

The pace of grammaticalization and the evolution of prepositional systems: Data from Romance

Benjamin Fagard, Alexandru Mardale

► **To cite this version:**

Benjamin Fagard, Alexandru Mardale. The pace of grammaticalization and the evolution of prepositional systems: Data from Romance. *Folia Linguistica*, De Gruyter, 2012, 46 (2), pp.303-340. halshs-01241407

HAL Id: halshs-01241407

<https://halshs.archives-ouvertes.fr/halshs-01241407>

Submitted on 10 Dec 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The pace of grammaticalization and the evolution of prepositional systems: Data from Romance

It is generally assumed that grammaticalization is a major process in linguistic change (Lehmann 1985, Marchello-Nizia 2006). It is also assumed that grammaticalization does not affect homogeneously all linguistic categories (Hopper & Traugott 2003), or even all languages: it has been said, for instance, that some languages are further down most grammaticalization clines than others and, thus, appear more grammaticalized than others (see, for French, Lamiroy 1999, 2001, 2003, 2011, Carlier 2007, De Mulder 2001, De Mulder & Lamiroy to appear; for English, König & Gast 2007). In this paper, we illustrate the uneven pace of grammaticalization in Romance by analysing a specific area of language, viz. simple and complex prepositions. Our goal is to establish a list of prepositions which are in actual use in five present-day Romance languages (to wit, French, Italian, Portuguese, Romanian and Spanish). In order to do so, we check the validity of existing lists against written and spoken corpora, and evaluate the degree of grammaticalization of each morpheme or construction, on the basis of (mainly) morpho-syntactic criteria. Additionally, and most importantly, a corpus-based approach makes it possible to observe these items' frequency (Bybee 2006). The result offers a clear picture of the degree of grammaticalization of prepositions in present-day Romance, showing that French seems indeed to be the most grammaticalized Romance language, followed by Italian, Spanish, Portuguese and Romanian.

Keywords: adpositions, Romance, diachrony, corpus, grammaticalization.

Benjamin Fagard
Lattice (CNRS — ENS)
1 rue Maurice Arnoux
92120 Montrouge
France
benjamin.fagard@ens.fr

Alexandru Mardale
INALCO, ECO – Section de roumain & SeDyL, UMR 8202 CNRS
65, rue des Grands Moulins
75013 Paris
France
alexandru.mardale@inalco.fr

1. Introduction

The study of linguistic evolution has changed much in the last twenty years or so, partly thanks to extensive research on grammaticalization. This phenomenon seems to play a major role in language change (Lehmann 1985, Marchello-Nizia 2006): grammatical elements are overwhelmingly the result of grammaticalization, in all natural languages.¹

It has been assumed, however, that grammaticalization does not affect homogeneously all linguistic categories (Hopper & Traugott 2003). It has even been said that, though grammaticalization seems to affect languages universally, it may do so in different ways. More specifically, the speed with which constructions undergo grammaticalization seems to vary from one language to the other, and even from one period to the other, in the same language. As a result, within a family of languages, one language may be further down most grammaticalization clines than others, and thus appear more grammaticalized. This seems to be the case for English with respect to other Germanic languages (with Dutch not far behind, cf. König & Gast 2007), and for French vis-à-vis other Romance languages: it has been shown that French is more grammaticalized than Italian, and Italian more grammaticalized than Spanish, by comparing for instance the paradigms of determiners, auxiliaries, or the expression of tense and mood (Lamiroy 1999, 2001, 2003, 2011, Carlier 2007, De Mulder 2001, De Mulder & Lamiroy to appear).²

The questions we address in this paper are the following: Is it really possible to compare the degree of grammaticalization of languages? Is this scale of grammaticalization valid for all areas of the language? Do the paradigms of adpositions in Romance confirm the trend mentioned above? In order to answer these questions, we first evaluate different ways of comparing the degree of grammaticalization of languages (Section 2), then propose a definition of adpositions (Section 3) and describe the adpositional system found in Latin (Section 4). We go on to describe the adpositional systems of present-day French, Italian, Portuguese, Romanian and Spanish (Section 5), on the basis of existing lists (e.g. Fagard & Mardale 2007, Fagard 2010). This description enables us to compare the degree of grammaticalization of adpositions in these languages. Finally, in order to check the validity of this scale of grammaticalization – based on a purely formal description – we compare in an onomasiological perspective the use of simple and complex adpositions in Romance, on the basis of modern (mostly spoken) corpora (Section 6), and conclude with suggestions for future research (Section 7).

2. Comparing the (degree of) grammaticalization of languages

We deal here with the first question asked in the introduction: Is it possible to compare the degree of grammaticalization of different languages? It is by now almost consensual to say that grammaticalization theory proposes an adequate framework for the description of language-internal changes involving the drift (to use Sapir's term) from lexicon to grammar. It is mostly used as a tool to describe the development of a given morpheme or construction, or the appearance of a new part of speech in a language, for instance determiners in Romance

¹ Or at least in all languages investigated to date in this perspective, to the authors' knowledge. Other phenomena include reanalysis and analogy, claimed to be "major mechanisms in language change" (Hopper & Traugott 2003: 69), and marginally borrowing, which is attested for adpositions (e.g. Romanian *başca* from Turkish, 'international' *via* from Latin, etc.).

² In a different theoretical framework, Gawelko (1999, 2001) also points out the outstanding position of French, "less redundant" than other Romance languages (1999: 24). As he recalls, Brøndal (1939: 343) already noted the fact that French had a more grammaticalized set of prepositions than all other Romance languages, though Gawelko argues against this idea (Ibid.: 31-33).

(Carlier 2007) or demonstrative adjectives vs pronouns in French (Marchello-Nizia 1995). Only recently have linguists started using this framework to compare the degree of grammaticalization of different languages. This new type of study entails a series of questions: we will try to determine in this contribution whether it is legitimate to compare languages in this way, and provide clues as to how such a comparison can be operationalized, i.e. how we can use a theoretical definition to measure the degree of grammaticalization on the basis of corpus data.

The first issue is whether such a comparison is legitimate. It is quite obvious that languages do not necessarily grammaticalize the same elements: what is grammatical in one language can be expressed lexically or not at all in another one (cf. evidentials (Aikhenvald 2004), to give just an example). This does not mean, however, that the latter is ‘less grammaticalized’. In order to establish a ranking according to the degree of grammaticalization of a language as a whole, we would need – ideally – to make a list of all possible grammatical features, and check for each language which ones are indeed the result of grammaticalization processes. Moreover, even if we can do so, it is far from obvious that all categories should be given the same weight. For instance, should a language with more TAM markers and less deictic markers be considered more or less grammatical than a language with opposite features? It seems quite impossible to answer. What does seem possible, though, is to compare related languages, one category at a time (see De Mulder & Lamiroy, to appear), or to compare the global tendency of related languages (Gawelko 2001: 397). For each grammatical category, we can evaluate and compare the degree of grammaticality of two languages.

The second issue – how to operationalize this kind of comparison – is also delicate. There are various ways to do so. In previous studies, the main factor called upon is the degree of paradigmaticization of a given word class (Lamiroy 2011, De Mulder & Lamiroy to appear), i.e. its size and obligatory vs optional nature (ibid.), following Lehmann’s (1995: 164) idea that a situation in which the speaker has a free choice between various elements on the basis of his communicative intentions is less grammaticalized than a situation in which his choice is systematically constrained.

Concerning adpositions, when comparing two languages, we can hypothesize that one of them will display at least one of the following features with respect to the other (a-d):³

- a) a greater number of simple adpositions;
- b) a more grammaticalized set of simple adpositions;
- c) a more specialized class of adpositions, cut off formally and distributionally from adverbs, conjunctions, preverbs, etc.⁴;
- d) a greater number of adpositions in general.

Is one of these features a relevant indication of the language’s (greater) degree of grammaticalization? In what follows, we will see that French is probably most grammaticalized among Romance languages, for all four features. It remains to be seen how we determine which adpositions are most grammaticalized; in order to do so, we need a precise definition of adpositions, which is given in the following section.

³ Note that we view adpositions as grammatical elements, though it is quite clear to us that grammaticality, in this respect, is a matter of degree: adpositions are less grammatical (as a whole) than, say, case marking on nouns or TAM marking on verbs, but more grammatical than nouns or verbs.

⁴ We mean here that, when used as an adposition, a given morpheme which may have been used earlier as a verb, an adverb or a noun will have a different distribution and different formal characteristics. This does not mean that it can no longer be used with its original function.

3. A working definition of adpositions

Traditional definitions of grammar often present grammatical elements as members of closed-class paradigms. Elements which do not fit in well because they annoyingly display features of two different categories, or not all features of one category, can then be described as ‘poor’ instances of the category in question: ‘pseudo’ adpositions, ‘improper’ adverbs, etc. In our view, a prototypical framework more adequately describes such classes (Company Company 2002): grammatical categories are then seen as heterogeneous paradigms containing elements which display the category’s prototypical features to various degrees. In the case of adpositions, we expect to find a core of prototypical adpositions, a set of less prototypical elements and finally a set of elements which share only some (proto)typical features of adpositions. The prototypical framework implies, besides, that these distinctions are not clear-cut, and that many elements can be shown to be in-between these subcategories.

Many linguists have worked on adpositions, and proposed a series of features which present partial overlap (among others, Rizzi 1988, Van Riemsdijk 1990, Zwarts 1997, Gaatone 2001, Di Meola 2000, Melis 2003, Abeillé et al. 2003, Cuniță 2004, Ciobanu & Nedelcu 2008, Mardale 2009, Hagège 2010). Drawing on these studies, we will take (i) as a working definition of adpositions. This definition consists of a series of features ranging from syntax (a, b, c) to semantics (d), morphology (e) and frequency (f).

(i) An adposition

- (a) is used as the head of an adpositional phrase
- (b) typically takes a noun phrase as its complement
- (c) syntactically relates its complement to a unit higher up in the clause hierarchy
- (d) semantically qualifies this relation
- (e) formally consists of a simple morpheme, and is short (no more than two syllables)
- (f) is frequent

A prototypical adposition is expected to present all features. Note that many adpositions will not satisfy all criteria, at least not in all contexts of use.⁵ Besides, the above features include inherent, empirically observable characteristics (e, f) and others which can only be found as the result of the linguist’s analysis, on the basis of previous theoretical assumptions. For this reason, the operationalization of this definition will have to rely heavily on observable features, i.e. (b), (d), (e) and (f) (and, to a lesser extent, (c)).

This definition is valid in synchrony. From a diachronic point of view, one would have to add that adpositions are sometimes borrowed, but are much more frequently the result of a grammaticalization process. Even though linguists might disagree as to what weight must be given to each factor, it seems consensual to say that the grammaticalization of a morpheme with respect to the source construction involves mainly but not always phonetic attrition, semantic weakening and morpho-syntactic loss of flexibility. These phenomena can operate on different types of elements and have, consequently, different outcomes: the grammaticalization of a construction may yield a complex adposition; that of a complex adposition may yield a simple adposition; finally, that of a simple adposition yields what we will call a ‘functional’ adposition, which may in turn grammaticalize further (see below, (ii)). This evolution is gradual, so that it is sometimes a complex matter to determine whether a

⁵ For instance, concerning feature (a), a functional adposition such as *to* in *I gave a book to Paul* may be analysed as not being the head of an adpositional phrase; concerning feature (b), an adposition such as *from* in *I ran from behind the school / from there* does not always take a noun phrase as its complement, but the adpositional phrase *behind the school* or the adverb *there*; the same can be shown for features (c)-(f).

given construction is a free construction or a complex adposition, or whether a simple adposition is functional or not.

We describe each subtype in section 5; our point here is that, if determining the degree of grammaticalization of a morpheme is sometimes complex, it is even more so if one is looking at a whole paradigm. Keeping to adpositions, there are languages with a very small subset of simple (highly grammaticalized) adpositions – sometimes only one, as in Tzeltal (Brown 2006; Grinevald 2006), Samoan, Burmese, Siamese, Lipke or Tagalog (none in Klamath, cf. DeLancey 2005, and a few other languages, cf. Hagège 2010: 68). Other languages have just a few (Longgu), a few dozens (Portuguese, cf. Jakubowicz Batoréo 2000: 496), or more (fifty or so adpositions, for instance, in English). From this point of view, we have to determine which Romance language has undergone most renewal, along two different lines of evolution: which language has (a) renewed its adpositional system the most, and (b) grammaticalized furthest its already existing adpositions? In order to answer question (a), we first need to describe the adpositional system of Latin: this is the object of the next section. We will base our answer to question (b) by measuring the degree of grammaticalization of each adposition with respect to the grammaticalization cline usually given for adpositions (see e.g. Lehmann 1985, Svorou 1994: 101 sqq., Fagard 2006):

(ii) complex adposition > simple adposition > functional adposition > case affix > \emptyset ⁶

4. Latin adpositions and paradigm renewal

Romance languages did not create their adpositional paradigms out of nothing: there were already adpositions in Latin. Were they comparable to paradigms of adpositions in Romance? In order to answer this question, it will be helpful to describe briefly the paradigm of adpositions in Latin.

4.1. Latin adpositions

Latin had quite a few adpositions, as the reader can judge by himself on the basis of Table 1, which lists a total of 80 different morphemes. However, as indicated in section 3, we expect these adpositions to be heterogeneous in form and function. Indeed, some Latin adpositions were also used as preverbs (for instance *per*, *ad*, *ab*, *ob*, *in*, *ex*, *cis*, *cum*, *de*, *sē*, *uls*, cf. Meillet 1948: §784, Leumann, Hoffmann & Szantyr 1965; for their preverbal uses, see Haverling 2003), while others were also used as adverbs (*adversus*, *ante*, *circa*, *circum*, *citra*, *contra*, *coram*, *extra*, *intra*, *infra*, *juxta*, *post*, *prope*, *propter*, *supra*), others yet as both adverbs and preverbs (*post*, *ante*, *sub*, *super*, *apud*). Some of these only had a limited functional resemblance to the rest of the category, having rare uses as adpositions, and not being fully grammaticalized: these adpositions were presumably recent additions to the paradigm, such as *coram* ‘in the presence of’, *causa* ‘because of’, *gratia* ‘thanks to’, *ergo* ‘because of’, *erga* ‘in front of, concerning, for’, *penes* ‘at the hands of’, *procul* ‘far from’, *circum* ‘near’, *tenus* ‘all the way to’, *fine* (and *fini*) ‘all the way to’, *foras* (and *foris*) ‘out of’ (Meillet 1948: 521).⁷ This

⁶ Note that this is but one possible outcome of the further grammaticalization of adpositions. Other outcomes include the grammaticalization of functional adpositions into determiners, such as French *de*, or into complementizers (Noonan 2007: 57), among other things (see Hagège 2010).

⁷ Note that these constructions are not equivalent and need not have undergone grammaticalization to the same extent or along the same lines. For instance, *gratia* or *causa* can be considered adpositions only when they combine with a noun in the genitive, then constituting an adjunct as a whole; the use of *foras* or *circum* is quite different, since these items function as adverbs, and thus as adjuncts, whether they combine with a noun or not. Their sole commonality, in our view, is that they are distributionally equivalent to adpositions, without having grammaticalized fully.

last class of adpositions is mainly composed of nouns in a fixed form, most frequently the ablative, which always govern a complement in the genitive, at least in the first stages of their grammaticalization, whereas some adpositions of the other subclasses govern more than one case (Sausy 1992: 177-178). Table 1 below illustrates the heterogeneity of Latin adpositions, on the basis of the criteria given in section 3, with a core class of only 19 morphemes (with frequent use, always preposed, cf. Hagège 2010: 117, and reduced phonetics), out of the 80 which have adpositional uses.

Type of adposition	Forms
'prototypical' adpositions ⁸	ab, ad, ante, apud, contra, cum, de, ex, in, infra, ob, per, post, prae, pro, sine, sub, super, trans
recent grammaticalizations (some used as adverbs and adpositions; less frequent)	absque, adversum, adversus, asuper, beneficio, causa, circa, circum, cis, citra, clam, coram, deintus, deretro, desub, desubtus, erga, ergo, extra, foras, foris, gratia, insuper, inter, iuxta, merito, palam, penes, praeter, procul, prope, propter, retro, secundum, subinde, subter, subtus, supra, tenus
atypical adpositions (either archaic, not completely grammaticalized or late; still less frequent)	abusque, adusque, cata, circiter, exadversum, exsuper, fine, fini, intra, intus, pone, pridie, propius, proxime, sē(d), secus, simul, uls, ultra, usque, versus, vitio

Table 1: Simple Latin adpositions.

Some of these morphemes disappeared more or less completely in the transition from Latin to Romance, such as *coram*, *ob*, etc. Others survived only locally, like *versus* in Gallo- and Italo-Romance. Only a few are still found in more or less all Romance languages, including *in*, *ad*, *de*, *per* and *pro*.

4.2. Loss and renewal

The transition from Latin to Romance resulted in a partial reduction of the original paradigm. We will not dwell here on this evolution, since we do not plan to account for the differentiation of Romance languages in diachrony, but to describe its results in synchrony.⁹ Suffice it to say that Romance languages replaced Latin adpositions by new ones mostly through grammaticalization processes. Some of these were local, e.g. Provençal *doumaci* 'thanks to' (lit. 'god-grace', i.e. "thank god"), French *chez* 'at the house of, in the works of', while others were quite widely spread, like Catalan, French, Provençal and Romansh *durant* (with various spellings in Romansh: *duront*, *durànt*, etc.), Italian, Spanish and Portuguese *durante* 'during', which is also present as a loan-word in Sardinian *duranti*: only Romanian seems to have no similar construction, using instead *în timpul* "in the time" followed by a noun phrase in the genitive.¹⁰ Another important phenomenon was the grammaticalization of simple Latin adpositions into functional adpositions, attested in all Modern Romance languages – even in Medieval Romance.

⁸ Some of these adpositions also display uses as preverbs: they might therefore be considered less prototypical (as adpositions).

⁹ For a more detailed description of this evolution, see Fagard (2010).

¹⁰ For some of these items, one might wonder whether they are sufficiently grammaticalized to qualify as full adpositions: they could also be considered as lexical items, which underwent not a process of grammaticalization but one of lexicalization. In our view, even for items which did not reach 'full' adpositional status, such as *oiant*, *veant* frequent in phrases like *oiant toz*, *veant toz* (lit. "(everyone) hearing, seeing"), there is a process of lexicalization, which can be considered as the first step of a grammaticalization process (Prévost & Fagard 2007).

5. Paradigms of adpositions in Romance

The result of the renewal sketched out in the previous section is that Romance languages all have adpositional systems which are partly based on Latin adpositions and partly the result of the more recent grammaticalization of elements from various categories. As predicted by the theoretical description in section 3, Romance adpositions are very heterogeneous, with different classes of adpositions in each language (cf. e.g. Melis 2003: 41-43, for French adpositions). This is because they do not all display the entire list of features outlined in (i): some are more prototypical than others. Besides, while it is generally possible to determine whether a given adposition is more or less grammaticalized than another one, it seems impossible to establish clear-cut category boundaries between very grammaticalized, less grammaticalized and hardly grammaticalized morphemes. As with other parts of speech, grammaticality is a matter of continuum rather than of discrete categories (Di Meola 2000: 144 sqq.).

5.1 Paradigms

In each language, a small class of adpositions always take a complement, which typically is a noun phrase but can also be an adpositional phrase, an adverb or even a whole clause.¹¹ These highly grammaticalized adpositions, which we call *functional* adpositions (cf. however Mardale 2009 for a stricter definition of functional adpositions), make up a very limited set in each Romance language (Table 2).

Spanish	French	Italian	Portuguese	Romanian
a, de, en, por, para, con	à, de, en, par, pour, avec, dans ¹²	a, da, di, in, con, per	a, de, em, por, para, com	a, de, la, pe, în, cu

Table 2: Highly grammaticalized adpositions in Romance.

These morphemes display a very high frequency, associated with almost all prototypical features of adpositions, except feature (a) in definition (i) (Section 3), which is not always satisfied. As a result of the fact that the definitional features of adpositions are not always present, membership of this category is a matter of degree: some of the adpositions listed above are more grammaticalized than others. For instance, one common feature of these highly grammaticalized adpositions is their greatly bleached meaning and corollary polysemy, but this is a gradual feature, as illustrated by the fact that it seems more applicable to French *à* and *de* than to *avec*. Another common feature, less prototypical of adpositions, is the fact that they have various ‘grammatical’ uses, introducing for instance the genitive or dative object (2),¹³ along with ‘predicative’ uses of *de* (3), the dative (4), the patient of a passive (5) or the introduction of a given verb’s complements (6a-c). We illustrate below a few of these features

¹¹ Cf. Melis (2003: 12-17), who considers these complements to be distributional equivalents of a noun phrase, or Hagège (2010: 58), who provides a hierarchy of more or less frequent terms governed by adpositions, from personal pronouns and noun phrases (most frequent) to subject-predicate associations (least frequent).

¹² We could add here the adposition *sur* (Table 3), which seems to have been undergoing a process of further grammaticalization since the 19th century, with increasing frequency and clear semantic bleaching, especially in a few specific contexts. For instance, *sur* introduces the complements of a growing number of verbs (see De Mulder & Vanderheyden, 2000).

¹³ Generally expressed by the adposition *de/di*, sometimes by *a/à*, as in example (2). Note that the semantic relation could be seen as genitive (a rough equivalent to *John’s car*) or dative (glossed as *she’s being friendly to us both*).

(for a more developed analysis, see Melis 2003). Paradoxically, in their most grammatical uses, these items end up behaving less prototypically as adpositions, and may actually develop into other grammatical morphemes such as complementizers, determiners, etc., as we saw with the grammaticalization cycles given above in (ii).

- (2) madame brumer c' est une copine à nous deux
 missis Brumer DEM be.PRS.3SG INDF.F.SG friend at we.OBJ two
 'Missis Brumer is a friend **of** the two of us.' (Clapi, *Bielefeld*, 225-230)
- (3) - não é perigoso, o javali?
 NEG be.PRS.3SG dangerous.M.SG DEF.M.SG boar.SG
 - ah! coitado (sic) **dos** bichos! perigosos somos nós!
 oh! poor.M.SG of.DEF.M.PL beast.PL dangerous.M.PL be.PRS.1PL we.SBJ
 '- Aren't boars dangerous?
 - Oh, poor beasts! We're dangerous, not them!' (Corpus do Português, 1989, Lisboa, Cartografia Portuguesa)
- (4) quindi dare a Tizio per togliere a Caio non c' è
 so give.INF at Titius to take.INF at Caius NEG there be.PRS.3SG
 nessun vantaggio
 NEG.INDF.M.SG advantage.SG
 'So if you give to Titius what you took from Caius, what is the use?' (Badip, RE22B)
- (5) Pues me gustan los libros que hablan sobre mujeres,
 Well me.OBJ please.PRS.3SG DEF.M.PL book.PL REL speak.PRS.3PL on woman.PL
 feministas y eso, escritos **por** mujeres y que
 feminist.F.PL and DEM.M.SG write.PST.M.PL by woman.PL and REL
 cuenten historias de mujeres.
 tell.PRS.SBJV.3PL story.PL of woman.PL
 'Well, I like books that talk about women, feminists and all that, written by women and which tell womens' stories.' (Crea, Grupo G 7, Magisterio, España)
- (6a) ne-am întâlnit **cu** emil la gară
 REFL.1PL.OBJ-have.PRS.1PL meet.PST with Emil at station.SG
 'We met Emile at the station.' (Ruxăndoiu, dialog *La Brăila, la vie*)
- (6b) m-am întâlnit **cu** ea și cu
 REFL.1SG.OBJ-have.PRS.1SG meet.PST with PRO.F.G.SG¹⁴ and with
 maică-sa
 mother.SG-POSS.F.3SG
 'I met her and her mother.' (Ruxăndoiu, dialog *Într-o cameră de hotel*)
- (6c) s-a întâlnit **cu** ea la: prefectură
 REFL.3SG-have.PRS.3SG meet.PST with PRO.F.G.SG at precinct.SG
 'He met her at the precinct.' (Ruxăndoiu, dialog *În redacție*)

¹⁴ We arbitrarily gloss the Romanian genitive-dative as 'G', for convenience, since they are homonymous in nouns.

Finally, as a rule¹⁵, they cannot be used as adverbs; compare (7a) with (7b):

- (7a) **Il marche à / Il marche à Paris.*
 PRO.M.SG.SBJ walk.PRS.3SG at PRO.M.SG.SBJ walk.PRS.3SG at Paris
 ‘*He walks in.’ ‘He walks in Paris.’
- (7b) *Il marche devant (moi).*
 PRO.M.SG.SBJ walk.PRS.3SG ahead PRO.1SG.OBJ
 ‘He walks ahead (of me).’

The degree of grammaticalization of these adpositions is reflected in their frequency, their degree of semantic bleaching and foremost the variety of their uses: the least grammaticalized (*cu, con, com, avec*) only have few ‘grammatical’, i.e. syntactically bound uses, typically the introduction of verb complements (ex. (6a-c)).

The next series of adpositions, in terms of degree of grammaticalization, is that of ‘lexical’ adpositions. They take a complement which is not introduced by a functional adposition; however, in some cases, this complement is not lexically realized. They make up a larger series, in each language, and display a lesser degree of grammaticalization. They are characterized by only *some* of the typical features of adpositions (cf. (i)). For instance, adpositions such as *côté, question*, etc. only take bare nouns as complements, to the exclusion of real NPs (8a-b):

- (8a) *J’ ai du mal à me stabiliser côté boulot...*
 I.SBJ have.PRS.1SG of.DEF.M.SG wrong at PRO.1SG.OBJ stabilize.INF side work
 ‘Concerning my job, it’s hard for me to be stable.’ (internet)¹⁶
- (8b) *Là question travail ça valait mieux que (inaud.)*
 there question work DEM.M.SG be.good.PST.3SG better than (inaudible)
 ‘At that point, as far as the job was concerned, it was better than...’ (Clapi, *At the dentist’s* n°1, 2003)

Other morphemes, such as *contro, derrière* can be used either as adpositions or as adverbs. Examples (9a-b) illustrate the difference between an adposition with no complement (9a, in which *contro* can have a complement: *chi vota contro questa decisione?* ‘who votes against this decision?’) and an adverb (9b, where *derrière* cannot have a complement; *derrière* ‘behind’ is used to mean ‘in the back of the field’, but ‘the field’ cannot be used as a complement in this case).

- (9a) *adesso non sto vedendo chi vota contro * #*
 now NEG be.PRS.1SG see.PROG.M.SG REL.SBJ vote.PRS.3SG against
- nessuno chi si astiene * * alzate le mani*
 NEG.INDF.M.SG who.SBJ REFL.3SG abstain.PRS.3SG raise.IMP.2PL DEF.F.PL hand.PL

¹⁵ There are exceptions, for instance stylistically marked (or coordinated) uses in which the complement can be absent. Some of these adpositions can have uses with a ‘zero’ or ‘null’ complement, in particular *pour* and *avec* in French.

¹⁶ Written August 3rd, 2005; accessed September 2010 (<http://www.yabiladi.com/forum/stabiliser-cote-boulot-1-677162.html>).

gli astenuti prego
 DEF.M.PL abstainer.M.PL please
 ‘Now I can’t see who’s voting against. No one. Who abstains? Anyone abstaining please raise your hand.’ (Badip, MC111B)

(9b) nan nan nan/ joue **derrière** joue **derrière** \ qu’
 NEG NEG NEG play.IMP.2SG behind play.IMP.2SG behind COMP

on perde pas l’ ballon d’ entrée là\
 INDF.3SG.SBJ lose.PRS.SBJV.3SG NEG DEF.M.SG ball of entrance there
 ‘Come on, stay in the back, stay in the back, we don’t want to lose the ball right away now.’
 (Clapi, *Video games – soccer*, Aix, 549-550)

Some of these adpositions are only used in specific genres, literary or other: this is the case of French *jouxte* ‘near, next to’ (old, lit. or technical term), Spanish *cabe* ‘near’ (poetic), Italian *avverso* ‘against’ (old, technical) and many others. Listing simple adpositions exhaustively without including morphemes which are only rarely used as such is therefore quite problematic. Another problem is to decide what makes a simple adposition ‘simple’: in some cases, though the univerbation of a complex adposition has phonetically taken place, as shows its lack of variation, the resulting morpheme is still a complex one – in writing. This is the case for French *à travers*, *par-devers*, *par-dessus*, *par-dessous*, etc., which never allow insertions or modifications, as illustrated by the comparison of (10a-b). This justifies in our view the inclusion of *à travers* and the like in the paradigm of simple adpositions; the fact that it is quite impossible to gloss independently *à* and *travers* is also an indication that the univerbation has indeed taken place.

(10a) À côté de lui → à son côté
 at side of him at POSS.M.3SG side
 ‘By his side.’

(10b) à travers lui → *à son travers
 through him at POSS.M.3SG (intranslatable)
 ‘Through him.’

Taking into account all these elements, it can be shown that simple adpositions form a large set, as can be seen in Table 3 below.

	Spanish	French	Italian	Portuguese	Romanian
‘prototypical’ adpositions (never used with another adposition introducing their complement, except in Italian with personal pronouns;	ante, bajo, contra, desde, durante, entre, excepto, hacia, hasta, incluso, mediante, salvo, según, sin, sobre, tras	(mis) à part, à travers, après, avant, chez, concernant, contre, d’après, d’entre, de par, depuis, derrière, dès, devant, durant, entre, envers, environ, excepté, jusqu’à,	circa, dopo, durante, eccetto, entro, fra, malgrado, mediante, salvo, secondo, senza, su, tra, tramite, tranne, verso	após, consoante, dèsde, durante, excepto, mediante, perante, salvo, segundo, sem, sob, sobre	(pe) lângă, sub, în, spre, cu, peste, fără, între, până, contra, grație, datorită, mulțumită, drept, versus, via, bașca, dintre, pentru, prin, printre,

'frequent' use in all registers) ¹⁷		malgré, outre, par devers, par-delà, par-derrrière, par-dessous, par-dessus, par-devant, parmi, passé, pendant, sans, sauf, selon, sous, sur, vers			după, înspre, împotriva
adpositions sometimes used intransitively ¹⁸	junto	hors, en face, en dessous	appresso (a), attraverso (a), contro (a), davanti (a), dentro (a, da, di), dietro (a), dinanzi (a), fin(o) (a / da), fuori (da, di), innanzi (a), intorno (a), lungo (a), oltre (a), presso (a), rasente (a), sopra (a), sotto (a), vicino (a)	até, conforme, fora, trás	după, peste, sub
literary, specialized or not completely grammaticalized adpositions	allende, cabe, cuando, empós, entro, menos, pro, so, tipo, versus, via	attendu, côté, ès, hormis, jouxte, jusque, lez, moyennant, niveau, nonobstant, plein, question, suivant, touchant, vu	appo, avverso, giusta, nonostante	afora, dês, menos, senão, tirante, visto, tipo	gen, tip ¹⁹

Table 3: Lexical adpositions in Romance (simple and complex but unanalysable morphemes, e.g. *à travers*, *jusqu'à*)

Finally, there is in each language a large series of complex constructions which take on syntactic and semantic functions similar to those of simple adpositions. These constructions, which we will call complex adpositions, can follow various patterns (Melis 2003: 106 sqq.), but generally consist of a lexical nucleus, generally a noun, preceded and/or followed by (most often functional) adpositions, e.g.

¹⁷ Note that some of these morphemes can also be used as adverbs, for instance.

¹⁸ This category needs to be refined, as it contains elements which are mostly transitive and others which are mostly *intransitive*. This is no simple task, since these variations can be diatopical (for instance French *en face (de)*), diachronic (e.g. Modern Spanish *detrás de* vs Medieval Spanish *tras*), etc.

¹⁹ Adpositions such as Romanian *gen*, *tip*, French *genre*, Italian, Portuguese or Spanish *tipo* 'like, for instance' are actually closer to hedges. However, from a distributional point of view, they present some similarities with adpositions – more or less depending on language and period – which is why we included them in this account.

[Adposition (+ Determiner) + Noun/Adverb/Adjective/Participle/Infinitive/... + Adposition].

The most frequent patterns observed in Romance languages are the following:

[Adposition + Noun + Adposition],

[Adposition + Determiner + Noun + Adposition],

[Adverb + Adposition]

Note that the lexical item which forms the core of the construction (noun, adverb, etc.) generally loses its referential potential or part of its meaning once it is integrated in a fixed construction. Different types of complex adpositions are illustrated in Table 4 below, which does not aim at exhaustivity, neither for each construction nor for the types of constructions.

Pattern	Spanish	French	Italian	Portuguese	Romanian
Adposition + Noun + Adposition	a camino de, a través de, de acuerdo con	à cause de, à côté de, à défaut de	a malgrado di, a causa di, a confronto di, da parte di	a despeito de, a fim de, de redor de	cu privire la, în curs de, în ajun de, în comparație cu, în decurs de, în legătură cu, în loc de, în relație cu, la nivel de, pe bază de, pe timp de
Adposition + Determiner + Noun + Adposition	al revés de, en el interior de	au-dessus de, au sein de, du côté de, au moyen de	all'opposto di, all'interno di	no meio de, no fundo de	(structurally impossible)
Adverb + Adposition	encima de, debajo de, delante de	en-dessous de, à côté de, en-dehors de	accanto a, addosso a, attorno a	abaixo de, acerca de, acima de, cerca de	în afară de, alături de, departe de, vizavi de, înainte de, împreună cu

Table 4: Recurrent patterns of complex adpositions in Romance.

A specific characteristic of Romanian is that, owing to the existence of a partial case system, complex adpositions can follow other patterns, illustrated in Table 5 below.

Schema	<i>Spatial constructions</i>	<i>Other constructions</i>
Adposition + Noun + Determiner (+ Genitive)	în urma, în mijlocul, în limita, în jurul, la marginea, la nivelul, pe parcursul, pe raza, pe marginea, pe malul, în fața	în baza, în numele, în privința, în vederea, din cauza, din momentul, pe vremea
(Adposition +) Adverb + Determiner (+ Genitive)	în afara, dedesubtul, înapoia	

Table 5: Other types of complex adpositions in Romanian

5.2 Adpositional categories and pace of grammaticalization

The data presented above clearly indicate that Romance languages have quite similar categories of adpositions. However, there are differences. For instance, the number of simple adpositions varies greatly from one language to the other, with many more adpositions in French (around 40) than in all others (from 13 to 23). Types of complex adpositions also vary: in Spanish and Portuguese, the pattern [(Adposition +) Adverb + Adposition] is very frequent, much more than in French or Italian, which display other patterns; in Romanian, other

patterns are frequent on account of the case system, and the class of complex adpositions thus seems much less grammaticalized (or lexicalized). We could, provisionally, claim French to be the most grammaticalized language for all four criteria given in section 2, with a very large set of simple adpositions; besides, it has relatively rigid patterns of complex adpositions²⁰. It is followed by Italian, which has a large set of simple adpositions but many adpositions displaying various types of variation, such as the presence of *di* when the complement is a pronoun, or the presence of the adposition *a* which governs the complement, as in *dietro all'albero* “behind (at) the tree”. Italian also has patterns of complex adpositions. Next come Spanish and Portuguese, with small sets of simple adpositions, and recurrent patterns of complex adpositions. Finally, Romanian adpositions are quite clearly less grammaticalized, since they do not ensure alone the syntactic link with their complement: this link is also the result of case marking. Thus, although Romanian has relatively many simple adpositions, they are all to be considered discontinuous morphemes, and therefore less grammaticalized. The existence of more varied patterns of complex adpositions confirms this status of Romanian. Before reaching any conclusion as to the degree of grammaticalization of these classes of adpositions, we will have to take other elements into account.

6. An onomasiological approach

6.1. A ‘basic’ set of adpositions?

We believe it is necessary to go beyond a purely paradigmatic view of the language; as pointed out a.o. by Schøsler (2008), usage is more important than the lexical distribution itself²¹. For this reason, we decided to study the use of Romance adpositions with an onomasiological perspective. In order to do this, we propose to analyse the expression of a subset of contrastive spatial configurations given in Figures a-h by simple and/or complex adpositions.



Figure a (IN(SIDE))



Figure b (OUT(SIDE))



Figure c (ON)

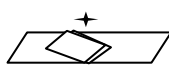


Figure d (ABOVE)



Figure e (UNDER)



Figure e' (UNDER)²²



Speaker
Figure f (IN FRONT OF)



Speaker
Figure g (BEHIND)

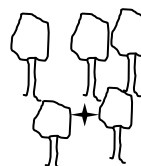


Figure h (BETWEEN / AMONG)

²⁰ We believe this to be an indication of their greater degree of grammaticalization – the less a pattern is flexible, the more entrenched and grammaticalized it is.

²¹ Gawelko (1999: 20) also stresses the need to go beyond purely formal criteria for the classification of Romance languages.

²² This type of spatial configuration was left aside, on account of the fact that no Romance language distinguishes between Figures e and e' (i.e. under with vs without contact).

We have seen in section 3 that a simple adposition is more grammatical than a complex one. Our working hypothesis, which is admittedly intuitive, is as follows:

The more a language uses simple (rather than complex) adpositions to express ‘basic’ configurations, the more its adpositional system is grammaticalized.

The global degree of grammaticalization of a language could then be measured by the number of basic spatial configurations which it expresses with a grammatical(ized) item or construction. Some authors have claimed the existence of semantic universals (e.g. Wierzbicka 1988), or written that “the simplest spatial notions are topological – concepts of proximity, contiguity, containment” (Piaget & Inhelder 1956, apud Levinson et al. 2003: 485). This claim is linked to the much discussed idea that, among spatial adpositions, those expressing configurations such as IN, ON and UNDER are learned first in language acquisition (Johnston & Slobin 1979, Hickmann 2007)²³. However, our position is quite different: for one thing, it seems that children do not always acquire spatial adpositions first (Morgenstern & Sekali 1997: 210, Hallan 2001: 115), and that the order of acquisition of new words might be better accounted for by frequency (cf. Bybee & Hopper 2001). Besides, as Evans & Levinson put it, “languages differ so fundamentally from one another at every level of description (sound, grammar, lexicon, meaning) that it is very hard to find any single structural property they share” (2009: 429); more to the point, “languages differ enormously in the concepts that they provide ready-coded in grammar and lexicon” (Ibid.: 435)²⁴. The idea of conceptual universals would therefore require hard evidence.

What we have in mind here is not that there are basic spatial concepts, or semantic primitives, but that, as was demonstrated for colors by Berlin & Kay (1969), there are probably some spatial configurations which are more often expressed by languages than others. This is what seems to indicate the study on adpositions by Levinson et al. (2003²⁵), which investigates the question of spatial categories, trying to find out whether there are implicational hierarchies in this domain. They show that, in the languages of their sample, simple adpositions follow a hierarchy: while some spatial configurations (such as AT, IN, ON) are expressed by adpositions in most or all languages, others are rarely expressed by (simple) adpositions (such as SPIKED, HANGING OVER, DISTRIBUTED OVER). The result of their study is thus an implicational hierarchy with more or less basic spatial configurations (Ibid.: 510):

AT < IN < ON, UNDER < OVER, NEAR < ON-TOP < ATTACHED < INSIDE < SPIKED,
HANGING, DISTRIBUTED OVER

This hierarchy is to be interpreted as follows: if there is a simple adposition for ON-TOP, there will be simple adpositions for all configurations to the left (i.e. AT, IN, ON, UNDER, OVER, NEAR). In other words, simple adpositions tend to refer to simple spatial configurations, more specifically topological relations,²⁶ while more complex adpositions tend to encode intrinsic and relative relations; this phenomenon is widely attested, in languages which are typologically and geographically far apart (Fortis & Fagard 2010: IV). This can be illustrated

²³ Of course, it is quite impossible to know what concepts underlie these adpositions. This is why we prefer to work with spatial configurations.

²⁴ The authors argue against two opposite claims – that “closed-class items reveal a recurrent set of semantic distinctions” (Talmy 2000) or that “open-class items (like nouns) are (...) less cross-linguistically variable” (Gentner & Boroditsky 2001) –, considering that “neither of these views seems correct, for both ends of the spectrum are cross-linguistically variable” (Evans & Levinson 2009: 436).

²⁵ Using a booklet of drawings illustrating various topological relations (spatial relations corresponding to those expressed by English *in, on, under, over, near, against, inside, on top of, in the middle of...*), Levinson et al. analyse the way adpositions are used in different languages to express topological relations. Their survey includes Basque, Dutch, Ewe, Lao, Lavukaleve, Tiriyo, Trumai, Yéli Dnye and Yukatek.

²⁶ Generic adpositions – such as simple adpositions in languages which have only one such adposition – are quite different, as we saw.

with the opposition between adpositions used for topological configurations (e.g. coincidence, enclosure, cf. Herskovits 1986: 55) and others: in Romance languages, complex adpositions and adpositional phrases are indeed used for specific spatial configurations, while simple spatial adpositions have a wider range of uses; some spatial configurations, such as *far from*, *all the way to*, are never expressed by simple adpositions (Fagard 2010: 271, 284).

6.2. The case of Romance

We thus decided to adopt an onomasiological approach in order to check the validity of our hypothesis. We tried to find out how Romance languages use simple and complex adpositions for the spatial configurations illustrated above (Figures a-h). If French is more grammaticalized than other Romance languages, we would expect it to use simple adpositions more systematically for these semantic relations (cf. Sections 2, 3 and 5.1).

The matter should be quickly resolved, then: all one needs to do is find the adposition used, in each language, for each of those spatial configurations. However, it is no simple matter to identify the default construction used to express a given meaning, for various reasons. There are frequently various possible means of expression; there is often an opposition between a very frequent simple adposition which rarely has a spatial meaning (see Section 5.1: a higher degree of grammaticalization is linked to a greater semantic bleaching, and thus to fewer spatial uses), and a much less frequent complex adposition which is more systematically spatial. Levinson et al. “abstracted away from these additional codings” (ibid: 488), but we think it is essential to take them into account. Of course, there remains the question of which one is most natural; we tried to find this out on the basis of written and spoken corpora, which are listed in the reference section.

We met quite a few problems on the way. Thus, though Portuguese *sob* appears with a spatial meaning in written texts (11a), this meaning is quite rare in spoken Portuguese (11b – even there, a somewhat lexicalized example), where abstract uses are much more common (11c). Actually, even spatial situations give rise to abstract interpretations with such adpositions (11d).

- (11a) Enquanto conversavam, captavam detalhes um do
 while discuss.PST.3PL perceive.PST.3PL detail.PL INDF.M.SG of.DEF.M.SG
 outro; **sob a mesa** pequena, os joelhos se
 other.M.SG under DEF.F.SG table.SG small.F.SG DEF.M.ML knee.PL REFL.3
 esbarravam ocasionalmente.
 meet.PST.3PL occasionally
 ‘While they were talking, they perceived details one of the other; under the small table, their knees sometimes met.’ (CdP, 19:Fic:Br:Garcia:Silencio)
- (11b) Se você não sai falando, não sai lutando
 if you NEG exit.PRS.3SG talk.PROG.M.SG NEG exit.PRS.3SG wrestle.PROG.M.SG
 corpo a corpo com o livro **sob** o braço, ele
 body.SG at body.sg with DEF.M.SG book.SG under DEF.M.SG arm.SG PRO.M.SG.SBJ
 termina ignorado e é evidente que o leitor
 finish.PRS.3SG ignore.PST.M.SG and be.PRS.3SG clear COMP DEF.M.SG reader.SG
 não vai desejar comprar um livro a respeito do
 NEG go.PRS.3SG want.INF buy.INF INDF.M.SG book.SG at respect.SG of.DEF.M.SG

qual nunca ouviu falar.
 which.M.SG never hear.PST.3SG talk.INF
 ‘If you don’t go and talk, wrestling bodily with your book under the arm, he ends up forgotten and clearly no reader will want to buy a book he never hear anyone talk about.’ (CdP, 19Or:Br:Intrv:ISP)

- (11c) Sempre pensei que teria de tomar esta opção, **sob**
 always think.PST.1SG COMP hold.COND.3SG of take.INF DEM.F.SG option.SG under
 o risco de cair em situações desagradáveis.
 DEF.M.SG risk.SG of fall.INF in situation.PL disagreeable.PL
 ‘I always thought I would have to take this option, or risk falling into disagreeable situations.’
 (CdP, 19Or:Pt:Intrv:Jrnl)

- (11d) Houve a candidatura de Otelo... a última bandeira **sob**
 have.PST.3SG DEF.F.SG candidacy.SG of Otelo DEF.F.SG last.F.SG banner.SG under
 a qual se acolheu a última (sic) unidade...
 DEF.F.SG which.SG REFL.3SG receive.PST.3SG DEF.F.SG last.SG unity.SG
 ‘Otello was a candidate – the last banner under which we formed a unified front...’ (CdP, 19Or:Pt:Intrv:Web)

The same can be said of Spanish *ante*, though the difference with *delante de* is less obvious: *ante las cámaras* seems to be associated with the meaning ‘on TV’ (12a), and *delante de las cámaras* with ‘in front of the cameras’ (12b). The association of simple adpositions with abstract meanings even in (potentially) spatial contexts is also perceptible for *tras* (12c).

- (12a) A partir de hoy y todos los jueves, cuando el
 at leave.INF of today and ALL.M.PL DEF.M.PL Thursday.PL when DEF.M.SG
 programa debate termine **ante las cámaras** de
 programm.SG debate.SG finish.PRS.3SG in.front.of DEF.F.PL camera.PL of
 televisión, después de la hora y media que saben
 television.SG after of DEF.F.SG hour.SG and half.F.SG REL know.PRS.3PL
 ustedes que dura, se em inaugurará se
 PRO.formal.2PL COMP last.PRS.3SG REFL.3 be(gin) inaugurate.FUT.3SG REFL.3
 empezará celebraremos un coloquio que se emitirá
 begin.FUT.3SG hold.FUT.1PL INDF.M.SG debate REL REFL.3 emit.FUT.3SG
 por Radio Cadena.
 through Radio Chain
 ‘Starting today, on Thursdays, when the debates on TV are finished, after an hour and a half, as you know this is how long it lasts, there will be a discussion emitted on Radio Cadena.’
 (Crea, 05/02/87, TVE 1, España)

- (12b) - Debes estar como un flan, digo, nerviosa.
 must.PRS.2SG be.INF like INDF.M.SG pudding.SG say.PRS.1SG nervous.F.SG

- No. Bueno, un poquito. Después de tanto tiempo, siempre,
 NEG good INDF.M.SG bit.M.SG after of so.much.M.SG time.SG always

pero nada, en cuanto hable contigo dos minutos
 but nothing in as.much.M.SG talk.PRS.SBJV.1SG with.PRO.2SG two minute.PL

ya se me pasa.
 already REFL.3 REFL.1SG pass.PRS.3SG

- Ponerse **delante de las cámaras, de los focos,** de
 put.INF.REFL.3 in.front of DEF.F.PL camera.PL of DEF.M.PL projector.M.PL of

pronto, después de tanto tiempo
 quick.M.SG after of so.much.M.SG time

- Sí, pero te olvidas.
 yes but REFL.2SG forget.PRS.2SG

- Sí, sí.
 yes yes

‘- You must feel real uncomfortable, I mean, nervous.

- No. Well, a little bit. After so much time, always, but no, just talking to you for two minutes and I feel fine.

- To find yourself in front of the cameras, of the projectors, after so much time...

- Yes, but you forget.

- Yeah.’ (Crea, Barcelona, 06/03/91, TVE 1, España)

(12c) Guerrero, que ha estado gobernado por una estirpe de
 Guerrero REL have.PRS.3SG be.PST.M.SG govern.PST.M.SG by INDF.F.SG stock.SG of

caciques asesinos y corruptos, y cuando no están ellos
 cacique.PL murderous.M.PL and corrupt.M.PL and when NEG be.3PL they.M.PL

dejan el poder tras el trono, y siguen
 leave.PRS.3PL DEF.M.SG power.SG behind DEF.M.SG throne.SG and continue.PRS.3PL

gobernando los caciques.
 govern.PROG.M.SG DEF.M.PL cacique.PL

‘[Take] Guerrero, which has been governed by a class of caciques, murderous and corrupt – and when they do not govern themselves, they feign to leave the power and go on governing from behind the throne (i.e. the curtains).’ (Crea, Cámara de Senadores de México)

Another problem arises with the distinction between simple and complex adpositions, which as we noted earlier is not always clear-cut. Thus, Portuguese *fora de* ‘outside’ is sometimes used as a simple adposition (13), though it then has an abstract meaning:

(13) As raves começaram a acontecer aqui depois de dois ou três
 DEF.F.PL rave.party.PL begin.PST.3PL at happen.INF here after of two or three

anos de acontecerem na Europa. **Fora o eletrônico** que
 year.PL of happen.INF.3PL in.DEF.F.SG Europe outside DEF.M.SG electronic.SG REL

não é de pista – o disco do Agentss é
 NEG be.PRS.3SG of dance.floor.SG DEF.M.SG disc.SG of.DEF.M.SG Agentss be.PRS.3SG

de 81 e nos anos 80 tinha até gravadora de
 of 81 and in.DEF.M.PL year.PL 80 have.PST.3SG including cutter of

música eletrônica por aqui.
 music.SG electronic.SG by here

‘Raves started here two or three years after Europe. Except electronic music which you don’t hear on dance floors – The disc of the Agentss is from 81 and in the 80’s they even had a cutter of electronic music around here.’ (CdP, Alexandre Matias)

Similarly, though it is possible to use *davanti* ‘in front (of)’ as a simple adposition (e.g. 243.000 occurrences of *davanti la casa* on the internet – Google search, 07/09/2010), the most frequent construction of *davanti* is as a transitivized adverb (859.000 occurrences of *davanti alla casa*, id.). In the Badip database, we found *no* occurrence of *davanti* used as a simple spatial adposition, but many occurrences as a transitivized adverb (as in *è qui davanti* “(s)he’s (lit. here) in front”). Another example is Italian *dietro*, which seems less clearly adverbial, with almost one third of adpositional uses (of the type *dietro la casa* “behind the house”), as in (14b).

(14a) si prende il bambino si mette **davanti** al
 REFL.3 take.PRS.3SG DEF.M.SG child.SG REFL.3 put.PRS.3SG in.front at.DEF.M.SG

lavandino
 washbowl.SG

‘You take the child and put him in front of the washbowl.’ (Badip, F A 2 164 A)

(14b) ma sa che Max riceve le persone dietro
 but know.PRS.3SG COMP Max receive.PRS.3SG DEF.F.PL person.PL behind

la porta
 DEF.F.SG door.SG

‘But do you know that Max receives people behind his door?’ (Badip, F E 15 7 B)

Some constructions are subject to little or no variation. These include complex adpositions with (very rare) uses as simple adpositions, as in Spanish (*por encima* + *N*, corresponding to fewer than 10 occurrences in the whole Crea database (as in 15a), i.e. much rarer than (*por encima de N* (more than 500 occurrences). The same can be said of *dentro de* (we found less than a dozen adpositional occurrences of *dentro N*, 4 occurrences of *por dentro de N*, but 3.000 occurrences of *dentro de N*), as in (15b).

(15a) una lamparita y te la pones **encima el**
 INDF.F.SG lamp.DIM.SG and REFL.2SG PRO.F.SG put.PRS.2SG above DEF.M.SG

ordenador
 computer.SG

‘[You buy] a small lamp and you put it on your computer.’ (Crea, conversación familiar, 10/08/91, España)

- (15b) Puede que pongas el detalle de una alfombra
 can.PRS.3SG COMP put.PRS.SBJV.2SG DEF.M.SG detail.SG of INDF.F.SG rug.SG
- roja **dentro** la **casa**, que te da muchísimo
 red.F.SG inside DEF.F.SG house.SG REL PRO.2SG.OBJ give.PRS.3SG much.SUP
- calor en un ambiente.
 heat.SG in INDF.M.SG environment
- ‘You could add the detail of a red rug in a house, and it adds a lot of warmth to your environment.’ (Crea, LP4, Bolivia)

Another difficulty yet is that complex adpositions typically display a higher degree of morphological variation, which is often associated with semantic variation, such as French *en-dessous* and *au-dessous* ‘below’, between which the semantic distinction is delicate. This is particularly obvious in Portuguese, with the alternations *em cima de / acima de / por cima de*, which are associated to semantic variation, as in (16a-b), where *em cima de* clearly means ‘on’ (with contact) and *acima do* ‘above’ (without contact).

- (16a) fui fazer uma visita à casa e encontrei
 be.PST.1SG do.INF INDF.F.SG visit.SG at.DEF.F.SG house.SG and find.PST.1SG
- um monte de papéis **em cima do sofá**
 INDF.M.SG mountain.SG of paper.PL in top.SG of.DEF.M.SG sofa.SG
- cheio de poeira.
 full.M.SG of dust.SG
- ‘I went home to look around and found a huge pile of papers on top of the sofa, full of dust.’
 (CdP, 19Or:Br:Intrv:ISP)

- (16b) INF Não conheço bicho nenhum com pêlo (...) que
 NEG know.PRS.1SG bug.SG NEG.INDF.M.SG with hairs.SG REL
- ande **acima do chão**.
 walk.PRS.SBJV.3SG on.top of.DEF.M.SG floor.SG
- INQ2 O moscardo não tem pêlo também?
 DEF.M.SG horsefly.SG NEG have.PRS.3SG hairs.SG too
- INF Não. O moscardo também não tem pêlo.
 NEG DEF.M.SG horsefly.SG too NEG have.PRS.3SG hairs.SG
- ‘- I don’t know of any hairy bug that can walk on the ground.
 - Doesn’t the horsefly have hairs, too?
 - No, the horsefly doesn’t have hairs, either.’
 (CdP, Cordial: ALC40)

The examples above illustrate some of the difficulties we encountered. We tried to take into account all elements of semantic and morpho-syntactic variation in order to provide an adequate representation of the way Romance languages use adpositions to express these spatial configurations. We present in Tables 6-13 the results of our investigation, i.e. the default adposition or construction used to denote the corresponding configuration, along with less frequent or semantically marked adpositions, marked with a sharp sign (“#”, here and in the following tables as well).

First, all languages examined here have simple adpositions for IN, INSIDE, as shown in Table 6, in which we indicate the relative frequency²⁷ of all adpositions in spoken language corpora, based on the criteria illustrated by examples (11) to (16). This is expected, given the scale we presented above. The existence, next to these simple forms, of more precise and much less frequent complex adpositions is also expected.

Relative frequency (per million words)	Spanish	French	Italian	Portuguese	Romanian
> 2500	en	#en, dans ²⁸	in	em	în
> 250	dentro de		dentro	dentro de	
< 50	#al (/ hacia el / en el / por el...) interior de	#à l'intérieur de, #dedans	#dentro a/di	#por dentro de, #para dentro de	#în interiorul +G, #în sânul +G, din sânul +G

Table 6: Romance adpositions denoting the spatial configuration IN / INSIDE.

The same can be said of the spatial configuration of support: all languages have simple adpositions for ON, as shown in Table 7, but Ibero-Romance stands apart, with an adposition that is underspecified for inclusion and support (with contact). These adpositions are specifically used to express spatial configurations involving contact; as we will see below, situations where there is no contact are generally expressed by complex adpositions.

Relative frequency (per million words)	Spanish	French	Italian	Portuguese	Romanian
> 5000	en			em	pe
> 1000	sobre	sur	su	sobre	

Table 7: Romance adpositions denoting the spatial configuration ON.

The case of UNDER is slightly different: although all languages have simple adpositions, it seems that at least Portuguese *baixo* and Spanish *bajo* rather take on abstract uses, which is not necessarily the case of the other simple adpositions listed in Table 8. Relative frequencies indicate a specialization of French *sous* and Romanian *sub*, with complex adpositions clearly much less frequent.

Relative frequency (per million words)	Spanish	French	Italian	Portuguese	Romanian
> 100	bajo	sous	sotto	sob	sub
> 20	debajo de		sotto a	debaixo de, abaixo de	
> 15			al di sotto di		

²⁷ Relative frequencies are computed on the basis of, on the one hand, automatic counts of adpositions in spoken language corpora and, on the other hand, a hand-count of the proportion of non-ambiguous spatial uses, on an arbitrary selection, for each adposition. The thresholds indicated vary with the frequencies of each group of adpositions. The corpora are listed in the reference section.

²⁸ Though the adposition *en* can be used with a spatial meaning, its most frequent uses are rather abstract (see De Mulder 2008: 284 and De Mulder & Amiot, unpublished manuscript).

< 10	por debajo de	en-dessous de, au-dessous de, dessous	al di sotto	em baixo de, por debaixo de	dedesubt, dedesubtul
------	---------------	---------------------------------------	-------------	-----------------------------	----------------------

Table 8: Romance adpositions denoting the spatial configuration BELOW.

For AMONG / BETWEEN, Romance languages all have simple adpositions; as illustrated in Table 9, they also use complex adpositions, though their frequency is much lower. Relative frequencies show a greater specialization in French, with a greater difference between *entre* and the other options for expressing loose insertion.

Relative frequency (per million words)	Spanish	French	Italian	Portuguese	Romanian
> 500	entre	entre	tra, fra	entre	
> 100	dentro de			no meio de	între
< 50		parmi, au sein de	in mezzo a	dentre	printre
< 10	por entre, por dentro (de)		nel mezzo di		în sânul, din sânul

Table 9: Romance adpositions denoting the spatial configuration AMONG / BETWEEN.

All other spatial configurations examined here automatically trigger the use of complex adpositions in one or more languages. Thus, IN FRONT OF is expressed by complex adpositions in Romanian and Italian, as illustrated in Table 10. Besides, Spanish *ante* and Portuguese *perante* tend to be associated with abstract meanings, so that the default spatial forms are rather *delante de* and *dianete de*, respectively.

Relative frequency (per million words)	Spanish	French	Italian	Portuguese	Romanian
> 100	#ante	devant	davanti a, #davanti	dianete de, #perante	în fața +G
> 50	delante de		#di fronte a	#em frente (de/a)	
> 10		#en face (de)		#defronte de	
< 10	#por delante de	en face			în fruntea +G, înaintea +G, în capul +G

Table 10: Romance adpositions denoting the spatial configuration IN FRONT OF.

Note that these adpositions are not exactly equivalent, among other things because the importance of the intrinsic orientation of the complement is variable, as illustrated by the spatial configurations below (Vandeloise 1991: 97-98). According to Vandeloise, it is possible to say *le râteau est devant le crabe* ‘the rake is in front of the crab’ only in Figure i, while in Figure j *devant le crabe* would have to apply to the net (*le filet est devant le crabe*); however, *en face de* would have to apply to the net in both figures.

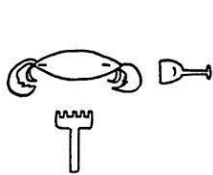


Figure i

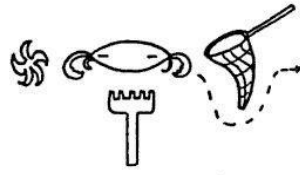


Figure j

The spatial configuration BEHIND is also expressed by complex adpositions in some of these languages: Spanish, Portuguese and Romanian, as illustrated in Table 11, without the variation noted for IN FRONT OF. The case of Spanish is particularly interesting, given that Medieval Spanish had a simple adposition (*tras*) with this meaning (Fagard 2010: 249), and now uses it as a complex adposition in the construction (*por*) *detrás de*. French is the only language with a frequent simple adposition and no variation²⁹.

Relative frequency (per million words)	Spanish	French	Italian	Portuguese	Romanian
> 50	<i>detrás de</i>	<i>derrière</i>	<i>dietro</i>	<i>atrás de</i>	
> 10			# <i>dietro a</i>	# <i>detrás de</i> , # <i>por detrás de</i>	
< 5	# <i>por detrás de</i>		# <i>al di dietro di</i>	# <i>para detrás de</i>	<i>în spatele</i> +G, <i>înapoia</i> +G, <i>îndărătul</i> +G

Table 11: Romance adpositions denoting the spatial configuration BEHIND.

The spatial configuration ABOVE is expressed by complex adpositions everywhere; only French, Italian and Spanish use both simple and complex forms, as illustrated in Table 12. Note that in Italian, the default form is simple (*sopra*), while in Spanish and French the default form is complex (*encima de*, *au-dessus de*). The high relative frequency of adpositions based on *cima* in Spanish and Portuguese is linked to their use as indicative of support with contact.³⁰

Relative frequency (per million words)	Spanish	French	Italian	Portuguese	Romanian
> 100	# <i>encima</i>				
> 50	<i>encima de</i>			<i>acima de</i> , <i>em cima de</i>	
> 10	# <i>por encima de</i>	<i>au-dessus de</i>		<i>por cima de</i>	<i>asupra</i> +G
< 5		# <i>au-dessus</i>	<i>sopra</i> , # <i>sopra a</i> , <i>al di sopra di</i>	<i>em sobre</i>	<i>deasupra</i> +G

Table 12: Romance adpositions denoting the spatial configuration ABOVE.

Finally, the spatial configuration OUTSIDE is typically expressed by complex adpositions in all languages, as illustrated in Table 13: although some simple forms exist (French *hors*, Italian *fuori*, Portuguese *fora*), they all have very marked uses, occurring only in fixed constructions

²⁹ As far as grammaticalized expressions are concerned. Of course, French also has lexical means of expressing this semantic relation, such as *dans le dos de*, lit. “in the back of”.

³⁰ We did not exclude these uses, because it is not easy to disambiguate them without extralinguistic context.

such as Italian *fuori città* ‘out of town’, *fuori le mura* ‘beyond the city walls’, *fuori Roma* ‘outside of Rome’, and have low relative frequencies.

Relative frequency (per million words)	Spanish	French	Italian	Portuguese	Romanian
> 50	fuera de		fuori di, #fuori da	fora de	în afara +G
> 10		hors de, en-dehors de	#fuori		
< 5	#por fuera de, #al exterior de	#hors	#all’infuori di, al di fuori di	#fora	în exteriorul +G

Table 13: Romance adpositions denoting the spatial configuration OUTSIDE.

6.3 Discussion

This brief description seems to confirm – at least in the five languages we examined – the existence of spatial configurations of different types, some basic, others less so. There are first ‘basic’ spatial configurations such as inclusion and support.³¹ There are also various types of more complex spatial configurations, including projective configurations (*in front of*, *behind*) or configurations associated with more than one semantic feature (ABOVE being associated for instance with both superiority on the vertical axis and absence of contact). These configurations may be expressed by simple adpositions, but are consistently expressed by complex adpositions. A gradient of grammaticalization can be elaborated on the basis of how many of these configurations a language expresses with simple adpositions. From this point of view, French seems indeed to be the most grammaticalized, followed by Italian, Spanish, Portuguese and Romanian, in that order.

7. Conclusion

After a brief look at the framework of grammaticalization theory and at Latin adpositional paradigms, we tried to describe the uses of adpositions in Romance, from two points of view which we consider complementary. In all Romance languages examined here, the systems of adpositions result from the grammaticalization of various (mostly noun-based) constructions, lexical items and already grammaticalized adpositions – all three processes being instances of grammaticalization, and involving the same basic phenomena, but with different results: complex adpositions, simple adpositions and ‘functional’ adpositions.

The grammaticalization of new simple adpositions (on the basis of participles, adverbs or nouns) is found in all Romance languages, as well as that of adpositions into functional adpositions; these adpositions are mostly grammaticalized on the basis of one Latin adposition, with a few exceptions: Romanian *la < illac ad*, Italian *da < de ad*. Finally, patterns of complex adpositions are also found in all Romance languages. We showed that 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.

³¹ As well as the relation of *contact*, which we did not investigate but is sometimes expressed by functional adpositions (*à*, *a*, *la* ‘at’), in all Romance languages. For this adposition, too, it seems that French displays more grammaticalized uses than other Romance languages; see Goyens, Lamiroy & Melis (2002: 302).

We then confirmed this trend on the basis of an onomasiological study, comparing our data with spoken databases in order to make sure that they reflected actual language use. This approach yielded similar results, and thus seems to confirm the observations already made on the basis of other parts of the language: as far as adpositions are concerned, if our approach is valid, French is indeed the most grammaticalized Romance language, followed by Italian, Spanish, Portuguese and Romanian. Many questions remain; for instance, does our hypothesis remain valid for other domains than space? Are spatial configurations really representative?

References

- Abeillé, Anne, Danièle Godard, Olivier Bonami & Jesse Tseng. 2003. The syntax of French *à* and *de*: an HPSG analysis. In Patrick Saint-Dizier (ed.), *Syntax and the semantics of prepositions*. Dordrecht: Kluwer, 133-144.
- Aikhenvald, Alexandra Y. 2004. *Evidentiality*. Oxford: Oxford University Press.
- Berlin, Brent & Paul Kay. 1969. *Basic color terms: their universality and evolution*. Berkeley / Los Angeles: University of California Press.
- Brøndal, Viggo. 1939. L'originalité des prépositions du français moderne. In Albert Sechehaye et al. (eds.), *Mélanges de linguistique offerts à Charles Bally*. Genève: Georg et cie, 337-346.
- Brown, Penelope. 2006. A sketch of the grammar of space in Tzeltal. In Stephen Levinson & David Wilkins (eds.), *Grammars of space*. Cambridge: Cambridge University Press, 230-72.
- Bybee, Joan & Paul Hopper (eds.). 2001. *Frequency and the emergence of linguistic structure*, Amsterdam / Philadelphia: John Benjamins.
- Bybee, Joan. 2006. From usage to grammar: The mind's response to repetition. *Language* 82/4, 711-733.
- Carlier, Anne. 2007. From Preposition to Article. The grammaticalization of the French partitive. *Studies in Language*, 31, 1, 1-49.
- Ciobanu, Fulvia & Isabela Nedelcu. 2008. Prepoziția. In Valeria Guțu Romalo (coord.), *Gramatica limbii române*. Vol. 1, *Cuvântul*. București: Editura Academiei Române, 607-630.
- Company Company, Concepción. 2002. Grammaticalization and category weakness. In Ilse Wischer & Gabriele Diewald (eds.), *New Reflections on Grammaticalization*. Amsterdam / Philadelphia: John Benjamins, 201-215.
- Cuniță, Alexandra. 2004. Prepoziția, abordări vechi și noi în lingvistica românească. In Gabriela Pană Dindelegan (ed.), *Tradiție și inovație în studiul limbii române*. București: Editura Universității din București, 79-85.
- De Mulder, Walter & Lamiroy Béatrice. To appear. Stages of Grammaticalization: the position of French among the other Romance languages. In Bert Cornillie & Hubert Cuyckens (eds.), *New Reflections on Grammaticalization*. Amsterdam / Philadelphia: John Benjamins.
- De Mulder, Walter. 2001. La Linguistique Diachronique : grammaticalisation et sémantique du prototype. *Langages* 130, 8-32.
- De Mulder, Walter. 2008. *En* et *dans* : une question de « déplacement » ? In Olivier Bertrand, Sophie Prévost, Michel Charolles, Jacques François & Catherine Schnedecker (eds.), *Discours, diachronie, stylistique du français – Etudes en hommage à Bernard Combettes*. Bern: Peter Lang, 277-291.
- De Mulder, Walter & Dany Amiot, Unpublished manuscript. L'insoutenable légèreté de la préposition *en*.
- De Mulder, Walter & Anne Vanderheyden. 2000. Vers une sémantique diachronique cognitive ? Réflexions sur l'évolution de la préposition *sur*. In Annick Englebert, Michel Pierrard, Laurence Rosier & Dan Van Raemdonck (eds.), *Actes du XXIIe Congrès International de Linguistique et de Philologie Romanes*, vol. II. Tübingen: Max Niemeyer, 129-238.
- DeLancey, Scott. 2005. Adpositions as a non-universal category. In Zygmunt Frajzyngier, Adam Hodges & David Rood (eds.), *Linguistic diversity and language theories*. Amsterdam: John Benjamins, 185-202.
- Di Meola, Claudio. 2000. *Die Grammatikalisierung deutscher Präpositionen*. Tübingen: Stauffenburg.
- Evans, Nicholas & Stephen C. Levinson. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32, 429-492.

- Fagard, Benjamin. 2006. La grammaticalisation en question – du latin aux langues romanes modernes, *Modèles linguistiques* XXVII-1/53, 91-110.
- Fagard, Benjamin. 2010. *Espace et grammaticalisation: l'évolution sémantique des prépositions dans les langues romanes*. Sarrebrück: Editions Universitaires Européennes.
- Fagard, Benjamin & Alexandru Mardale. 2007. Systèmes prépositionnels des langues romanes: la notion de partie du discours en diachronie. In David Trotter (ed.), *Actes du XXIVème Congrès International de Linguistique et Philologie Romanes*, Vol. 1. Tübingen: Niemeyer, 91-105.
- Fortis, Jean-Michel & Benjamin Fagard. 2010. Space in Language. *Leipzig Summer School of Typology*, August 14-28, 2010. Available on-line: <http://htl.linguist.univ-paris-diderot.fr/jmfortis.htm>.
- Gaetone, David. 2001. Les prépositions : une classe aux contours flous, *Travaux de linguistique* 42/43, 23-31.
- Gawelko, Marek. 1999. Classification des langues romanes et pouvoir explicatif des critères. *Zeitschrift für romanische Philologie* 115, 20-44.
- Gawelko, Marek. 2001. Sobre la tendencia analítica de algunas lenguas romances: español, italiano y portugués. *Revista española de lingüística*, 31 (2), 393-412.
- Gentner, Dedre & Lera Boroditsky. 2001. Individuation, relational relativity and early word learning. In Melissa Bowerman & Stephen Levinson (eds.), *Language acquisition and conceptual development*. Cambridge: Cambridge University Press, 215-256.
- Goyens, Michèle, Béatrice Lamiroy & Ludo Melis. 2002. Déplacement et repositionnement de la préposition à en français. *Linguisticae Investigationes* 25(2), 275-310.
- Grinevald, Colette. 2006. The expression of static location in a typological perspective. In Maya Hickmann & Stéphane Robert (ed.), *Space in languages: linguistic systems and cognitive categories*. Amsterdam / Philadelphia: John Benjamins, 29-58.
- Hagège, Claude. 2010. *Adpositions*. Oxford: Oxford University Press.
- Hallan, Naomi. 2001. Paths to prepositions? A Corpus-Based Study of the Acquisition of a Lexico-Grammatical Category. In Joan L. Bybee & Paul Hopper (eds.), *Frequency and the emergence of linguistic structure*, Amsterdam / Philadelphia: John Benjamins, 91-120.
- Haverling, Gerd. 2003. On prefixes and actionality in classical and late Latin, *Acta linguistica Hungarica* 50(1-2). Kiefer, Ferenc & Wolfgang Dressler (eds.), Budapest: Akadémiai kiadó, 113-135.
- Hickmann, Maya. 2007. Static and dynamic location in French: Developmental and cross-linguistic perspectives. In Michel Aurnague, Maya Hickmann & Laure Vieu (eds.), *The categorization of spatial entities in language and cognition*. Amsterdam: John Benjamins, 205-231.
- Hopper, Paul & Elizabeth Traugott. 2003 [1993]. *Grammaticalization*. Cambridge: Cambridge University Press.
- Jakubowicz Batoréo, Hanna. 2000. *Expressão do espaço no português europeu. Contributo psicolinguístico para o estudo da linguagem e cognição*. Coimbra: Fundação Calouste Gulbenkian.
- Johnston, Judith & Dan Slobin. 1979. The development of locative expressions in English, Italian, Serbo-Croatian and Turkish. *Journal of Child Language* 6: 529-545.
- König, Ekkehard & Volker Gast. 2007. *Understanding English-German Contrasts*. Berlin: Schmidt.
- Lamiroy, Béatrice. 1999. Auxiliaires, langues romanes et grammaticalisation. *Langages*, 135, 33-45.
- Lamiroy, Béatrice. 2001. Le syntagme prépositionnel en français et en espagnol: une question de grammaticalisation ? *Langages*, 143, 91-106.
- Lamiroy, Béatrice. 2003. The External Possession Structure and Grammaticalisation in Romance and Germanic languages. In: D'Hulst, Yves & Martine Coene (eds.). *From NP to DP. The expression of Possession in Noun Phrases* (Vol. II). Amsterdam: John Benjamins, 257-281.
- Lamiroy, Béatrice. 2011. Degrés de grammaticalisation à travers les langues de même famille. *Bulletin de la Société de Linguistique de Paris* 19, 167-192.
- Lehmann, Christian. 1985. Grammaticalization: synchronic variation and diachronic change. *Lingua e Stile* 20(3): 303-318.
- Lehmann, Christian. 1995. *Thoughts on Grammaticalization*. München, Newcastle: Lincom Europa.
- Leumann, Manu, Johann Hofmann & Anton Szantyr. 1965. *Lateinische Grammatik*. München: C. H. Beck.

- Levinson, Stephen, Sérgio Meira & The Language and Cognition Group. 2003. 'Natural concepts' in the spatial topological domain – adpositional meanings in crosslinguistic perspective: an exercise in cognitive typology. *Language*, 79(3): 485-516.
- Marchello-Nizia, Christiane. 1995. *L'évolution du français : ordre des mots, démonstratifs, accent tonique*. Paris: A. Colin.
- Marchello-Nizia, Christiane. 2006. *Grammaticalisation et changement linguistique*. Bruxelles: De Boeck-Duculot.
- Mardale, Alexandru. 2009. *Les prépositions fonctionnelles du roumain. Études comparatives sur le marquage casuel*. Paris: L'Harmattan.
- Meillet, Antoine. 1948. *Linguistique historique et linguistique générale*. Paris: Champion.
- Melis, Ludo. 2003. *La préposition en français*. Gap / Paris: Ophrys.
- Morgenstern, Aliyah & Martine Sekali. 1997. L'acquisition des premières prépositions chez un enfant francophone. *La préposition : une catégorie accessoire ? Faits de langues* 9: 201-210.
- Noonan, Michael. 2007. Complementation. In Timothy Shopen (ed.), *Language Typology and Syntactic Description: Complex constructions*. Cambridge: Cambridge University Press, 52-150.
- Piaget, Jean & Bärbel Inhelder. 1956 [1948]. *The child's conception of space*. London: Routledge / Kegan Paul.
- Prévost, Sophie & Benjamin Fagard. 2007. Présentation: Grammaticalisation et lexicalisation: la formation d'expressions complexes. *Langue française* 156: 3-8.
- Rizzi, Luigi. 1988. Il sintagma preposizionale. In Lorenzo Renzi & Giampaolo Salvi (eds.), *Grande Grammatica di consultazione*, Bologna: Il Mulino, 507-531.
- Sausy, Lucien. 1992. *Grammaire latine complète*. 8th ed. Paris: Fernand Lanore.
- Schøsler, Lene. 2008. L'expression des traits manière et direction des verbes de mouvement. Perspectives diachroniques et typologiques. In Elisabeth Stark, Roland Schmidt-Riese & Eva Stoll, (eds.). *Romanische Syntax im Wandel*. Tübingen: Gunter Narr, 113-132.
- Svorou, Soteria. 1994. *The grammar of space*. Amsterdam / Philadelphia: John Benjamins.
- Talmy, Leonard. 1972. *Semantic Structures in English and Atsugewi*. Dissertation, University of California, Berkeley.
- Talmy, Leonard. 2000. *Toward a Cognitive Semantics*, vol. 1 & 2. Cambridge: MIT Press.
- Vandeloise, Claude. 1991. *Spatial prepositions: a case study from French*. Chicago: University of Chicago Press.
- Van Riemsdijk, Henk. 1990. Functional Prepositions. In Harm Pinkster & Inge Genee (eds.), *Unity in diversity*, Dordrecht: Foris, 229-241.
- Wierzbicka, Anna. 1988. *The Semantics of Grammar*. Amsterdam / Philadelphia: John Benjamins.
- Zwarts, Joost. 1997. Lexical and Functional Properties of Prepositions. In Dagmar Haumann & Stefan Schierholz (eds.), *Lexikalische und grammatische Eigenschaften präpositionaler Elemente*. Tübingen: Max Niemeyer Verlag, 1-18.

Corpora

Romanian: Ruxăndoiu and CORV.

Portuguese: Corpus do Português, Davies, Mark and Michael Ferreira. (2006-) (45 million words, 1300s-1900s). Available online at <http://www.corpusdoportugues.org>.

Spanish: Crea (<http://corpus.rae.es/creanet.html>) and Corpus del español, Davies, Mark. (2002-) (100 million words, 1200s-1900s). Available online at <http://www.corpusdelespanol.org>.

Italian: Corpus Badip (<http://badip.uni-graz.at/>).

French: Clapi (<http://clapi.univ-lyon2.fr/>) and Valibel (<http://www.uclouvain.be/valibel-corpus.html>).