



HAL
open science

Les prépositions complexes du roumain - approche sur corpus

Benjamin Fagard, Alexandru Mardale

► **To cite this version:**

Benjamin Fagard, Alexandru Mardale. Les prépositions complexes du roumain - approche sur corpus. Revue Romane, 2019. halshs-02194109

HAL Id: halshs-02194109

<https://shs.hal.science/halshs-02194109>

Submitted on 25 Jul 2019

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.

version auteur avant épreuves

Les prépositions complexes du roumain – approche sur corpus

Benjamin Fagard et Alexandru Mardale

Lattice (CNRS, ENS, Paris Sorbonne Nouvelle ; PSL & USPC) / Inalco (SeDyL, UMR 8202)

Abstract

In this paper, we investigate complex adpositions in Romanian, on the basis of a corpus study. Our goal is to provide an extensive list of complex adpositions and analyze the productivity of various patterns. In order to do so, we selected a number of candidate sequences, applied semantic and morpho-syntactic tests, and evaluated their frequency. Our results confirm the importance of two patterns: [Preposition + Noun + Determiner] followed by a genitive, and [Preposition + Noun + Preposition]. Though there are a few other patterns (such as [(Preposition +) Adverb (+ Preposition)]), these two are the most productive, making up a core of relatively frequent sequences with low variability, which pattern with simple prepositions.

Keywords: complex adpositions, corpus, grammaticalization, lexicalization, Romanian

Introduction

Malgré certaines études qui ont fait date (voir notamment Lehmann 1985), il existe relativement peu de travaux sur les locutions prépositionnelles – du moins en comparaison du nombre maintenant assez considérable d'études sur les prépositions simples, et ce dans divers cadres théoriques et avec diverses approches méthodologiques (voir, entre autres, Tseng 2013). Ceci est particulièrement vrai pour le roumain ; les grammaires se contentent au mieux d'en donner un rapide aperçu (GA 1954/1963, DSL 1997, GALR 2005/2008, GR 2013), et on ne trouve que quelques articles dédiés à cette question (voir notamment Ciobanu 1957, 1961, Găitănar 1965, 1999, Căpățână 2000, Duda 2002, Petrescu 2005, Cuniță 2007, Vlasin 2015). Nous proposons ici de faire un premier point sur la catégorie, sur la base d'une étude sur corpus, avec pour but d'évaluer l'envergure et la cohésion interne de la catégorie : quels sont les principales matrices de construction des locutions prépositionnelles en roumain ? Peut-on proposer une liste des principales constructions ?

Au sein des langues romanes, le roumain constitue un cas particulièrement intéressant en ce qui concerne les prépositions complexes, dans la mesure où les vestiges de système casuel entraînent la coexistence de deux sous-systèmes. Dans le premier, le régime de la locution est introduit par la préposition *de* (1). Dans le second, il est marqué à l'aide du génitif, soit avec une morphologie synthétique (par exemple la désinence *-(l)or* en (2), *-(e)i* en (3)), soit avec une morphologie analytique (la préposition fonctionnelle *a*, employée notamment quand le complément est précédé d'un élément invariable, comme en (4)). Il faudra également prendre

en compte le fait que ces alternances s’accompagnent, dans certains cas, d’un changement de sens (cf. Găitănaru 1999 : 12, Petrescu 2005 : 201, Vlasin 2015), particulièrement clair entre (3) où le sens de la locution est spatial et (4-5) où elle introduit une approximation (voir plus bas en section 3.).

- | | |
|---|---|
| (1) <i>în lipsă de</i> + accusatif « faute de » : | <i>în lipsă de mijloace</i>
en faute de moyen.PL.ACC
(faute de moyens) |
| (2) <i>în lipsa</i> + génitif : | <i>în lipsa mijloacelor</i>
en faute.DET moyen.PL.G
faute des moyens
(faute de moyens) |
| (3) <i>în jurul casei</i> | en tour.DET maison.PL.G
(autour de la maison) |
| (4) <i>în jurul a</i> 100.000 de lei | en tour.DET à 100 000 de leu.PL.ACC
(environ 100 000 lei) |
| (5) <i>în jur de</i> 100.000 de lei | en tour de 100 000 de leu.PL.ACC
(environ 100 000 lei) |

Une dernière question à laquelle nous tenterons de répondre est donc de savoir dans quelle mesure cette coexistence de deux sous-systèmes peut avoir constitué un frein au figement des locutions prépositionnelles.

Dans un premier temps, nous passons en revue les principaux résultats des travaux existants sur le sujet (section 1). Nous exposons ensuite la méthode que nous avons employée pour contrôler et étendre les listes existantes (section 2). Enfin, nous présentons (section 3) et discutons (section 4) les résultats de notre étude sur corpus.

1. Prépositions et locutions prépositionnelles en roumain : définition et approches théoriques

1.1. Les prépositions simples du roumain

On considère généralement les prépositions comme une classe de mots fermée, composée d’éléments ayant des degrés de grammaticalisation différents. L’inventaire de la classe est plutôt stable en diachronie, avec un nombre limité d’éléments nouveaux l’intégrant sporadiquement (Ciobanu 1957, GA 2005/2008).

Comme dans beaucoup de langues, les prépositions du roumain sont des mots invariables (entre autres, Mardale 2009, 2011, 2013). On peut les définir d’une manière similaire à celle adoptée pour les autres catégories grammaticales, à savoir comme des têtes (lexicales) ayant la possibilité de se combiner avec différents types de compléments (généralement de nature nominale), pour former des groupes prépositionnels. La plupart des prépositions fonctionnent soit comme des prédicats indépendants exprimant des sens lexicaux, c’est-à-dire circonstanciels (notamment le lieu, le temps, la manière, la cause, etc.) – mais pas seulement (cf. la préposition *cu* « avec ») –, soit elles sont sous-catégorisées ou, autrement dit, régies par

un prédicat verbal (*a se baza pe* « se baser sur »), nominal (*încrederea în* « la confiance en »), adjectival (*mândru de* « fier de »), adverbial (*dincolo de* « de l'autre côté de ») ou interjectionnel (*vai de* « pauvre de », litt. « hélas de »).

Certaines prépositions sont des éléments vides de sens lexical, qui apparaissent pour des raisons purement grammaticales, plus précisément pour marquer des relations casuelles dans des contextes morpho-syntaxiques et parfois pragmatico-sémantiques spécifiques (voir plus haut (4), et ci-dessous (6-7)).

(6) *mama a doi copii*

mère.DET à deux enfant.PL.ACC

(la mère de deux enfants)

(7) *(i-)a chemat pe copii*

CL.ACC.3PL-avoir.PRS.3SG appeler.PST.PTCP MDO enfant.PL.ACC

(il a appelé les enfants)

Sur la base de ces différentes propriétés, on distingue généralement trois classes de prépositions (simples) : les prépositions dites *lexicales* (libres, non régies par un prédicat), les prépositions dites *semilexicales* (régies par un prédicat) et les prépositions dites *fonctionnelles* (imposées par différentes contraintes du contexte) (Geană 2008, 2009, 2011, Mardale 2009, GR 2013).

Les prépositions roumaines ont également la particularité d'imposer des régimes casuels différents à leur complément. De ce point de vue, on peut distinguer trois classes de prépositions dont le régime casuel est clairement visible lorsque leur complément est un pronom (GA 2005/2008) : celles qui imposent l'*accusatif* (*pentru / cu / lângă / despre / la mine* « pour / avec / à côté de / de / chez moi »), celles qui imposent le *génitif* (*înaintea / înapoia / împotriva mea* « devant / derrière / contre moi »), celles qui imposent le *datif* (*grație / datorită / mulțumită mie* « grâce à moi »). Quelle que soit leur classe, les prépositions ont toutes en commun le rôle d'élément de relation, c'est-à-dire de relier deux constituants ou bien d'une même catégorielle (p. ex., deux noms *Petre cu Monica* « Pierre avec Monique ») ou bien de catégories différentes (p. ex., un verbe et un nom *vorbește cu Ana* « (il) parle avec Anne ») (Cuniță 2007).

1.2. Les locutions prépositionnelles : première définition

Tout comme les prépositions (simples), les locutions prépositionnelles ont, elles aussi, le rôle de relier deux constituants dans une structure donnée (*apă în loc de vin* « de l'eau à la place du vin », *vorbește în numele directorului* « (il) parle au nom du directeur »), et elles peuvent imposer un régime casuel différent à leur complément, plus précisément l'*accusatif* (*față de / cu privire la mine* « par rapport à moi / à mon égard ») ou le *génitif* (*în jurul / în locul meu* « autour de moi / à ma place »). À la différence des prépositions simples, il n'y a pas de locution prépositionnelle imposant le datif.

Au-delà de ce fonctionnement commun sur la base duquel on range d'ailleurs dans la même catégorie grammaticale locutions et prépositions (simples), les premières se distinguent des secondes d'abord par leur structure : les locutions sont des regroupements a priori figés et cohésifs de plusieurs mots tandis que les prépositions sont des lexèmes uniques, avec souvent

un gabarit phonétique réduit (1 à maximum 3 syllabes) (GALR 2005/2008, GR 2013). Par ailleurs, on distingue locutions prépositionnelles et prépositions complexes (Cuniță 2007, Mardale 2013), ces dernières étant des formes composées (et parfois morphologiquement contractées) de deux ou de trois prépositions simples (*de la, de pe, de lângă, de pe la, de pe lângă, din (de + în), dintre (de + între), printre (pre + între)*). Elles expriment ensemble un sens lexical compositionnel (et non pas figé, comme c'est le cas des locutions), par exemple la séparation d'avec un point de contact comme dans *a luat cartea de pe birou* « (il) a pris le livre sur le bureau » (litt. *de sur* le bureau).

Comme nous l'avons dit plus haut, la structure des locutions prépositionnelles est aussi l'un des critères de définition et de délimitation de la classe. Qu'est-ce qu'on considère comme une locution prépositionnelle et sur quels critères ?

1.3. Tests

Les études ayant essayé de répondre à ces questions pour le roumain (Găitănaru 1965, 1999, Ciobanu 1961, DSL 1997, Căpățână 2000, Duda 2002, Petrescu 2005, Cuniță 2007, Vlasin 2015) s'accordent sur le fait que les locutions prépositionnelles sont des combinaisons comportant au moins une préposition simple (lexicale ou fonctionnelle) et d'autres catégories grammaticales (nom, adverbe, verbe).

Elles sont donc considérées comme un ensemble cohérent et figé qui est à priori l'équivalent sémantique et distributionnel d'une préposition simple (Ciobanu 1957, 1961, Avram 2001).

On peut ainsi employer (8) à la place de (9) ou l'inverse :

- (8) *am vorbit în numele meu*
 avoir.PRS.1 parler.PST.PTCP en nom.DET mien.G
 (j'ai parlé en mon (propre) nom)
- (9) *am vorbit pentru mine*
 avoir.PRS.1 parler.PST.PTCP pour moi.ACC
 (j'ai parlé pour moi)

Cette équivalence sémantique et fonctionnelle n'est en revanche pas toujours possible ou, quand elle existe, elle est souvent associée à d'autres contraintes, p. ex., des contraintes casuelles (régime génitif vs. accusatif, comme en (8) qui impose le génitif et en (9) qui impose l'accusatif), morphologiques (absence ou présence de déterminant avec le complément, comme en (10a-b) où le déterminant défini est requis et en (11a-b) où il est au contraire exclu), sémantiques (lecture référentielle ou non référentielle du complément, comme en (10) et (11) également).

- (10) a. *în locul turnului au plantat copaci*
 en lieu.DET tour.SG.DET.G avoir.PRS.3PL planter.PST.PTCP arbre.PL.ACC
 (à la place de la tour, ils ont planté des arbres)
- b. *la începutul anului*
 à début.SG.DET an.SG.DET.G
 (au début de l'année)
- (11) a. *în loc de turn au construit case*
 en lieu de tour.SG.ACC avoir.PRS.3PL construit.PST.PTCP maison.PL.ACC
 (à la place d'une tour, ils ont construit des maisons)

b. la început de **an**
à début de an.SG.ACC
(en début d'année)

On note aussi qu'il n'existe pas de préposition simple sémantiquement équivalente (en roumain) pour certaines locutions, dont les suivantes : *la dreapta* « à droite de », *la stânga* « à gauche de », *în josul* « en bas de », *în susul* « en haut de », *în locul* « au lieu de », *în ciuda*, *în pofida* « malgré, en dépit de », *cu tot / toată* « malgré tout(e) ».

De même, le degré de figement et de cohésion interne devrait être examiné de plus près et nuancé, puisque les exemples concrets sont très variés, avec des comportements morpho-syntaxiques différents, comme nous le verrons plus loin. Ceci implique donc différentes façons d'envisager la définition des locutions prépositionnelles et par conséquent l'existence d'inventaires plus ou moins larges selon les critères d'analyse adoptés.

Les critères retenus par ces études, et dont nous présentons ci-après une sélection, sont principalement d'ordre sémantique (A.a-b), lexical (B.a-b) et morpho-syntaxique (C.a-c) :

A.a. l'existence d'une différence sémantique de la base nominale entre l'emploi en tant que locution et l'emploi libre : *în fața omului* « face à / devant l'homme » vs. *fața omului* « la face de l'homme », *în vederea reușitei* « en vue de la réussite » vs. *vederea bătrânului* « la vision du vieil homme », *în mijlocul casei* « au milieu de la maison » vs. *mijlocul bătrânei* « les hanches de la vieille dame ».

A.b. la perte partielle de sens référentiel, perceptible notamment dans le type de relation génitivale : ainsi, dans *în mijlocul oamenilor* « parmi les hommes (litt. au milieu des hommes) », le nom « mijloc » n'a pas un sens (pleinement) référentiel, à la différence de *în mijlocul orașului* « au centre de la ville », *în mijlocul drumului* « au centre de la route », *în mijlocul casei* « au centre de la maison ».

B.a. la présence de mots qui n'existent plus en dehors de la structure locutionnelle (*potriva* « contre », *preajma* « proximité », *pofida* « ennui, mépris ») ou dont l'emploi est archaïque (*privața* « égard », *pilda* « exemple », *răspărul* « contrepoil », *toiul* « point culminant », *ciuda* « merveille, étonnement », *seama* « somme », *curmeziș* « diagonale, transversale », *prelargul* « étendue », *osebit* « différent »).

B.b. l'existence d'un néologisme ou d'un emprunt souvent calqué sur une locution prépositionnelle d'une autre langue : *detrimentul* « détrimement », *intermediul* « intermédiaire », *parcursul* « parcours », *prisma* « prisme », *virtutea* « vertu », *favoarea / defavoarea* « faveur / défaveur », *spezele* « dépenses », *decursul* « déroulement », *cherem* « faveur ».

C.a. pour certaines locutions prépositionnelles, la conversion en locution adverbiale (c'est-à-dire en emploi absolu, sans complément) est possible : *în afară de* « en dehors de » > *în afară* « en dehors », *pe parcursul* « pendant » > *pe parcurs* « pendant », *în mijlocul* « au milieu de » > *în mijloc* « au milieu », *în jurul* « autour » > *în jur* « à proximité ».

C.b. pour la plupart des locutions prépositionnelles, l'insertion d'un adjectif (qualificatif) modifiant le nom qui se retrouve dans sa structure est impossible : **pe parcursul lung al anului* « pendant le long parcours de l'année », **în favoarea exagerată a locuitorilor* « en la faveur exagérée des habitants », **la începutul bun al anului* « au bon début de l'année ».

C.c. la modification par un adjectif démonstratif, possible pour certaines séquences : *din pricina / cauza / aceasta* « de cette cause », *cu această excepție / condiție* « à cette exception / condition », *în ipoteza / scopul / contextul / sensul / cazul / timpul acesta / aceasta* « dans ce(tte) hypothèse / but / contexte / cas / pendant ce temps » ; *la nivelul / îndemnul acesta* « à ce niveau / sur ce conseil » ; en revanche, l'insertion d'un démonstratif n'est pas possible pour d'autres locutions : **în fața / spatele / ciuda / detrimentul / locul / susul / preajma / jurul / lumina / sânul / numele / centrul / dosul / mijlocul / vederea / lipsa / fruntea acesta / aceasta* « en ce(tte) face de / derrière / dépit / détrimement / lieu / haut / proximité / autour / lumière / sein / nom / centre / dos / milieu / but /

manque / front » ; **pe urma / seama / potriwa / socoteala / măsură / parcursul acesta/aceasta* « à ce(tte) suite / base / attente / compte / mesure / parcours » ; **la stânga / dreapta / sfârșitul / poalele acestea / aceasta / acesta* « à ce(tte/s) gauche / droite / fin / pieds ».

C.d. l'invariabilité (morphologique) des différents éléments (notamment nominaux, adjectivaux et verbaux) de la séquence. Par exemple, les bases nominales ne se déclinent généralement pas, c'est-à-dire qu'elles sont invariables, apparaissant soit au singulier (*pe parcursul anului universitar* « pendant l'année universitaire » vs. **pe parcursurile anului universitar*), soit pluriel (*pe spezele statului* « aux dépens de l'état » vs. **pe speza statului*) ; de même, les bases verbales ont une conjugaison figée qui ne varie pas en mode, en temps ou en personne, mais apparaissent par défaut avec une seule et unique forme, généralement la 3^{ème} du singulier (*în (ceea) ce privește* « en ce qui concerne » vs. **în (ceea) ce privea / a(i) privit* « en ce qui concernait / a(s) concerné »).

Cependant, il faut noter que ces phénomènes – l'évolution sémantique, l'augmentation de la cohésion interne et la perte d'autonomie grammaticale – sont progressifs : en conséquence, les critères sont rarement discrets et les résultats des tests ne sont pas toujours univoques, comme cela a été noté pour d'autres langues (voir notamment, pour le français, Adler 2001 et, pour le roumain et le français, par Cuniță 2007). Par exemple, pour la perte d'autonomie grammaticale, la majorité des structures [P Ndét] n'exclut pas les modifications (P Ndét poss), y compris pour des structures pourtant unanimement reconnues comme locutions prépositionnelles : *în fața / spatele / ciuda / pofida / locul / numele / preajma meu / mea...* « devant / derrière moi / à mon encontre / malgré moi / à ma place / en mon nom / à proximité de moi » ; *din partea / pe socoteala mea* « de ma part / à mon compte » ; *la stânga / dreapta mea* « à ma gauche / à ma droite ». À ce propos, Căpățână (2000 : 144-147) émet l'hypothèse que ce ne sont pas les adjectifs possessifs qui déterminent le substantif, mais des pronoms personnels au génitif précédés de locutions prépositionnelles. Il s'agirait donc d'un faux problème.

On notera par exemple le cas de *din cauza* « en raison de / à cause de ». Cette séquence permet certaines modifications (*din această cauză* « pour cette raison », *din cauzele acestea* « en raison de cela (litt. pour ces raisons) », *din toate aceste cauze* « pour toutes ces raisons », etc.) et ne présente pas d'opacité sémantique. Cependant, elle exclut d'autres modifications : **din cauza neprevăzută* « pour une raison imprévue ». En conséquence, elle est codée comme locution par certains chercheurs, comme séquence libre par d'autres, en fonction du test retenu (cf. aussi Petrescu 2005, Vlasin 2015).

Il existe bien quelques cas consensuels, pour lesquels les tests convergent, comme à *l'instar de* en français – *în preajma* « à la veille de, autour de, à proximité de », *în pofida* « en dépit de, malgré », *în ciuda* (idem, sur *ciuda* « jaloux »), *pe spezele* « aux dépens de (litt. sur les dépenses de) », *în detrimentul* « au détriment de ». Cependant, non seulement aucun des critères n'est suffisant pour délimiter la classe, mais même en croisant les critères on obtient difficilement une image claire de ce qu'est une locution prépositionnelle ; c'est pourquoi l'application de différents critères a abouti à des inventaires différents selon les auteurs.

1.4. Matrices et listes

Aucune des études mentionnées plus haut (Ciobanu 1957, 1961, Găitănar 1965, 1999, Căpățână 2000, Petrescu 2005, Cuniță 2007, Vlasin 2015) ne prétend fournir une liste

exhaustive des locutions prépositionnelles du roumain. Cependant, la plupart des locutions qu'elles recensent se retrouvent d'une étude à l'autre et ont comme base de départ la liste initialement établie par la GA (1963) et partiellement enrichie par Duda (2002) et par Petrescu (2005). Ces locutions correspondent à différentes constructions ou matrices, présentées dans le tableau 1 ci-dessous.

numéro de matrice	éléments constitutifs	version abrégée
1.	Préposition (simple/composée) + Nom (singulier ou pluriel, avec déterminant) + Génitif	PNDET
2.	Préposition + Nom (singulier, sans déterminant) + Préposition	PNP
3.	Préposition + Adverbe (+ Préposition)	(P)Adv(P)
4.	Préposition + Adjectif (+ Préposition)	PAdj(P)
5.	Préposition + Pronom relatif + / (Adverbe +) Verbe	PPro(Adv)V

Tableau 1 : Principales matrices de locutions prépositionnelles en roumain

Ces constructions sont inégalement représentées dans les listes existantes. Les plus fournies sont les deux premières, la moins fournie la dernière, comme l'illustre le tableau 2 ci-dessous. Ce tableau illustre également l'existence de quelques séquences présentant des variations par rapport aux matrices de base. On trouve ainsi, entre autres, *față de* « par rapport à », *față cu* « en présence de », *funcție de* « en fonction de » à côté de *în față de*, *de față cu*, *în funcție de*.

Matrice	Exemples
1. PNDET	<i>în / din față</i> « devant / face à », <i>din partea</i> « de la part de », <i>în privința</i> « à l'égard de », <i>în ciuda</i> « en dépit de », <i>în pofida</i> (idem), <i>în locul</i> « au lieu de / à la place de », <i>în / din / prin jurul</i> « autour de », <i>pe calea</i> « sur la voie de », <i>pe seama</i> « sur le compte de », <i>în / de-a latul</i> « au large de », <i>în vederea</i> « en vue de », <i>prin intermediul</i> « par l'intermédiaire de », <i>la îndemâna</i> « à la portée de », <i>pe parcursul</i> « pendant / le long de », <i>în răspărul</i> « pendant », <i>în schimbul</i> « en échange de », <i>în toiul</i> « en plein milieu de », <i>în numele</i> « au nom de », <i>în / (de) prin preajma</i> « à proximité de », <i>în prelargul</i> « au large de », <i>pe spezele</i> « aux dépens de », <i>de jur împrejurul</i> « tout autour de », <i>de primprejurul</i> (idem), <i>de-a latul</i> « contre », <i>de-a lungul</i> « le long de », <i>de-a curmezișul</i> « contre »
2. PNP	<i>de față cu</i> « en présence de », <i>în conformitate cu</i> « conformément à », <i>în funcție de</i> « en fonction de », <i>în raport cu / de</i> « par rapport au », <i>în caz de</i> « en cas de », <i>cu privire la</i> « à l'égard de », <i>o dată cu</i> « en même temps que », <i>în jur de</i> (≈ <i>în jurul</i>) « environ », <i>pe cale de</i> « en voie de », <i>în urmă cu</i> « il y a », <i>în decurs de</i> « au cours de », <i>în loc de</i> « au lieu de »
3. (P)Adv(P)	(<i>în</i>) <i>afară de</i> « en-dehors de, à l'exception de », <i>în afara</i> (idem), (<i>de</i>) <i>dincoace de</i> « de l'autre côté de », <i>împreună (și) cu</i> « (ensemble) avec », <i>în / din sus de</i> « en haut de », <i>întocmai ca</i> « tout comme », <i>cât despre</i> « en ce qui concerne »
4. (P)AdjP	<i>relativ la</i> « à propos de », <i>potrivit cu</i> « selon », <i>referitor la</i> « concernant », <i>privitor la</i> « en ce qui concerne », <i>contrar cu</i> « contrairement à », <i>indiferent de</i> « malgré », <i>conform cu</i> « conformément à », <i>cu tot (toată, toți, toate)</i> « malgré », <i>cu tot cu</i> (idem)
5. PPro(/Adv)V	<i>în ce privește</i> « en ce qui concerne », <i>în ceea ce privește</i> (idem), <i>cât privește</i> (idem)
Autres	<i>față de</i> « par rapport à », <i>față cu</i> « en présence de », <i>funcție de</i> « en fonction de »

Tableau 2 : Matrices et locutions prépositionnelles (listes non exhaustives)

Les corpus permettent d'ajouter quelques éléments à ces listes, comme nous le verrons dans la section suivante. Cela tient entre autres au fait que, comme le note Lehmann (ce volume), les

locutions prépositionnelles ne constituent pas une sous-classe complètement fermée ; nous y reviendrons.¹

2. Méthode : une analyse sur corpus

2.1. Corpus utilisés

Les locutions prépositionnelles ont bien souvent une fréquence relativement faible ; par exemple, en français, leur fréquence reste généralement inférieure à 0,1‰ (Fagard sous presse ; cela semble être vrai de manière générale pour les langues romanes, voir par exemple pour l'occitan Sibille, ce volume, pour l'italien Ganfi & Piunno, 2017). Il est donc important, pour repérer et tester ces constructions, d'avoir un corpus suffisamment vaste. C'est la raison pour laquelle nous avons utilisé en priorité, pour cette étude, le corpus roTenTen16. Il fait partie de la famille des corpus TenTen, que l'on pourrait appeler des « corpus sales » (Fagard 2015) : il s'agit d'une « photographie » des pages internet en roumain, avec un post-traitement permettant d'éliminer les doublons et d'avoir une lemmatisation et un étiquetage morpho-syntaxique relativement fiables. Les avantages sont la taille considérable (environ 2,6 milliards de mots compilés à partir de plus de 9 millions de documents), l'étiquetage, et l'existence d'une interface d'interrogation puissante (SketchEngine) ; la principale limite, le manque de contrôle sur les textes le composant. On peut penser que la taille du corpus suffit en partie à compenser ce défaut. Nous avons également eu recours à d'autres corpus, afin de contrôler ces résultats : nous avons pour cela utilisé les corpus oraux de Ionescu-Ruxăndoiu (IVLRA 2002) et de Dascălu-Jinga (CORV 2002). Ces corpus sont d'étendue limitée (environ 200 000 mots), mais ont été compilés par des linguistes et sont donc plus fiables que le corpus roTenTen16 ; ils constituent ainsi un complément utile. Afin d'exploiter les données de ces deux corpus, nous avons mis à profit les ressources mis à disposition par la plateforme TXM (<http://textometrie.ens-lyon.fr/>, cf. Heiden et al. 2010, Pincemin et al. 2010), qui permettent de faire des requêtes complexes à l'aide d'un langage CQP. Il serait intéressant, pour des études futures, d'inclure d'autres corpus intermédiaires, comme le corpus CoRoLa (<http://corola.racai.ro/>).

2.2. Méthode exploratoire

Afin de dresser une liste de locutions prépositionnelles en roumain à l'aide des corpus sélectionnés, nous avons procédé par étapes. Dans un premier temps, nous avons testé l'ensemble des locutions mentionnées en section 1.3. Nous avons ensuite cherché à compléter cette liste sur la base des matrices mentionnées en 1.2. Pour cela, nous avons dressé une liste exhaustive des séquences correspondant à ces matrices, dans le corpus roTenTen, à l'aide des requêtes morpho-syntaxiques (du type [préposition + nom + préposition]), puis sélectionné parmi les plus fréquentes les séquences présentant une équivalence distributionnelle au moins

¹ D'après cet auteur, cela se vérifie également pour les prépositions simples : « the linguistic doctrine of adpositions as a closed class is a myth ».

partielle avec une préposition simple. Cependant, cette approche n'est possible de manière productive que pour certaines matrices : pour d'autres, le nombre de séquences attestées est soit bien trop important, soit négligeable. Il est donc probable que notre liste de départ, bien qu'elle contienne plus de 200 séquences, ne soit pas exhaustive.

Une étape importante a été l'élimination manuelle du bruit, à savoir de toutes les séquences fortuites ou relevant de manière évidente d'autres constructions, par exemple parce que la préposition initiale de la séquence (pour les matrices PNP) dépend du verbe, comme en (12a), ou que la base varie – il s'agit alors, le plus souvent, d'une séquence fortuite qui n'est pas figée et n'a rien à voir avec une locution prépositionnelle, comme en (12b) :

(12) a. renunță **la funcția** sa **de** manager

renoncer.PRS.3SG à fonction.SG.DET.ACC POSS.F.3SG de manager
(il renonce à sa fonction de manager) (roTenTen16)

b. Cred că femeile se relaxează numai **cât privesc** lucruri frumoase

croire.PRS.1SG que femmes.PL.DET.N REFL.3.PL relaxer.PRS.3PL seulement combien regarder.PRS.3PL chose.PL.ACC belle.PL
(je crois que les femmes se détendent seulement quand elles regardent de belles choses)

Pour les cas conservés, y compris les constructions issues des listes de locutions, nous avons ensuite procédé à une série de tests, décrits en section 1.1 : insertion avant et après la base comme *sa* en (12a), flexion de la base comme *cât privesc* en (12b). Pour cette étape, nous avons de manière générale utilisé uniquement le corpus roTenTen16, les corpus oraux étant trop réduits. Elle suppose de faire des requêtes ciblées, pour chaque séquence, pour l'insertion avant et après la base, puis pour la modification de la base. Par exemple, pour la séquence *pe bază de* (matrice 2, PNP), nous avons formulé les requêtes suivantes : insertion avant la base (13a), après la base (13b), modification ou flexion de la base (13c). Pour la matrice 1 (PNdét+G), le test d'insertion après la base est un peu plus délicat, puisqu'il suppose d'identifier les insertions avant un complément au génitif, soit « manuellement » (13d), soit automatiquement, par exemple pour les noms communs masculins au génitif singulier (13e).

(13) a. "pe" [] "bază" "de"

b. "pe" "bază" [] "de"

c. "pe" "baz.*" "de"

d. "pe" "baza" [] ".*(i|or)a?"

e. "pe" "baza" [] [tag="Ncmsoy"]

Nous avons ensuite vérifié manuellement les résultats ainsi obtenus, afin d'éliminer le bruit, comme en (14a) où les guillemets ont été identifiés comme une insertion, ou en (14b) où la préposition *din* est répétée de manière erronée.

(14) a. sponsorizări substanțiale din « partea unor întreprinderi cu capital de stat »

financement.PL.N substantiel.PL de+en PART.DET un.PL.G entreprises.PL.N avec capital de état
(des financements substantiels de la part d'entreprises ayant un capital d'état) (roTenTen16)

b. Am primit din din partea tuturor angajaților un feedback minunat

avoir.PRS.1 recevoir.PST.PTCP de de PART.DET tout.PL.G engagé.PL.G un retour merveilleux

3. Résultats

La combinaison des requêtes sur corpus et des tests nous a permis (a) de repérer un certain nombre de candidats ; (b) de confirmer le statut de locution pour une soixantaine de séquences, y compris la plupart des séquences énumérées dans les listes existantes ; (c) d'exclure une partie des séquences issues des listes et du corpus ; (d) enfin, pour le reste – une cinquantaine de séquences –, d'établir qu'il s'agit de séquences présentant des indices de figement, sans être pour autant tout à fait figées. Nous reviendrons sur ce dernier groupe de constructions. Il nous faut cependant d'abord rendre compte des séquences exclues.

3.1. Exclusions

Avant de pouvoir classer les locutions prépositionnelles en fonction de leur degré de figement, nous avons dû opérer un certain nombre d'exclusions, pour diverses raisons, que nous exposons ci-après.

Ainsi, un petit nombre de séquences étaient totalement absentes du corpus roTenTen16, et pour quelques autres nous n'avons trouvé que quelques occurrences. Il nous a semblé raisonnable d'exclure les séquences présentant une fréquence relative inférieure à 0,1 occurrence par milliard de mots (soit au maximum environ cent occurrences dans un corpus de 2,6 milliards de mots). C'est le cas de certaines séquences, comme *în paguba* « au détriment de », qui sont recensées dans les listes de locutions prépositionnelles, mais n'apparaissent pas ou presque pas dans nos corpus.

Certaines séquences présentent une fréquence plus élevée et une variabilité faible voire nulle, mais n'ont pas une distribution typique de la classe des prépositions. Ainsi, la séquence *în sinea* (+ génitif) est proche de la matrice 1 (la base étant un pronom), mais le régime est un possessif comme en (15a) dans plus de 97% des occurrences de la séquence ; il y a peu d'occurrences du type de (15b). Ces emplois semblent trop marginaux (moins d'une centaine d'occurrences sur un corpus de 2,6 milliards de mots) pour classer la séquence comme locution prépositionnelle.

- (15) a. Nefiind orgolioasă am plâns **în sinea mea**
non-être.PTCP.PRS orgueilleux.F.SG avoir.PRS.1 pleurer.PTCP.PST.M.SG en PRON.REFL.DET POSS.1.F.SG
(n'étant pas fière de moi, j'ai pleuré intérieurement)
- b. sunt niște coduri secrete **în sinea acestor cuvinte**
être.PRS.3PL certain.PL code.PL.N secret.PL en PRON.REFL.DET DEM.PL.G parole.PL
([pour voir s'il y a des codes secrets cachés dans ces paroles])

On peut en dire autant, par exemple, des séquences *pe numele* « au nom de / prénommé » et *în felul* « de la manière », qui sont généralement référentielles, c'est-à-dire que le nom a un référent qui peut être explicité et repris anaphoriquement. De manière générale, le critère sémantique de référentialité semble être assez efficace pour distinguer les locutions

prépositionnelles figées des autres séquences (cf. Căpățână 2000, 2003, Cuniță 2007 et Stosic & Fagard, ce volume) ; cependant, il est difficilement quantifiable.

La séquence *de necaz* est assez atypique dans notre sélection ; elle est rarement suivie d'un régime (nominal ou autre), et la quasi-totalité de ses emplois sont adverbiaux, en particulier dans la locution *a face haz de necaz* « faire fi des difficultés, rire des ennuis » ou dans des expressions comme *a muri de necaz* « être très envieux / mourir d'envie ». Nous l'avons donc exclue.

Enfin, les séquences *la data de*, *în data de* « en date de » présentent un autre cas de figure. Elles alternent facilement avec une préposition simple, *la*, *în* ou (régionalement) *pe* (ces dernières étant largement préférées) : *la data de 1 septembrie* > *la / pe 1 septembrie* « le 1^{er} septembre », *în data de 10 august* > *în 10 august* « le 10 août ». On pourrait donc les considérer comme des locutions prépositionnelles. Cependant, leurs caractéristiques distributionnelles sont atypiques pour des prépositions, dans la mesure où elles sont presque exclusivement suivies d'une date, d'autres types de compléments étant quasiment exclus (on trouve cependant quelques occurrences comme *la data de înregistrare* « à la date d'enregistrement »). En outre, elles s'emploient principalement dans le registre formel-administratif. Pour ces raisons, nous ne les avons pas retenues dans notre liste de locutions prépositionnelles.

3.2. Locutions prépositionnelles : matrices et figement

Parmi les séquences que nous n'avons pas exclues, on observe différents degrés de figement. Il n'y a de corrélation parfaite ni entre matrice et figement, ni entre figement et fréquence. Les séquences analysées comme locutions prépositionnelles ont donc une fréquence très variable², et appartiennent à différentes matrices. Par définition, elles ont des emplois prépositionnels, c'est-à-dire qu'il y a équivalence distributionnelle avec une ou plusieurs prépositions simples, et présentent une variabilité limitée, c'est-à-dire qu'elles répondent négativement aux tests retenus dans la section précédente (insertion et modification / flexion impossibles ou limitées).

Pour les tests d'insertion et de modification, nous avons cherché à éliminer le bruit et à ne conserver que les occurrences pertinentes. On peut constater que certains modificateurs semblent compatibles avec un figement partiel : c'est le cas notamment d'adverbiaux tels que *chiar* « même », *doar* « seulement », qui peuvent également modifier des prépositions simples. Certaines séquences peuvent être modifiées par ces adverbes mais sans insertion : on trouve ainsi *chiar în fața* + *G* « juste en face » mais pas **în chiar fața*. Ceci est bien évidemment à interpréter comme un signe de figement fort.

Pour évaluer le degré de figement, nous avons fixé des seuils arbitraires, à 0,1%, 1% et 10% des occurrences. Le tableau 3 ci-dessous illustre les disparités entre matrices, pour trois degrés de figement : fort (0-0,1% maximum d'insertion et/ou de modification), modéré (0,1-1% insertion et/ou modification) et faible (1-10% d'insertion et/ou modification).

² Comme le note Sibille (ce volume), il n'y a pas de « corrélation évidente entre la fréquence d'emploi et le degré de figement », peut-être en raison de l'existence, dans la plupart des cas, d'un équivalent simple.

Figement	Matrice	Séquences concernées
fort : 0-0,1% maximum d'insertion et/ou de modification	1. PNDET (59 séquences)	<i>cu excepția, de dragul, de jur împrejurul, de pe urma, de-a latul, de-a lungul, din cauza, din pricina, după placul, fără voia, în afara, în calitate, în ciuda, în contra, în culmea, în dauna, în dosul, în dreptul, în fața, în favoarea, în folosul, în fruntea, în jurul, în latul, în lipsa, în locul, în lumina, în marginea, în numele, în ochii, în preajma, în prezența, în privirea, în răspărul, în schimbul, în seama, în serviciul, în slujba, în spatele, în susul, în temeiu, în urma, în vederea, în virtutea, la adresa, la cheremul, la data, la dreapta, la marginea, la sfârșitul, la stânga, pe măsura, pe potrivă, pe seama, pe socoteala, pe spezele, pe temeiu, prin prisma, sub raportul</i>
	2. PNP (4 séquences)	<i>în curs de, în decurs de, în jur de, în urmă cu, pe cale de</i>
	3. (P)Adv(P) (6 séquences)	<i>aproape de, cât despre, (de) dincoace de, în sus de, întocmai ca</i>
	4. PAdj(P) (1 séquence)	<i>cu tot</i>
	autres (2 séquences)	<i>față cu, la un loc cu</i>
modéré : 0,1-1% insertion et/ou modification	1. PNDET (25 séquences)	<i>cu ajutorul, cu scopul, cu voia, din partea, în ajunul, în contul, în decursul, în defavoarea, în detrimentul, în josul, în largul, în mijlocul, în pofida, în privința, în sânul, în scopul, în timpul, în toiul, la îndemâna, la poalele, pe baza, pe calea, prin intermediul, prin mijlocirea, sub rezerva</i>
	2. PNP (9 séquences)	<i>cu privire la, în calitate de, în caz de, în comparație cu, în conformitate cu, în materie de, pe bază de, pe data de, spre deosebire de</i>
	3. (P)Adv(P) (3 séquences)	<i>împreună cu, (în) afară de</i>
faible : 1-10% d'insertion et/ou modification	1. PNDET (13 séquences)	<i>de felul, de-a curmezișul, dinspre partea, în amintirea, în capul, în centrul, în cursul, în interesul, în lungul, în măsura, în perioada, în spiritul, la începutul</i>
	2. PNP (12 séquences)	<i>de față cu, în ajun de, în colaborare cu, în comun cu, în data de, în funcție de, în lipsă de, în loc de, în raport cu, la data de, o dată cu, pe cap de</i>
	3. (P)Adv(P) (2 séquences)	<i>cu tot cu</i>
	4. PAdj(P) (2 séquences)	<i>indiferent de, potrivit cu</i>
	5. PPro(/Adv)V	<i>cât privește</i>
	autres (1 séquence)	<i>față de</i>

Tableau 3 : Matrices et degrés de figement des locutions prépositionnelles en roumain

La plupart des séquences ont trait à la matrice 1 et 2, seules quelques séquences relevant d'autres matrices.

3.3. Figement partiel

Outre ces séquences que nous analysons comme des locutions prépositionnelles, on trouve des séquences partiellement figées. Cette variabilité peut être liée à des insertions avant ou après la base, ou à la modification ou flexion de la base. Ces séquences sont reproduites dans le tableau 4 ci-dessous. Certaines d'entre elles sont considérées comme des locutions prépositionnelles par la tradition grammaticale. Sans remettre en cause cette classification, on

peut noter au moins qu'elles présentent un degré de figement moindre par rapport aux locutions figurant dans le tableau 3.

Matrice	séquence	modifications		
		avant la base	après la base	de la base
1. PNDET	<i>cu numele</i>	-	+	-
	<i>după măsura</i>	-	+/-	+
	<i>în ideea</i>	+	+	-
	<i>pe parcursul</i>	+	-	-
	<i>pe placul</i>	-	-	+
	<i>pe urma</i>	-	-	+
2. PNP	<i>de acord cu</i>	+/-	+/-	-
	<i>în acord cu</i>	+	+/-	+/-
	<i>pe atât de</i>	-	+/-	+/-
	<i>în bază de</i>	+/-	-	+/-
	<i>în căutare de</i>	+/-	+/-	+
	<i>în concordanță cu</i>	+	+/-	-
	<i>în cuantum de</i>	+/-	+	+/-
	<i>în mijloc de</i>	+	-	+
	<i>la fel cu</i>	+/-	+	-
	<i>la început de</i>	+	+/-	+/-
	<i>la sfârșit de</i>	+	-	-
	<i>în raport de</i>	+	+/-	+
	<i>la data la</i>	+	+	-
3. (P)Adv(P)	<i>înainte cu</i>	NP	-	+/-
4. PAdj(P)	<i>conform cu</i>	NP	+	-
	<i>privitor la</i>	NP	-	+
	<i>referitor la</i>	NP	-	+
	<i>relativ la</i>	NP	-	+
	<i>contrar cu</i>	NP	+	+
5. PPro(/Adv)V	<i>în ce privește</i>	-	+	-
	<i>în ceea ce privește</i>	-	+/-	-
autres	<i>funcție de</i>	NP	+	-

Tableau 4 : Séquences présentant une variabilité élevée ('+' : plus de 10% des occurrences ; '+/-' : entre 1 et 10% ; '-' : moins de 1% ; 'NP' : non pertinent).

Par rapport aux séquences que nous avons identifiées comme des locutions prépositionnelles, ces dernières sont généralement moins fréquentes, ont des emplois moins prépositionnels, c'est-à-dire qu'il n'y a pas toujours équivalence distributionnelle avec une ou plusieurs prépositions simples, et répondent positivement à au moins une partie des tests retenus dans la section précédente.

4. Discussion – Les locutions prépositionnelles : une catégorie ?

Les tableaux présentés ci-dessus illustrent bien le problème général des locutions prépositionnelles : il s'agit d'une catégorie relativement peu homogène, comportant des séquences formées sur des matrices différentes, présentant des fréquences variées et n'ayant pas toujours le même fonctionnement. Cependant, on peut tout de même souligner l'existence d'un noyau assez bien identifiable. À la différence des autres langues romanes, ce noyau n'est

pas limité aux séquences [P (DET) N P] : en roumain, les matrices les plus productives sont les matrices 1 et 2 ([PNDET + génitif] et [PNP]). Comme on le voit dans les tableaux 3 à 5, ces matrices fournissent la plupart des locutions prépositionnelles figées et moins figées.

On peut donc penser que ce sont les matrices 1 et 2 qui permettent le renouvellement de la catégorie : il semble en effet qu'il y a un mouvement régulier, avec des séquences figées qui disparaissent (ce qui expliquerait l'absence, dans nos corpus, de certaines séquences identifiées comme locutions prépositionnelles dans les études antérieures) et d'autres qui sont en voie de figement. Pour confirmer cette hypothèse, il faudrait inclure dans une prochaine étude un volet diachronique.

Les autres matrices sont anecdotiques du point de vue du nombre d'éléments, du moins en comparaison des deux premières. Cependant, si l'on prend en compte la fréquence des locutions, on peut constater – notamment pour les matrices 3 et 4 – que la différence n'est pas si importante. Le tableau 5, ci-dessous, montre ainsi que la fréquence relative cumulée de ces deux matrices reste comparable à celle des deux premières, à la différence de la matrice 5, par exemple. Cependant, le nombre très limité de séquences relevant de ces matrices identifiées comme locutions prépositionnelles invite à se demander si ces dernières font système, à la différence des matrices 1 et 2, dont il est évident qu'elles peuvent servir de patron productif pour la création de nouvelles locutions (cf. Hoffmann 2004).

Matrice	Nombre de locutions	Fréquence relative cumulée (par million de mots)
1. P (simple/composée) + N (sg/pl, + dét)	97	1600,6
2. P + N (sg, - dét) + P	26	582,6
3. P + Adv (+ P)	11	200,8
4. P + Adj (+ P)	3	81,8
5. PPro(/Adv)V	1	0,3
autres	3	3,1
Total	140	2469,2

Tableau 5 : Matrices et fréquences relatives.

On peut se demander s'il n'y a pas, en roumain comme dans d'autres langues, une tendance à la réduction de ces constructions complexes, avec notamment la perte de la préposition initiale, illustrée par l'alternance entre *în afară de* et *afară de*. Cependant, d'une part cette tendance est encore trop faible pour que les séquences ainsi obtenues participent réellement au renouvellement du paradigme des prépositions simples, d'autre part il faut reconnaître qu'en roumain, les groupes comme *în afară de* sont plus fréquents et ressentis comme naturels, tandis que les séquences sans préposition initiale comme *afară de* sont plus rares et semblent vieillies.

On pourrait en outre s'interroger plus avant sur le statut prépositionnel de certaines des constructions retenues, notamment celles permettant d'introduire une approximation (*în jurul*, *în jur de* « environ », cf. les exemples (3-5) plus haut) ou celles qui semblent fonctionner comme des marqueurs de topique, comme *cât despre* « en ce qui concerne ». Ainsi, les séquences comme *în jurul*, par exemple, présentent une distribution proche de certains adverbes, notamment *aproximativ* / (fam.) *cam* « environ » – avec lesquels elles peuvent

d’ailleurs alterner – mais, à la différence des adverbes, leur complément est obligatoire, cf. (16).

(16) au cheltuit în jur de 1.800 de euro (/ *au cheltuit în jur de)
avoir.PRS.3PL dépenser.PST.PTCP en tour de 1800 de euro
(ils ont dépensé environ 1800 euros) (roTenTen16)

Conclusion

Nous avons cherché, dans cet article, à établir une liste de locutions prépositionnelles du roumain sur la base de tests principalement sémantiques et morpho-syntaxiques, appliqués aux occurrences d’un peu plus de 200 séquences dans trois corpus. Pour ce faire, nous avons eu recours principalement au corpus roTenTen16 (2,6 milliards de mots). Sans être exhaustive, la liste que nous dressons compte 140 locutions ; elle est un peu plus fournie que les listes existantes, mais surtout, à la différence des études précédentes, elle ne se limite pas aux locutions du type (Préposition + Nom) suivies du génitif – notre matrice 1, ou PNDET.

La liste que nous avons établie nous permet de confirmer l’importance de cette matrice, mais aussi de montrer l’importance relative d’autres matrices, aussi bien pour le nombre de locutions prépositionnelles (matrices 2 – PNP –, et 3 – (P)Adv(P)) que pour la fréquence cumulée (matrices 2, 3 et 4 – PAdj(P)). La productivité des matrices 1 et 2 est cependant plus forte, et ce sont surtout ces dernières, semble-t-il, qui contribuent au renouvellement de la catégorie. A ce propos, on peut d’ailleurs faire l’hypothèse que la catégorie des locutions prépositionnelles est plus ouverte que celle des prépositions simples, avec un renouvellement plus important, et que la coexistence des deux sous-systèmes (matrices 1 et 2, génitif et préposition fonctionnelle) ne semble pas avoir constitué un frein au figement des locutions.

Ces derniers points constituent précisément une piste de recherche future : pour les élucider, il conviendrait d’analyser ces séquences non seulement en synchronie, mais aussi en diachronie.

Abréviations

1	1 ^{ère} personne	N	nominatif
3	3 ^{ème} personne	PART	partitif
ACC	accusatif	PL	pluriel
CL	clitique	POSS	possessif
DEM	démonstratif	PRON	pronom
DET	déterminant	PRS	présent
F	féminin	PST	passé
G	génitif	PTCP	participe
M	masculin	REFL	réfléchi
MDO	marquage différentiel de l’objet	SG	singulier

Références bibliographiques

- Adler, S. (2001) : Les locutions prépositives : questions de méthodologie et de définition. *Travaux de linguistique*, 42-43, pp. 157-170.
- Avram, M. (2001) : *Gramatica pentru toți*. Ediția a III-a, Editura Humanitas, București.
- Căpățână, C. (2000) : Locuțiunile prepoziționale, in : Negomireanu, D. (éd.) : *Ion Coteanu. In memoriam*. Editura Universitaria, Craiova, pp. 84-89.
- Căpățână, C. (2003) : Locuțiunile prepoziționale, *Limba Română*, pp. 48-53.
- Ciobanu, F. (1957) : Observații asupra prepozițiilor și locuțiunilor prepoziționale din limba română. *Studii de gramatică*, 2, pp. 357-388.
- Ciobanu, F. (1961) : Unele aspecte ale corespondenței dintre elementele prepoziționale și cele conjuncționale, cu referire specială la locuțiuni. *Studii de gramatică*, 3, pp. 43-66.
- Cuniță, A. (2007) : Locuțiunile prepoziționale – instrumente cu încărcătură culturală ale (re)configurării referentului substantivului regim, in : Stan, C., R. Zafiu & A. Nicolae (éds.) : *Studii lingvistice. Omagiu profesoarei Gabriela Pană Dindelegan, la aniversare*, Editura Universității din București, București, pp. 67-75.
- Dragomirescu, Gh. N. (1963) : Problema locuțiunilor ca obiect al analizei gramaticale. *Limba Română*, XII, 6, pp. 618 – 625.
- DSL 1997 = Bidu Vranceanu, A., C. Călărașu, L. Ionescu-Ruxăndoiu, M. Mancaș, & G. Pană Dindelegan (1997) : *Dicționar general de științe. Științe ale limbii*. Editura Științifică, București.
- Duda, G. (2002), *Dicționar de locuțiuni ale limbii române*, Editura All Educational, București.
- Fagard, B. (2015) : *Pour une romanistique moderne : la microtypologie*. Mémoire de synthèse pour l'habilitation à diriger les recherches. Université François-Rabelais de Tours.
- GA (1954/1963) = Graur, A., M. Avram & L. Vasiliu (éds.) (1954/1963) : *Gramatica limbii române*. Editura Academiei Române, București.
- GALR (2005/2008) = Guțu-Romalo, V. (éd.) (2005/2008) : *Gramatica limbii române*. Editura Academiei Române, București.
- Ganfi, V. & V. Piunno (2017) : Preposizioni complesse in italiano antico e contemporaneo. Grammaticalizzazione, schematismo e produttività. *Archivio Glottologico Italiano*, CII,2, pp. 184-204.
- Găitănanu, Ș. (1965) : Descrierea locuțiunilor prepoziționale cu genitivul. *Limba Română*, 3-4, pp. 87-92.
- Găitănanu, Ș. (1999) : Criteriile deosebirii locuțiunilor prepoziționale cu genitivul de grupările libere, analizabile. *Limba și literatura română*, XXVIII, 2, pp. 12-14.
- Geană, I. (2008) : Prepozițiile semilexicale. Aspecte teoretice. *Al II-lea Simpozion internațional de lingvistică, Institutul de Lingvistică „Iorgu Iordan – Al. Rosetti”*, București, les 28-29 novembre 2008.

- Geană, I. (2009) : Din nou despre prepozițiile semilexicale. *Al IX-lea Colocviu al Catedrei de limba română, Limba română: controverse, delimitări, noi ipoteze*, București, les 4-5 décembre 2009.
- Geană, I. (2011) : Vechi și nou în interpretarea prepoziției. *Colocviul Internațional Limba română – abordări tradiționale și moderne, ediția a II-a*, Universitatea „Babeș-Bolyai”, Cluj-Napoca, les 6-7 mai 2011.
- GR (2013) = Pană Dindelegan, G. (éd.) (2013) : *The Grammar of Romanian*. Oxford University Press, Oxford.
- Heiden S., J.-P. Magué & B. Pincemin (2010) : TXM : Une plateforme logicielle open-source pour la textométrie – conception et développement, in : Bolasco, S., I. Chiari & L. Giuliano (éds.) : *Proceedings of the 10th International Conference on the Statistical Analysis of Textual Data*. Edizioni Universitarie di Lettere Economia Diritto, Rome, pp. 1021-1032.
- Hoffmann, S. (2004) : Are Low-Frequency Complex Prepositions Grammaticalized? On the Limits of Corpus Data – and the Importance of Intuition, in : Lindquist, H. & C. Mair (éds.) : *Corpus Approaches to Grammaticalization in English*. John Benjamins, Amsterdam & Philadelphia, pp. 171-210.
- Lehmann, C. (1985) : Grammaticalization: Synchronic Variation and Diachronic Change. *Lingua E Stile*, 20, pp. 303-318.
- Mardale, A. (2009) : *Les prépositions fonctionnelles du roumain : études comparatives sur le marquage casuel*. L'Harmattan, Paris.
- Mardale, A. (2011) : Prepositions as a semilexical category. *Bucharest Working Papers in Linguistics*, XIII, 2, Editura Universității din București, București, pp. 35-50.
- Mardale, A., I. Giurgea & C. Dobrovie-Sorin (2013) : Adnominal Prepositional Phrases, in : Dobrovie-Sorin, C. & I. Giurgea (éds.) : *A Reference Grammar of Romanian*, Volume 1 : *The Noun Phrase*. John Benjamins, Amsterdam, pp. 549-584.
- Petrescu, M. (2005) : Despre locuțiunile prepoziționale cu genitivul. *Annales Universitatis Apulensis. Series Philologica*, 6.3. (<http://philologica.uab.ro/index.php?pagina=pg&id=17&l=ro>)
- Pană Dindelegan, G. (1997) : Din nou despre statutul prepoziției. Cu referire specială la prepoziția PE. *Limba română*, XLVI, 1-3, pp. 165-174.
- Pincemin B., S. Heiden, M.-H. Lay, J.-M. Leblanc & J.-M. Viprey (2010) : Fonctionnalités textométriques : Proposition de typologie selon un point de vue utilisateur, in : Bolasco, S., I. Chiari & L. Giuliano (éds.) : *Proceedings of the 10th International Conference on the Statistical Analysis of Textual Data*. Edizioni Universitarie di Lettere Economia Diritto, Rome, pp. 341-353.
- Tseng, J. (éd.) (2013) : *Prépositions et postpositions, approches typologiques et formelles*. Hermès Lavoisier, Paris.
- Vlasin, V. A. (2015) : Câteva considerații asupra statutului morfologic al locuțiunilor prepoziționale cu genitivul, in : Pomian, I. & N. Mocanu (éds.) : *Înspre și dinspre Cluj, contribuții lingvistice: omagiu profesorului G. G. Neamțu la 70 de ani*. Editura Scriptor, Cluj-Napoca, pp. 645-657.

Annexe : analyse des séquences étudiées.

Dans les tableaux (i-viii) ci-dessous, nous détaillons les résultats obtenus à partir du corpus roTenTen16, en particulier : l'insertion avant la base (en pourcentage des occurrences prépositionnelles, colonne « insertion avant »), après la base (en pourcentage des occurrences prépositionnelles, colonne « insertion après »), la modification de la base (en pourcentage des occurrences prépositionnelles, colonne « modification ») et la fréquence relative par million de mots (colonne « modification »). Nous présentons les résultats par matrice, par degré de figement et/ou par motif d'exclusion (pour les séquences non retenues). L'abréviation « np » signifie que la mesure est non pertinente.

séquence	fréquence				
<i>cu excepția</i>	27,5	<i>în folosul</i>	3,6	<i>în temeiul</i>	3,5
<i>de dragul</i>	8,5	<i>în fruntea</i>	7,5	<i>în urma</i>	85,9
<i>de pe urma</i>	14,6	<i>în jurul</i>	64,8	<i>în vederea</i>	54,5
<i>de-a latul</i>	0,6	<i>în latul</i>	0,1	<i>în virtutea</i>	2,9
<i>de-a lungul</i>	65,9	<i>în lipsa</i>	4,9	<i>la adresa</i>	34,3
<i>din cauza</i>	176,7	<i>în locul</i>	15,3	<i>la cherecul</i>	0,9
<i>din pricina</i>	9,1	<i>în lumina</i>	3,3	<i>la data</i>	37,2
<i>după placul</i>	0,4	<i>în marginea</i>	0,5	<i>la dreapta</i>	1,4
<i>fără voia</i>	0,7	<i>în numele</i>	12,4	<i>la marginea</i>	12,0
<i>în afara</i>	30,1	<i>în ochii</i>	5,8	<i>la sfârșitul</i>	91,4
<i>în calitate</i>	0,9	<i>în preajma</i>	9,1	<i>la stânga</i>	1,0
<i>în ciuda</i>	18,4	<i>în prezența</i>	9,1	<i>pe măsura</i>	4,1
<i>în contra</i>	0,7	<i>în privirea</i>	0,2	<i>pe potrivă</i>	0,1
<i>în culmea</i>	0,6	<i>în răspărul</i>	0,0	<i>pe seama</i>	14,4
<i>în dauna</i>	1,9	<i>în schimbul</i>	5,8	<i>pe socoteala</i>	0,2
<i>în dosul</i>	0,2	<i>în seama</i>	0,8	<i>pe spezele</i>	0,1
<i>în dreptul</i>	8,6	<i>în serviciul</i>	1,0	<i>pe temeiul</i>	0,5
<i>în fața</i>	7,0	<i>în slujba</i>	4,6	<i>prin prisma</i>	10,2
<i>în favoarea</i>	16,3	<i>în spatele</i>	34,8	<i>sub raportul</i>	0,5
		<i>în susul</i>	0,2		

Tableau (i) : Figement élevé – matrice 1 (PNDET) (modification, insertion impossibles).

séquence	insertion avant	insertion après	modification	fréquence
<i>cu ajutorul</i>	0,01-0,1%	0	0	98,1
<i>cu scopul</i>	0,1-0,5%	0	0	3,7
<i>cu voia</i>	0,5-1%	0	0	0,8
<i>din partea</i>	0,01-0,1%	0	0,01-0,1%	136,7
<i>în ajunul</i>	0,1-0,5%	0	0	1,3
<i>în contul</i>	0,01-0,1%	0	0	3,7
<i>în decursul</i>	0,5-1%	0	0	4,0
<i>în defavoarea</i>	0,01-0,1%	0	0	1,7
<i>în detrimentul</i>	0,01-0,1%	0	0	3,6
<i>în josul</i>	0	0	0	0,7
<i>în largul</i>	0,01-0,1%	0	0	1,5
<i>în mijlocul</i>	0,1-0,5%	0	0	18,8
<i>în pofida</i>	0,01-0,1%	0	0	4,5
<i>în privința</i>	0	0,1-0,5%	0	23,3
<i>în sânul</i>	0,5-1%	0	0	2,5
<i>în scopul</i>	0,01-0,1%	0	0	15,1
<i>în timpul</i>	0,1-0,5%	0	0	75,6
<i>în toiul</i>	0,01-0,1%	0	0	1,1
<i>la îndemâna</i>	0,01-0,1%	0,01-0,1%	0	6,8
<i>la poalele</i>	0,01-0,1%	0	0	7,1
<i>pe baza</i>	0,01-0,1%	0	0	69,6
<i>pe calea</i>	0,5-1%	0	0	14,6
<i>prin intermediul</i>	0	0	0,1-0,5%	81,2
<i>prin mijlocirea</i>	0,01-0,1%	0	0	1,0
<i>sub rezerva</i>	0,01-0,1%	0	0	1,1

Tableau (ii) : Figement modéré – matrice 1 (PNDET).

séquence	insertion avant	insertion après	modification	fréquence
<i>de felul</i>	1-3%	0	0	1,9
<i>de-a curmezișul</i>	0	0	1-8%	0,3
<i>dinspre partea</i>	0	0	1-8%	0,1
<i>în amintirea</i>	0	0	1-8%	1,4
<i>în capul</i>	0	0	1-8%	1,8
<i>în centrul</i>	1-3%	0,1-0,5%	0,5-1%	13,0
<i>în cursul</i>	1-3%	0	0	11,5
<i>în interesul</i>	0	0	1-8%	4,5
<i>în lungul</i>	1-3%	0	0	0,6
<i>în măsura</i>	0	1-6%	0	0,7
<i>în perioada</i>	0,5-1%	0,5-1%	0,1-0,5%	10,0
<i>în spiritul</i>	0,01-0,1%	0	1-8%	3,4
<i>la începutul</i>	0,01-0,1%	0	1-8%	45,8

Tableau (iii) : Figement faible – matrice 1 (PNDET).

Matrice	séquence	insertion avant	insertion après	modification	fréquence
2. PNP	<i>în curs de</i>	0	0	0	9,2
	<i>în decurs de</i>	0	0	0	2,4
	<i>în jur de</i>	0	0	0	12,7
	<i>în urmă cu</i>	0	0	0	32,0
	<i>pe cale de</i>	0	0	0	7,5
3. (P)Adv(P)	<i>aproape de</i>	np	0	0	75,4
	<i>cât despre</i>	25-90%	0	0	7,3
	<i>de dincoace de</i>	0	0	0	0,2
	<i>dincoace de</i>	np	0	0	0,4
	<i>în sus de</i>	0	0	0	0,3
	<i>întocmai ca</i>	np	0	0	1,1
4. PAdj(P)	<i>cu tot</i>	0	0	0	13,7
autres	<i>de jur împrajurul</i>	0	0	0	0,3
	<i>față cu</i>	np	0	0	0,6
	<i>la un loc cu</i>	0	0	0	0,5

Tableau (iv) : Figement élevé – autres matrices.

Matrice	séquence	insertion avant	insertion après	modification	fréquence
2. PNP	<i>cu privire la</i>	0,01-0,1%	0,01-0,1%	0,1-0,5%	131,8
	<i>în calitate de</i>	0,1-0,5%	0,1-0,5%	0	22,4
	<i>în caz de</i>	0	0	0	34,4
	<i>în comparație cu</i>	0	0,1-0,5%	0	7,0
	<i>în conformitate cu</i>	0,5-1%	0,1-0,5%	0,01-0,1%	26,0
	<i>în materie de</i>	0,01-0,1%	0,1-0,5%	0,1-0,5%	9,7
	<i>pe bază de</i>	0,1-0,5%	0,01-0,1%	0,01-0,1%	47,9
	<i>pe data de</i>	0,01-0,1%	0	0,5-1%	26,5
	<i>spre deosebire de</i>	0	0,01-0,1%	0	20,0
3. (P)Adv(P)	<i>afară de</i>	0	0,01-0,1%	0	9,1
	<i>împreună cu</i>	np	0,1-0,5%	0	82,8
	<i>în afară de</i>	0	0,1-0,5%	0	10,1

Tableau (v) : Figement modéré – autres matrices.

Matrice	séquence	insertion avant	insertion après	modification	fréquence
2. PNP	<i>de față cu</i>	1-3%	1-6%	0	1,0
	<i>în ajun de</i>	1-3%	0	0	0,2
	<i>în colaborare cu</i>	1-3%	0,1-0,5%	1-8%	12,6
	<i>în comun cu</i>	0	1-6%	0	1,6
	<i>în data de</i>	0	0	1-8%	23,5
	<i>în funcție de</i>	0,1-0,5%	0,5-1%	1-8%	59,8
	<i>în lipsă de</i>	1-3%	1-6%	0,5-1%	0,6
	<i>în loc de</i>	1-3%	0,01-0,1%	1-8%	11,3
	<i>în raport cu</i>	0,01-0,1%	0,5-1%	1-8%	13,3
	<i>la data de</i>	0,01-0,1%	0,1-0,5%	1-8%	60,0
	<i>o dată cu</i>	1-3%	0	0	4,8
	<i>pe cap de</i>	0	0	1-8%	4,3
3. (P)Adv(P)	<i>cât privește</i>	0	1-6%	1-8%	0,3
	<i>cu tot cu</i>	0	0	1-8%	13,8
4. PAdj(P)	<i>indiferent de</i>	np	1-6%	0,1-0,5%	66,3
	<i>potrivit cu</i>	np	1-6%	0	1,9
autres	<i>față de</i>	np	1-6%	0	2,0

Tableau (vi) : Figement faible – autres matrices.

Matrice	séquence	insertion avant	insertion après	modification	fréquence
1. PNDET	<i>cu numele</i>	0	9-25%	0	5,7
	<i>după măsura</i>	0	0,1-0,5%	10-650%	0,1
	<i>în ideea</i>	25-90%	40-60%	0	0,5
	<i>pe parcursul</i>	7-15%	0	0	41,2
	<i>pe placul</i>	0	0	10-650%	6,6
	<i>pe urma</i>	0	0	10-650%	8,5
2. PNP	<i>în concordanță cu</i>	7-15%	1-6%	0	3,7
	<i>în cuantum de</i>	0,1-0,5%	9-25%	0,1-0,5%	2,2
	<i>în mijloc de</i>	25-90%	0	100+%	0,1
	<i>la fel cu</i>	0,01-0,1%	9-25%	0	4,7
	<i>la început de</i>	7-15%	0,5-1%	1-8%	4,4
	<i>la sfârșit de</i>	25-90%	0	0	4,4
4. PAdj(P)	<i>conform cu</i>	np	9-25%	0	3,7
	<i>privitor la</i>	np	0,1-0,5%	100+%	3,6
	<i>referitor la</i>	np	0,1-0,5%	100+%	27,3
	<i>relativ la</i>	np	0	10-650%	1,6
5. P + Pron. rel. + / (Adv +) Vb	<i>în ce privește</i>	0,1-0,5%	9-25%	0,1-0,5%	5,6
	<i>în ceea ce privește</i>	0,1-0,5%	1-6%	0,1-0,5%	40,3
autres	<i>funcție de</i>	np	40-60%	0	1,7

Tableau (vii) : Séquences non figées.

Matrice	séquence	statut
1. PNDET	<i>de la potriva, de preajma, de primprejurul, de prin preajma, în colindul, în curmezișul, în necazul, în paguba, în pica, în pilda, în pizma, în plinul, în potriva, în prada, în prealargul, în prelargul, în sinele, la potriva</i>	5. fréquence trop faible
2. PNP	<i>în apropiere de, în pagubă de</i>	
3. (P)Adv(P)	<i>împrejur de</i>	
autres	<i>de necaz, din punctul de vedere, drept numele</i>	
1. PNDET	<i>în felul, în sinea, pe numele</i>	6. distribution non prépositionnelle
2. PNP	<i>cu privire de, de acord cu, în acord cu, în bază de, în căutare de, în contextul în, în raport de, la data la, pe atât de</i>	
3. (P)Adv(P)	<i>înainte cu</i>	
4. PAdj(P)	<i>contrar cu</i>	
1. PNDET	<i>din dreptul, din fața, din josul, din jurul, din mijlocul, din preajma, din spatele, din susul, prin dreptul, prin jurul, prin mijlocul, prin preajma, prin spatele</i>	7. allomorphe
3. (P)Adv(P)	<i>cât și despre, din sus de, împreună și cu</i>	

Tableau (viii) : Autres séquences – matrice et motif de l'exclusion.

Adresses des auteurs :

Benjamin Fagard

Lattice – Ecole Normale Supérieure

1 rue Maurice Arnoux

92120 Montrouge

Alexandru Mardale

INaLCO

65 Rue des Grands Moulins

75013 Paris