

# Applications du corpus parallèle RomCro dans les recherches contrastives - les valeurs de l'article zéro dans les langues romanes

Original Scientific Paper

**Gorana Bikić-Carić**

*Département d'études romanes, Chaire de linguistique romane  
gbcaric@ffzg.unizg.hr<sup>1</sup>*

Dans le présent article nous nous consacrons aux emplois prototypiques et non prototypiques des articles (notamment de l'article zéro) dans cinq langues romanes (français, espagnol, portugais, italien, roumain). Dans cette approche contrastive nous nous servons du corpus parallèle multilingue et multidirectionnel RomCro, que nous développons depuis 2019 et que nous présentons brièvement à cette occasion. Nous comparons les articles dans les cinq langues romanes citées en fonction de l'expression du rapport entre le nom et le référent. Les comparaisons des phrases alignées montrent une claire différence entre le français et les autres langues quant à l'emploi de l'article zéro. C'est pourquoi nous nous sommes concentrée sur les valeurs de l'article zéro en espagnol, portugais, italien et roumain exprimées par les articles défini, indéfini et partitif en français.

Mots-clés : linguistique contrastive, article grammatical, emploi (non)prototypique, langues romanes, corpus parallèle RomCro

---

<sup>1</sup> Gorana Bikić-Carić est professeure associée à la Chaire de linguistique romane. Dans ses travaux elle se consacre principalement à la linguistique contrastive (français, espagnol, portugais, italien, roumain, croate), concernant les domaines nominal et verbal. Elle est coordinatrice du projet de développement du corpus parallèle multilingue et multidirectionnel RomCro.

CROSB: <https://www.bib.irb.hr/pregled/profil/14582>

## 1. Introduction

Dans le présent article nous voudrions mettre en évidence les valeurs de l'article zéro en comparant quatre langues romanes (espagnol, portugais, italien, roumain) avec le français. Pour une pareille approche contrastive il est très utile d'avoir à sa disposition des phrases alignées, à savoir une même phrase en différentes langues. C'est pourquoi nous nous sommes servie dans notre recherche du corpus RomCro. Ce corpus est le résultat d'un projet parrainé par l'Université de Zagreb et la Faculté de philosophie et lettres que nous développons depuis 2019. En effet, comme nous nous intéressons depuis des années à la linguistique contrastive des langues romanes et du croate, nous nous sommes rendu compte de la nécessité d'un corpus parallèle qui correspondrait à nos besoins. Depuis la création du corpus, nous l'avons utilisé dans plusieurs recherches portant sur les domaines nominal et verbal. Après un bref aperçu des résultats de nos recherches antérieures, nous comparerons les phrases alignées dans un sous-corpus créé pour cette occasion avec le but d'approfondir nos connaissances sur les valeurs de l'article zéro.

## 2. Corpus parallèle RomCro

RomCro est un corpus parallèle multilingue et multidirectionnel qui contient des textes littéraires écrits en français, espagnol, italien, portugais, roumain et croate. Toutes les langues sont représentées avec des textes originaux et leurs traductions. Nous développons ce corpus depuis novembre 2019 dans le cadre d'un projet conçu à la Chaire de linguistique romane du Département d'études romanes, Faculté de philosophie et lettres de l'Université de Zagreb en Croatie.<sup>2</sup> En ce moment, le corpus comprend une trentaine de romans du XX<sup>ème</sup> et du XXI<sup>ème</sup> siècles et environ 16 millions de mots. Il est composé de

---

<sup>2</sup> Nous voudrions remercier nos collègues Mme Bojana Mikelenić et Mme Metka Bezljaj pour leur travail toujours efficace et leurs contributions innovantes, de même que nos étudiants qui ont participé à la composition du corpus et aux recherches où il a été utilisé.

phrases alignées et il est lemmatisé et annoté morphologiquement, ce qui signifie qu'il peut donner des informations non seulement sur les traductions d'un mot, mais aussi sur l'emploi des catégories grammaticales. Il est accessible sur la plateforme Sketch Engine et, en version non lemmatisée, sur la plateforme ELRC (European Language Resource Coordination) 17, sous licence CC-BY-NC-4.0 (Bikić-Carić / Mikelenić / Bezljaj 2023). Le corpus est prévu exclusivement pour l'utilisation académique et non commerciale. Afin de protéger les droits d'auteur, l'ordre des phrases est aléatoire et il n'est pas possible de récupérer un texte dans sa totalité. Nous voudrions souligner le fait que c'est le seul corpus où sont présentés cinq langues romanes et le croate, ce qui le rend incontournable pour des recherches dans le domaine de linguistique contrastive romane en partant d'une langue slave, le croate, mais il peut être très utile aussi pour les romanistes en général, de même que pour les traducteurs ou les professionnels de l'enseignement.

Nous avons utilisé le corpus RomCro, depuis sa création, dans plusieurs recherches sur la détermination du nom (Bikić-Carić / Bezljaj 2020, Bikić-Carić 2020, Bikić-Carić 2022, Bikić-Carić / Bezljaj 2023) et sur l'emploi de l'infinitif (Bezljaj / Bikić-Carić 2021). De même, nous avons envisagé des utilisations du corpus dans la thèse de doctorat de notre doctorante Mme Metka Bezljaj et dans les mémoires de master de plusieurs de nos étudiants. Comme nous allons le montrer, nous trouvons la comparaison des phrases alignées très utile dans le but de mieux comprendre non seulement les différences et similarités entre les langues, mais aussi le fonctionnement de chacune d'entre elles.

### **3. Les articles**

Le rôle de la détermination est de désigner le mode d'introduction d'un nom dans le discours. Dans chaque langue il existe plusieurs manières de déterminer un nom, mais nous avons choisi de nous concentrer sur une catégorie grammaticale, l'article, dont c'est la fonction primaire.

Le nom, avant d'être employé dans le discours et d'être actualisé, ne représente qu'un concept et n'a pas de référent. Guillaume (1975: 89) distingue le *nom en puissance* et le *nom en effet*. Une fois le nom employé dans le discours, il faut faire le choix: ou bien le nom continue à ne pas avoir de référent et c'est un concept qu'il évoque (en général, c'est l'article zéro), ou bien le référent vient d'être introduit dans le discours et il est considéré comme « indéfini » (l'article est indéfini), ou bien le référent est considéré comme assez connu (déjà introduit dans le discours ou connu généralement) pour que le nom soit accompagné de l'article défini. Weinrich (1989: 205) explique le rôle des articles par le fait qu'il est plus facile pour le destinataire, et plus économique, d'insérer le message dans sa vision du monde s'il sait qu'il a à sa disposition, la pré-information, la post-information ou aucune des deux.

Pourtant, il existe des auteurs qui affirment, et nous sommes d'accord avec eux, que dans certaines situations l'article défini n'exprime rien, mais tout simplement permet au nom de fonctionner dans la phrase. Leeman (2004: 33), qui appelle cet emploi « intensionnel », l'explique par le fait que le déterminant n'actualise pas le nom (ne lui fait pas désigner un référent), mais permet au nom d'apparaître grammaticalement dans l'énoncé pour ne désigner que le concept (*L'eau est indispensable à la vie*).

Comme il est bien connu, l'article comme catégorie grammaticale n'existait pas en latin classique. Il s'est formé en latin vulgaire, à partir du démonstratif ILLE (ou, exceptionnellement, IPSE) et du numéral UNUS. D'après Posner (1996: 126), l'article a fait ses débuts dans les langues romanes en exprimant l'anaphore, et ce n'est que plus tard que son emploi s'est étendu sur les entités uniques, les noms abstraits, la possession inaliénable, l'emploi générique.

Voici les formes des articles dans les cinq langues romanes analysées ici :

## Les articles définis

	masculin sg.	féminin sg.	masculin pl.	féminin pl.
Portugais	o	a	os	as
Espagnol	el	la (el)	los	las
Français	le (l')	la (l')	les	les
Italien	il (l', lo)	la (l')	i (gli)	le
Roumain	-ul, -le	-a	-i	-le

L'article peut se contracter avec une préposition (*du, des, au, aux* en français), ce qui est particulièrement fréquent en portugais (p. ex. *em+o > no, por+o > pelo*). Seul le roumain ne se prête pas à la contraction pour la simple raison que l'article défini est postposé au nom (ou à l'adjectif si celui-ci précède le nom) et fonctionne comme suffixe (*profesor - profesorul*)<sup>3</sup>.

## Les articles indéfinis

	masculin sg.	féminin sg.	masculin pl.	féminin pl.
Portugais	um	uma	uns	umas
Espagnol	un	una	unos	unas
Français	un	une	des (de <sup>4</sup> )	des (de)
Italien	un (uno)	una (un')	-	-
Roumain	un	o	niște	niște

## Les articles partitifs

	masculin sg.	féminin sg.	masculin pl.	féminin pl.
Français	du	de la	-	-

<sup>3</sup> C'est un trait qui caractérise l'aire linguistique balkanique.

<sup>4</sup> La forme *de* s'emploie si l'adjectif précède le nom (p. ex. *de beaux jours*)

Italien	del (dello, dell')	della (dell')	dei (degli)	delle
---------	-----------------------	---------------	-------------	-------

En italien il est considéré que l'article indéfini n'a pas de pluriel, mais que les formes *dei*, *degli*, *delle* sont le pluriel de l'article partitif (*Enciclopedia dell'Italiano*).

Même en français, où la forme *des* (*de + les*) a remplacé *uns*, *unes* au XIII<sup>ème</sup> siècle, il est possible de trouver des points communs entre *des* et *du / de la* non seulement du point de vue morphologique, mais aussi syntaxique. D'après Leeman (2004:140), dans certains cas *des* désigne un pluriel plutôt massif que comptable, ce qui le rapproche de l'article partitif (*Cette nourriture suffit pour deux chiens, mais non pour trois chiens/\*mais non pour des chiens*). Enfin, *des* et *du / de la* sont tous deux victimes de la règle de cacophonie (*On manque de ~~du~~ lait / On manque de ~~des~~ pommes*).

#### 4. Le nom et le référent

Dans notre approche de la détermination du nom nous nous concentrerons sur le rapport entre le nom et le référent. Pour mieux expliquer nos hypothèses et conclusions, nous avons emprunté le terme *emploi prototypique* à la linguistique cognitive. En effet, Langacker (1987: 17) précise que l'appartenance d'un élément à une catégorie est traitée de manière graduelle. Plus une entité possède de traits associés au mot, plus elle a vocation à être désignée par ce mot. L'entité qui est le prototype est l'élément central d'une catégorie. Dans le glossaire de son livre *Foundations of Cognitive Grammar*, Langacker (1987: 492) définit le prototype comme l'unité qui est la plus saillante et la plus représentative de sa catégorie. En comparant les articles dans les cinq langues examinées, nous avons essayé de trouver les concepts prototypes exprimés par les articles, de même que ceux qui s'en sont éloignés ou qui représentent des particularités dans chacune des langues romanes analysées.

Les explications dans les grammaires générales consultées (*Nueva gramática de la lengua española*, *Gramática do português*, *Le bon usage*, *Gramatica limbii române*, *Enciclopedia dell'Italiano*), concernant l'emploi des articles, ont certains points communs pour toutes les langues citées : introduction d'un nouvel élément (article indéfini), reprise d'un élément connu (article défini), absence de référent (article zéro), quantité indéterminée d'un référent massif (article partitif). Nous considérons ces emplois comme prototypiques. Il faut ajouter une particularité du roumain, qui est la disparition de l'article défini si le nom est introduit par une préposition (*A ieșit din garajă* / *Il est sorti du garage*), excepté s'il s'agit de la préposition *cu* / *avec* et si le nom est suivi d'un complément déterminatif.

Voici comment nous avons classifié l'emploi des articles en fonction de la relation entre le nom et le référent, en distinguant les emplois prototypiques et ceux qui s'éloignent du prototype (Bikić-Carić, à paraître):

### I. Le référent n'est pas concret (article prototypique : article zéro)

En général cela veut dire que le nom exprime des caractéristiques d'un autre élément dans la phrase (*table de bois*). C'est l'emploi prototypique, et l'article est zéro.

Bouchard (2002: 275) souligne que dans les syntagmes nominaux du type *N de N* le deuxième nom n'a pas de déterminant (*un uniforme de général*, *une montre de dame*) parce qu'il n'a pas de référent, ce qui est très évident dans un exemple comme *la queue de cheval de Juliette* - il serait absurde de demander de quel cheval il s'agit. Guillaume l'explique par une transition incomplète du nom en puissance vers le nom en effet (Guillaume 1975: 283).

À part cet emploi prototypique de l'article zéro, où il représente le référent zéro, nous sommes d'avis qu'il y a d'autres situations, plus éloignées de ce prototype, où le référent n'est pas concret : ce sont l'emploi générique (le nom ne représente pas un référent concret, mais

tous les référents d'une espèce) - les articles prototypiques sont l'article défini (*Les lions sont des carnivores*) et l'article indéfini au singulier (*Un lion défend son territoire*) - de même que l'emploi virtuel (on ne sait pas si le référent existe), avec l'article indéfini comme prototypique (*Ils cherchent une aide-soignante qui parle espagnol et chinois*). Nous pouvons y ajouter les noms abstraits qui représentent des concepts (*L'amour est indéfinissable*), avec l'article défini comme prototypique.

II. Le référent est concret et isolé (articles prototypiques : défini ou indéfini)

Si le référent est concret et isolé comme entité, l'article donne des informations quant à son caractère connu ou inconnu (au moins pour le locuteur).

Les emplois de l'article défini pour le référent connu et l'article indéfini pour le référent inconnu sont panromans. Pourtant, nous avons remarqué certaines différences entre les langues :

a) l'adjectif possessif en français par rapport à l'article défini dans les autres langues romanes

Dans une comparaison des langues romanes (Bikić-Carić 2020), nous nous sommes rendu compte d'une présence de l'adjectif possessif plus élevée en français par rapport aux autres langues. Cette conclusion a inspiré une autre recherche (Bikić-Carić, Bezljaj 2022), où nous nous sommes concentrées sur ce détail, en comparant le français et l'espagnol (comme représentant des autres langues romanes). Dans un sous-corpus nous avons comparé les segments ayant l'article défini dans l'original (11 655 en français et 14 595 en espagnol) avec leur traduction. Les résultats ont confirmé l'hypothèse de départ : l'article défini dans l'original espagnol a été traduit par un adjectif possessif dans 225 segments, tandis que seuls 6 articles définis dans l'original français ont été traduits comme adjectifs possessifs en espagnol. Nous avons trouvé une confirmation de cette différence dans Alloa / de Torres (2005: 66).



b) l'article indéfini au pluriel est plus fréquent en français par rapport aux autres langues romanes

L'article est souvent un élément indispensable pour la grammaticalité de la phrase en français, ce qui n'est pas le cas dans les autres langues romanes. C'est surtout évident quand l'équivalent de l'article indéfini au pluriel en français est l'article zéro dans les autres langues. Teyssier compare l'article indéfini et l'article zéro avec les noms au pluriel dans plusieurs langues romanes (Teyssier 2004: 160). D'après l'auteur, les formes espagnoles et portugaises *unos, unas / uns, umas* ne sont pas vraiment des articles, mais désignent une quantité indéterminée : *unos libros, uns livros*. Il en va de même en roumain, où il est possible d'utiliser le quantitatif indéterminé *niște* : *niște cărți* ou seulement *cărți*. À la différence du français (*J'ai des livres*), en italien, bien qu'on puisse dire *ho dei libri*, il suffit de mettre le nom au pluriel : *vende libri*. En espagnol, portugais et roumain, en règle générale, le nom au pluriel suffit : *tengo libros, tenho livros, am cărți*. Quant à la forme *unos/unas* en espagnol, Sarmiento et Esparza (1993: 41) la voient comme actualisateur de deux unités ou plus, non déterminées, sans préciser la quantité (*Llegaron unos libros, pero no los miré*). Mais cette forme peut avoir le sens partitif (Sarmiento / Esparza 1993: 42), dans les contextes où on sous-entend la référence à une partie de l'ensemble (*Por favor, sírveme unas patatas*) ; le sens de quantité réduite (*Pasaremos unos días en la playa*) ; une valeur emphatique très marquée (*Te lo he dicho muchas veces: son unos vagos = muy vagos*) ; la valeur d'adverbe de quantité, signifiant « approximativement » (*Sólo ha gastado unas dos mil pesetas*). Il existe une différence sémantique entre *Compré libros* - un nombre non déterminé, et *Compré unos libros* - quelques-uns, un petit nombre de livres spécifiques.

Nous reviendrons vers cette particularité du français en examinant les différentes valeurs de l'article zéro.

## c) autres différences

Nous voudrions ajouter d'autres observations qui néanmoins ne sont pas caractéristiques d'une langue particulière. Comme nous l'avons mis en évidence (Bikić-Carić 2020), dans la comparaison d'une même phrase il est possible de détecter une approche différente de la question du référent du nom, que nous avons appelée « changement de perspective ». Dans une langue l'attention peut être attirée sur la description (*bendiciendo las aspas de un ventilador que susurraba entre lectores adormecidos*) et dans une autre sur la détermination du nom par le complément qui suit (*en benisnant les ailes du ventilateur qui bourdonnait au milieu des lecteurs endormis*). Ensuite, la différence peut résider dans le choix de l'article pour représenter l'espèce (*como un lobo huele la sangre / comme le loup flaire le sang*) ou tout simplement il peut s'agir d'un choix délibéré de la part du traducteur pour donner un autre point de vue (*dijo la voz a mi espalda / dit une voix derrière moi*).

## III. Le référent est massif

Pour exprimer une quantité indéterminée d'un référent massif, qui n'est pas une entité isolée et n'a pas de contours spécifiques, le français et l'italien ont recours à l'article partitif. Cette construction (préposition *de* en français ou *di* en italien + article défini) existait dans d'autres langues romanes, mais ce n'est qu'en français et en italien qu'elle s'est transformée en catégorie grammaticale. Pourtant, l'article partitif est plus courant en français où il peut aussi désigner une matière (*C'est du café*). Par contre, en italien son emploi est restreint et souvent facultatif: *c'è (del) riso nella dispensa?* (*Enciclopedia dell'Italiano*). Dans les autres langues romanes son équivalent est, le plus souvent, l'article zéro<sup>5</sup>.

---

<sup>5</sup> Il faut pourtant mentionner la possibilité, en roumain, de dire *Am baut bere* (*J'ai bu de la bière*), mais aussi *Am baut niște bere*, où *niște* a un sens partitif.

## 5. L'article zéro

Dans nos recherches sur les rapports entre le nom et le référent nous avons mis en évidence, comme nous l'avons fait observer plus haut, des emplois prototypiques des articles, mais aussi ceux qui s'éloignent du prototype. Cette fois-ci nous voudrions nous concentrer sur les valeurs de l'article zéro.

Bien évidemment, il n'est pas facile d'envisager une recherche informatisée où l'objectif serait de trouver un élément qui, justement, ne figure pas dans une phrase. Pourtant, en comparant la même phrase dans plusieurs langues, il s'est avéré que l'article zéro dans une langue peut avoir des équivalents dans une autre qui renseigne sur sa valeur. Comme notre but n'était pas d'obtenir des données sur le nombre d'occurrences, mais de trouver des exemples qui puissent expliquer les équivalences en question, nous avons dû réduire le nombre de segments analysés. En effet, même s'il est relativement facile de donner une consigne concernant les objectifs envisagés, tous les résultats doivent être interprétés personnellement, ce qui représente des limites. Néanmoins, nous croyons que nous avons réussi à former un sous-corpus qui offre des exemples représentatifs.

Pour les besoins de cet ouvrage, nous avons décidé, d'un côté, de trouver des exemples dans 500 segments écrits originalement en français avec leurs traductions dans les autres langues, et de l'autre côté, d'adopter la démarche inverse, à savoir partir de l'original dans une autre langue romane (en l'occurrence, 500 segments écrits en espagnol) et comparer les résultats avec la traduction en français et les autres langues. À partir de ce que nous venons de présenter, reprenons la question du rapport entre le nom et son référent :

- 1) l'article zéro a sa valeur prototypique (le nom n'a pas de référent particulier : *C'est une table de bois*)
- 2) l'article zéro remplace l'article qui a « disparu » pour des raisons qui relèvent des règles et non des cas particuliers :

- l'article défini disparaît après une préposition en roumain (*A ieșit din garajul / Il est sorti du garage*)
  - l'article partitif et l'article indéfini au pluriel disparaissent après la préposition *de* en français (*La route est couverte de de la neige ; La pelouse est couverte de des feuilles*)
- 3) l'article zéro a la valeur non prototypique de l'article défini (le nom ne représente pas un référent concret et connu)
  - 4) l'article zéro a la valeur non prototypique de l'article indéfini au singulier (le nom ne représente pas un référent concret et inconnu)
  - 5) l'article zéro a la valeur non prototypique de l'article indéfini au pluriel (les référents ne sont pas présentés comme unités isolées)
  - 6) l'article zéro a comme équivalent l'article partitif (le référent est massif, ce qui est la valeur prototypique de l'article partitif ; pourtant, comme le référent existe, ce n'est pas la valeur prototypique de l'article zéro)

Dans cette contribution nous nous concentrerons sur les valeurs non prototypiques énumérées ci-dessus, à savoir les emplois de l'article zéro dans les langues romanes examinées où son équivalent en français est l'article défini, indéfini ou partitif.

Parmi les exemples trouvés dans les segments examinés, nous en avons choisi quelques-uns qui nous paraissent représentatifs pour notre propos. Nous voudrions ajouter que nous avons dû écarter de nombreux exemples dans lesquels au moins une des cinq langues analysées présentait une traduction trop différente pour pouvoir être comparée.

Comme nous l'avons indiqué, nous nous sommes servie d'exemples originaux en espagnol ou en français (la première phrase indique la langue du texte original).

- 1) l'article zéro a comme équivalent l'article défini en français

Dans les exemples suivants, il est évident que le nom ne représente

pas un référent concret (ce qui renverrait à l'emploi prototypique de l'article défini, mais, en principe, excluait l'article zéro). Il s'agit de noms qui représentent un concept (p. ex. *la beauté* - mais, avec article zéro : *belleza, bellezza, beleza, frumusețe*) ou l'espèce (*les enterrements* - mais, avec article zéro : *entierros, funerali, enterros, înmormântări*<sup>6</sup> ; *les femmes* - mais, avec article zéro : *mujeres, donne, mulheres, femei*<sup>7</sup>).

Tout en elle respirait <i>la beauté, la jeunesse</i> et <i>la vie</i> .	Todo en ella respiraba <i>belleza, juventud</i> y <i>vida</i> .	Trasudava <i>bellezza, giovinezza</i> e <i>vita</i> .	Tudo nela respirava <i>beleza, juventude</i> e <i>vida</i> .	Era toată numai <i>frumusețe, tinerețe</i> și <i>viață</i> .
---	---	---	--	--

—Estoy harto de <i>entierros</i> , Daniel.	— Je suis fatigué <i>des enterrements</i> , Daniel.	Sono stanco di <i>funerali</i> , Daniel.	Estou farto de <i>enterros</i> , Daniel.	— Sunt sătul de <i>în mormântări</i> , Daniel.
--	---	--	--	--

—De <i>mujeres</i> , y de otros menesteres mundanos, bastante más que usted.	— J'en sais plus que vous sur <i>les femmes</i> et sur le monde.	Di <i>donne</i> e di questionii mondane, parecchio più di lei.	De <i>mulheres</i> , e de outros misteres mundanos, bastante mais que o Daniel.	— Despre <i>femei</i> , ca și despre alte treburi lumești, știu destul de mult față de dumneata.
--	--	--	---	--

## 2) l'article zéro a comme équivalent l'article indéfini au singulier en français

Même si la fréquence de l'article indéfini en français par rapport aux

<sup>6</sup> Notons, toutefois, qu'en roumain l'article défini est exclu puisque le nom est introduit par une préposition.

<sup>7</sup> Idem.

autres langues romanes est plus évidente au pluriel qu'au singulier, nous sommes d'avis que dans les deux cas il s'agit d'une tendance du français à spécifier un référent comme unité, tandis que dans les autres langues c'est plutôt un concept sans contours spécifiques (*une femme, un trésor, un trésorier* - mais, avec article zéro : *mujer, tesoro, tesorero / donna, tesoro, tesoriere / mulher / tesouro / tesoureiro / femeie, tezaur, trezorier*). Par conséquent, ce nom peut représenter une caractéristique du nom principal (*gueule d'un zombie* - mais, avec article zéro : *cara de zombi / faccia da zombie / cara de mortovivo / mutră de zombie*). La différence entre un référent conçu comme une unité en français et comme un concept dans les autres langues est aussi évidente dans l'exemple suivant : *une idée* - mais avec article zéro : *idea / idea / ideia / idee*.

Penélope, la preciosa Penélope, era <i>mujer</i> y por tanto <i>tesoro</i> , no <i>tesorero</i> .	Penélope, l'adorable Penélope, était <u>une</u> <i>femme</i> , donc <u>un</u> <i>trésor</i> , mais <u>un</u> <i>trésor</i> ne fait pas <u>un</u> <i>trésorier</i> .	Penélope, la bella Penélope, era <i>donna</i> e pertanto <i>tesoro</i> , non <i>tesoriere</i> .	A Penélope, a bela Penélope, era <i>mulher</i> e portanto <i>tesouro</i> , não <i>tesoureiro</i> .	Penelope, frumoasa Penelope, era <i>femeie</i> și, ca atare, <i>tezaur</i> , iar nu <i>trezorier</i> .
--	--	---	--	--

— ¿Tiene <i>idea</i> de quién era ese hombre?	— Avez-vous <u>une</u> <i>idée</i> de l'identité de cet homme ?	«Ha <i>idea</i> di chi fosse quell'uomo?»	Tem <i>ideia</i> de quem era esse homem?	— Aveți <i>idee</i> cine era omul ăsta?
---	--	---	--	--

A midi, il avait rendez-vous avec Angie et il ne tenait pas à débarquer devant sa fille avec la <i>gueule d'un zombie</i> .	Había quedado a mediodía con Angie y no quería presentarse ante su hija con <i>cara de zombi</i> .	A mezzogiorno aveva appuntamento con Angie e non voleva presentarsi con la <i>faccia da zombie</i> .	Ao meio-dia, tinha um encontro marcado com Angie e não queria aparecer à frente da filha com uma <i>cara de morto-vivo</i> .	La prânz urma să se întâlnească cu Angie și n-ar fi vrut să apară în fața fetei cu o <i>mutră de zombie</i> .
---	--	--	--	---

### 3) l'article zéro a comme équivalent l'article indéfini au pluriel en français

Comme nous l'avons fait observer, l'article *des* partage quelques traits avec l'article *du / de la*. Nous pouvons y ajouter un autre point commun : ils ont tous les deux l'article zéro comme équivalent fréquent dans les autres langues romanes analysées ici. Nous y voyons une différence entre la présentation des référents comme unités isolées (avec article indéfini) ou comme un ensemble (avec article zéro). Voici quelques-uns des très nombreux exemples où l'article indéfini au pluriel ne se rencontre qu'en français (même dans les cas où le nom a la fonction de sujet et qu'il se trouve à sa place habituelle) :

L'endroit était convivial avec <i>des bougies</i> partout, <i>des coussins multicolores</i> et quelques bibelots hindous.	Era un sitio acogedor, con <i>velas</i> por todas partes, <i>cojines multicolores</i> y algunos objetos indios.	Era un appartamento accogliente, costellato di <i>candele e cuscini variopinti</i> e rallegrato da qualche soprammobile indiano.	O local era agradável, com <i>velas</i> por toda a parte, <i>almofadas multicolores</i> e alguma estatuária hindu.	Locul era foarte primitiv, plin de <i>lumânări decorative, pernuțe multicolore</i> și bibelouri indiene.
---	---	--	--	--

Un petit bébé dans la maison et tout est chamboulé : partout <i>des biberons</i> , <i>des paquets</i> de couches, du lait deuxième âge.	Un bébé en casa y todo acaba patas arriba: por todas partes <i>biberones</i> , <i>paquetes</i> de pañales, leche de continuación.	Con una neonata in casa, il disordine è garantito: dappertutto <i>biberon</i> , <i>pacchi</i> di pannolini, scatole di latte in polvere.	Um pequeno bebé na casa e tudo se altera e complica: por toda a parte há <i>biberões</i> , <i>pacotes</i> de fraldas, leite para bebés.	Apariția unui bebeluș întoarce toată casa cu fundul în sus: peste tot sunt întinse <i>biberoane</i> , <i>pachete</i> de scutece, cutii de lapte pentru sugari.
---	---	--	---	--

—En tiempos de guerra ocurren cosas que son muy difíciles de explicar, Daniel.s	— En temps de guerre, il se passe <i>des choses</i> qui sont très difficiles à expliquer.	«In tempo di guerra accadono cose orribili, Daniel, molto difficili da spiegare.	Em tempo de guerra acontecem coisas que são muito difíceis de explicar, Daniel.	— În vreme de război se petrec <i>lucruri</i> foarte greu de explicat, Daniel.
---	---	--	---	--

Quizá. Aquí, como ve, sólo vendemos <i>libros</i> .	Ici, comme vous voyez, nous vendons seulement <i>des livres</i> .	Qui, come vede, ci limitiamo a vendere <i>libri</i> .»	Talvez. Aqui, como vê, só vendemos <i>livros</i> .	Poate. Aici, precum vedeți, vindem numai <i>cărți</i> .
---	---	--	--	---

<i>Simios</i> es lo que llegan a las aulas.	<i>Des singes</i> , oui, voilà ce que j'ai dans mes cours.	<i>Scimmie</i> , in classe non ho altro che <i>scimmie</i> .	<i>Símios</i> , é o que me aparece nas aulas.	Numai <i>maimuțe</i> ajung în aulă.
---	--	--	---	-------------------------------------

Nosotros somos <i>amigos</i> de Julián.	Nous sommes <i>des amis</i> de Julián.	Noi siamo <i>amici</i> di Julián.	Nós somos <i>amigos</i> do Julián.	Noi suntem <i>prieteni</i> cu Julián.
---	--	-----------------------------------	------------------------------------	---------------------------------------



Pour résumer, l'article zéro comme équivalent de l'article défini au singulier et au pluriel, de même que de l'article indéfini au singulier, indique des emplois non prototypiques de ceux-ci, qui se résument à la non-existence d'un référent concret, soit qu'il s'agisse d'un concept (*la beauté ; un trésor*) ou d'une espèce (*les femmes ; une femme*). De plus, l'article zéro comme équivalent de l'article indéfini au singulier et au pluriel peut signifier que le nom représente une caractéristique d'un autre élément de la phrase (*gueule d'un zombie / cara de zombi ; nous sommes des amis de Julián / somos amigos de Julián*), même si, en français, l'article indéfini concrétise les référents de ces noms.

#### 4) l'article zéro a la valeur de l'article partitif

Bien évidemment, dans les langues où l'article partitif n'existe pas, l'équivalent est l'article zéro. Mais même en italien, où il est possible de trouver des emplois avec article partitif, l'article zéro est plus courant. Voici quelques exemples :

Trouver <i>de</i> <i>la</i> <i>drogue</i> à San Francisco n'était pas très compliqué.	Encontrar <i>droga</i> en San Francisco no era muy difícil.	Trovare <i>droga</i> a San Francisco non era molto difficile.	Encontrar <i>droga</i> em São Francisco não era muito difícil.	La San Francisco nu era foarte greu să găsești <i>droguri</i> .
Avec <i>de</i> <i>la</i> <i>chantilly</i> !	Con <i>nata</i> .	"Sì, con la <i>panna</i> ", risponde lei.	Com <i>chantilly</i> .	Cu <i>cremă</i> <i>chantilly</i> !
Nunca me decía adónde iba. Casi siempre traía <i>dinero</i> o <i>joyas</i> .	Il ne me disait jamais où il allait. Il rapportait presque toujours <i>de</i> <i>l'argent</i> ou des bijoux.	Non mi diceva mai dove andava e tornava quasi sempre con <i>denaro</i> o gioielli.	Nunca me dizia onde ia. Trazia quase sempre <i>dinheiro</i> ou <i>jóias</i> .	Nu-mi spunea niciodată unde se ducea. Mai întotdeauna aducea <i>bani</i> sau bijuterii.

## 6. Conclusion

À partir de ce que nous venons de présenter, voici quelques conclusions :

### I. Valeurs prototypiques et non prototypiques

I.a) L'article zéro exprime, comme valeur prototypique, que le nom n'a pas de référent concret

Nos exemples montrent qu'il est possible d'y inclure l'expression d'un concept (*belleza; mujer, tesoro*), qui pourtant a l'article défini (*la beauté*) ou indéfini (*une femme, un trésor*) en français. De même, l'expression d'une caractéristique du nom principal peut être considérée comme valeur prototypique de l'article zéro (*de zombi*), bien que le français emploie un article indéfini (*d'un zombie*). Même l'expression de l'espèce peut se rapprocher de l'emploi prototypique, puisqu'il n'y a pas de référent concret (*mujeres ; les femmes*).

Ces emplois confirment notre classification de départ, où nous avons ajouté, dans la rubrique sur le nom qui n'a pas de référent concret, l'espèce (même si en général l'article zéro n'est pas associé à son expression) et les noms abstraits qui expriment des concepts (*belleza, bellezza, beleza, frumusețe*).

Ce sont des valeurs prototypiques de l'article zéro, mais des valeurs non prototypiques de ses équivalents en français (articles défini ou indéfini).

I.b) L'article zéro est employé pour exprimer des valeurs qui ne lui sont pas prototypiques

L'article zéro remplace souvent l'article partitif en français, ce qui, à notre avis, s'éloigne de son emploi prototypique, où il signifie qu'un nom n'a pas de référent concret. En effet, un référent massif, même s'il n'a pas de contours et ne représente pas des unités isolées, peut être considéré comme un référent concret et non comme un concept.

## II. Les différents emplois des articles montrent les différentes approches du référent

Comment peut-on comparer *Leo libros* et *Bebo café* ? Le français distingue bien *Je lis des livres* et *Je bois du café*, à savoir des référents distincts et un référent massif. Pourtant, dans les autres langues romanes ce n'est pas le cas. Dans ces langues cette différence s'estompe, au profit de la non-particularisation du référent (les livres ne sont pas précisés, le café - et sa quantité - non plus). Cela ressort aussi dans d'autres exemples de notre recherche, comme *des coussins multicolores* (représenté comme unités isolées) et *cojines multicolores* (représenté comme un ensemble).

De même, nous trouvons que la différence entre *Nous sommes des amis de Julián* et *Nosotros somos amigos de Julián* réside elle aussi dans une approche différente : dans les langues avec l'article zéro dans cette construction, *amigos / amici / amigos / prieteni* représente plutôt une caractéristique du sujet, tandis qu'en français (*des amis*) le référent est concrétisé. Un autre exemple similaire est *Des singes, oui, voilà ce que j'ai dans mes cours* (mais : *Simios / Scimmie / Simios / maimuțe*).

Ajoutons que l'absence de l'article défini ou indéfini a été plus fréquente dans les stades antérieurs des langues romanes, ce qui peut signifier que le français est allé plus loin dans la diversification des informations données sur le rapport entre le nom et le référent. Le français a, semble-t-il, renforcé l'emploi de l'article, puisque dans ces situations, à la différence de ce qui a lieu dans les autres langues romanes citées, l'article devient souvent un élément indispensable pour qu'une phrase soit grammaticalement correcte. Mais ces différences sont possibles parce que le nom en question peut être vu sous différents angles et porter ainsi des valeurs différentes. Ces quelques traits, à notre avis, passent en général plutôt inaperçus si nous nous bornons à l'étude d'une seule langue romane.



## 7. Bibliographie (y compris communications aux colloques):

Alloa, Hugo / de Torres, Silvia Miranda (2005). *Hacia una lingüística contrastiva: francés – español*, Córdoba: Comunic-arte.

Bezljaj, Metka / Bikić-Carić, Gorana (2021). Le choix entre l'infinif et une forme conjuguée après les verbes d'opinion dans cinq langues romanes, communication au colloque *Considérations philologiques en contexte français et francophone*, Skopje, du 19 au 20 novembre 2021 (article soumis à l'évaluation)

Bikić-Carić, Gorana (2020). Quelques particularités dans l'expression de la détermination du nom - comparaison entre cinq langues romanes, in: *Recherches sur les langues romanes et leurs variétés, număr special Studia Universitatis Babeş-Bolyai* [coord. Adrian Chircu], *Philologia*, 65 (LXV), N° 4, Cluj-Napoca, Universitatea Babeş-Bolyai, pp. 39-54.

Bikić-Carić, Gorana (2022). Le corpus parallèle RomCro et ses applications, communication au colloque *XX<sup>e</sup> rencontre internationale des romanisants « Défis des études romanes dans la troisième décennie du XXI<sup>e</sup> siècle »*, Université Palacký d'Olomouc, du 25 au 26 novembre 2022.

Bikić-Carić, Gorana (à paraître). *Zanimljivosti o romanskim jezicima u dijakroniji i sinkroniji (uz usporedbu s hrvatskim)*, Zagreb: FF Open Press.

Bikić-Carić, Gorana / Bezljaj, Metka (2020). Construcción de un corpus multilingüe y su aplicación en el análisis contrastivo de los artículos, communication au colloque *XLIX Simposio de la Sociedad Española de Lingüística*, Universitat Rovira i Virgili, Tarragona, 21-24 de enero de 2020.

Bikić-Carić, Gorana / Bezljaj, Metka (2023). Neke specifičnosti upotrebe određenog člana u romanskim jezicima (s posebnim naglaskom na francuski i španjolski), in: *70 ans d'études romanes culturelles, linguistiques et littéraires à l'Université de Sarajevo* [coord. Edina Spahić et al.], Sarajevo: Univerzitet u Sarajevu - Filozofski

fakultet, pp. 15-27.

Bikić-Carić, Gorana / Mikelenić, Bojana / Bezljaj, Metka (2023). Construcción del RomCro, un corpus paralelo multilingüe. *Procesamiento del Lenguaje Natural*, 70. Sociedad Española para el Procesamiento del Lenguaje Natural, pp. 99-110.

Bouchard, Denis (2002). *Adjectives, Number and Interfaces - Why Languages Vary*, Amsterdam: Elsevier.

*Enciclopedia dell'Italiano* [https://www.treccani.it/enciclopedia/frasi-completive\\_\(Enciclopedia-dell'Italiano\)](https://www.treccani.it/enciclopedia/frasi-completive_(Enciclopedia-dell'Italiano)) (consulté le 12 mars 2023)

Grevisse, Maurice / Goosse, André (2008). *Le bon usage*, Bruxelles : Éditions De Boeck Université.

Guillaume, Gustave (1975). *Le problème de l'article et sa solution dans la langue française*, Paris: Librairie A.-G. Nizet, Québec: Les Presses de l'Université Laval.

Guțu-Romalo, Valeria et al. (2008). *Gramatica limbii române*, București: Editura Academiei Române.

Langacker, Ronald Wayne (1987). *Foundations of Cognitive Grammar, Vol. I.*, Stanford: Stanford University Press.

Leeman, Danielle (2004). *Les déterminants du nom en français : syntaxe et sémantique*, Paris: Presses Universitaires de France.

Posner, Rebecca (1996). *The Romance Languages*, Cambridge: Cambridge University Press.

Raposo, Eduardo Paiva et al. (2013). *Gramática do português*, Lisboa: Fundação Calouste Gulbenkian.

Real Academia Española / Asociación de Academias de la Lengua Española (2009). *Nueva gramática de la lengua española*, Madrid: Espasa Libros.

Sarmiento, Ramón / Esparza, Miguel Angel (1993). *Los determinantes*, Madrid: Sociedad General Española de Librería.

Teyssier, Paul (2004). *Comprendre les langues romanes*, Paris: Chandeigne.

Weinrich, Harald (1989). *Grammaire textuelle du français*, Paris: Alliance Française, Didier/Hatier.



## Primjene paralelnog korpusa RomCro u kontrastivnim istraživanjima - vrijednosti nultog člana u romanskim jezicima

U ovom se članku istražuju prototipne i neprototipne uporabe članova u pet romanskih jezika (francuskom, španjolskom, portugalskom, talijanskom i rumunjskom). Posebno nas zanima odnos između nultog člana u španjolskom, portugalskom, talijanskom i rumunjskom s jedne strane, te njegovih ekvivalenata u francuskom (određenog, neodređenog i partitivnog člana). Budući da u kontrastivnim istraživanjima važnu ulogu imaju korpusi, opisujemo paralelni računalni korpus RomCro, koji stvaramo od 2019., a sastoji se od izvornih tekstova i njihovih prijevoda na pet navedenih romanskih jezika i hrvatskom. Dijelom tog korpusa poslužili smo se i u ovom istraživanju, gdje uspoređujemo odnose između imenice i referenta s obzirom na to kako ih izražava član. Navodimo prototipne članove za svaki od tih odnosa (referent nije konkretan - prototipni član je nulti; referent je konkretan i izdvojen - prototipni članovi određeni, odnosno neodređeni, referent je masivan - u francuskom i talijanskom prototipni član je partitivni). U skladu s tim istaknuli smo razlike između francuskog i ostalih navedenih romanskih jezika. Na temelju usporedbi odgovarajućih primjera iz korpusa, naveli smo nekoliko

zaključaka: nulti član izražava svoju prototipnu vrijednost (referent nije konkretan), što pretpostavlja neprototipne vrijednosti određenog člana i neodređenog člana u jednini kao njegovih ekvivalenata u francuskom; nulti član izražava svoju neprototipnu vrijednost ako mu je ekvivalent u francuskom partitivni član; nulti član kao ekvivalent neodređenom članu u množini u francuskom izražava različite pristupe referentima (koje se u francuskom prikazuje kao izdvojene jedinice, a u ostalim jezicima kao cjelinu). Ovi nas rezultati usporedbe navode na zaključak kako su te razlike moguće zahvaljujući tome što imenica u sebi sadržava više vrijednosti, od kojih se s pomoću člana istakne samo jedna.

Ključne riječi: kontrastivna lingvistika, gramatički član, (ne)prototipna uporaba, romanski jezici, paralelni korpus RomCro

## **Applications of the parallel corpus RomCro in contrastive researches - values of zero article in Romance languages**

This article investigates the prototypical and non-prototypical uses of articles in five Romance languages (French, Spanish, Portuguese, Italian and Romanian). We are particularly interested in the relationship between the zero article in Spanish, Portuguese, Italian and Romanian on the one hand, and its equivalents in French (definite, indefinite and partitive article). Since corpora play an important role in contrastive research, we describe the parallel corpus RomCro, which we have been creating since 2019, and which consists of original texts and their translations into the five mentioned Romance languages and Croatian. We used a part of that corpus in this research, where we compare the relations between the noun and the referent with regard to how they are expressed by the article. We list the prototypical articles for each of these relations (the referent is not concrete - the prototypical article is zero; the referent is concrete and isolated - the prototypical articles are definite or indefinite, the referent is massive - in French and Italian

the prototypical article is partitive). Accordingly, we have highlighted the differences between French and the other mentioned Romance languages. Based on the comparison of the corresponding examples from the corpus, we have stated several conclusions: the zero article expresses its prototypical value (the referent is not concrete), which assumes the non-prototypical values of the definite article and the singular indefinite article as its equivalents in French; the zero article expresses its non-prototypical value if its equivalent in French is the partitive article; the zero article as the equivalent of the indefinite article in plural in French expresses different approaches to referents (which are presented as separate individuals in French, and as a whole in other languages). The results of these comparisons lead us to the conclusion that these differences are possible thanks to the fact that the noun contains multiple values, of which only one is highlighted using the article.

Key words: contrastive linguistics, grammatical article, (non-) prototypical use, Romance languages, parallel corpus RomCro