i> ‘outside’, <i>au delà de</i> ‘beyond’, <i>au dessous de</i> ‘below’, <i>au dessus de</i> ‘above’, <i>au devant de</i> ‘in front of’, <i>auprès de</i> ‘near’

Their relative frequency is much higher than it was in Medieval French. While the most frequent CAs in Medieval French barely reached 20 occ. per million (*au lieu de*, BFM), they can be at least three times as frequent in Modern French, as illustrated in Table 8.

Table 8: Relative frequency of a few CAs, in Old and Modern French

corpus	sequence	relative frequency (occ. per million)
BFM (11 th -15 th c.: 3 million words)	<i>au lieu de</i> ‘instead of’	23.3
	<i>au moyen de</i> ‘by means of’	6.2
Frantext (after 1968: 54.2 million words)	<i>au lieu de</i> ‘instead of’	68.2
	<i>à travers</i> ‘through’	133.8
	<i>à l’instar de</i> ‘like’	5.3
	<i>faute de</i> ‘for lack of’	19.2

Some of the older CAs grammaticalize into simple adpositions. In some cases, this simplification is found even in Standard French, and the corresponding constructions have an even higher frequency, as is the case for *à travers* ‘through’, or *par (la) faute de / à faute de > faute de* (Table 8). In other cases, this simplification is limited to some contexts, i.e. mainly to informal or regional French, for instance *en face de > en face*, *du côté de > du côté > côté* (Fagard 2008), *par rapport à > rapport à* (24-25).

24. c’ est-il vrai, monsieur Joigneau, ce qu’ on
 this.M.SG be.PRS.IND.3SG.IT true sir Joigneau this.M.SG which.ACC.SG one
 raconte, **rapport à** la mère Daigne?

say.PRS.IND.3SG about at the.F.SG mother Daigne
 ‘is it true, Mr. Joigneau, what they say ’bout Mother Daigne?’ (Roger Martin du Gard, *Vieille France*, 1933, p. 1030)

25. Je me souviens très bien maintenant. Mon taxi faillit
 I me.ACC.SG remind.PRS.IND.1SG very well now my cab fail.PST.IND.3SG
 vous écraser, il y a une quinzaine de jours,
 you.ACC.SG run-over.INF it there have.PRS.IND.3SG one.F.SG fortnight of day.PL
en face la gare du Nord (...)
 in face the.F.SG station of.the.M.SG North
 ‘I remember perfectly now. My cab almost ran you down, two weeks ago, in front of the Gare du Nord’
 (Raymond Queneau, *Le Chiendent*, 1933, p. 43)

6 Variation in the language family

In this section, we provide an overview of CAs in Romance languages, focusing first on common features (6.1), then on variation (6.2): as we will see, there are both.

6.1 Common features

There seems to be a set of core patterns of CAs which are found in all Romance languages. As illustrated by Table 9, the most wide-spread patterns besides the main pattern ([PREP₁ (DET) N PREP₂]) are [PREP₁ (DET) ADV PREP₂] and [N/ADV PREP₂]. Among these 6 patterns (3*2 possibilities: PREP₁ / zero, DET / zero, N/ADV), it could be argued that the main patterns are actually [PREP₁ N PREP₂], [PREP₁ ADV PREP₂], [ADV PREP₂] and [PREP₁ DET N PREP₂], since the [N PREP₂] pattern is less frequent, generally resulting from the loss of the initial P, while in the [PREP₁ DET ADV PREP₂] pattern the adverb is nominalized, and could or should be analyzed as a noun (the same could be said of all complex adpositions in which the lexical nucleus is not originally – or not considered to be – a noun, but takes the definite article). Finally, there are very few ‘other’ patterns. Note that most of these patterns are already found in Medieval Romance, in all languages for which medieval texts are available (Fagard 2006).

Table 9: Patterns of Complex Adpositions in Romance languages (C: Catalan, I: Italian, O: Occitan, P: Portuguese, R: Romanian, S: Spanish).

		Pattern			
		[PREP ₁ N/ADV PREP ₂]	[N/ADV PREP ₂]	[PREP ₁ DET N/ADV PREP ₂]	other
LEXICAL NUCLEUS	NOUN	all Romance languages		FCI?OPRS	FP
	ADVERB			?(FCOR)	FCPS
	VERB	FCOPR	FCPR	C	FCPS
	OTHER	CP	IP?R	C	?

The four main patterns, PNP, PAdvP, AdvP and PDetNP, also seem to be the most productive and those which yield the most frequent complex adpositions. There is also a striking lexical homogeneity: for quite a few CAs, patterns are instantiated almost identically across languages. Indeed, in many cases, the very same noun, adverb or verb form is used as nucleus of a CA *and* follows similar patterns in different Romance languages, as illustrated in Table 10.

Table 10: A few Complex Adpositions in Romance – similarities and differences.

Lexical Nucleus	Portuguese	Spanish	Catalan	Occitan	French	Italian	Romanian
'base'	<i>na base de, com base em, à base de</i>	<i>a base de</i>	<i>a base de</i>	<i>a basa de</i>	<i>à base de</i>	<i>a base di, sulla base di</i>	<i>pe baza (+ GEN)</i>
'cause'	<i>por causa de</i>	<i>por causa de</i>	<i>a causa de</i>	<i>a causa de</i>	<i>à cause de</i>	<i>per causa di</i>	<i>din cauza (+ GEN)</i>
'outside'	<i>fora de</i>	<i>fuera de</i>	<i>fora de, en/al defora de</i>	<i>en/al defòra de</i>	<i>en-dehors de</i>	<i>all'infuori di</i>	<i>în afară de</i>

These three lexical nuclei illustrate quite well on the one hand the proximity between languages, since the same lexical nucleus appears in all 7, but also the possible variations, for instance the absence or presence of Prep₁, with Ibero-Romance standing out, or the use of the genitive instead of Prep₂ in Romanian. In the next subsection, we further illustrate these sources of cross-linguistic variation.

6.2 Variation

There is indeed a great deal of variation across Romance languages. Variation can be found in the *lexical* elements which form the core of CAs, and which vary along with the rest of the lexicon; it is also present in the respective proportions of [PREP₁ (DET) N PREP₂] and [ADV P] patterns; finally, the degree of grammaticalization of CAs, as measured by the way they react to syntactic tests, seems uneven. We address these three types of variation in sequence.

6.2.1 Lexical variation

Some lexical nuclei are less wide-spread than others, on account of grammaticalizations which took place after the differentiation of Romance. For instance, Late Latin **bassius* 'lower than' acquired adverbial (and later on prepositional) uses only in parts of Romania, with Portuguese *baixo*, Spanish *bajo*, French and Occitan *bas* 'low' > 'down' in Ibero-Romance and Gallo-Romance, respectively. The same can be said of the grammaticalization of Latin **monte* and **valle*, with corresponding adpositions mainly in Gallo-Romance: French *en amont de, en aval de*, Catalan *al damunt de, davall de*, Occitan *daval de, *en amont de* (Portuguese equivalents, *a montante de* 'upstream' and *a jusante de* 'downstream', seem to be a later creation). More locally, individual languages may stand out with the grammaticalization of a new CA, for instance Portuguese *à beira de* '(very) close to, on the verge of' on the noun *beira* (probably from *ribeira* < (Lat) *ripariam*, on *riparius* 'which frequents the banks of

rivers’), or Catalan *a la vora de* and French *à l’orée de* with the same meaning, but a different etymology: (Lt) *ora* ‘border, edge’, plural of *os, oris*, n. ‘mouth’.

6.2.2 Patterns

The main differences as far as patterns are concerned are (i) the importance of the AdvP pattern in Ibero-Romance, (ii) the dialectal variation in Italo-Romance and the use of *di* for personal pronouns: Italian *senza di me* ‘without (lit. of) me’ / *senza questo* ‘without this’, and finally (iii) the complex system of CAs in Romanian, with an alternation between Prep₂ and genitive marking.

The importance of the AdvP pattern in Ibero-Romance is quite obvious (Fagard 2006, 2010); they are described in Spanish and Portuguese as “relational adverbs” (*advérbios relacionais*, Raposo et al. 2013: 1500), on account of their distributional characteristics, i.e. mainly their capacity to appear either alone (in the form [Adv]), or with a complement ([AdvP [SN]]), as in (26-27).

26. o comboio já está perto da estação (Portuguese)
 DET.M.SG train already be.PRS.IND.3SG near of.DET.F.SG train.station
 ‘the train is already near the train station’ (Raposo et al. 2013, *ibid.*)

27. o comboio já está perto
 DET.M.SG train already be.PRS.IND.3SG near
 ‘the train is already nearby’

Even if the pattern is found in other languages, there is a much more systematic correspondence between adverbial and adpositional uses in Ibero-Romance than in other Romance languages. Table 11 illustrates the systematicity of the Adv/CA pattern in Portuguese, and the fact that both Italian and French differ in this respect. While there is a perfect systematicity in Portuguese, for these examples at least, it is not the case in French and Italian. Indeed, in most instances there is either no variation (e.g. for ‘after’, in both languages) or various types of alternation between adverbial and prepositional uses (e.g. with an optional Prep₂ as in Italian, or a CA construction as in French).

Table 11: Adv and AdvP/CA uses in Portuguese, Italian and French¹⁰.

Meaning	Adverbial uses			Adpositional uses		
	Portuguese	Italian	French	Portuguese	Italian	French
‘above’	<i>acima</i>	<i>sopra</i>	<i>dessus</i>	<i>Adv + de</i>	<i>Adv (+ a/di)</i>	CA
‘below’	<i>debaixo</i>	<i>sotto</i>	<i>dessous</i>			(<i>au-dessus de, en-dessous de</i>)
‘far’	<i>longe</i>	<i>lontano</i>	<i>loin</i>			<i>Adv + de</i>
‘before’	<i>antes</i>	<i>prima</i>	<i>avant</i>		<i>Adv + di</i>	ADV = PREP
‘near’	<i>perto</i>	<i>vicino</i>	<i>près</i>		<i>Adv (+ a)</i>	
‘after’	<i>depois</i>	<i>dopo</i>	<i>après</i>		ADV = PREP	

¹⁰ Note that French *près (de)* ‘close (to)’, mainly has uses as a CP. It also has adverbial uses; these are generally found in specific constructions such as *tout près* ‘very close’, *assez près* ‘quite close’, but there are examples of adverbial *près* alone (e.g. *C’est passé près!* ‘That was close!’ – lit. *that went close by*, of bullets –, Cavanna, 1979, *Les Russkoffs*, Frantext).

This table also illustrates the variation in Italian, with the same form displaying uses as an adverb or an adposition, but also CA constructions, e.g. *sopra (a/di) + NP*, for which a corpus search (itTenTen and Repubblica, SketchEngine) shows the existence of various possibilities: *sopra di + pronoun* (28), *sopra a + Adv/NP* (29), *al di sopra di* (30), *sopra + NP* (31) (with widely different frequencies: *sopra* is above 20 occurrences per million for prepositional uses only, while all other constructions are well below 5 occurrences per million). This variation is partly constrained by dialectal, sociolinguistic or register-related considerations.¹¹

28. **Sopra di** noi torreggiano i palazzi. (Italian)
 above of us tower.PRS.IND.3PL the.M.PL building.PL
 ‘The skyscrapers loom above us’ (Repubblica corpus, SketchEngine, #26298006)
29. La casa è bassa, un solo piano **sopra a** quello terreno
 the.F.SG house be.PRS.IND.3SG low.F.SG one.M.SG only.M.SG floor above at that.M.SG ground
 ‘The house is low, only one floor above the ground floor’ (Repubblica corpus, SketchEngine, #30558684)
30. Un testimone **al di sopra di** ogni sospetto
 a witness at.the.M.SG of above of all.M.SG suspicion
 ‘A witness above suspicion’ (Repubblica corpus, SketchEngine, #873531)
31. **Sopra** la sua voce si sono sentite distinte
 above the.F.SG POSS.F.SG voice REFL be.PRS.IND.3PL hear.PST.PART.F.PL distinct.F.PL
 due signore che si raccontavano i fatti loro.
 two lady.PL REL.PRO REFL tell.PST.IND.3PL the.M.PL fact.PL POSS.PL
 ‘Above his voice one could hear distinctly those of two ladies who were telling each other their own stories’ (Repubblica corpus, SketchEngine, #21425897)

Romanian naturally stands out on account of its case-marking system, which allows for a pattern found in no other Modern Romance language: the [PREP₁ N + GEN] pattern, in which the link with the complement is marked not by an adposition, but by case marking (Găitănanu 1999, Petrescu 2005, Vlasin 2015). The [PREP₁ N PREP₂] pattern is also found (Ciobanu 1957, Căpățână 2003), and in some cases both constructions are possible for a given CA. This may entail a difference in meaning, as can be seen in examples (32-35) for *în afara + GEN / în afară de* ‘outside of, except’ and *în jurul + GEN / în jur de* ‘around’.

32. *inclusiv în afara Uniunii Europene* (Romanian)
 even in outside.DET.F.SG Union.GEN European.GEN
 ‘even outside the European Union’ (roTenTen16)
33. *Ce altceva mai sculptați, în afară de portrete?*
 what else still sculpt.IND.PRS.2PL in outside of portrait.PL
 ‘what else do you make, besides portraits?’ (roTenTen16)

¹¹ This is also true for other Romance languages, of course, with e.g. the alternation between French *dessous la table* and *sous/en-dessous de la table*, or *en face la gare* and *en face de la gare* (the first variant being clearly substandard, in both cases). But it is quite frequent in Italian, possibly for historical reasons, the standard language having emerged only recently in comparison with other (national) Romance languages.

34. *Protestul va dura în jur de două ore.*
 protest.DET will.PRS.3SG last.INF in turn of two.F.PL hour.PL
 ‘the protest will last around two hours’ (roTenTen16)
35. 20 aprilie 2016, în jurul orei 06.00
 20 april 2016 in turn.DET hour.GEN 06:00
 ‘On April 20th, 2016, at around 6 am’ (roTenTen16)

Most frequently, though, CAs follow only one pattern. The [PREP₁ N + GEN] construction is clearly the most productive one, with close to a hundred items, for instance *cu excepția* ‘except’, *din cauza* ‘because of’, *în ciuda* ‘despite’. The [PREP₁ N PREP₂] pattern is less productive, with less than thirty items, among which *în (de)curs de* ‘during’, *în jur de* ‘around’, *în urmă cu* ‘ago’ (Fagard and Mardale 2019).

6.2.3 Degree of grammaticalization

All languages, it seems, evolve in part through grammaticalization processes. These result in a progressive renewal of grammatical systems. Thus, in the evolution from Latin to Modern Romance languages, there has been a steady renewal of grammatical items, including TAM markers, demonstratives, pronouns, and even the emergence of new paradigms of grammatical items, for instance determiners. In this respect, it has been argued that, among Romance languages, French is the furthest along the grammaticalization cline (Lamiroy 1999, De Mulder 2001, Fagard 2010, 2011). It has for instance developed the most extensive paradigm of determiners, including indefinite plural and partitive uses (Carlier 2007). This could also be said of the adpositional system:

French grammaticalized further its renewed stock, creating more simple adpositions than the other languages and more rigid patterns of complex adpositions, followed by Italian, then Spanish, Portuguese and Romanian (Fagard and Mardale 2012: 335)

Of course, the tendency of CAs to grammaticalize and thus form less complex – and eventually simple – adpositions is probably universal, and examples could be found easily in the other Romance languages, e.g. Catalan *a la vora de* > *vora de* > *vora* ‘close to’, Italian *per tramite di* ‘by way of’ > *tramite* ‘through’. However, French CAs, in comparison with other Romance CAs, seem to have both emerged more consistently as a class, and – possibly as a consequence – to have contributed more constructions available to further grammaticalization into simple adpositions. One possible indication, among others (see e.g. Stosic and Fagard 2019, Hüning et al., this volume, Stosic, this volume), of the greater degree of grammaticalization of CAs in French – more specifically of their greater internal coherence (Lehmann 2002) – is the possible alternation with the possessive, which is almost never possible in French once the CA is grammaticalized, but often remains possible in other Romance languages, as illustrated by examples (36-39)¹². As noted in the *Nueva gramática de la lengua española* (2009, I:55), for some scholars, the possibility of using the possessive automatically excludes a sequence from the group of CAs; however, the authors (ibid.) plead for a more flexible approach, considering that these sequences may display

¹² In general, Romance languages display individual exceptions to this tendency; for instance, in Portuguese, the grammaticalization of *a pesar de* ‘in spite of’ > *apesar de* ‘despite’ led to the loss of the alternate construction with the possessive (initially *a seu pesar* ‘in his spite’) (Lima 2019b).

‘semilexicalization’ (*semilexicalización*). The possibility of modifying such sequences may evolve with time: for instance, in Medieval and Classical French, insertion was still possible even with demonstrative determiners, as illustrated in (40) – which would be translated in Modern French by *à cause de cela* ‘because of that’, while [#]*à cette cause* can no longer alternate with an adpositional use of *à cause de* (*j’ai raté mon examen à cause de la fête* ‘I failed my exam because of the party’ → **à cette cause* ‘for this reason’).

36. (French) **en-dessous de lui** ‘below him’ → **en son dessous*

37. (Portuguese) **ao seu lado / ao lado dele** ‘next to him’, **a teu respeito** ‘about you’

/ **a respeito de ti, na sua esteira / na esteira dele* ‘in his wake’

38. (Italian) **per mia colpa** ‘because of me’ / **per colpa di me*

39. (Catalan) **davall de mi / davall meu** ‘beneath me’, **davall nostre** ‘below us’, **sobre meu** ‘above me’

40. en quoy ne se peürent accorder et, à ceste cause,
in REL.PRO.ACC.SG NEG REFL can.PST.IND.3SG agree.INF and at this.F.SG cause
partit très mal content d’ eulx de la court
leave.PST.IND.3SG very bad.M.SG happy of they.3PL.ACC of the.F.SG court
‘on which question they could not find an agreement, and, for that reason [lit. *at this cause*], he left
the court feeling very mad at them’ (Philippe de Commines, *Mémoires*, 15th c., vol. 7, p.7)

(Medie

A thorough contrastive study should check whether this is indeed the case, i.e. whether there is a clear contrast between French and other Romance languages.

7 Conclusion

Systems of CAs are fairly well-established in Romance languages, with a few productive patterns providing constructions with high frequencies, which sometimes end up grammaticalizing into simple adpositions. There are CAs in all Romance languages, and they appear to be quite homogeneous, with some variation, however, in terms of lexical nuclei and patterns. There seem to be areal tendencies, with a greater importance of the AdvP pattern in Ibero-Romance, of the P(Det)NP patterns in French, of variation in Italian, and the existence of specific patterns in Romanian.

There is evidence of the emergence of CAs as a class between Medieval and Modern Romance. Indeed, the data available for French and other Romance languages point to the existence of individual constructions in Medieval Romance. It is also quite clear that CA patterns were much more limited in number, frequency and systematicity in Medieval Romance than they are now. After that, there has been a steady rise of CAs. This can be documented, for instance, from Old French to Modern French: more constructions appear, gaining in frequency and internal coherence, and progressively yielding systematic patterns which become available for the creation of new CAs (*grammaticalization by analogy*, see Meillet 1912, Hoffmann 2004). It remains to be seen whether there has been, as hypothesized by Van der Horst (2013, cf. Hüning 2014¹³), a global emergence of CAs in the languages of

¹³ Van der Horst actually shows that these strong relations between European languages are only a few centuries old (“Ze zijn er niet altijd geweest, die vaste verbindingen. Duizend jaar geleden, in de periode van het Oudnederlands, waren ze er nog niet of nauwelijks. Het is pas in de laatste paar eeuwen dat ze opduiken en de

Europe in the 19th c. – one feature among others indicating a Sapirian *drift* of central average European, or more exactly western European languages, as a result of a more general *cultural drift* (ibid.: 182). If it is the case, we should try to understand whether this phenomenon is linked to cultural trends, to language contact, or to typological shifts (from synthetic Latin to analytic Romance).

Abbreviations

wind mee krijgen. Maar dan gaat het ook hard.” “These fixed connections have not always existed. A thousand years ago, during the period of Old Dutch, they did not exist, or barely. It is only in the last two centuries that they appeared and gained strength. But they took off quickly” [our translation]; Van der Horst 2013: 131, quoted in Hüning 2014: 439).

ABL	ablative	M	masculine
ACC	accusative	NEG	negation
ADV	adverb	NOM	nominative
AUX	auxiliary	PL	plural
COMP	complementizer	POSS	possessive
DAT	dative	PRO	pronoun
DEF	definite	PRS	present
DET	determiner	PTCP	participle
DOM	differential object marking	REFL	reflexive
F	feminine	REL	relative
GEN	genitive	SBJ	subjunctive
IND	indicative	SG	singular
INDF	indefinite	VOC	vocative
INF	infinitive		

References

- Abeillé, Anne, Danièle Godard, & Annie Delaveau. In press. *Grande Grammaire du français*. Paris: Actes Sud.
- Adler, Silvia. 2001. Les locutions prépositives: questions de méthodologie et de définition, *Travaux de linguistique* 42–43(1). 157–170.
- Adler, Silvia. 2008. French compound prepositions, prepositional locutions and prepositional phrases in the scope of the absolute use. In Dennis Kurzon & Silvia Adler (eds.), *Adpositions. Pragmatic, semantic and syntactic perspectives*, 17–35. Amsterdam & Philadelphia: John Benjamins.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad, & Edward Finegan. 1999. *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education Limited.
- Borillo, Andrée. 1997. Aide à l'identification des prépositions complexes de temps et de lieu. *Faits de langue* 9. 173–184.
- Borillo, Andrée. 2000. Degrés de grammaticalisation des prépositions composées de lieu. *Actes du CERLICO, La grammaticalisation: (dé)motivation et contrainte*. Rennes: Presses universitaires de Rennes.
- Borillo, Andrée. 2002. Il y a prépositions et prépositions. *Travaux de Linguistique* 42–43. 141–155.
- Bybee, Joan, & Paul Hopper (eds.). 2001. *Frequency and the emergence of linguistic structure*. Amsterdam & Philadelphia: John Benjamins.
- Bybee, Joan. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press.
- Căpățână, Cecilia. 2003. Locuțiunile prepoziționale [Complex prepositions]. *Limba Română* [Romanian language] XIII(1). 48–53.
- Cifuentes Honrubia, José Luís. 2003. *Locuciones prepositivas – sobre la gramaticalización preposicional en español*. Alicante: Universidad de Alicante.
- Ciobanu, Fulvia. 1957. Observații asupra prepozițiilor și locuțiunilor prepoziționale din limba română. *Studii de gramatică* [Grammar studies] 2. 357–388.
- Diewald, Gabriele, & Marijana Kresić. 2010. Ein übereinzelsprachliches kontrastives Beschreibungsmodell für Partikelbedeutungen. *Linguistik online* 44. http://www.linguistikonline.de/44_10/diewaldKresic_a.html (accessed 2 July 2013).
- Fagard, Benjamin. To appear. Prépositions et locutions prépositives. In Christiane Marchello-Nizia, Bernard Combettes, Sophie Prévost, & Tobias Scheer (eds.). *Grande grammaire historique du français*. Berlin & New York: De Gruyter.
- Fagard, Benjamin. 2009. Prépositions et locutions prépositionnelles : un sémantisme comparable ? *Langages* 173. 95–113.

- Fagard, Benjamin. 2012. Prépositions et locutions prépositionnelles : la question du renouvellement grammatical. *Travaux de linguistique* 64(1). 161–189.
- Fagard, Benjamin, & Alexandru Mardale. 2007. Systèmes prépositionnels des langues romanes : la notion de partie du discours en diachronie. In David Trotter (ed.), *Actes du XXIVe Congrès International de Linguistique et de Philologie Romanes*, 91–104. Tübingen: Niemeyer.
- Fagard, Benjamin, & Alexandru Mardale. 2012. The pace of grammaticalization and the evolution of prepositional systems: Data from Romance. *Folia Linguistica* 46(2). 303–340.
- Fagard, Benjamin, & Alexandru Mardale. 2019. Les prépositions complexes du roumain – approche sur corpus, *Revue Romane* 54(1). 176–203 (10.1075/rro.00020.fag).
- Fagard, Benjamin, Paola Pietrandrea, & Julie Glikman. 2016. Syntactic and semantic aspects of Romance Complementizers. In Kasper Boye & Petar Kehayov (eds.), *Complementizers in European Languages*, 75–130. Berlin & New York: De Gruyter.
- Fagard, Benjamin, & Walter De Mulder. 2007. La formation des prépositions complexes : grammaticalisation ou lexicalisation ? *Langue française* 156. 9–29.
- Găitănaru, Ștefan. 1999. [Criteria for distinguishing complex prepositions with the genitive from free, analyzable sequences]. *Limba și literatura română* [Romanian language and literature] XXVIII(2). 12–14.
- Hoffmann, Sebastian. 2005. *Grammaticalization and English Complex Prepositions. A corpus-based analysis*. London & New York: Routledge.
- Horst, Joop van der. 2013. *Taal op drift: lange-termijnontwikkelingen in taal en samenleving*. Amsterdam: Meulenhoff.
- Hüning, Matthias, Elena Smirnova, & Anatol Stefanowitsch. This volume. Complex adpositions in three West Germanic Languages: German, Dutch, and English.
- Hüning, Matthias. 2014. Over complexe preposities en convergentie. [About complex prepositions and convergence]. In Freek Van de Velde, Hans Smessaert, Frank Van Eynde, & Sara Verbrugge (eds.), *Patroon en argument. Een dubbelfeestbundel bij het emeritaat van William Van Belle en Joop van der Horst* [Pattern and argument. A double festschrift on the occasion of the retirements of William Van Belle and Joop van der Horst], 433–445. Leuven: Universitaire Pers Leuven.
- Koch, Peter. 2002. Das Französische in typologischer und historisch-vergleichender Sicht. In Ingo Kolboom, Thomas Kotschi, & Edward Reichel (eds.), *Handbuch Französisch*, 36–58. Berlin: Schmidt.
- Kurzon, Dennis, & Silvia Adler (eds.). 2008. *Adpositions. Pragmatic, semantic and syntactic perspectives*. Amsterdam & Philadelphia: John Benjamins.
- Leeman, Danielle. 2007. De la préposition à la locution prépositionnelle. *Modèles linguistiques* 55, 715.
- Lima, José P. 2014. Grammaticalization of complex prepositions in European Portuguese. In José P. Lima, *Studies on Grammaticalization and Lexicalization*, 163–183. Munich: Lincom Europa
- Lima, José P. 2019a. On grammaticalized complex prepositions in Portuguese: Deployment, shift, redundancy, complementation. In Benjamin Fagard, José P. Lima, & Dejan Stosic (eds.) 2019. *Les prépositions complexes dans les langues romanes*. Special issue of *Revue Romane* 54(1). Amsterdam & Philadelphia: John Benjamins, 126–140.
- Lima, José P. 2019b. Preposições complexas em Português [Complex prepositions in Portuguese]. In Ernestina Carrilho, Ana Martins, Sandra Pereira, & João P. Silvestre (eds.), *Estudos linguísticos e filológicos oferecidos a Ivo Castro* [Linguistic and philological studies offered to Ivo Castro], 1181–1211. Lisboa: Centro de Linguística da Universidade de Lisboa. [<https://repositorio.ul.pt/handle/10451/39619>]
- Marchello-Nizia, Christiane Bernard Combettes, Sophie Prévost, & Tobias Scheer (eds.). In prep. *Grande grammaire historique du français*. Berlin & New York: De Gruyter.
- Meillet, Antoine. 1912. L'évolution des formes grammaticales. *Scientia, Rivista di scienza* 12(26). (reprinted in Antoine Meillet. 1958. *Linguistique Historique et Linguistique générale*, Paris: Champion, 130–149).

- Melis, Ludo. 2003. *La préposition en français*. Paris: Ophrys.
- Petrescu, Maria. 2005. Despre locuțiunile prepoziționale cu genitivul. *Annales Universitatis Apulensis. Series Philologica*, 6.3.
- Piunno, Valentina, & Vittorio Ganfi. 2019. Usage-based account of Italian Complex Prepositions denoting the Agent. In Benjamin Fagard, José P. Lima, & Dejan Stosic (eds.), *Les prépositions complexes dans les langues romanes*. Special issue of *Revue Romane* 54(1), Amsterdam & Philadelphia: John Benjamins, 141 – 175.
- Raposo, Eduardo P., Maria F. Nascimento, Maria A. Mota, Luísa Segura, & Amália Mendes. 2013. *Gramática do Português* [Portuguese grammar]. Vols. I & II. Lisboa: Fundação Calouste Gulbenkian.
- Schoonjans, Steven. 2013. Modal particles: Problems in defining a category. In Liesbeth Degand, Bert Cornillie, & Paola Pietrandrea (eds.), *Modal Particles and Discourse Markers. Categorization and Description*, 133–162. Amsterdam & Philadelphia: John Benjamins.
- Shears, Fred. 1922. *Recherches sur les Prépositions dans la Prose du Moyen Français (XIV^e et XV^e siècles)*. Paris: Honoré Champion.
- Sibille, Jean. 2019. Les prépositions complexes en occitan languedocien: premières explorations. In Benjamin Fagard, José P. Lima, & Dejan Stosic (eds.), *Les prépositions complexes dans les langues romanes*. Special issue of *Revue Romane* 54(1), Amsterdam & Philadelphia: John Benjamins, 62–92.
- Squartini, Mario. 2013. From TAM to discourse: The role of information status in North-Western Italian *gia* ‘already’. In Liesbeth Degand, Bert Cornillie, & Paola Pietrandrea (eds.), *Modal Particles and Discourse Markers. Categorization and Description*, 163–190. Amsterdam & Philadelphia: John Benjamins.
- Stosic, Dejan. 2012. *En passant par* : une expression en voie de grammaticalisation?, *CORELA* / Special issue / *Langue, espace, cognition*. (<http://corela.edel.univ-poitiers.fr/index.php?id=2844>)
- Stosic, Dejan. This volume. Complex adpositions in Slavic languages: the case of Serbian.
- Stosic, Dejan. To appear. Les prépositions complexes du français. *Carnets de grammaire*.
- Vlasin, Veronica. 2015. Câteva considerații asupra statutului morfologic al locuțiunilor prepoziționale cu genitivul [A few considerations about the morphological status of complex prepositions with the genitive]. In Ionuț Pomian & Nicolae Mocanu (eds.), *Înspre și dinspre Cluj, contribuții lingvistice: omagiu profesorului G. G. Neamțu la 70 de ani* [To and from Cluj, linguistic contributions: A tribute to professor G. G. Neamțu on the occasion of his 70th birthday]. Editura Scriptor, Cluj-Napoca, 645–657.

Corpora

- French: Valibel, PFC, Clapi, Frantext, frTenTen; Italian: Badip, itTenTen;
- Portuguese: Corpus do Português (Davies & Ferreira 2006), ptTenTen11, CRPC (Corpus de referência do português contemporâneo), CETEM-Público;
- Romanian: CORV (Dascălu-Jinga 2002), Ruxândoiu, Romanian web corpus;
- Spanish: CREA, CORDE, Corpus del Español (Davies 2002), esTenTen.

Grammars and dictionaries (selection)

- Anglade, Joseph. 1921. *Grammaire de l'ancien provençal*. Paris: Klincksieck.
- Avram, Mioara. 1960. *Evoluția subordonării circumstanțiale cu elemente conjuncționale în limba română* [Evolution of circumstantial clauses with conjunctive elements in Romanian]. Bucharest: Editura Academiei Republicii Populare Române.
- Badia Margarit, Antoni. 1962. *Gramática catalana*. Madrid: Gredos.
- Blumenthal, Peter, & Achim Stein. 2002. *Tobler-Lommatzsch, Elektronische Ausgabe*. Stuttgart: Steiner.

- Buridant, Claude. 2000. *Grammaire nouvelle de l'ancien français*. Paris: Sedes.
- Coutinho, Ismael. 1977. *Gramática histórica* [Historical grammar]. Rio de Janeiro: Ao Livro Técnico.
- De Mauro, Tullio, Federico Mancini, Massimo Vedovelli, & Miriam Voghera. 1993. *Lessico di frequenza dell'italiano parlato*. Milano: Etaslibri.
- De Rochegeude, Henri Pascal. 1819. *Essai d'un glossaire occitanien pour servir à l'intelligence des poésies des troubadours*. Toulouse: Benichet Cadet.
- Densusianu, Ovid. 1938. *Histoire de la langue roumaine*. Paris: Leroux.
- Diccionario Medieval Español*. 1986. Salamanca: Universidad Pontificia de Salamanca.
- Dicționarul explicativ al limbii române* [Explanatory dictionary of Romanian](DEX). 1975. Bucharest: Academia Republicii Socialiste Româna.
- Ernout, Alfred, & François Thomas. 1964. *Syntaxe latine*. Paris: Klincksieck.
- Etienne, Eugène. 1980. *Essai de grammaire de l'ancien français (IX^e–XIV^e siècles)*. Genève: Slatkine Reprints. [Original edition: 1895, Nancy, Paris: Berger-Levrault].
- Euromosaic – The production and reproduction of the minority language groups in the European Union. 1996. Luxembourg: Office for Official Publications of the European Communities. [<http://www.lavplu.eu/central/bibliografie/euromosaic1.pdf>]
- Fabra, Pompeu. 2008 [1933], *Gramàtica catalana* [Catalan Grammar]. Barcelona: Institut d'Estudis Catalans.
- Gougenheim, Georges. 1974. *Grammaire de la langue française du seizième siècle*. Paris: Picard.
- Gramatica Limbii Române* [Romanian Grammar]. 2005. Bucharest: Editura Academiei Române.
- Jensen, Frede. 1994. *Syntaxe de l'ancien occitan*. Tübingen: Max Niemeyer.
- Migliorini, Bruno. 1966. *Storia della lingua italiana*. Firenze: Sansoni.
- Moignet, Gérard. [1973] 1984. *Grammaire de l'ancien français*, 2nd edn. Paris: Klincksieck.
- Moll, Francesc de Borja. 1952. *Gramática histórica catalana*. Madrid: Gredos.
- Musso, Giancarlo. 2004. *Gramática astèsan-a*. Torino: Gioventura Piemontèisa.
- Penny, Ralph J. [1993] 2014. *Gramática histórica del español*. Barcelona: Ariel Lingüística.
- Raposo, Eduardo P., Maria F. Nascimento, Maria A. Mota, Luísa Segura, & Amália Mendes. 2013. *Gramática do Português*. Vols. I & II. Lisboa: Fundação Calouste Gulbenkian.
- Real Academia Española y Asociación de Academias de la Lengua Española. 2009. *Nueva gramática de la lengua española*. Madrid: Espasa.
- Rohlf, Gerhard. [1949] 1954. *Grammatica storica della lingua italiana e dei suoi dialetti*. Torino: Einaudi [Historische Grammatik der Italienischen Sprache und ihrer Mundarten. Bern: Francke].
- Tekavčić, Pavao. 1972. *Grammatica storica dell'italiano. II : Morfosintassi*. Bologna: Il Mulino.
- Väänänen, Veikko. 1981 [1963]. *Introduction au latin vulgaire*. Paris: Klincksieck.
- Vocabolario degli accademici della Crusca*. 1866. Accademia della Crusca.
- Wolff, Alexandre (ed.). 2014. *La langue française dans le monde*. Paris: Nathan.