Arbeitspapier Nr. 127

Proceedings of the VI Nereus International Workshop
“Theoretical implications at the syntax/semantics interface in Romance”

Sofiana Chiriacescu (ed.)
Proceedings of the VI Nereus International Workshop sponsored by:

Universität zu Köln

The publication of these proceedings is being funded by:

Universität Konstanz
Proceedings of the VI NEREUS INTERNATIONAL WORKSHOP: THEORETICAL IMPLICATIONS AT THE SYNTAX/SEMANTICS INTERFACE IN ROMANCE

SOFIANA CHIRIACESCU (ED.)
Table of contents

Preface

Martin Becker (Universität zu Köln)
Competing zones: The zero article in Old French and its loss in the history of the French language ................................................................. 1

M. Teresa Espinal (Universitat Autònoma de Barcelona) / Sonia Cyrino (Universidade Estadual de Campinas)
Preverbal bare nominals in Brazilian Portuguese ........................................ 19

Susann Fischer (Universität Hamburg)
Unaccusatives, existentials and the definiteness restriction: Explorations of the syntax-semantic interface .................................................. 33

Chiara Gianollo (Universität zu Köln)
Latin aliquis as an epistemic indefinite.......................................................... 55

Klaus von Heusinger (Universität zu Köln) / Sofiana Chiriacescu (Universität zu Köln)
Discourse effects of differential object marking in Romanian ...................... 83

Manuel Leonetti (Universidad de Alcalá)
On contrastive readings in the interpretation of NPs/DPs .......................... 99

Sarah Schwellenbach (Universität zu Köln)
The syntax-semantics interface of avertive and proximative in Romance .... 117

Melanie Uth (Universität zu Köln)
On the interpretation of preverbal subjects in Spanish in contexts of narrow information focus ................................................................. 135
Preface

This volume is a collection of papers presented at the VI NEREUS (http://gerlin.phil-fak.uni-koeln.de/kvh/forschung/Nereus/) international workshop on the topic *Theoretical Implications at the Syntax-Semantics Interface in Romance*, held at the University of Köln on November 9th and 10th, 2012.

Language is a system of communication in which grammatical structures function to express meaning in context. While all languages can achieve the same basic communicative ends, they each use different means to achieve them, particularly in the divergent ways that syntax, semantics and pragmatics interact across languages. The sixth NEREUS international workshop brought together experts in the fields of syntax and semantics, who explore the interaction between form, meaning, and communicative function in Romance languages from both a diachronic and a synchronic perspective. The aim of the conference was to describe the division of labour between syntax and semantics and the nature of the interface between them. The languages discussed ranged from Catalan and Spanish over French and Brazilian Portuguese to a multitude of different Italo-Romance varieties.

Although some authors present at the workshop could not include their presentation in this volume, we would like to thank all the contributors and participants at the workshop for their talks and interesting discussion.

We would like to particularly thank the University of Cologne and the “Cologne Center for Language Sciences (CCLS)” for contributing to the success of the workshop. Finally, many thanks to Stephanie Eßer for her editing assistance.

Köln, November 11th 2013

Sofiana Chiriacescu
Competing zones: The zero article in Old French and its loss in the history of the French language

Martin Becker
martin.becker1<at>uni-koeln.de

1. Introduction

In this article, I will discuss the development of the zero article from Old French to modern French with particular focus on its semantic functions and, especially, on its underlying abstract “value”. I will begin with an overview of the residual contexts of occurrence of the zero article in contemporary French. I will then provide an outline of how the zero article worked in the Old French system of determination. My aim is not purely descriptive, however: I want to make a case for a unitary semantic account of the “value” of the zero article, which has to date only been attempted by Gustave Guillaume in his very idiosyncratic framework of “psycho-mechanic theory” (see Guillaume 1975).

In the second part of my article, I would like to address some developments from Late Old French to early modern French which gave rise to a profound restructuring of the “ideal” medieval determiner system. I will therefore focus on the domains of overlap and competition between the zero article and the ascendant determiners, especially the definite and indefinite article as well as the nascent partitive article. Finally, in an empirical section, I will identify some trends in the changes observed, as well as the underlying principles that can be established from the historical texts of the Frantext corpus, comprising the period from Middle to early modern French. A short summary will round off my contribution.

2. The zero article in contemporary French

The linguistic standard grammar of contemporary French, La grammaire méthodique du français edited by Riegel et al. (1999: 163-167), provides a detailed overview of the remaining contexts in which the zero article still prevails. The most important contexts listed by Riegel et al. are as follows:

- with proper nouns:

  (1) Jean achète des livres.
      John buy.1SG books
      ‘John buys books.’

- in coordinative structures, but only

  a) in exhaustive (“totalizing”) coordinations comprising all members of a category: parents et enfants (‘parents and children’), civils et militaires (‘civilians and soldiers’) etc. – usually the members of the coordination constitute converse antonymies of role pairs occurring in frames such as ‘parenting’, ‘citizenship’ and the like.
2 Competing zones: The zero article in Old French and its loss in the history of the French language

b) in identifying coordinations:

(2) Mon collègue et néanmoins ami Presskopf
‘My colleague and nevertheless friend Presskopf’

- in vocatives:

(3) Garçon, un demi!
‘Waiter, a half!’

- in appositions with a facultative status:

(4) Gérard, (un) excellent français, s’engagea dans l’armée.
‘Gerard, (an) excellent Frenchman served in the army.’

- frequently in predicative complements:

a) in subject complements

(5) Je suis médecin/étudiant.
‘I am a doctor/a student.’

as well as

b) in object complements:

(6) On l’a élu président.
‘He was elected president.’

- in noun complements with attributive function:

(7) bijou en or (‘jewellery made of gold’), un coiffeur pour dames (‘a ladies’ hairdresser’), un poste de professeur (‘a post as a teacher’), un homme sans scrupules (‘an unscrupulous man’).

It should be noted that some prepositional clauses lack a determiner, especially when they function as manner adverbials; the zero article is particularly common with the prepositions sans (‘without’), avant (‘before’), après (‘after’) and so on; however, most of these adverbial expressions are highly lexicalized, as for instance: aller à pied (‘go on foot’), avancer avec lenteur (‘go slowly’) and sur terre (‘on earth’).

Moreover non-determination is characteristic of verb-object collocations which are partly or fully fixed verbal expressions. In particular functional verbs with a generic meaning such as faire (‘to do’), rendre (‘to give back’) and prendre (‘to take’) form strongly lexicalized patterns with non-determined noun bases, for example: faire peur (‘to inspire fear’), rendre hommage (‘to pay respect to’), prendre place (‘take place’).

This general picture is fully in line with stage III of Greenberg’s grammaticalization scale for articles (see Greenberg 1978:62ff.): Apart from the high degree of lexicalization of the zero article, this stage is characterized by the fact that bare plurals are no longer available in modern French. This aspect is one of the typological differences that distinguish modern French from other Romance languages. Compare the following examples:
We can conclude that the articles in modern French are highly grammaticalized as they act as noun markers in all kinds of syntactic environments. In particular, they indicate the morphological features of gender and number: the latter is all the more important in view of the fact that the plural suffix -s is no longer phonetically perceptible in modern French (in fact, it ceased to be pronounced in late Old French). In modern French, the article ceased to indicate the referential status of the noun and turned into a marker of predetermination which has the primary role of encoding morphological information.

3. The Old French system

3.1 The zero article in Old French

In contrast to modern French, the zero article was a key element of the Old French system of determination. Before providing a semantic account of the zero article, I will first discuss its functional role in Old French in greater detail.

A survey of typical zero-article contexts is provided by Moignet (1976) and, more recently, by Claude Buridant (2000) in his Grammaire nouvelle de l’ancien français. However, language historians classify the different occurrences by falling back on textual categories or discourse traditions as they put forward explanatory concepts such as definition, proverb, maxim, and so on in order to account for the use of the zero article. However, the different readings may in fact be subsumed under more general semantic labels: The zero article plays a crucial role when it comes to signalling “generic reference”. In contexts of generic reference, bare nominal phrases denote a particular kind as follows:

a) Kinds are the subject of definitions, but they are also evoked in proverbs or maxims. In all these contexts, the predications revolve around a kind whose typical properties are focused. A case in point is example (9) – a definition:

(9) Cocodrille est uns animaus a .iiiij. piez et de jaune color  
    Crocodile be.3SG an animal with 4 feet and of yellow colour  
    ‘The crocodile is an animal with 4 feet and yellow colour.’  
    (Tresor, V, Buridant 2000: 110)

b) Another interesting example quoted by Buridant is a comparison (see example (10)). In this example, the property of swiftness is associated with its prototypical bearers, the sparrowhawk and the swallow.

(10) Plus est isnels que esprever nè arunde  
    More is quick than sparrowhawk or swallow  
    ‘Quicker than a sparrowhawk or a swallow.’  
    (Roland, 1492, Buridant 2000: 110)
c) In classifications, as in example (11), a referent is classified as an instance of a category or a kind. The speaker identifies the subject with the woman-kind:

(11) \textit{Bien i pert que vos estes fame, Qui se corroce quant ele ot/Nelui qui bien feire li lot}

'It becomes apparent that you are a woman given that a woman is angry when she notices that someone gives her wise advice.'

(Yvain, 1654-56, Buridant 2000: 110)

A second basic value of the zero article in Old French is the marking of non-specific reference: The typical environment of the NP is modal which means that the NP is in the scope of a modal operator. Particularly prominent operators are the

- the negation operator (=NEG-operator):

(12) \textit{Dex ne fist terre qui enves lui n’apende}

‘God does not create a world which does not depend on him (his will).’

(CourL., AB, 17, Buridant 2000: 111)

- the conditional operator IF/SI:

(13) \textit{Et s’il est hom qui li face nul tort}

‘And if there is someone who causes him whatever damage.’

(CourL, AB, 22, Buridant 2000: 111)

- the interrogative operator (Q-operator):

(14) \textit{A dons soz ciel ne roi ne conte Que eüst an ma fille honte ...?}

‘Is there a king or a count under the sun who would feel ashamed of my daughter?’

(Erec, 533-34, Burdiant 2000: 112)

Another frequent intensional context is constituted by subjunctive relative clauses which display the form \(N_i + que + \pro_{N_i} + V_{(subj)}\). This very typical subjunctive acts as an operator which combines all-quantification with domain-widening and has the meaning of “whatever \(N\)”. The semantics of this expression can be formalized as follows:

for all \(x \in [[N]]^w\) and \(w \in \MB_{\text{epistemic}}(w_0, w)\): \(P(x)\).

The formula indicates that the predication holds for all \(x\)-individuals not only in the real world (in this case, the relevant domain of quantification, would have been \(w_0\)), but in all epistemically accessible (= imaginable) worlds:

(15) \textit{Arme qu’il ait ne vaut un paresis}

‘Whatever kind of army might exist, it is not worth a penny.’

(ChevalerieV, C, 638, Buridant 2000:111)
Finally, another salient operator is the negative polarity item *onques* which also displays a domain-widening effect: *onques + negation (ne): ‘ever’ in all possible worlds ($w \in MB_{\text{epistemic}}$):*

\begin{equation}
\text{car onques hom de son age ne fu plus amês}
\end{equation}

given that ever human being of his age not be.3.PRS.IND. more loved
de ses homes
‘given that absolutely none of his age was ever more beloved by his fellow men’

(CdC 37, Heinz 1982: 141)

The zero article can also be found with undetermined plural nouns indicating an indefinite quantity of individuals as in the following example:

\begin{equation}
Tant i fui que j’oï venir Chevaliers, ce me fu avis;
\end{equation}

So often there be.1.PST that I hear.1.PST come knights, this me be.3.PST opinion
‘I was always there when I heard the arrival of knights.’

(Yvain, 478-479, Burdiant 2000: 115)

The zero article also appears with undetermined mass-denoting nouns: in these contexts, it profiles an indefinite quantity withdrawn from a non-countable entity such as water, wine, bread, beer and so on:

\begin{equation}
Si mangierent pain et burent cervoise
\end{equation}

Then eat.3.PST bread and drink.3.PST beer
‘The ate bread and drank beer.’

(Queste, 211, 5, Moignet 1976: 118)

The zero article is very common in noun phrases with a unique reference:

a) The nouns phrases can denote natural phenomena: the earth, the sun, the moon, God, the devil and so forth.

\begin{equation}
il sembla que terre fondist.
\end{equation}

it seem.3.PRS.IND that earth melt.3.IMP.SUBJ
‘it seemed as if the earth melted’

(CdC, 28, Heinz 1982: 219)

b) finally, it is the norm for proper names:

(20) Roland, Oliver, Charles

To complete the picture, we should mention the lack of determination in several other contexts:

- vocatives:

(21) Sire cumpaninz, amis, nel dire ja.
‘Sir, companion, friend’

(Rol. 1113, Moignet 1976: 109)
- lexicalized verb-noun-collocations:

(22) avoir foi, avoir paor, faire folie, faire honte, faire mal, faire retor, espousier fame (MiraclesND, 3, I, 19), dire folie/verité, oïr messe/escolter messe (Partonopeu, 8318)
‘to have faith’, ‘to be afraid of’, ‘to feel ashamed’, ‘to do harm’, ‘to return’, ‘to marry a woman’, ‘to tell stupidities/the truth’, ‘to listen to a mass’

- in idiomatic prepositional expressions like venir a cort, aler a cort, partir de cort, estre a cort (Mort Artu), where ‘court’ is a stereotypical place-role in a medieval frame (something like ‘social centre of kings and knights’);

(23) venir a cort, aler a cort, partir de cort, estre a cort (Mort Artu)
‘come to the court’, ‘go to the court’, ‘leave the court’, ‘be at the court’

- with idiomatic prepositional adverbials:

(24) a droit, a tort, de gré, par amor, par foi
‘unjustly, on purpose, for love, by faith’

With this general picture in mind, we can now turn to the semantics of the zero article in Old French.

3.2 An account of the semantics of the zero article in Old French

This section seeks to address the key question of whether it is possible to pinpoint a basic and all-encompassing abstract meaning of the zero article.

The zero-article can be conceived as a paradigmatic option (or value) within the paradigm of determiners consisting of items such as the definite article, the indefinite article, the partitive article – and the zero article. The functions of these items within the determiner system are distributed in a different way in each stage of the language system. The system that we are able to reconstruct on the basis of the texts from the 12th and 13th century is a snapshot of a transitional period between an “ideal” system A and a nascent system B. We will begin by analysing the semantics of the zero article with regard to the “ideal” Old French system.

Following Heim (1990), we can firstly say that a predicate expression such as ‘horse’ comes out of the lexicon with a situation (or, alternatively, with a world variable w) and has the meaning of ‘to be a horse in a world w or at a world-time index I’. (Note that for our purposes no distinction between worlds and situations is required). This idea is captured by the following formula:

\[ [[\text{horse}]] = \lambda x. \lambda w. \text{horse}(x)(w) \]

This representation of the predicative noun ‘horse’ corresponds to its intension which can be defined as a function from worlds (indexes) into sets of individuals. My claim is that the semantics of the zero article can be described as a function which takes a noun and yields its intension. As already mentioned, the intension of a predicative noun denotes the set of individuals in each world to whom the predicate applies. With this general characterization in mind, I will now undertake a closer examination of the different readings of the zero article in the original French system proposed so far:

In definitions, comparisons and other generic predications the undetermined noun refers to kinds or prototypical properties. As Chierchia (1998: 348ff.) points out there is a correspondence between kinds “generally seen as regularities that occur in nature” (ibid.) and the property of being N (for instance, the property of being a horse) and belonging to the N-
(e.g. the horse)-kind (see also Chierchia 2008: 452).

In definitions such as “cocodrille est” – see example (9) – the speaker refers to the crocodile-concept (the kind of being a crocodile or the crocodile property) which is defined taxonomically within a conceptual hierarchy (est uns animaus) and by evoking the typical (necessary and sufficient) properties, e.g. the colour of the skin and other definitional features. In comparisons like “swifter than a swallow”, a prototypical property of the kind-concept is evoked in order to highlight the maximal degree of a certain property under scrutiny. Finally, in classifications as in example (11), “X is N”, the noun X is attributed the property of being N (“to be a woman”).

This description can also be captured by a more Montague-style analysis. We depart from the idea that in all these contexts the zero-article acts a function from worlds (or world-time-indexes) into sets of individuals. This function is instantiated by the context that determines the world (or world-time) indexes relevant for the evaluation of the NP. We apply this basic insight to the different contexts of occurrence:

- Definitions describe the typical instances of a category and are therefore propositions in the scope of the genericity operator GEN. The genericity operator is a function which only picks up those worlds in which the individuals, the members or instances of a category, appear in their typical realization, according to the definitional properties of that category. The instantiation of the underlying proposition at these “typical” indexes is provided by elements of the context, and principally by the predicates of a definition. These predicates allow for a kind-level reading as in the example of “to be an animal”, “to whinny” or “to be extinct”.

- Comparisons are based on the prototypical properties of categories. The prototypical “sparrowhawk” of example (10) is not found in all worlds - in some, sparrowhawks may have lost their capacity of flying swiftly (or they may never have possessed it). Here again we are looking at worlds in which categories display their ideal or most characteristic properties. In modal semantics, these specific worlds are characterized as “stereotypical worlds”. Knowledge about prototypical properties or typical instances of a category is part of our conceptual knowledge and constitutes the basis for our operations of categorisation as demonstrated by prototype semantics.

- A somewhat similar rationale can be developed to account for verb-noun collocations, as for example, espouser femme, escolter messe, faire honte etc. In these cases, the direct object is a role-category in a stereotypical frame, for instance, in a “marriage”-frame or in an “action type”-frame. The noun and the predicate of verb-noun-collocations share the same world-time index (indexes) as they constitute a complex predicate which can be instantiated at whatever index (or set of indexes) is determined by a specific discourse context. We can summarize the semantics of these generic readings using the following formula:

$$\text{GEN} (\lambda x. \lambda w. \text{horse} (x)(w)), \text{with } w \in MB_{\text{stereotypical}}(w_0, w), \text{with } w: (\text{stereo})\text{typical worlds;}$$

The zero article also appears in contexts in which the noun phrase has a non-specific reading. Here again, the zero article acts as a function which yields the intension of the noun in question. In contexts with a non-specific reading, the world variable of the noun phrase is in the scope of a modal operator. Some operators are particularly prominent, for example, the negation operator, the question operator, the conditional operator IF and intensional verbs such as to look for or to wish. The world variable of the noun is instantiated at an index (or a
set of indexes) in accordance with the semantics of the operator which takes the noun phrase in its scope. Modal operators do not anchor noun phrases in the real world, but with respect to a set of possible worlds:

- the negation operator instantiates the semantic content of the noun phrase in indexes which exclude the real world;
- the conditional operator picks up indexes that are most similar to the real world and in which the restrictions laid down in the protasis hold;
- finally, intensional verbs like ‘to wish’ or ‘to desire’ anchor the noun phrases in indexes in which the wishes or the desires are fulfilled; at the same time, these wish- or desire-worlds must be maximally similar to the real world given that we do not normally take into account fantastic or supernatural worlds if we are pursuing rational interests and intentions.

The following formal representation captures the semantics of noun phrases in the scope of an intensional operator:

\[ \text{NEG} (\lambda x. \lambda w. \text{horse} (x)(w)), \text{ with } w \neq w_0; \]
\[ \text{IF} (\lambda x. \lambda w. \text{horse} (x)(w)), \text{ with } w = \text{protasis-worlds most similar to } w_0; \]
\[ \text{INTENSIONAL VERBS} (\lambda x. \lambda w. \text{horse} (x)(w)), \text{ with } w = \text{wish-worlds most similar to } w_0; \]

The combination of the zero article with plural nouns like “knights” (bare plurals as in “we saw knights”) and with mass-denoting nouns like “water” (“they drank water”) is a special case. In contrast to the readings we have analysed so far, they do not necessarily occur in intensional contexts. Here the zero article acts principally as a marker of indeterminacy. With an undetermined quantity of individuals (“knights”) or of non-discrete entities (like “water”, “bread” and “wine”), the denotation – i.e. the set of individuals or the portion of a mass – remains unspecified. The bare (mass or plural) noun in itself can denote kinds (the knight-kind or the water-kind) in Old French and can therefore be compared with the situation in modern English (cf. “Gold is rare”, “dogs are widespread” in Chierchia 1998: 363). However, in episodic contexts, the noun variable is instantiated in the real world \( w_0 \) (as in \( j'oï venir chevaliers \) with \( t < t_0 \)) and the denotational set turns out to be undetermined, given that neither its quantity nor its composition are specified. We can formalize the semantics of indeterminate noun phrases (knights, water) as:

\[ \lambda x. \lambda w. \text{knights/water} (x)(w_0 \text{ or in other } w) = \text{undetermined in } w_0 \text{ or in any other } w. \]

This article does not look in detail at the interpretation of the zero article with proper names and unique referents. The semantics of proper names pose particular issues: one possible solution could follow Kripki’s idea that proper names are rigid designators, which is to say that they denote the same referent in all possible worlds (irrespective of whether the referent really exists in that possible world or not). The proper name “Charles” would then yield the individual constant Charles in every possible world. This is, however, a line of thought that we will not pursue here.

To summarize the results of our analysis, we can say that the semantics of the zero article can be captured by positing two different basic values: for most readings – the generic or kind reading, the unspecific reading in intensional contexts – the zero article acts as a function which yields the intension of the noun in question. In these contexts the noun is in the scope of a modal operator (the genericity-operator, the \text{NEG}-operator, an intensional predicate and so on) with the effect that the noun (phrase) does not denote an individual or a set of individuals in the real world \( w_0 \). However, when the zero article combines with plural
nouns (bare plurals) or mass nouns, it does not anchor the noun (phrase) in a world apart from the real world $w_0$ but marks that the denotational set (in $w_0$ or in any other world) remains undetermined.

An interesting parallelism which cannot be explored in this article, but which should nevertheless be mentioned in passing, is the fact that the zero article seems to have at least partly the same function in the DP-domain as the subjunctive mood in the IP-domain. The zero article parallels the subjective in that it appears in the scope of intensional operators or marks the denotational indeterminacy of the description in its domain.

The subjunctive occurs, for instance, in the scope of (strong) intensional predicates like “want” and it marks – at least in some contexts – the denotational indeterminacy of the proposition (which is equal in the domain of the IP to a set of indexes in which the proposition holds). For example:

- The subjective selected by strong intensional predicates – their focus lies on wish-worlds which are most similar to $w_0$:

(25) $\textit{dex veult que cest fet soit seuz ...}$

God want that this fact be.3.PRS.SUBJ know.PTCP

‘God wants this fact to be known...’

(ndchar, 4696)

- Indeterminacy of the denotational set: the composition of the denotational set is not specified. A case in point: the world-individual pairs consisting of the world indexes and the set of individuals in the extension of whoever

(26) $\textit{car quiconques tiegne teus eritages, la juridicions en}$

given that whoever have.3.PRS.SUBJ of this kind heritage the jurisdiction of it

$\textit{apartient au seigneur}$

belong.3.PRS.IND to master

‘given that whoever has a heritage of this kind, its jurisdiction belongs to the master’

(beaumacb, 156.016-156.025)

Space constraints prevent us from further exploring the parallelism between the zero article in the DP-domain and the subjunctive in the IP-domain more carefully. At any rate, there is a striking parallel concerning the marking of the referential status of the NP by the zero-article and the vericonditional status of propositions marked by the subjunctive mood. It must be stressed, however, that the subjective progressively loses its relevance as a marker of intensionality in developments from Old to modern French.

4. The emergence of a new determiner system

New developments in the domain of the determiner system were already on the horizon by the end of the 13th century, and would lead to a profound restructuring of the original Old French system. The texts of the 13th century reveal some interesting overlaps and conflicting domains which mirror the competition between the zero article and other determiners of the paradigm. A very early document analysed by Carlier and Goyens (1998) is a translation of Jean d’Antioche which dates back to the year 1282. The document casts light on the increasing ousting of the zero article from its entrenched functional domains. What is more, the zero article competes with three different items, namely the definite, the indefinite and the partitive article, which are all advancing due to their progressive grammaticalization. The definite
article evolves particularly swiftly at the end of the Old French period. A closer examination of the data is provided below.

a) The definite article encroaches upon the most typical contexts of occurrence of the zero article. As example (27) reveals, the definite article occurs in definitions, and hence becomes compatible with generic readings:

(27) *Li exorde au commencement si doit avoir mout de* sentences et d’auctoritez (...).

‘The exordium at the beginning has to be composed of many wise judgement and be based on authority.

(27) *Li exorde au commencement si doit avoir mout de* sentences et d’auctoritez (...).

‘The exordium at the beginning has to be composed of many wise judgement and be based on authority.

(Jean d’Antioche chap. XVI, in: Carlier & Goyens: 84)

Consequently, the definite article, also shows up in maxims, proverbs and other types of generic predications. A case in point is (28), in which a modalized (deontic) predication deals with the concept of loyalty:

(28) *La leauté doit l’en toz jorz amer* every day love

‘You shall love loyalty every day.’

(Charroi, 442, Buridant 2000: 113)

It comes as no surprise either that the definite article also competes with the zero article in comparisons as in example (29):

(29) *Si cum li cerfs s’en vait devant les chiens* in front of the dogs

‘He flees like the deer from the dogs.’

(Roland, 1874)

Finally, the definite article also arises in combination with unique referents:

(30) *Ainz que li solez liet* Before the sun rise.3.PRS.IND

‘Before the sun rises.’

(Parise, 2020, Buridant 2000: 112)

This article will not offer a detailed discussion of the semantic and discourse-pragmatic peculiarities of the definite article (see von Heusinger 2002). However, some standard accounts can offer useful information on the relationship between the zero article and the definite article.

The semantics of the definite article is captured by the iota-operator which combines the notions of existence and of uniqueness. Its semantics can be translated as “there is one and only one x”. Combined with a noun (phrase), the definite article forms a definite expression which – according to Wespel (2008: 3) – “refers unambiguously relative to a certain type of informational domain”. In other words: a definite expression picks up the unique or maximal referent from a relevant domain of entities (see e.g. Himmelmann 1997, von Heusinger 2002

---

1 Exordium is a technical term which denotes the introductory part of a discourse.
and Wespel 2008). From a discourse-pragmatic point of view, the definite article signals that both the speaker and the hearer are familiar with a determined discourse referent. Definiteness as familiarity has been conceptualized by Heim in terms of a file card metaphor: the definite article is conceived of as an instruction to pick up a file card which has already been set up and which contains all the information already known by the speaker and the hearer. This file card is updated with the information provided by the predicational content of the assertion (cf. Heim 1982: 276).

The further grammaticalization of the definite article is reflected in its occurrence in generic contexts. As already discussed above, the expression of generic reference is not per se the semantic value of the zero article or any determiner, but ensues from the interaction between the determiner in question and the specific context. The determiners elaborate the concept of generic reference in a particular way, in accordance with their abstract semantics and by interacting with salient elements of the context (e.g. kind-level predicates). This amounts to saying that the zero article and the definite article elaborate the same underlying basic concept (of generic reference) in different ways: the zero article (as already shown) focuses on the intensionality of nominal expressions by presenting kind-denoting items as concepts and/or properties, while the definite article draws on the notion of uniqueness and/or maximality. However, there are (at least) two different ways of accounting for the profiling of a generic reading:

One line of interpretation takes the definite expression “the horse” in a generic predication such as “the horse has a tail” to be a “generic individual” which represents the whole class or category. This “generic individual” could be equated with the prototype, conceived as the best representative (or instance) of the category (see Kleiber 1990). A different interpretation is put forward by Chierchia (1998: 381). The definite expression “the horse” (our example) denotes the “totality of horses” or the group containing all horses. So whereas the prototype interpretation capitalizes on the notion of uniqueness – there is only one best representative of the category –, Chierchia’s account highlights the feature of maximality (the maximal individual). In this context, Chierchia emphasizes the “mass-like behavior” of singular definite generics: this property is confirmed by the fact that definite generics display singular morphology and do not allow for numerals (e.g. * The horse is four/many). Chierchia formalizes this insight by integrating a mass-feature in his description (see Chierchia 1998: 380)

the horse: THE (MASS(horse)) = g(t MASS(horse)) (see Chierchia 1998:380)

The formula captures the idea that singular definite generics refer to the maximal group (tota-operator in combination within a group-function g) containing horses or the totality of horses. Consequently, the intension of a singular kind can be defined as a function “from worlds into a collective or totality of some sort” (Chierchia 1998: 382). The intension of the singular kind “the horse” can, therefore, be spelled out as follows:

THE (MASS(horse) = \^g(t MASS(horse)) = \lambda w[g(t MASS(horse,w))]

To sum up, we can say that according to this interpretation, definite descriptions conceptualize the notion of kind in group or collective (“totality”) perspective by highlighting its “class-as-one” character (Chierchia 1998: 381).

b) The indefinite article, though much less grammaticalized than the definite article, is another element that began to compete with the zero article in late Old French. Sieglinde Heinz (1982) studied the advance of the indefinite article with particular reference to Villehardouin’s Conqueste de Constantinople, an emblematic text for the study of
linguistic developments in ordinary 13th century prose. She and other authors (e.g. Foulet 1965, Moignet 1976 and Buridant 2000) cast light on the different contexts in which the indefinite article competes with the zero article:

First, the indefinite article can also be employed in order to express genericity, but the extension of its original uses is confined to comparisons (in contrast to the general availability of the definite article in generic contexts):

(31) *Plus est isnels que uns falcuns*

More be.3.PRS.IND swift than a falcon

‘He is swifter than a falcon.’

(Roland, 1874, Buridant 2000: 112)

The indefinite article surfaces in predicative structures as well as:

(32) *al Vernas, qui... ere uns Greus qui se tenoit a  to Varnas who... be.3.PRS.IMP a Greek who PRO.REFL.HOLD.BE.3.PRS.IMP to als them

‘to Varnas who ...was a Greek who was at their side’

(CdC 403, Heinz 1982: 219)

(33) *la Serre,qui ere une citez forz et riche*

La Serre which be.3.PRS.IMP a town strong and rich

‘La Serre which was a strong and rich city.’

(CdC 280, Heinz 1982: 219)

Again, space constraints preclude a detailed account of the developments surrounding the indefinite article: the reader is thus referred to the seminal work of Elisabeth Stark (2006) on Indefiniteness and textual coherence. However, I will briefly outline some of the major stages of its trajectory: the indefinite article is an innovation of the Romance languages and starts as a partitive construction with a cardinal value: *unos* in Latin referred to one element of a given set. In a first stage of evolution, the indefinite article marked countability and profiled an entity as a single unit. In a second stage which reaches its peak in the Middle Old French period, the indefinite article evolved into a marker of specific indefinite reference. The indefinite article indicates that the hearer is not familiar with a discourse referent newly introduced by the speaker into the discourse. Therefore, the indefinite article serves to introduce a new discourse referent or – to resume the file card metaphor – to set up a new file card.

The indefinite article conceptualizes the notion of genericity (but still restricted to comparisons) by singling out an arbitrary but representative member of the category in question. A similar explanation can be put forward for predicative structures such as “x was an N, e.g. a Greek”: The individual who is the subject of the predication is presented as an arbitrary member of the category described by the noun phrase.

To complete the picture we must also look at a third competing item: it is not until the 13th century, in the period of later Old French, that the partitive construction (de + article) surfaces. The partitive construction occurs with entity-denoting plural nouns (e.g. *des chevaliers:* “knights”) or with abstract nouns (e.g. *du vin:* “wine”). It competes with the zero article particularly in one special domain, that of quantificational indeterminacy. A quite early and probably the most quoted example is taken from the Holy Grail story:
In this early stage, the partitive article \((du + vin)\) only occurs jointly with a noun phrase which refers to a determined and closed set of entities (the whole). In example (34), the noun phrase is determined by a relative clause. The speaker makes it clear that he refers to a well-defined quantity of wine (the specific portion of wine which is not cloudy). The notion of quantificational indeterminacy associated with the partitive article is conceptualized by profiling the part-whole relation and combining it with the notion of maximality. The part-whole relation is indicated by the preposition \(de\) \((x_{\text{part}} \text{ from } N_{\text{whole}})\), whereas the definite article highlights the maximal quantity of the substance delimited by an explicit description (the relative clause which specifies the non-cloudy, hence poison-free, glass of wine).

These developments reveal an interesting process of specialisation of the determiner system against the backdrop of the original system in which the zero article played a central role: in the nascent new system, the kind reading is progressively expressed by definite and indefinite noun phrases, with the former being more grammaticalized than the latter. The second basic function of the zero article, the marking of quantificational indeterminacy is progressively accomplished by the emerging partitive construction which, on the threshold of the 14th century, still requires a determined quantity of reference (the “totality” of \(N\) from which \(x_{\text{part}}\) is taken). As yet, he have not studied a further, particularly interesting, domain of competition between the zero article and the indefinite article – the large number of intensional contexts in which the noun phrase yields a non-specific reading: this is the focus of the next section of the article.

5. Evolutive tendencies in intensional contexts

This section looks at the advance of the indefinite article in intensional contexts with a non-specific reading, a development that would take several centuries to be completed. A very early example, which attests to the changes underway is quoted by Carlier & Goyens (1998). The passage we will comment on is taken from Jean d’Antioche’s translation of the \textit{Rhetorica ad Herennium} and dates back to the end of the 13th century (approx. 1282). For ease of analysis, I quote the whole passage:

\[\text{(34)} \quad \text{Et, quant il se virent en tele tempeste en grant peril, si voerent que, se il arrivassent en sauveté a cel port qu’il desiroient, il feroient sacrefixe d’un toreau a celui deu qui seroit illeques. Il lor avint qu’en cel port ou il arriverent, si estoit le temple de cele Diane a qui n’aferoit pas sacrefier torel. Cil, qui mesconoissoient cele loy, quant il eissirent dou vaissel, si sacrifierent le toreau, si com il avoient voé. (JA chap. LX, Carlier & Goyens 1998: 90)}\]

(‘And when they came in this storm in great danger, they swore that if they arrived alive at that port of destination, they would sacrifice a bull in honour of the god who would be the local divinity. They happened to arrive at a port with a temple dedicated to Diana to whom it was not convenient to sacrifice bulls. Those who did not know that law, when they disembarked the vessel, sacrificed the bull as they had sworn.’)

In this discourse sequence, the noun \(toreau\) (‘bull’) occurs three times in successive sentences, and, what is more, with different types of determiners: in the first sentence the
noun, though in the scope of an intensional operator (the conditional operator), occurs with an indefinite article. In the second sentence, the translator falls back on the zero article and in the last one he turns to the definite article. How can we explain this distribution of the determiners?

To begin with, the zero article *sacrifier torel* appears in a stereotypical frame, comparable to other common frames of the medieval world such as *epouser femme* (‘to marry a woman’) and to *dire messe* (‘to say a mass’). This systematic use of the zero article is best accounted for by a conceptual explanation: the bare noun phrase *torel* can be interpreted as a kind-instantiation of the role-slot **THEME** (the object of the sacrifice). When the speaker evokes a certain frame (e.g. the sacrifice or the marriage frame), the theme-slot is filled in by a generic and stereotypical value – e.g. a kind like *femme* as in the marriage frame or *toreau* in the sacrifice frame.

Turning to the indefinite article, I would like to argue that its presence cannot be explained by its abstract semantics given that it does not introduce a referent whose existence is presupposed – the noun phrases of the successive sentences do not refer to the same individual. Rather, its use can only be explained by an extended discourse function: the indefinite article introduces the noun *bull* into the discourse and makes it salient as a relevant referent for the discourse to come. In other words, in contrast to the zero article, the indefinite article evolved into a kind of prominence-marker which indicates the status of the noun as a discourse theme. The listener can infer from the indefinite article that the individual denoted by the noun will be picked up again as the discourse unfolds in order to provide more information on it by means of further predications. Whether or not the referent will be individualized in the discourse is irrelevant, however note that the first sentence does not deal with a specific bull in *w₀*). What is central here is the fact that the bull-role is upgraded to a relevant theme of the discourse and its instantiation can be realized by whatever bull may be specified by the context. To sum up, we can say that whereas the zero article marks a noun as a generic instantiation within a frame, the indefinite article signals the status of the noun as being a relevant (not necessarily individualized) referent of the discourse. The starting point of these developments is the pragmatic use of the definite article in intensional contexts. Given the oddity of an existential presupposition in these contexts, the only available interpretation is one of prominence (in whatever possible world is at stake).

More general trends in the development of the opposition between the zero article and the indefinite article can be observed in the texts of the Frantext Corpus dating from between the 15th and the 17th century.

We studied different contexts created by the French verb *chercher* (‘to look for’) which is characterized by the fact that it can display an intensional reading whenever the referent of the theme-role may not exist in the actual world but only in wish or believe-worlds (e.g. “to look for a unicorn”).

In the 15th century, the zero article is still very frequent, particularly with plural nouns (bare plurals). Take the following example:

(35) *Il cherche delices etaises, et le royaume des cieulx*  
He look for.3.PRS.IND pleasure and comfort and the kingdom of the heavens  
‘He strives for pleasure and comfort and the kingdom of heaven...’  
(Alain Chartier, Le livre de l’espérance, 1429, page 34)

In this intensional context, *delices* and *aises* are an object of quest whose existence is, at least, questionable. The NP refers to a type-concept and indicates the kind of question the person is engaged in. Compare example (35) with (36): the latter denotes an undetermined portion taken from the category DELICES. In this non-intensional context created by the expression of impact *être gagné de*, the speaker selects the indefinite form *des* (de + les).
(36) Parvenus à Narbonne, ils furent tellement gagnés des delices que le plantureux terroir de ceste ville fournit largement ....

(Bénigne Poissenot, Esté, 1583, p. 60)

‘Having arrived at Narbonne, they were so overwhelmed by the pleasure that the rich land of this town provided them ...’.

(Olivier Serres, Le Théâtre d’agriculture et mesnage des champs, t.2, 1603, p. 502)

In singular NPs we encounter a systematic contrast between nouns, which denote a discrete entity on the one hand and the class of abstract nouns on the other. With the former group, a general tendency towards a token-interpretation marked by the indefinite article un/une (an instance of type N) prevails. In (37), the individual in question (Ulespeigel) is looking for exactly one place to sleep. Apart from the entity-type (lieu is a discrete entity), we may argue that the place-role (ung lieu) is rendered prominent by the indefinite article. As in the case of the bull (see above) the place-role (or place variable) is resumed in the further context and is filled in by a specific value (ung jardin d’ung juif).

(37) Ule spiegle s'en alla ung jour à la dedicasse ave c sa mere, où il s'en yvra au disner et il chercha ung lieu pour paisiblement dormir affin que aulcune personne ne luy fist mal. Lors il s'en alla derriere ung jardin d'ung juif (...).

‘Uelespiegel went with his mother to a ceremony where he got drunk when having dinner. So looked for a place to sleep in peace... Some time later he went behind a Jew’s garden.’

(Anonyme, Ulenspiegel, 1530, p. 155, Histoire 7)

Example (38) demonstrates another similar case:

(38) et tousaultres chiefz de guerre cherchoient une honneste yssue,

and all other masters of war look for.3.PRS.IMP an honorable way out by specific arrangement

‘and all the other warlords tried to achieve a gentlemen’s agreement by a specific arrangement’

(Philipp de Commynes, Mémoires t.3, 1489, livre VIII, page 220)

In this passage, the nobles are negotiating an honorable peace treaty, which is in line with a specific judicial procedure (yssue par appoinctement). Once again, the referent issue par appoinctement is made salient by the indefinite article and announces the relevance of this referent for the further discourse, which deals with the results of those negotiations.

Finally, the examples of the Frantext Corpus show that the zero article still occurs with abstract nouns, but also with some role nouns which designate a type concept. The verb-object combination chercher + undetermined object noun refers to an abstract quest-type and not to an individualized (or discrete) object-entity. This is illustrated by the abstract quest-type chercher verité (‘search for the truth’) in example (39). In this example what is indicated is only the quest-type, with no further singularization of any specific instance of truth:
(39) (ces maraulx sophistes) lesquels en leurs disputations ne cherchent
these corrupt sophists who in their discussions not look for.

verité, mais contradiction et débat.
truth but contradiction and debate.

‘These corrupt sophists who do not search for the truth in their discussions, but for contradictions and dispute.’

(Rabelais, François, Pantagruel, 1542, p. 217, CHAPITRE XVIII, Comment un grand clerke de Angleterre vouloit arguer contre Pantagruel et fut vaincu par Panurge.)

Many attestions of the zero article associated with quest-types can be found in the texts. However, the author who provides the most examples of predications revolving around quest-types is the theologian Jean Calvin. He provides an exhaustive inventory of all possible kinds of quest, for instance, chercher ignorance (‘ignorance’), chercher preuve de loin (‘evidence from far’), chercher assurance de leur justification (‘the certainty of their justification’), chercher justice en la Loy (‘justice in the Law (of God)’), chercher seurté (‘certainty’), chercher salut (‘redemption’), chercher vie et immortalité (‘life and immortality’) and so on. One example is selected for illustration purposes:

(40) Si on cherche ignorance pour ne savoir que c'est de Dieu
If pro.imp look for. ignorance in order not know that it be.

‘If you try to be ignorant in order not to know that it comes from God ...’

(Calvin, Jean, Institution de la religion chrestienne : livre second, 1560, p. 37, livre II, chapitre II)

In the early classical French period, the determinerless verb-object combinations have already become highly lexicalized. In the last example (41) taken from the playwright Pierre Corneille, the direct object NP refers to a social role in the marriage frame (chercher Femme) denoting the kind concept FEMME (‘woman’). The speaker’s exhortation does not consider an individual or a set of individuals (a certain number of candidates) but focuses on the kind (the stereotypical role) within a social frame.

(41) Qu’il cherche fille ailleurs, et pour moy de ma part
That he look for. woman elsewhere and for me of my part

J'attendray du destin quelque meilleur hazard
I wait.of destiny some better chance

‘So he should look for a woman somewhere else, and for my part I will wait for a better time in my life.’

(Corneille Pierre, Mélite ou les Fausses lettres,1633, acte V, scène 5e, p. 122)

6. Conclusion
As we have seen, the semantics of the zero article did not change over time, nor was it bleached in its abstract value. However historical evidence has shown that the functional range of the zero article diminished as it was relegated to residual contexts by competing elements of the determiner system. The definite and the indefinite article in particular gained ground and succeeded in progressively ousting the zero option in the majority of contexts. The definite article extended its functional domain at a very early stage (already in late Old French), in particular by enabling kind readings in generic predications as well as in comparisons. The indefinite article required more time to evolve into a fully-fledged determiner, with the process beginning in late Old French and coming to its completion in the 17th century. The intrusion of the indefinite article into intensional contexts has clearly been a
crucial step in the process of its grammaticalization. In these intensional contexts, the indefinite article was introduced as a discourse-functional device designed to give prominence to the noun and mark it as a thematic element. The prominent role of the noun became apparent through its systematic resumption in the unfolding discourse at hand. Note that the highlighted noun denoted a discourse-relevant individual, not necessarily a specific in w₀.

The zero article maintained its original function of highlighting the intension of a nominal expression. However, the functional range it covered was progressively taken over by the other competing determiners. The expression of the notion of kind made it clear that the different determiners provided alternative ways of conceptualizing one and the same basic concept. The interaction between the abstract and underspecified meaning of the determiners, on the one hand, and the specific contexts in which they occurred (whenever necessary, enriched with encyclopedic knowledge), on the other, turned out to be of paramount importance.

However, a great deal of research ahead before we can gain a full understanding of the complex developments and the profound restructuring of the determiner system from Old to modern French.

7. References


Preverbal bare nominals in Brazilian Portuguese

M. Teresa Espinal & Sonia Cyrino
teresa.espinal<at>uab.cat
sonia.cyrino<at>gmail.com

1. Introduction*

In this paper we aim at examining the interpretation of preverbal bare nominals (BNs) in Brazilian Portuguese (BrP) on the basis of its distribution and meaning. We will introduce some interesting comparisons between BrP data and other Romance languages, mainly French, on the one hand, and Catalan and Spanish, on the other.

We will claim that preverbal BNs are assigned two different interpretations in correlation with different syntactic structures: a definite kind interpretation and a definite plural generic. The former interpretation is constrained at the syntax-semantics interface by a DP structure with a null D and no specification for syntactic Number. The latter is constrained by a full DP structure with a null D specified for plural Number. We will also correlate the DP interpretation with the type of predicate it combines with (i.e., individual-level, kind-level or stage level) and the type of sentential context (generic/habitual or episodic).

2. Preliminary data

Let us start by considering the data in (1) to (5) from Müller (2002):

(1) O brasileiro é trabalhador.  
the.SG Brazilian is hardworking
‘Brazilians are hardworking.’

(2) Os brasileiros são trabalhadores.  
the.PL Brazilian.PL are hardworking
‘Brazilians are hardworking.’

(3) Um brasileiro é trabalhador.  
a Brazilian is hardworking
‘Brazilians are hardworking.’

* This research has been funded by the following research grants: Spanish Ministerio de Economia y Competitividad (FFI1011-23356), Generalitat de Catalunya (2009SGR-1073, ICREA Acadèmia award), Brazilian CNPQ-Conselho Nacional de Desenvolvimento Científico e Tecnologico (Research Grant 303006/2009-9), and CAPES-Spanish Ministerio de Educación Cultura y Deportes (research grants CAPES/DGU-305/13 and 0214. 2013-2014).

For this presentation we partly rely on previous work developed by the authors (Cyrino & Espinal 2011, 2012), and also joint research in collaboration with Olga Borik (Borik & Espinal 2012, 2013; Borik, Cyrino & Espinal 2012).
(4) Brasileiro é trabalhador.
Brazilian is hardworking
‘Brazilians are hardworking.’

(5) Brasileiros são trabalhadores.
Brazilian.PL are hardworking
‘Brazilians are hardworking.’

All these examples show an i(ndividual)-level predicate, that introduces a property of the subject, a meaning which is not bindable to spatio-temporal or event variables (Ladusaw 1994 Kratzer 1995). Examples (1) to (3) show overt determiners, whereas (4) and (5) do not. Examples (1) to (3) are generic statements: (1) and (2) express genericity at the DP level (what is known as D-genericity), whereas example (3) express genericity at the sentential level (what is known as I-genericity: Krifka et al. 1995). Example (1), with an overt DP in subject position refers to a kind: o brasileiro denotes the atomic kind/class referred to by means of the noun, without considering the members of this class (a reading that has been referred to in the English literature as the definite generic the N construction, Carlson 1977; or as the singular generic construction, Chierchia 1998). By contrast, example (2), by means of the overt DP, expresses a generic definite plural reading, according to which reference is made to the maximal sum of individuals that instantiate the brasileiro kind, without considering the kind itself. Finally, example (3) shows an indefinite generic DP in subject position, which is conceived as a heimian indefinite participating in a generic quantification.

The question we want to address in this article is what sort of genericity, if any, do the sentences in (4) and (5) express. Examples (4) and (5) introduce bare nominals (henceforth BNs) in preverbal position, also in combination with the same i-level predicate we have just introduced: ser trabalhador. We claim that the BN in (4) is ambiguous between a definite kind reading (as in (1)) and a maximal sum interpretation (as in (2)); the BN in (5) also denotes a maximal sum of individual objects, but its interpretation is different from the bare plural kind interpretation in English (contra Dobrovie-Sorin & Pires de Oliveira 2008).

An interesting piece of additional data that is crucial to understand the above set of data is given in (6):

(6) (a) Os brasileiro é trabalhador.
the.PL Brazilian is hardworking
‘Brazilians are hardworking.’

(b) Os brasileiros são trabalhadores.
the.PL Brazilian are hardworking.PL
‘Brazilians are hardworking.’

Notice that the examples in (6), which are both well-formed, show lack of nominal agreement between the D and the N, and (6a) also shows lack of agreement between the head of the DP and the V.

In view of all these data our claim is that (4), in one of its interpretations has the same structure and meaning as (1). But example (4) shares with (5) and (6) the same interpretation that (2) has. We will refer to these two interpretations as the definite kind and the generic definite plural readings, respectively.

In Section 3 we first consider what (1) and (4) have in common, and what the requirements for a definite kind interpretation are. In Section 4 we focus on the syntactic structure common to (2), (4), (5) and (6), and we advance an analysis of how the interpretation of preverbal bare nominals in (4) and (5) is constrained by the syntactic structure assigned to them.
3. Definite kind interpretation

We assume, following previous studies, that the canonical structure for nominal expressions in BrP projects Number and Determiner over the Noun (Longobardi 1994, Zamparelli 1995, Chierchia 1998).

(7) \[ [\text{DP} \text{D} [\text{NumP} \text{Num} [\text{NP} \text{N}]]] \]

This full DP structure is to be postulated for the data in (1) to (3). It constrains an entity reading, either an object entity or a plural sum of entities. However, the example in (1) suggests that a smaller structure is also available (de Swart & Zwarts 2009), one in which no syntactic Number projection is postulated. Consider the structure in (8).

(8) \[ [\text{DP} \text{D} [\text{NP} \text{N}]] \]

In recent work (Espinal 2010; Borik & Espinal 2012, 2013) we defend the hypothesis that a definite kind generic interpretation is constrained by definite DPs with no Number involved. Why? Because Number is a function that turns properties of kinds (the denotation of the Noun) into properties of objects (singular or plural), and when reference is made to whole classes, without considering their members, no consideration of number whatsoever is involved in the nominal expression and, therefore, no Number intervenes in its syntactic structure. This hypothesis is specially interesting because we have observed that languages that have Determiners (null or overt), such as Russian, English or Spanish do not differ with respect to the possibility of having definite kinds: they all have them and in all of them no number is involved.

In BrP definite kinds can both be expressed by means of an overt or a covert D, as illustrated in (4) (in one of its interpretations) and in (9) in which a preverbal DP or BN combines with a k(ind)-level predicate, which selects for kinds in argument position.

(9) (a) *A jaguatirica é comum na selva amazônica.*

*the jaguar is common in-the jungle amazonic*

‘The jaguar is common in the Amazonic jungle.’

(b) *Jaguatirica é comum na selva amazônica.*

*Jaguar is common in-the jungle amazonic*

‘The jaguar is common in the Amazonic jungle.’

It should be noticed at this point that predicates that are usually said to require kind-referring objects (e.g., *inventar* ‘to invent’) seem to show a different behaviour, since they do not allow the omission of the definite article. This is illustrated in the following example:

(10) Jack Dorsey inventou *(o) Twitter, Steve Jobs inventou *(a) Apple, Mark Zuckerberg inventou *(o) Facebook, J.K. Rowling inventou *(o) Harry Potter,…*  


We think that the contrast between the data in (9), where the overt article is not obligatory, and the data in (10), where it is, is due to the fact that the reference of the object argument in (10) is not really equivalent to a kind, as already postulated by Beyssade (2005) for French. Her proposal is that objects of verbs like *invent* refer to concepts or prototypes, which are semantically and ontologically distinct from kinds.

Our reasons to support this proposal for BrP are the following: (a) objects of *invent*-type verbs are the only cases in this language where the article with a kind referring expression are really obligatory, (b) these objects cannot be antecedents for personal pronouns *ele/elas* in
Preverbal bare nominals in Brazilian Portuguese

BrP (e.g., ‘Jack Dorsey invented the Twitter and then improved it.’), as opposed to all the other kind referring expressions (e.g., ‘As pessoas temem a jaguatirica, porque ela é feroz, ‘People fear the jaguar because it is ferocious’), and (c) these nominal expressions cannot be replaced by definite plurals, which are expected to denote the maximal sum of individuals that instantiate the kind (e.g., ‘Inventou o telefone’ (S)he invented the telephone’ vs. ‘Inventou os telefones’ lit. ‘Invented the telephones’, as opposed to (A) ‘baleia está em extinção ‘the whale is on the verge of extinction’ vs. (As) ‘baleias estão em extinção ‘the whales are on the verge of extinction’).

It should furthermore be remarked that, since the structure in (8) does not involve Number, it is predicted that numerical expressions are incompatible with definite kinds (Borik & Espinal 2012, 2013). This prediction is borne out by the data: when numerical expressions specify arguments of k-level predicates, they can only license a taxonomic interpretation, as illustrated in (11).¹

(11) Dois dodós desapareceram no século XVII.
‘Two dodos disappeared in the 17th century’

These examples illustrate that if Number is imposed on nominal expressions that refer to species, the output can only be accepted when a taxonomic/subkind reading is intended. Furthermore, the data in (12) illustrate the fact that numerals are preferably combined with lexical items such as tipo, classe ‘type, class’ in BrP, to derive a subkind interpretation. These lexical items, by contrast, cannot directly combine with kind-referring expressions, as illustrated in (12b).

(12) (a) Dois tipo(s) de baleia(s) desapareceram da superfície da terra.
‘Two types of whale disappeared from the surface of the earth.’

(b) *Tipo(s) de baleia(s) desapareceram da superfície da terra.
‘Type of whale disappeared from the surface of the earth.’

The conclusion to be drawn from this section is that definite kinds require a D projection, but do not show any evidence for the presence of Number. The syntactic realization of definite kinds allows a null D. Furthermore, we conclude, that a kind-referring expression is incompatible with Number, both syntactically and semantically.

4. Generic definite plural interpretation

In BrP, canonical syntactic and semantic arguments have a DP structure with a determiner that can be null. Consider (13), which combines examples (2) and (5) from Müller (2002), given above.

(13) Os brasileiros são trabalhadores.
‘Brazilians are hardworking’

What is interesting is that, if we add examples (4) and (6) to this paradigm, repeated in (14) and (15) for convenience, we observe that specification of Number on the Noun is not required.

¹ We assume that reference to subkinds is built on number, either singular or plural (Borik & Espinal 2012, 2013).
(14) Brasileiro é trabalhador.
   Brazilian is hardworking
   ‘Brazilians are hardworking.’

(15) (a) Os brasileiro é trabalhador.
       the.PL Brazilian is hardworking
       ‘Brazilians are hardworking.’
(b) Os brasileiro são trabalhadores
       the.PL Brazilian are hardworking,PL
       ‘Brazilians are hardworking.’

Note also that, although the preferable English translation for the subject expression in all these examples is a bare plural, as in ‘Brazilians are hardworking’, this translation does not reflect adequately the meaning to be attributed to the BrP examples. Besides, it should be noted that the corresponding example in another Romance language like Spanish would be *Los brasileños son trabajadores* lit. the Brazilians are hardworking, with an overt definite plural article in subject position. In all these BrP examples and also in the Spanish translation the meaning of the subject refers obligatorily to the maximal sum of individuals of the Brazilian kind for which it is claimed that they are hardworking.

An initial evidence for the claim that the articleless version of (13) has a null plural determiner, could come from the fact that the only way to refer back to these nominal expressions in preverbal position in the following discourse is by means of *eles.*

(16) (Os) brasileiros são trabalhadores. Eles se preocupam com o futuro.
    the Brazilian are hardworking they self worry with the future
    ‘Brazilians are hardworking. They are concerned about the future.’

The same applies to the expressions with an overt D in (15) and the BN in (14), as illustrated in (17).

(17) (a) Os brasileiro é trabalhador. Eles se preocupam com o futuro.
       the.PL Brazilian is hardworking they self worry with the future
       ‘Brazilians are hardworking. They are concerned about the future.’
(b) Brasileiro é trabalhador. Eles se preocupam com o futuro.
    Brazilian is hardworking they self worry with the future
    ‘Brazilians are hardworking. They are concerned about the future.’

Interestingly, these examples suggest not only that the antecedent of the pronoun is a DP, but also that it is a plural DP, as proved by the ill-formedness of *ele* in the following discourse. This correlation further suggests that the postulated null D for an example such as (14) is to be specified for plural number.

Moreover, it is important to notice a subject-predicate asymmetry. In contrast to (16) and (17), the bare plural and the bare singular occurring in predicate position are not DPs, but smaller nominal expressions specified for Number, for which we would postulate a NumP structure (Dobrovie-Sorin et al. 2006). Evidence for this claim is that in the following discourse *eles* is not possible, as shown in (18) and (19).

---

2 A potential problem to this line of reasoning might come from English: bare plurals in this language are also referred back by means of a third person plural pronoun. This problem led in the early nineties to the syntactic proposal according to which English has a weak D and bare plurals in subject position are moved to D at LF; this movement would guarantee their argumental status. See Longobardi (1994, 2001, 2005).
Preverbal bare nominals in Brazilian Portuguese

(18) (Os) Brasileiros são trabalhadores, mas não costumavam ser *eles.
   the Brazilians are hardworking but not used be them
   ‘Brazilians are hardworking, but they didn’t use to be.’

(19) (a) Os brasileiro é trabalhador, mas não costumavam ser *eles.
   the.PL Brazilian is hardworking but not used be them
   ‘Brazilians are hardworking, but they didn’t use to be.’

(b) Brasileiro é trabalhador, mas não costumavam ser *eles.
   Brazilian is hardworking but not used be them
   ‘Brazilians are hardworking, but they didn’t use to be.’

It should be noted that these sentences become well-formed in BrP when VP ellipsis applies, as in (20) and (21).\(^3\)

(20) Os brasileiros são trabalhadores, mas não costumavam ser.
   the Brazilians are hard-working but not used be
   ‘Brazilians are hardworking, but they didn’t use to be.’

(21) (a) Os brasileiro é trabalhador, mas não costumavam ser.
   the.PL Brazilian is hard-working but not used be
   ‘Brazilians are hard-working, but they didn’t use to be.’

(b) Brasileiro é trabalhador, mas não costumavam ser.
   Brazilian is hard-working but not used be
   ‘Brazilians are hard-working, but they didn’t use to be.’

Besides that, it is interesting to compare these examples with what happens in other Romance languages that make a distinction between different types of clitics depending on the antecedent. One of these languages is Catalan, which distinguishes between third person accusative clitics el, la, els, les, which also require a DP antecedent, and a neuter clitic ho, which must be linked to a NP or AP in predicate position. The translation of (18) and (19) into Catalan will make use of the pronoun ho, thus suggesting again that the antecedent of the pronoun is not a DP: Els brasilers són treballadors, però no ho solien ser.

We have, furthermore, independent evidence that supports the hypothesis that the antecedent of eles in BrP is a DP.\(^4\) Therefore, we conclude that trabalhadores in (20) and

---

3. On VP ellipsis in BrP, see Cyrino and Matos (2002).

4. BrP third person weak pronouns, like accusative clitics in European Portuguese (EP), Spanish and Catalan, require that their antecedents be a DP. In order to support this claim, we consider the contrasts in (i) and (ii). While the antecedent of the accusative neuter clitic o ‘it.ACC’ in EP can be a CP (ia), this cannot be the case of ele ‘it’ in BrP (ib). First, the neuter clitic does not exist in the language anymore. Second, it can only be replaced by a null object (ib) (Cyrino 1997). Furthermore, the contrast between examples (ib) and (iib) show that in BrP ele cannot refer back to a CP, but only to a DP structure. This prediction is further borne out by the example in (iii).

(i) a. O João pergunta [CP se a Maria vem], mas o Pedro não sabe. [EP]
   it.ACC knows
   ‘João is asking whether Maria will come, but Pedro doesn’t know it.’

b. O João pergunta [CP se a Maria vem], mas o Pedro não sabe Ø /*ele. [BrP]
   knows Ø it
   ‘João is asking whether Maria will come, but Pedro doesn’t know it.’
trabalhador in (21) cannot be a DP, whereas the preverbal brasileiros in (13) and brasileiro in (14) are DPs.

Summing up, in BrP morphosyntactic Number may be instantiated only in the D head (cf. example (6)). In the possible case of number agreement between the D and the N (cf. example (2)), we assume number copying from D to N (dotted line) (cf. Costa & Figueiredo Silva 2006, Simioni 2007). Elsewhere, both overt and covert articles are possible. Therefore, the postulated syntactic structure for examples (2), (4), (5) and (6) is represented as in (22).

(22)

\[
[\text{DP } \text{D}_{[±\text{PL}]} \text{ [NumP } [±\text{PL} ] \text{ [NP N ]] ]}
\]

In other words, our analysis of the example in (5) is that the bare plural in preverbal position is like the full DPs in (2) and (6), with a null D specified for plurality. These nominal expressions do not convey a kind reading, because they are definite plurals. However, they have a generic interpretation as a result of combining with an i-level predicate. The same interpretation would apply in combination with a k-level predicate.

On the other hand, our analysis of the example in (4) is that it can be associated with two structures (namely, (7) and (8)), and denote either a generic definite plural or a definite kind. The first possibility derives from the fact that in BrP copy of Number on Noun is optional at spell-out. The second possibility derives from the fact that a Number category/feature is not required for definite kind interpretation.

On the interpretive side, recall also that in Romance BNs (in particular, bare plurals) do not allow a kind denotation, but an existential interpretation. Therefore, we claim (contra Dobrovie-Sorin & Pires de Oliveira 2008) that the generic reading of examples such as (4) (in the second interpretation it may have) and (5) cannot rely on Chierchia’s down operator.\(^5\) Accordingly, our conclusion is that if in the structure (7) a null article specified for Number occurs as the head of D, it cannot refer to a kind, but to a generic definite plural.

(ii) a. O João comprou \[\text{DP o livro}], mas não o leu. [EP]
the João bought the book but not it.ACCUS read
‘João bought the book but did not read it.’

b. O João comprou \[\text{DP o livro}], mas não leu \Ø/ele. [BrP]
the João bought the book but not read \Ø it
‘João bought the book but did not read it.’

(iii) O João pergunta \[\text{CP [DP a ora]}], mas o Pedro não sabe \Ø/*ela. [BrP]
the João asks the hour but the Pedro not knows \Ø it
‘João is asking the time, but Pedro doesn’t know it.’

\(^5\) Recall that kind-referring readings of bare plurals rely on Chierchia’s (1998) down operator, which applies to a set of pluralities and yields the maximal intensional plurality in that set.

See also Dobrovie-Sorin (2012), who postulates that the (im)possibility of kind-referring bare plurals depends on whether Number is (valued and) interpreted either on D (in Romance) or on little n (in English). According to this author: (i) when a (null) D is not marked for Num (English) it can function as the Down operator, and (ii) when a (null) D is marked for Num (Romance) it cannot function as an intensional maximalizing operator, because Number features on D trigger referentiality. However, Dobrovie-Sorin cannot account for the sort of generic reading of BrP bare plurals in subject position (e.g., Sambas são populares).

The proposal developed by Borik and Espinal (2013) for Spanish, based on the hypothesis that definite plurals in this language are not the counterpart of bare plurals in English, and therefore do not denote kinds but generic definite plurals can be extended to BrP. The only relevant difference that exists between BrP and other Romance languages (Catalan, French, Spanish) is that the former allows nul articles where the rest require overt articles.
5. Further preverbal BNs

So far we have considered preverbal BNs that combine with i-level and k-level predicates and license a generic interpretation, either because no Number is expressed at the DP level, or because a plural definite DP is coerced to a generic interpretation at the time its semantic composition with the predicate takes place.

In this section, we will focus on three additional types of BNs in preverbal position in BrP. What these BNs have in common is that they combine with s(tage)-level predicates, whose meanings contain a bindable spatio-temporal or event variable (Ladusaw 1994, Kratzer 1995).

We will first consider an example of a preverbal BN in a habitual / generic context (23), in which case the interpretation for the BN corresponds to a definite plural in a bijection relationship with a plurality of events. Second, we will examine the case of a plural BN in an episodic context, and we will describe the various interpretations it may be associated with: definite plural, indefinite and a totality quasi-quantificational reading (24). And, finally, we will focus on a preverbal BN unspecified morphophonologically for plural number occurring in an episodic context (25); we will see that what makes interesting this third type of BNs is that they are the only ones that require a contrastive or a pair list reading (Pires de Oliveira & Rothstein 2011).

(23) Bandido rouba banco.
    bandit rob.PRES.3SG bank
    ‘Bandits rob banks.’

(24) Deputados visitaram salas de aula.
    representative.PL visit.PAST.PERFECTIVE.3PL classrooms
    ‘Representatives visited classrooms.’

(25) %Deputado visitou salas de aula.
    representative visit.PAST.PERFECTIVE.3SG classrooms
    ‘Representatives visited classrooms.’

What these BNs share with the ones exemplified in (4) (in one of its interpretations) and (5) is that they refer to the maximal set of individuals that have the property denoted by the Noun. As before, our argumentation relies on the assumption that they all have the DP structure in (7) and that, therefore, preverbal BNs are not real BNs, but DPs with null Ds specified for plural Number. Since Number is essential for extensionalizing the property denoted by the Noun, these BNs all have a plural interpretation.

Let us now consider each one of the examples individually.

Example (23) is a generic sentence. More specifically, it is a habitual sentence, with the verb in the present tense (cf. Müller 2004, Munn & Schmitt 2005), in spite of the fact that there is no adverbial quantifier (like sempre ‘always’) overtly expressed. The verb of the sentence, roubar ‘to rob’, is a stage-level predicate, for whose subject an existential interpretation is expected. However, the subject in (23) is not existential since it is interpreted as the external argument of a generic/habitual sentence. The BN subject is salient and has a discourse referent; it is ‘semantically equivalent’ to a plural definite description. Furthermore, as expected for bare habituals, the plurality of individuals referred by bandido ‘bandits’ stands in a bijection relationship with the plurality of events denoted by the verb in the habitual aspect. According to what we have just claimed, and following Ferreira (2005) and de Swart (2006), we postulate that the semantic representation of the sentence in (23) looks like (26), which expresses that the DP instantiated by bandido denotes a maximal sum of individuals which is in a bijection relationship with a plurality of events of the roubar ‘to rob’ type.
In contrast to (23), the sentence in (24) is neither generic nor habitual. Note that the verb is morphologically marked for past tense and perfective aspect, and, semantically, it denotes an s-level predicate. Due to the properties of the verb, the subject is expected to denote not a kind but a plurality of objects. It may have a definite interpretation (as in the previous examples) or an indefinite reading. It allows for a definite interpretation, in a sort of contrastive/list interpretation, in an appropriate context, similar to quem visitou salas de aula na escola, deputados ou senadores? Deputados visitaram salas de aula ‘who visited classrooms in the school, representatives of senators? Representatives visited classrooms.’

The question that must now be answered is how we can account for the possible indefinite interpretation and for the occurrence of this BN in subject position. Syntactically, we assume for (24) that the BN is a DP with a null D. We also propose that this null D can be specified by an operator that encodes lack or absence of definiteness. We label this operator DE for deprivative. In parallel terms to the syntactic structure in (27) for French and Italian indefinite DPs, we postulate for the BrP bare plural in (24) the structure in (28), which differs from (27) in having a null D and an abstract deprivative DE. Notice that in both structures the plurality of the determiner has to be copied on the N.

6 Note that this operator occurs overtly in Romance languages like French, Italian, Portuguese and Spanish, in processes of lexical derivation (défaire, disfare, desfazer, deshacer ‘undo’). But, we also find it in languages like French and Italian for the expression of indefiniteness within the DP. Consider (i) and (ii).

(i) FRENCH
Des garçons sont venus.
DE.the.PL boys are come
‘Some boys have come’.

(ii) ITALIAN
Degli studenti stanno giocando fuori.
DE.the.PL students are playing outside
‘Some students are playing outside.

French des and Italian dei/degli/delle have been claimed to be indefinite articles, but what is more important is that they are complex determiners (Gross 1967, Kayne 1977, Ishane 2010) and encode plural semantics (Farkas 2006). In French and Italian, de- only applies to [+PL]. Furthermore, it should be noted that following the marker de-, which triggers the indefinite interpretation, morphological plurality is obligatorily specified on the Determiner, as the following examples illustrate.

(iii) FRENCH
*De garçons sont venus.
of boys are come

(iv) ITALIAN
*De bambini stanno giocando fuori.
of children are playing outside
Semantically, we postulate that this abstract deprivative operator cancels the definiteness and specificity of the head of D and, furthermore, syntactically licenses the null D by a spec-head relationship. The output interpretation is that the whole DP instantiated by deputados ‘representatives’ in (24) may denote a non-maximal sum of individuals.

Sentence (24) also allows a different interpretation. Given an appropriate discourse, it permits a universal reading for the BN in preverbal position. In this case, the BN expresses a generalization restricted to the actual set of individuals that have to be considered on a particular occasion (Condoravdi 1992).

(29) A. Ontem, houve uma festa na escola. (‘Yesterday, there was a party at school.’)
   B. (a) Deputados visitaram salas de aula.
       representative.PL visit,PAST.PERFECTIVE.3PL classrooms
       ‘All the congressmen visited classrooms.’
   (b) Representatives at that school at that time visited classrooms.

The BN in preverbal position in (29Ba) is not anaphoric to any previous nominal expression. Furthermore, it should be noted that the sentence does not make an existential assertion. Example (29Ba) is rather an assertion about the totality of the contextual relevant representatives whose existence in the actual world seems to be presupposed. This reading arises because the nominal expression is evaluated with respect to a context that entails the descriptive content of the BN. That is, in (29Ba), deputados has a functional reading because the context makes salient that we are referring to all the representatives that were at the school party mentioned in the discourse context specified in (29A). In this particular case the output interpretation for the preverbal bare plural refers to a maximal sum of individuals.

Let us now consider example (25). This example is felicitous only when associated with a contrastive interpretation for the BN (cf. Müller 2002, Pires de Oliveira and Rothstein 2011). This is what the symbol % means. Furthermore, it should be noted that in this example deputado can only be interpreted as referring to a plural set of individuals. See the test in (30).

(30) (a) *Deputado visitou salas de aula. Ele ficou contente.
    representative visited classrooms he got happy
   (b) Deputados visitaram salas de aula. Eles ficou/ficaram contente(s).
      representatives visited classrooms they got,SG/PL happy,PL
      ‘Representatives visited classrooms. They were happy.’

Furthermore, deputado in (25) must be interpreted in contrast to an alternative set. In this respect it should be pointed out that sentence (25) cannot be the answer to quem visitou salas de aula ‘who visited classrooms?’, but could be the reply to a discourse context such as quem visitou salas de aula na festa da escola, deputado ou senador? ‘who visited classrooms at the school party, representatives or senators?’. According to this test, the BN in preverbal position would be diagnosed as Contrastive Focus rather than Focus (Cohen and Erteschik-Shir 2002).

The preverbal BN in (25) could also be a Contrastive Topic in relation to a context where we know that there are congressmen and senators at the party, and somebody claims that congressmen visited classrooms. In this context, uttering (25) relates deputado to a contrasted set specified in the context.

Being either a Contrastive Focus or a Contrastive Topic, deputado must be interpreted as referring to a specific set of individuals and cannot have an indefinite singular interpretation. For an indefinite interpretation either an overt indefinite determiner should be used (as in um deputado visitou salas de aula ‘a representative visited classrooms’), or a bare plural (as in (24)). In other words, semantically, the interpretation of the BN deputado presupposes the existence of a maximal sum of individuals that have the property denoted by the noun:
deputado is an entity-type expression and constrains a maximal interpretation for the discourse referent.

To sum up our analysis of BNs in preverbal position, we postulate a full DP structure (see (22)), with a null D. This null D is morphosyntactically specified for plural Number, which explains the discourse relationships with eles/elas. Semantically, preverbal BNs introduce discourse referents. When the D is definite the BN denotes a maximal entity or a maximal sum, while when the definiteness of the D is cancelled by an adjoined deprivative operator, the BN denotes a non-maximal sum.

6. Final question

Why should a language with definite and indefinite articles, and singular and plural number specification (see (1)-(3)) make use of null Ds (among which we have postulated a null DE operator) in order to express what could be expressed with non-empty morphemes? Do the sentences in (1)-(6) encode the same meaning, as Müller’s (2002) English translations appear to suggest?

Our answer to these questions relies on the assumption that different semantic interpretations are basically constrained by morphosyntactic differences. With this perspective in mind, let us consider again sentences (1) to (6).

First, our analysis predicts that sentences (2), (5) and (6) are not ambiguous. We have shown that the difference these examples show is neither syntactic nor semantic; it is only due to the way Number is morphophonologically spelled out in BrP. The preverbal subjects of these sentences have the structure in (22), repeated in (31), which encodes a definite plural (since the D is specified for plurality) that is interpreted as generic at the time of combining with an i-level predicate.

\[(31) \quad [\text{DP} \ D_{\text{[i-PL]}} [\text{NumP} \ [\pm\text{PL}] \ [\text{NP} \ N ]]] \]

Second, although not discussed in this paper, we would like to claim that sentence in (3) is not ambiguous either. It is associated with sentential genericity, due to the indefinite singular being translated as a variable bound by a generic operator, as in (32).

\[(32) \quad \text{GEN} (\text{Brazilian}(x); \text{Hardworking}(x)) \]

Third, the sentence in (1) is not ambiguous either. It is a generic statement in which an i-level predicate combines with a DP structure (the so-called definite singular construction) that has the structure in (8), repeated here in (33). This structure corresponds to a definite kind interpretation.

\[(33) \quad [\text{DP} \ D \ [\text{NP} \ N ]] \]

Finally, the sequence in (4), with no overt number specification on the Noun and no overt article is the only one that we have discussed to be ambiguous. It can be associated with two distinct structures, namely (33) and (34), and for each of these structures a different meaning is to be obtained, namely the definite kind reading and the definite plural (maximal sum) reading, the latter conveying a generic interpretation at the time the DP is composed with the i-level predicate.
7. Conclusions

We have shown what types of DPs can express generic reference in combination with i-level and k-level predicates in BrP: DPs with an overt/null D not specified for Number, full DPs with an overt/null D specified for plural Number, and an indefinite singular under the scope of GEN.

We have also shown that preverbal BNs in BrP are full DPs with a null D that can be either specified for Number or remained unspecified for Number, in the case of a definite kind interpretation. Preverbal BNs in combination with s-level predicates can also express generic reference in bare habituals in BrP.

Our analysis of indefinite plurals supports Ishane’s (2010:24) claims: “that determination and (in)definiteness (…) depend on portions of structure and not on the presence of a determiner”, and that “bare common nouns with different readings have different edges”.

Finally, our analysis of BNs in BrP suggests that Chierchia’s (1998) Blocking Principle, which restricts covert applications of type-shifts (i.e., if a certain shift is lexicalized it cannot apply covertly), is not applicable to BrP. This language shows covert applications of the iota operator that corresponds to the definite article and, therefore, the Blocking Principle should not be considered a universal constraint.

8. References


Dobrovie-Sorin, Carmen, Tania Bleam & M.Teresa Espinal. 2006. Bare nouns, number and types of


Espinal, Maria Teresa & Louise McNally. 2011. Bare nominals and incorporating verbs in Catalan and Spanish. Journal of Linguistics 47. 87-128.


Pires de Oliveira, Roberta & Susan Rothstein. 2011. Bare singular noun phrases are mass in Brazilian Portuguese. Lingua 121, 2153-2175.


Unaccusatives, Existentials and the Definiteness Restriction: Explorations of the Syntax-Semantic Interface

Susann Fischer
susann.fischer<at>uni-hamburg.de

1. Introduction*

Since Milsark (1974, 1977) who documented this phenomenon for English it has been a well-known fact that certain positions or contexts within sentences require a noun phrase occurring there to have a particular value [-Def] or to be interpreted as having such a value. Restrictions of this kind are termed definiteness restrictions (DR) and have also been found and discussed with respect to different Romance languages (Belletti 1988, Leonetti 2008 among others). See the following examples from English and French1 illustrating that definite (or strong) determiners are excluded in unaccusative (1b, 1d) and existential (2b, 2d) constructions, whereas indefinite (or weak) determiners are allowed, thus exhibiting a clear-cut DR.

(1) (a) There arrives a train ModE
(b) *There arrives the train
(c) Il arrive un train ModF
(d) *Il arrive le train

(2) (a) There is a dog in the garden ModE
(b) *There is the dog in the garden
(c) Il y a un chien dans le jardin ModF
(d) *Il y a le chien dans le jardin

The DR has been discussed within different components of grammar: syntax, semantics, and pragmatics as well as a combination thereof. However, until now no explanation has been able to entirely account for the phenomenon. One of the problems seems to be that existential constructions have been paired up with unaccusative constructions (Belletti 1988, Perlmutter 1978, Lyons 1999, Carnie & Harley 2005, among others) or unaccusatives with existentials (Carnie & Harley 2005, Leonetti 2008) as if they both belonged to the same type of sentence/construction.

This paper is very much work in progress, it is an attempt to bring together the information about the DR concerning the two constructions that is scattered in different papers. New data (from some old and some new Romance languages) will be provided showing that there are more differences than similarities between these constructions. On the basis of these data and the results of the comparison between the two constructions I will argue that the DR in

---

* I would like to thank the audience at the VI NEREUS workshop: Theoretical implications of the syntax-semantics interface in Romance, held at the University of Cologne 9th until 11th November 2012, for critical questions and helpful comments, and the organizers Sofiana Chiriacescu and Klaus von Heusinger for teaching us to draught a perfect Kölsch. Special thanks go to Manuel Leonetti and Tanja Kupisch whose interest and expertise on “definiteness” has been a constant challenge to learn more and see things not only from a syntactic point of view, and to Xavier Villalba for challenging my view on Old and Modern Catalan existentials. All errors are of course mine.

1 In order to show which examples stem from which language I will use the following abbreviations: ModE = Modern English, OCat = Old Catalan, ModCat = Modern Catalan, OF = Old French, ModF = Modern French, ModIta = Modern Italian, OSp = Old Spanish, ModSp = Modern European Spanish, LAS = Latin American Spanish,
unaccusatives is a syntactic restriction and has to be explained in syntactic terms, whereas the DR in existentials is a semantic/pragmatic restriction that has to be explained by referring to the semantics of the constructions and the pragmatics of the context.

Chapter two will give a short overview of the variability of the DR in unaccusatives and existential in four different Romance languages. Chapter three will summarize the differences between the two constructions concerning structure and semantics. Chapter four will discuss different analyses that have been proposed over the last years in order to account for the DR in the two constructions and different languages. In chapter five new data will be presented showing that the explanations put forward so far cannot explain the diachronic development of existential constructions and neither the variability concerning the elements triggering the DR in two varieties of Spanish. It will be shown that additional features, e.g. animacy and word-order, need to be included in order to better understand the DR in existentials, whereas this is not necessary for unaccusatives. Chapter six will summarise and discuss the findings attempts an explanation.

2. The DR in Romance

The DR shows a variable behaviour in the different Romance languages. See the following examples from four Romance languages (French, Spanish, Catalan and Italian) of unaccusative (3, 5, 7 and 9) and existential constructions (4, 6, 8, and 10).

(3) (a) Il est arrivé une fille.
    there is arrived a girl
    ‘There has a girl arrived.’

(b) *Il est arrivé la fille.

(4) (a) Il y a une fille dans le jardin.
    there LOC has a girl in the garden
    ‘There is a girl in the garden.’

(b) *Il y a la fille dans le jardin.

The French examples exhibit the DR in accusative (3a,b) as well as in existential (4a,b) constructions. French being a non-null-subject language needs the expletive il whenever the NP is located postverbally. Furthermore, the verb is conjugated with être (to be), illustrated in example (3). Existential constructions consist of the expletive il, a locative clitic y and the verb avoir (to have) illustrated in example (4).

(5) (a) Ha llegado una chica.
    has arrived a girl
    ‘There arrived a girl.’

(b) Ha llegado la chica.

(6) (a) Hay una chica en el jardín.
    has a girl in the garden
    ‘There is a girl in the garden.’

(b) *Hay la chica en el jardín.

Spanish shows the DR only in existentials (6b). Since Spanish is a null-subject-language no expletive is needed when the subject NP is in a postverbal position. The unaccusative verb (5a) and (5b) is conjugated with haber (to have). The existential construction shows the lexicalised hay (there is/there are). From a diachronic perspective, the origin of the verb-form
hay is an incorporation process of the locative \( y \) (there) – derived from the Latin adverb \( ibi \) – into the third person singular form of the verb \( haber \) (to have).

(7) (a) \( \text{Va arribar una noia.} \) \hspace{1cm} \text{ModCat}
    \hspace{1cm} \text{PAST arrive a girl}
    \hspace{1cm} \text{‘There arrived a girl’}
(b) \( \text{Va arribar la noia.} \)

(8) (a) \( \text{Hi ha una noia al jardí.} \) \hspace{1cm} \text{ModCat}
    \hspace{1cm} \text{LOC has a girl in the garden}
    \hspace{1cm} \text{‘There is a girl in the garden’}
(b) \( \text{Hi ha la noia al jardí.} \)

Catalan does not present a definiteness restriction, neither in unaccusatives (7) nor in existential constructions (8). Like Spanish it is a null-subject-language and therefore it does not need an expletive when NP subjects are located postverbally. Like French, Catalan has a clitic locative \( hi \) (there).

(9) (a) \( \text{È arrivata una ragazza.} \) \hspace{1cm} \text{ModIta}
    \hspace{1cm} \text{is arrived a girl}
    \hspace{1cm} \text{‘There arrived a girl’}
(b) \( \text{È arrivata la ragazza.} \)

(10) (a) \( \text{C’è una ragazza in giardino.} \) \hspace{1cm} \text{ModIta}
    \hspace{1cm} \text{LOC is a girl in garden}
    \hspace{1cm} \text{‘There is a girl in the garden’}
(b) \( \text{C’è la ragazza in giardino.} \)

Italian looks similar to Catalan in that neither in unaccusatives (9) nor in existential constructions (10) the DR applies. Like Spanish and Catalan it is a null-subject language thus postverbal subjects do not need to be completed by an expletive. Like the other Romance languages the existential construction makes use of a clitic locative \( ci \).

The definite postverbal subjects in Spanish, Catalan and Italian without the definiteness restriction has often been termed “free inversion” and has been interpreted as a characteristic of null-subject languages, thus explaining why French as a non-null-subject-language shows the DR, in contrast to e.g. Italian or Catalan. However, something else must be at stake in existentials, since even though Catalan, Italian and Spanish are null-subject-languages only Spanish shows the DR.

3. Structural and semantic properties…

The data in chapter two displayed a difference between existentials and unaccusatives concerning the DR. In the following the structural and semantic properties of the two constructions will be presented and discussed.

3.1 ... of unaccusatives

Unaccusatives belong to the class of intransitive verbs. It is generally assumed that they include a single argument within their subcategorization frame, namely the internal argument of the VP as illustrated in (11) (cf. Belletti 1988, Perlmutter 1978 and others).
Unaccusatives vs. Existentials

(11) Unaccusatives

\[
\begin{array}{c}
\text{VP} \\
| \\
\text{V'} \\
\text{2} \\
\text{V°} \\
\text{DP}
\end{array}
\]

Thus, the postverbal subject or rather the syntactic pivot of these constructions is an object in disguise, and the theta role which is assigned to this argument is never [AGENT] but rather [THEME]. This is the reason why they have also been discussed together with psych verbs (Belletti & Rizzi 1988).

A number of structural criteria have been proposed in order to detect unaccusativity. The most well-known criteria is auxiliary selection in languages that use two different temporal auxiliaries (have and be) for analytic past/perfect verb forms: e.g. German: *ich bin gefallen* (I have fallen); French: *je suis tombé*; Italian: *sono caduto*). It has also been proposed that be encodes aspectuality (Sorace 2000, Grewendorf 1989). More specifically it was proposed that unaccusative verbs tend to express a telic change of state or location (12), while verbs choosing have tend to express an agentive activity which does not involve telicity (13).

(12) (a) La fille est arrivée  
(b) La ragazza è arrivata  
ModF  
ModIta  
the girl is arrived  
‘The girl has arrived’

(13) (a) La fille a parlé  
(b) La ragazza ha parlato  
ModF  
ModIta  
the girl has spoken  
‘The girl has spoken’

In Italian (and also Southern German\(^2\)) some verbs can be constructed with either *essere* or *habere* (14) denoting an aspectual difference.

(14) (a) Le mele sono marcite  
(b) Le mele hanno marcito al sole  
ModIta  
ModIta  
the apples are rotten  
the apple have rotted in-the sun  
‘The apples are rotten’  
‘The apples have been rotting in the sun’

*Sono marcite* expresses a completed process, i.e. that the apples are actually rotten, whereas *hanno marcito* denotes the duration of the action.

Another criterion is passivisation. Under the assumption that the crucial characteristic of the passive is the absorption of the subject theta-role, it is clear that no such absorption is

---

\(^2\) In Frankonian and Swabian the difference between using *be* vs. *have* with the following (and many other) verbs clearly denote an aspectual difference:

(i) Ich bin gestanden. vs. Ich habe gestanden.  
I is.1SG stood  
‘I stood.’  
I have.1SG stood  
‘I have been standing.’

(ii) Ich bin in den Raum getanzt. vs. Ich habe drei Stunden getanzt.  
I is.1SG in the room danced  
‘I danced into the room.’  
I have.1SG three hours danced  
‘I have danced for three hours.’
possible in the case of unaccusatives (15).

(15) (a) *The man was arrived  
         (b) *La fille a été arrivée  
         (c) *La ragazza è arrivata  
         (d) *La chica ha sido llegado  
         (e) *La noia ha estat arribada  

The literature on –er nominalisation has established the so-called external argument generalization, i.e. –er nominals typically denote the external argument of the underlying predicate, irrespective of the specific thematic role the argument has (Rappaport Hovav, Levin 1992, Schäfer 2008). Since the one argument in unaccusative verbs is the internal argument and not the external these verbs do not allow –er nominalisation (16).

(16) (a) *Arriver  
         (b) *Arrivateur  
         (c) *Arrivatore  
         (d) *Llegador  
         (e) *Arrivador  

En cliticisation in French (17) and ne cliticisation in Italian (18) is another criterion that has been proposed in order to detect unaccusativity. It was claimed that the partitive clitics en/ne can only cliticise out of the object position (Burzio 1986, Müller 2000). This being so, the behaviour of these clitics is used to test whether the quantified argument was an internal (17a, 18a), or external argument (17b, 18b).

(17) (a) Trois en arriveront demain.  
         three PART arrive.FUT tomorrow  
         ‘Three of them will arrive tomorrow.’  
         (b) *Trois en paraitront demain.  
         three PART appear.FUT tomorrow  
         ‘Three of them will appear tomorrow.’

(18) (a) Ne sono arrivato molti  
         PART are arrived many  
         ‘Many (people) have arrived’  
         (b) *Ne hanno telefonato molti  
         PART have called many  
         ‘Many (people) have called.’

In Catalan, however, both the internal (19a) and the external (19b) argument can be substituted by the partitive clitic en. Thus it is obvious that the partitive clitic in Catalan cannot detect the internal argument. Since all other tests also hold for the Catalan unaccusatives, it might be due to the categorical status of the partitive clitic in this language (cf. Bonet 1991).
Summarizing it can be said that unaccusatives denote a telic change of state or location and that the criteria applied show that the syntactic pivot of unaccusative constructions seems to be the internal argument and not the external argument, and that this argument never carries the thematic role [AGENT].

3.2 ... of existentials

Existential sentences are used to express a proposition about the existence or presence of someone or something. They primarily introduce a new referent into the discourse, that is why they are sometimes also called *thetic* sentences, i.e. purely thematic, or topicless, or alternatively if they are assumed to have a theme-rheme structure, the theme is hypothesized to be a location rather than the referent of the pivot (Leonetti 2008, McNally 2011).

Concerning the underlying structure of existentials different views are held in the literature. In English *be* is a copula and as such it appears that the single argument is a complement (cf. Lyons 1999: 238). As a matter of fact, the intransitive verbs which occur in existential sentences have sometimes been analysed as “unaccusatives” (Perlmutter 1978, Burzio 1986), i.e. they subcategorise for an NP-complement but do not assign Case (see also Maling 1988: 170). Under this view the structure is like the structure in (11) for unaccusatives.

In general however, copular verbs such as *to be, être, essere, and haber* are said to be defective, they are not capable of assigning a theta role (for a different view see Löbel 2009). Moro (1991, 1997) building on Heggie (1988) argues for a small clause analysis where the small clause consists of an NP/PP and an expletive (20), and where the expletive needs to move to SpecIP always (20).

Languages differ in the way they express existential meanings. In many languages an expletive subject is required like English *there* or French *il*, however there are many languages without the need of an expletive subject, e.g. Spanish and Catalan. Existential constructions do not always contain a verb\(^1\), if there is one it is often homophonous with a verb meaning ‘to be’ or ‘to have’.

\(^1\) Russian doesn’t need a copula in present tense sentences.
Italian uses c’è, 3SG of the verb esserci “to be.loc”, French il y a and Catalan hi ha use the 3SG of the verb avoir/haver "to have" together with a locative clitic, Spanish makes use of the element hay which shows an incorporated locative together with the 3SG verbal form of haber (to have). Although Italian, Catalan and French use the locative clitic in other contexts as well, it is generally assumed that its use in existential constructions is lexicalised. Thus, what all existential sentences seem to include is such a locative, either as an expletive (English there), or as a clitic (Italian ci, Catalan hi, French y) or incorporated into the verb (Spanish hay). The presence of such a locative is the reason why existential sentences are often grouped and discussed together with locative constructions (e.g. Freeze 1992, for a different view see Leonetti 2008). Thus, with the exception of English, unaccusative constructions – in contrast to existential constructions - do not include locative elements in the languages investigated here; ignoring for the moment how semantically bleached these elements are in the existential constructions.

Another difference between existentials and unaccusatives is that existential constructions do not express a telic change of state or location but as mentioned before the existence and presence of someone or somebody. It seems correct to claim that only atelic events are possible in existential constructions. Milsark (1977: 12) argues that those adjectives which are permitted in the coda of existential constructions denote states (21a) and (22), i.e. they are state-descriptive predicates in Milsark’s (1977) terminology, and stage-level predicates in more modern terminology. According to McNally (2011: 1845) this fact seems to be uncontroversial. In her analysis (McNally 2011: 1846) she proposes that the coda serves to restrict the spatio-temporal parameters within which the referent of the pivot is instantiated. Under this view individual level predicates are ruled out because they lack the ability to provide the necessary spatio-temporal restriction.

(21) (a) There is a man drunk.  
(b) *There is a man intelligent.

(22) (a) Il y a un homme ivre.  
(b) Hay un hombre borracho.  
(c) C’è un home ubriaco  
(d) Hi ha un home borratxo.

Carnie & Harley (2005: 46) argue on the basis of standard tests for presuppositionality, that presupposition of the process component of the bipartite event structure in telics is forced by the assertion of the endpoint component and that presupposed elements are not pragmatically consistent with existentials. In their view telic predicates are therefore ruled out in existential constructions (Carnie & Harley 2005: 47).

Summarizing the facts so far, the structure and semantics of existentials seem to vary in the different languages and also compared to unaccusative constructions: telic predicates in unaccusatives vs. atelic predicates in existentials, the use of a locative elements in existential constructions, whereas unaccusatives (with the exception of English) do not include any locative element. As for the similarity in both constructions the pivot never denotes the thematic role [AGENT]. As for unaccusatives the available tests have shown that the argument is never the external but the internal argument, whereas in existential sentences this is not clear. The tests for unaccusativity can not be applied, since copular verbs do not subcategorize for internal and/or external arguments (cf. Lohndal 2006, Löbel 2009). The pivot of existentials seems to be a small clause or a NP adjoined to the VP (see Moro 1997, Leonetti 2008) and not the internal argument of the copula verb.
Thus, in order to explain the DR we need to take into account the syntax, the argument structure and the semantics of existential and unaccusative constructions, looking more at their differences than at their similarities.

4. Previous analyses

Many analyses have been advanced in order to explain the DR in unaccusatives and existentials. In the following a short overview will be given.

4.1 Unaccusatives

For a long time, the definiteness effect in Modern English and Modern French unaccusatives has been explained with Diesing’s (1992) Mapping Hypothesis, i.e. syntactic positions are directly mapped to semantic interpretation and the observed definiteness restriction in (23b) and (24c) are due to the low VP-internal position of the subjects. These low subjects are mapped to nuclear scope and therefore get a weak existential reading.

\[
\begin{align*}
(23) \quad & (a) \text{ There arrives suddenly [} \text{VP a train]} \quad \text{ModE} \\
& (b) \text{ A train arrives suddenly} \\
& (c) \text{ *There arrives suddenly [} \text{VP the train]} \\
& (d) \text{ The train arrives suddenly}
\end{align*}
\]

\[
\begin{align*}
(24) \quad & (a) \text{ Il arrive souvent [} \text{VP une fille]} \quad \text{ModF} \\
& \text{ there arrives often a girl} \\
& \text{ ‘There arrives often a girl.’} \\
& (b) \text{ Une fille arrive} \\
& \text{ a girl arrives} \\
& \text{ ‘A girl arrives.’} \\
& (c) \text{ *Il arrive souvent [} \text{VP la fille]} \\
& \text{ there arrives often the fille} \\
& (d) \text{ La fille arrive} \\
& \text{ the girl arrive} \\
& \text{ ‘The girl arrives.’}
\end{align*}
\]

It has been also shown for Modern Catalan (25) and Modern Spanish (cf. Vallduví 2002, Ordoñez 1998, Sheehan 2004 for many Spanish examples) that in VS orders the verb and postverbal subject may be interrupted by adverbials.

\[
\begin{align*}
(25) \quad & (a) \text{ Ahir va tornar a l’India [un funcionari]} \quad \text{ModCat} \\
& \text{ yesterday PAST return to India a official} \\
& \text{ ‘Yesterday, a functionary returned to India.’} \\
& (b) \text{ I el cap de setmana arribarà tranquilament [la Nuria …]} \\
& \text{ and the end of week arrive.FUT untroubled the Nuria} \\
& \text{ ‘And on the weekend, Nuria will calmly arrive’}
\end{align*}
\]

Under Diesing’s Mapping Hypothesis these subject should get a weak existential reading when being inside VP. However, in Catalan and Spanish definite subjects in the postverbal position also follow the adverbials and are presumably in a VP internal position, nevertheless they don’t seem to get a weak existential reading.

The SV/VS alternation without any DR that we see in (26) and (27) has been often called free inversion and was argued to be a characteristic of null-subject-languages.
Under the general concept of minimalism where movement needs to be triggered and is a “last resort” option in order to check of uninterpretable features, free inversion is no longer an acceptable explanation. In order to explain the lack of the DR in null-subject-languages Barbosa (1995) and Alexiadou & Anagnostopoulou (1998) proposed that the DR only applies in those languages where an XP needs to check of the EPP feature in $I^0/T^0$, i.e. in non-null-subject-languages. By contrast in null-subject-languages the EPP is checked by the verb, by a head ($X^0$) and that is why no DR applies. As a consequence of the verb checking the EPP in null-subject-languages all preverbal subjects need to be analysed as topics. Sheehan (2004) (see also Silva-Villar 1998) building on the idea of a split in how to check of the EPP and the relation to the DR, specified the argument further. They argued that the possibility to check of the [EPP] on $T^0$ in non-null-subject-languages is dependent on the expletive: if an expletive is introduced the definiteness restriction applies, if no expletive is introduced no DR applies. Looking at the old strata of English⁴ (28) and also French⁵ (29) we notice that none of the explanations hold.

(28) (a) & *come a culur se briht as þeah ha bearnde of heouene* OE
   and come a dove so bright as though it burnt of heaven
   ‘And there came a dove from heaven so bright as though it burnt.’
   (Margarete 89.564, cited in Biberauer & Roberts 2003)

---

⁴ In the following many European Spanish, Latin American, Old Spanish, Old and Modern Catalan examples are translated into English including a definite DP in unaccusative and/or existential constructions. The English examples including definite DPs are all ungrammatical, nevertheless a free translation into English is included.

⁵ All examples of Old and Middle English are taken – if not indicated otherwise – from the York-Toronto-Helsinki Parsed Corpus of Old English Prose (YCOE), a morpho-syntactically annotated corpus of more than 1.5 million words of prose text that was created by Ann Taylor, Anthony Warner, Susan Pintzuk and Frank Beths and from the Penn-Helsinki Parsed Corpus of Middle English, 2nd ed. (PPCME2) that was edited by Anthony Kroch and Ann Taylor. More detailed information on the YCOE and PPCME2 are available at [http://www-users.york.ac.uk/~lang22/YCOE/](http://www-users.york.ac.uk/~lang22/YCOE/).

Unaccusatives vs. Existentials

(b) *and cuman yfele men and bereafian hine*
and came evil men and bereave him
‘and there came evil men and bereaved him’

(c) *þonne he cume*
when he comes
‘when he comes’

(d) *þere ben 5 provinces*
LOC are 5 provinces
‘Five provinces are there ....’

In (28a/b) no expletive is used in the unaccusative construction as if Old English were a null-subject-language, nevertheless the DR applies. In (28b) and many other sentences pronominals are attested thus the status of Old English as being a null-subject-language is not clear. However, in the very first texts expletive drop is the only available means in unaccusative constructions. *þere* (there) used as an expletive and not as a locative adverbial appears only from the 15th century onwards (cf. Ingham 2001). In (28d) *there* needs to be interpreted as a deictic element, a locative.

(29) (a) *quant il durent ariver une turmente*
when it shall arrive a storm
‘when a storm will arrive’

(b) *en toteneis est arrivez plusurs reis*
in now is arrive some kings
‘and now there arrive some kings’

(c) *en walcres arriva rois antiaumes*
in w. arrive kings antiaumes
‘in Walcres arrive Antiaume kings’

In (29a) we see an expletive and the DR applies, in (29b) an adverbial precedes the unaccusative verb and the DR applies, and in (29c) a locative is located in front of the unaccusative verb and also the DR applies.

If Old French and Old English are analysed as null-subject-languages the verb checks the EPP and thus no DR should apply (cf. Alexiadou & Anagnostopoulou 1998). However, if they are analysed as non-null-subject-languages the expletive can obviously be dropped (28, 29) and we should then find sentences where the DR does not apply. In Old English (28) and Old French (29), however, we see the DR applying independent of whether the expletive is dropped or not (and so far I didn’t find any counter-examples).

Traditional approaches analysed Old French (Boucher 2003) and Old English (Kroch 2002) as a null-subject-language. However, it has been argued by Roberts (1995), Rinke & Meisel (2009) that Old French did not exhibit the stereotypical syntactic behaviour of a null-subject language. Fischer (2008, 2010) and Zimmermann (2012) argue that Old French is not
a null-subject-language but a non-null-subject-language. The same has been proposed for Old English (Fischer et al. 2000, and Fischer 2010) “Old English did not allow referential pro drop and only limited expletive pro drop” (Fischer et al. 2000: 39).

In Fischer (2008, 2010), I have proposed that Old French and Old English are indeed non-null-subject-languages and that is why the DR always applies independent of whether an expletive is introduced or not. But since Old English and also Old French show object shift and VSO orders (Biberauer & Roberts 2003, Zaring 1998) the inverted subject can be outside VP (30-31) and is therefore close enough to the verb to check the EPP feature, thus no expletive needs to be introduced.

(30) (a) (...) þei shuld no meyhier haue
OE
they should no mayor have
‘they were not allowed to have a major’
(Chronicles 62,23)

(b) (...) þæt he mehte his feorh generian
OE
that he want his live save
‘(...) that he wanted to save his life’
(Or 2.5.48.18)
(cited according to Biberauer & Roberts 2003)

(31) (a) (...) por coi avès vos ce fait
OF
why have you this made
‘(...) why have you made this’
(Ar148-53)

(b) (...) les gens qui ont accoustumé a ce faire,
OF
the people who have accustomed to this make
‘(...) the people who are accustomed to do this.’
(StL189)
(cited according to Zaring 1998: 321)

In those sentences where an expletive is introduced the postverbal subject is in a different position, presumably inside the VP, in any case not close enough to the verb to check of the EPP. It is in these sentences that an expletive is needed to check the EPP in I°/T°. Independent of whether it is the expletive or the postverbal subject to check the EPP the DR applies, in contrast to null-subject-languages where the verb itself checks the EPP and no DR applies. Under this view (Fischer 2010) the DR in unaccusatives underlies strictly syntactic conditions.

4.2 Existentials
Many theories have been proposed in order to account for the DR in existential constructions. We find “syntactical” approaches (Freeze 1992, Safir 1985, Moro 1997) next to pragmatical (Abbott 1993 & 2005, Lumsden 1988), semantic (Milsark 1974, Keenan 1987) and a combination thereof (Leonetti 2008, McNally 2011).

The syntactic approaches w.r.t. English (Freeze 1992, Moro 1997) are similar to those proposed for unaccusatives. Thus there constructions are derived from their copular sentences by specific transformational rules.
(32) (a) There is a burglar in the basement
(b) Il y a un voleur dans le sous-sol
(c) Hay un ladrón en el sótano
(d) C’è un ladro nel sotterraneo
(e) Hi ha un lladre en el soterrània

‘There is a burglar in the basement’

In languages other than English this analysis is difficult to maintain (already under a GB analysis) since different verbs are used for existential and other copular sentences. This shows that these constructions do not share the same underlying structure. In minimalist terms the assumption runs into difficulties for English as well, since the numeration set of sentences with the expletive (there) and without the expletive is different and therefore the two sentences cannot compete for spell-out.

In pragmatic approaches (e.g. Abbott 1993: 44) existential constructions are considered to represent thetic sentences introducing a hearer-new referent into the discourse. (32) can open up a conversation whereas (33) is far more restricted, since it cannot serve to introduce a new referent or a new situation into the discourse in an “out-of-the-blue” contexts:

What happened?

(34) (a) Well there was a burglar in the basement.
(b) Bon, il y a un voleur dans la cave.
(c) Pues, había un ladrón en el sótano.
(d) Beh, c’è un ladro nel sotterraneo

(35) (a) *Well, a burglar was in the basement.
(b) * Bon, un voleur était dans la cave.
(c) *Pues, un ladrón estaba en el sótano.
(d) Beh, c’è un ladro nel sotterraneo

This shows that (32, 34) are existential constructions whereas (33, 35) should not be analysed as such.

Milsark (1974, 1977) who presented the first extensive study on the semantics of existentials, provides a typology for determiners. He classifies the different determiners and noun phrases as either strong or weak. According to him, the weak determines (a, many, some) involve cardinality, whereas the strong one (the, every, all) are quantificational. Milsark (1977, see also Lyons 1999: 240) analyses there be as an existential quantifier which has to bind a variable.
(36) (a) There is a man in the garden
(b) *There is the man in the garden
(c) The man is in the garden

(36a) makes a claim of existence about the entity described by the complement phrase, (36c) doesn’t. The DR applies in (36b) since strong determiners or noun phrases are themselves quantificational, i.e. a definite noun phrase complement does not provide a variable that is free for binding. The weak determiner in (39a) however is not necessarily quantificational and therefore provides a variable that the existential quantifier can bind.

Semantic-pragmatic approaches like e.g. Leonetti (2008) and McNally (2011) also include the syntactic component into their considerations. Leonetti (2008) thoroughly investigates the different possible structures that have been discussed under the heading existentials and provides data showing that there are at least three types of *there*-constructions (37).

(37) (a) Proper existential: There is a dog in the garden.  
(b) Eventive existential: There is a dog barking.  
(c) Enumerative existential (list): There is Peter, Mary, and Paul.

Proper existential – the only true existentials – in Leonetti’s (2008) view (see also Milsark 1974:20) always allow for a coda that provides a specification of their location, or other relevant relation to the discourse. He argues that indefinites in proper existentials are compatible with the locative coda within the same VP or dislocated, whereas definites need the coda to be dislocated at all costs. The fact that definite DPs are acceptable with a narrow focus reading when the coda is dislocated has already been noticed by Belletti (1988). See the Italian (37) and Catalan (38) examples created according to Leonetti (2008).

(38) (a) C’è una statua di Michelangelo, in Piazza della Signoria  
(b) ??C’è la statua di Michelangelo in Piazza della Signoria

(39) (a) Hi ha una estàtua de Joaquín Suárez, a la Plaça de l’Avenida Agraciada  
(b) Hi havia la estatua de Joaquín Suárez, a la Plaça de l’Avenida Agraciada.

(38a) and (39a) with an prosodic break, indicated by a comma, are true existentials, the sentences in (37b) and (38b) are not according to Leonetti (2008). The definite articles in these sentences do not license a coda. This seems like a convincing argument, even though I will show (section 5) that it is even more complex than the distribution proposed here. Furthermore, it needs to be noted that there is a difference in acceptance between Italian and Catalan. Catalan are less strict in allowing a definite DP without the dislocated coda, the five

7 The dislocation of the coda in these examples is marked by the comma indicating a prosodic break.
native speakers I checked (from Barcelona) rated these kind of examples (39b∗) as grammatical.

Another piece needs to be added to the jigsaw of existentials and this is word-order. Italian and Catalan unlike Spanish allow VOS / VXS with a narrow focus on the postverbal subject (see Leonetti 2008, Vallduví 1993, 2000) which is exactly the word-order and the position in which definite DPs are allowed. VSO orders are excluded in Modern Catalan (Vallduví 2000) and in Italian when the subject is definite (Belletti 1988, Leonetti 2008). This can be seen in existentials but also in other sentence types and was also discussed by Kaspers-Mahne et.al. (2004: 572ff) for subject inversion in French.

It is a well-known fact that word-order plays an important role for information structure, even in languages that are no longer considered to be “topic prominent languages” (Lehmann 1975, see also Calboli 1978). What we see in Catalan is exactly this: definite DPs are allowed in existentials since they do not immediately follow the verb, the canonical word-order in Catalan being VOS (Vallduví 1993, 2002). This postverbal subject position is presumably outside VP and thus escapes the weak existential interpretation, and it seems plausible to assume that this position is a grammaticalised topic position (cf. Lehmann 1976, Li & Thompson 1975) where however in Modern Catalan the subject can be/is located.

5. Some more data

In the following some new data of historical Catalan and historical European Spanish and Modern Latin American Spanish will be provided showing that word-order and its connection to information structure is in fact one of the main ingredients for the explanation of the DE.

5.1 Some new data of modern languages

Building on Abbott’s (2005) intuition that the weak-strong distinction cannot be thought of as a binary opposition, instead it should rather be considered a scale, Zielke (2012) carried out a study on the acquisition of the DR and found a tripartite distinction.

\[(40)\] weak expressions > less strong expressions > strong expressions
- indefinite article - definite article - definite article
- weak quantifier referent: inanimate referent: animate, unique identifiable
- zero article, bare noun - strong quantifiers - strong quantifiers
- - - proper names - possessives

The focus of her study was the acquisition of the DR in L2 European Spanish and L2 Latin American Spanish by L1 German and L1 Turkish speakers by means of an acceptability judgment task. Her main goal was to show that L1 Germans outperform the L1 Turkish group since the DR plays out similarly in German and Spanish but differently in Turkish. As expected by Zielke (2012), the L1 German group outperformed the L1 Turkish group on this phenomenon due to positive L1 transfer. However, in my view the far more interesting result is the tripartite distinction concerning the expressions triggering the DR that was manifested in the control group of Latin American Spanish speakers compared to the European Spanish speakers.

The native speakers of Latin American Spanish (LAS) only showed the DR in strong expressions, whereas the native speakers of European Spanish (ES) showed the DR both with less strong and strong expressions (Table 1).
Table (1): The weak strong distinction in European and Latin American Spanish (Zielke 2012: 30)

<table>
<thead>
<tr>
<th></th>
<th>weak expressions</th>
<th>strong expressions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>less strong</td>
<td></td>
</tr>
<tr>
<td></td>
<td>expressions</td>
<td>strong expressions</td>
</tr>
<tr>
<td>ES</td>
<td>no DE</td>
<td>DE</td>
</tr>
<tr>
<td>LAS</td>
<td>no DE</td>
<td>DE</td>
</tr>
</tbody>
</table>

What the two languages seem to share is that strong expressions that are specified for [+animate] can never occur in existential sentences. Zielke (2012) therefore argues for a tight connection between definiteness and animacy. See the following example of Latin American Spanish where the animate definite DP is forbidden (41a), but the inanimate definite DP is accepted by the speakers (41b). Both sentences are unacceptable in European Spanish.

(41) (a) *Hay todas las personas de la empresa
has all the people of the press
‘There are all the people from the press’
(b) Hay todas las películas
has all the films
‘There are all the films’

(Zielke 2012: 88)

5.2 Some new data of old languages

Looking at Old Spanish and Old Catalan we see that no change is detectable concerning the two constructions. Already in Old Spanish the DR is active (42-44) in existentials only.

(42) Entre essas compañas de casa, avié un omne bueno, que perdiera el viso;
among these neighbours of house has a man good that lost the radiance
‘Among these neighbours of the house, there was a good man that lost his glamour;’
(BerCEO: Martirio de San Lorenzo. Espasa-Calpe. ZIV eIIId 19/3)

(43) En Coloña la rica, cabeza de regnado, / avié un monasterio de Sant Piedro.
in colony the rich head of reign was a monastery of Saint Peter
‘In the rich colony, head of the reign, was a monastery of Saint Peter.’
(BerCEO: "Milagros de Nuestra Señora" § 160)

(44) (...) ca en l’otro ay cinquenta enxiemplos et en éste ay ciento.
since in the’other are fifty examples and in this are hundred
‘Since in the other there are fifty examples and in this one there are hundred.’
(DON JUAN MANUEL: "El Conde Lucanor". Tercera Parte: Escusación de Patronio al Conde Lucanor. Pp 288)

8 For the Old Spanish different texts: Milagros de Nuestra Señora and Martirio de San Lorenzo by Gonzalo BerCEO (Date of composition around 1252); El Conde Lucanor by DON JUAN MANUEL (Date of composition around 1335); General Estoria, Las siete partidas, Setenario by Alfonso X (Date of composition around 1270); Calila e Dimma (anonymous date of composition between 1195-1207); Mio Cid (anonymous dated around 1207). Thanks to Mario Navarro who checked the texts for existentials and unaccusatives and who created the excel table for further investigation.
Sentence (42) shows an indefinite animate DP, (43) an indefinite inanimate DP and in (44) we see an inanimate plural DP.

In unaccusatives no change seems to have happened from Old Spanish to Modern Spanish. Already in the Old Spanish texts unaccusative sentences without the DR are attested (45-46).

(45) (...) que llegue la sennal de la otra estrella;  
that arrive the signal of the other star  
‘... that there arrived the signal of the other star;’

(Alfonso X, Folio 97R)

(46) Destroyda es Moab. & las sus cibdades. subieron los enemigos sobrelos.  
destroyed is M. & the his townsfolk ascended the enemies over.them  
‘Moab and his townsfolk are destroyed. There ascended the enemies to them.’

(Alfonso X, Folio 97V)

In both unaccusative constructions a definite DP is used. In (45) it is an inanimate DP whereas in (46) it is a definite animate plural DP.

The Catalan data is not as clearcut, we get a deviant pattern. On the one hand, concerning the existential construction we detect a change compared to Modern Catalan, on the other hand no change is attested concerning the unaccusative construction.

In Modern Catalan, as discussed in section three, existentials are accepted by native speakers of Catalan that allow definite DPs not only with a dislocated coda but also without a dislocated coda. In Old Catalan this doesn’t seem to be the case. No definite DP has been attested in any of the three Old Catalan texts

9 and neither in the corpus CICA

10 that we examined with respect to existential and unaccusatives. Thus, all existentials showed a strict DR with inanimate (50) and animate DPs (47-49) in contrast to Modern Catalan.

(47) Fèlix --dix l’ermità--, en una terra havia un rey qui molt amava justícia  
Felix say the’ermit in a land was a king who much loves justice  
‘Felix says the ermit, there was a king in a land who loved justice very much.’

(Llull, Felix 30, 5-6)
(48) *E entorn de tots los dessusdits havia molts falcons*, OCat
and round of all the said had many hawks
‘And around of all the said there were many hawks.’
(Metge, Somni 18, 6-8)

(49) *car aquí havia dos cavallers qui volien defendre per batallal a honor* since here were two knights who wanted defend for battle the honour
*de la duquessa* of the duchess
‘since here there were two knights who wanted to defend by battle the honour of the duchess.’
(Curial 67, 4-5)

(50) *encare.t prec que, si y ha alcunes coses,* still.2sg.acc plea that if has some things
‘I ask you now that if there are some things.’
(CICA_Diàlegs St. Grego, 61)

Another change that took place from Old to Modern Catalan, is the fact that word-order changed more dramatically than for example in Spanish. In Old Catalan all different word-orders were possible including VSO (Fischer 2010, to appear), in Modern Catalan word-order got fixed on VOS (Vallduví 2002), also allowing SVO but never VSO. Thus, I conclude that only when word-order got fixed on VOS, the DR is no longer obligatorily in Catalan existentials, assuming that the reason for this is the fact that postverbal subjects in Catalan are allowed in a position where they can receive a topic and/or Narrow Focus interpretation in contrast to Modern Spanish where the postverbal subject position does not receives a narrow focus or topic interpretation, but always a broad focus interpretation.

No change has been attested over the centuries concerning the unaccusative constructions. Already in the Old Catalan definite DPs (51) and proper names (52) are possible. Thus, no DR applies in unaccusatives.

(51) *Tantost e sens triga vengueren Jacob e Curial* immediately and without delay arrived Jacob and Curial
‘There arrived Jacob and Curial immediately and without delay.’
(Curial 72/2)

(52) *Mas aprés dos dies (...), arribà lo prior, e tots eren tornats sinó Ypòlit.* more after two days arrived the prior, and all were returned but Y.
‘After two days (...), there arrived the prior, and all others had returned except for Y.’
(CICA_Martorell, Tirant, 729, 7).

For a better overview of where the DR has been attested in the different languages over the centuries, two tables which include the different factors (null-subject-language (NSL), non-null-subject-language (NNSL) weak DPs, strong-inanimate DPs, strong-animate DPs) mentioned before are given below.
### Table (2): Definiteness Restrictions in Unaccusatives

<table>
<thead>
<tr>
<th></th>
<th>Weak Expressions</th>
<th>Strong Expressions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>indefinite articles, weak quantifiers, zero articles &amp; bare nouns</td>
<td>definite articles + inanimate strong quantifiers + inanimate</td>
</tr>
<tr>
<td></td>
<td></td>
<td>definite article + animate, strong quantifiers + animate Proper names</td>
</tr>
<tr>
<td>NNSL OEng</td>
<td>no DR</td>
<td>DR</td>
</tr>
<tr>
<td>NNSL ModEng</td>
<td>no DR</td>
<td>DR</td>
</tr>
<tr>
<td>NNSL OFrench</td>
<td>no DR</td>
<td>DR</td>
</tr>
<tr>
<td>NNSL ModFrench</td>
<td>no DR</td>
<td>DR</td>
</tr>
<tr>
<td>NSL Europ.Span</td>
<td>no DR</td>
<td>no DR</td>
</tr>
<tr>
<td>NSL LAS</td>
<td>no DR</td>
<td>no DR</td>
</tr>
<tr>
<td>NSL OKat</td>
<td>no DR</td>
<td>no DR</td>
</tr>
<tr>
<td>NSL ModKat</td>
<td>no DR</td>
<td>no DR</td>
</tr>
</tbody>
</table>

In unaccusatives there is a clear-cut difference between non-null-subject-languages and null-subject-languages. The DR applies in all non-null-subject-languages whereas it does not apply in null-subject languages, irrespective of whether the strong DPs are animate or inanimate.

### Table (3): Definiteness Restrictions in Existentials

<table>
<thead>
<tr>
<th></th>
<th>Weak Expressions</th>
<th>Strong Expressions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>indefinite articles, weak quantifiers, zero articles &amp; bare nouns</td>
<td>definite articles + inanimate strong quantifiers + inanimate</td>
</tr>
<tr>
<td></td>
<td></td>
<td>definite article + animate, strong quantifiers + animate Proper names</td>
</tr>
<tr>
<td>NNSL OEng</td>
<td>no DR</td>
<td>DR</td>
</tr>
<tr>
<td>NNSL ModEng</td>
<td>no DR</td>
<td>DR</td>
</tr>
<tr>
<td>NNSL OFrench</td>
<td>no DR</td>
<td>DR</td>
</tr>
<tr>
<td>NNSL ModFrench</td>
<td>no DR</td>
<td>DR</td>
</tr>
<tr>
<td>NSL Europ.Span</td>
<td>no DR</td>
<td>no DR</td>
</tr>
<tr>
<td>NSL LAS</td>
<td>no DR</td>
<td>no DR</td>
</tr>
<tr>
<td>NSL OKat</td>
<td>no DR</td>
<td>no DR</td>
</tr>
<tr>
<td>NSL ModKat</td>
<td>no DR</td>
<td>no DR</td>
</tr>
</tbody>
</table>

Concerning the existentials the picture is not as clear-cut. It seems more factors need to be included in explaining the occurrence of the DR. First, we need to differentiate between strong animate DPs and strong inanimate DPs, and as has been shown with respect to the Catalan data, it seems we also need to include word-order since it is directly connected to information structure in our interpretation of the data.

### 5. Attempting a summary

Summarizing the findings on unaccusatives and existential constructions concerning the different languages presented here, we get a complex picture concerning existentials and a less complex picture concerning unaccusatives.

It seems correct to claim that whether the DR applies in unaccusatives or not is dependent on whether a language is a null-subject-language or not. Independent of whether an expletive is introduced or not, independent of whether an animate DP or an inanimate DPs is used, and also independent on word-order the DR applies in non-null-subject-languages and it does not apply in null-subject-languages. It has been argued that the reason for this is the fact that in
null-subject-languages it is the verb that checks the EPP, thus that postverbal subjects can be in different positions outside the VP, thus escaping weak existential interpretation (Fischer 2010). Thus, I propose that the DR in unaccusatives is a syntactic phenomenon that needs to be explained by syntactic features.

Concerning the existentials the picture is complex. The DR applies irrespective of whether a language is a non-null-subject-language or not. It has been shown that the DR in existentials is far more dependent on pragmatic and semantic features than in unaccusatives. The data presented here has shown that the weak - strong distinction is not sufficient to explain the occurrences of the DR in Latin American Spanish compared to European Spanish. In order to capture the difference the semantic feature [+/- animate] needs to be included (cf. Zielke 2012). Furthermore, word-order and its connection to information structure, plays an important role in explaining the DR. Data was presented showing that in Old Catalan the DR applied obligatorily in all existential sentences. Only since word-order has been fixed on VOS existential sentences are attested in which the DR does not apply. It seems to be correct that the reason for this lies in the different information structure that Modern Catalan has compared to Old Catalan.

Thus, against Lyons (1999: 46) claim that the DR “is more likely to be a semantic or pragmatic constraint than a syntactic one“, I suggest that the DR in unaccusatives is a syntactic restriction and has to be explained in syntactic terms, whereas the DR in existentials is a syntactic, semantic and pragmatic restriction that has to be explained by referring to the semantics of the constructions, the pragmatics of the context and the word-order, i.e. information structure of the language investigated.

6. References


Latin *alicuis* as an epistemic indefinite

Chiara Gianollo

chiara.gianollo<at>uni-koeln.de

1. Introduction

The Latin indefinite marker *alicuis* ‘some’ is continued by Romance forms like It. *alcuno*, Fr. *aucun*, Sp. *algún*, Pt. *algum*, which derive from the univerbation of *alicuis* with *unus* ‘one’ (*ALICUNUM*). *Alicuis* undergoes a typologically rare and theoretically puzzling diachronic development in Romance: in Latin it displays features typical of specific indefinites, whereas one of the functions (in some languages the only function) of the Romance continuations in the singular is that of a negative polarity item. That is, Romance derivatives of Latin *alicuis* are licensed in contexts, like the scope of negation, which were precluded to Latin *alicuis* and which, more in general, are known to be incompatible with specific readings of indefinites.

In this paper I lay the basis for a diachronic investigation of this development, by better defining its initial state: I focus on the distribution of *alicuis* in Classical Latin and try to establish in which sense it can be considered a specific indefinite at that stage. I argue that, in fact, the scopal behavior and the meaning contribution of *alicuis* with respect to the parameter of speaker’s knowledge suggest that it belongs to the class of epistemic indefinites (e.g. Alonso-Ovalle & Menéndez-Benito 2003, Jayez & Tovena 2006), which are characterized by (cross-linguistic diversified) context-sensitivity: they give rise to variable meaning effects according to the syntactic-semantic context they interact with. I further discuss how the (minimal) widening of the domain of quantification brought about by this type of indefinites can represent a first step on a diachronic cline leading to the Romance situation.

I proceed as follows: in Section 2, I present an overview of the Romance outcomes. In Section 3 I state more precisely the diachronic challenge posed by the development of *alicuis*, by formulating the problem in terms of Haspelmath’s (1997) semantic map. In that section I also discuss the notion of specificity used in Haspelmath (1997) and successive work on Latin indefinites. In Section 4 I discuss epistemic indefinites, focusing on their context-sensitivity and on their relation with other types of indefinites. Section 5 presents evidence supporting the classification of Classical Latin *alicuis* as an epistemic indefinite, and showing the first signs of its diachronic development. Section 6 summarizes the main conclusions.

2. The Romance outcomes of *alicuis*

The Latin indefinite *alicuis* had pronominal (*alicuis, aliquid*) and adjectival (*aliqui, aliqua, aliquod*) forms. ¹ It is attested throughout Latinity, but in the late non-written language it gave rise to a new form *ALICUNUS* by univerbation with *unus* ‘one’. The form as such is never actually attested in Latin texts but is reconstructed from the Old Romance varieties, which continue it with e.g. It. *alcuno*, Fr. *aucun*, Sp. *algún*, Pt. *algum*. Ibero-Romance also continues

¹ *Alicuis* is an interrogative-based indefinite, from Proto-Indo-European interrogative/indefinite stem *kʰi-; the first morpheme originates from (the same stem of) Lat. *alisus* ‘other’, yielding the original meaning ‘some or other’. When adjectival/adnominal, it is found predominantly in pre-nominally position, but a post-nominally order is also possible. Marouzeau (1922:165) gives the following figures for Classical Latin: in Cicero’s orations *alicuis* is found 280 times pre-nominally and 55 post-nominally; in Caesar 22 times pre-nominally and 8 post-nominally.
the masculine and neuter accusative of alius with e.g. Sp. alguien, Pt. alguém < ALIQUEM, Sp. and Pt. algo < ALIQUOD.

As we will see in Section 3, recent studies on Latin accept – with some provisos – Haspelmath’s (1997:253-254) characterization of alius as a ‘specific unknown’ indefinite, that is, as an indefinite that can appear in a ‘prototypical realis’ sentence (affirmative declarative sentence with a predicate in the perfective past or ongoing present, cf. Haspelmath 1997:39). The ‘unknown’ component consists in the conveyed message that the speaker cannot further identify the referent of the indefinite. Haspelmath (1997:254) presents (1) as an example of specific unknown alius:

\[(1) \quad \text{Late Latin (Vulg. Luc. 8.46):}\]

\[\text{tetigit me alius, nam ego novi virtutem de me exisse.}\]

‘Someone did touch Me, for I was aware that power had gone out of Me’

The most conspicuous feature of Latin alius, typically described in the grammars, consists in its complementary distribution with respect to other indefinites: according to the traditional descriptions, alius is found in positive contexts, whereas in hypothetical contexts the bare form quis ‘any’ is used, and in negative contexts the Latin negative-polarity series (quispiam, quisquam, ullus) fills in (cf. Kühner & Stegmann 1912:635, Ernout & Thomas 1951:194, Traina & Bertotti 1965:186-189). I will come back to the distribution of Latin alius in Section 3 and 5. For now let it suffice to note that alius is very successful from a diachronic perspective: it is amply continued in Romance, unlike the just mentioned Latin indefinites in complementary distribution with it, which have no Romance outcome.

Turning to Romance, a very cursory overview of the contexts of distribution immediately hints at profound diachronic changes in the continuations of alius. Only the plural can be interpreted specifically (2.a), and must be interpreted above negation (2.b), whereas the singular has become a negative polarity item (NPI) in all the languages considered here.

According to certain analyses of Negative Concord (Zeijlstra 2004, Penka 2011), French aucun is actually a negative indefinite (NI), i.e. its predominant licensing domain is the scope of clausal negation (3.a) and it can occur as negative elliptical answer, where it is ‘self-licensing’ (3.b). The plural only survives in the archaic plain or specific d’aucuns ‘some, certain’ (and as negative indefinite with pluralia tantum like vacances ‘holidays’, frais ‘fees’).

\[(2) \quad \text{Italian:}\]

(a) \text{Franco vuole comprare alcune libri nel negozio qui vicino}  
‘Franco wants to buy some books in the nearby shop’ (specific books possible)

(b) \text{Franco non vuole comprare alcune libri}  
‘There are some books that Franco does not want to buy’  
* ‘Franco does not want to buy any book’

\[\text{In spoken French (the variety which is gradually losing Negative Concord) aucun can occur without the negative marker ne in sentences with a negative meaning, thus behaving like e.g. German niemand ‘no one’(I am grateful to Fabienne Martin for the judgments):}\]

(i) \text{Il y en a aucun qui mange ton gâteau}  
‘No one is eating your cake’
(3) **French:**
   (a) *Aucun étudiant n’est venu*
       ‘No student came’
   
   (b) A: *Avez-vous une objection?*  B: *Aucune*
       ‘A: Do you have an objection?  B: None’

In Spanish and European Portuguese the NPI-uses of *algún / algum* must be accompanied by nominal inversion (4.a-c), a fact that Martins (2012) explains as a way to build the NPI in the syntax. Inversion is possible also in Italian and French, where it yields an emphatic strengthening of the negation (or of anti-additive elements like It. *senza*, Fr. *sans* ‘without’), which is absent in Spanish and Portuguese.

(4) **Spanish:**
   (a) *no toma precaución alguna*
       ‘He doesn’t take any precaucion’

   Portuguese (Martins 2012:2):
   (b) *Nao vive aqui animal algum*
       ‘No animal lives here’

   Portuguese (Martins 2000:211):
   (c) *A Maria não comprou algum livro*
       ‘Maria didn’t buy any book’

Spanish and European Portuguese differ with respect to the contexts where they license this NPI. Portuguese being more restrictive, in that *algum* behaves like a strong NPI, while Spanish post-nominal *algún* can also appear in weak downward-entailing contexts. Martins (2012) attributes these differences to an ongoing process of change, and shows how Portuguese represents a more advanced stage, as it used to be like Spanish until the end of the 18th century. Martins (2000:210-211, 219) further presents examples from Old Spanish and Old Portuguese showing that the continuations of *aliquis* could appear under negation without inversion. Pending further diachronic investigation on this aspect, it is interesting to notice that the fixation of the inverted order for the complex ‘noun + NPI’ happened only in those languages where the singular also has another, non-NPI use. The hypothesis that the syntactic mechanism of negative inversion proposed by Martins (2012) was grammaticalized as a disambiguating strategy cannot be further pursued here, but seems worth exploring.

Another important piece of evidence could come from a diachronic study of Catalan: Catalan also continues *aliquis* with *algún*, but the modern language does not have the NPI use (Vallduví 1994, Montserrat Batlloiri p.c.), but only the specific (Haspelmath 1997:258) and epistemic one. ³ However, Par (1923:46-47) shows that in the 14th century language of Bernat Metge *alcu / alcun* was regularly used as NPI indefinite under negation. Moreover, Par remarks that this use is still completely grammatical in his contemporary language if *algún* is postponed to the noun. Thus, the disappearance of *algum* from the context of direct negation in Catalan seems a relatively recent phenomenon, and this language as well witnesses a stage where nominal inversion becomes obligatory. ⁴

When pre-nominal (5.a), Spanish *algún* and Portuguese *algum* show a so-called ‘epistemic

---
³ Some Catalan varieties (e.g. Balearic) also have *qualque*, which Wheeler et al. (1999:134) describe as synonymous with *algun*.
⁴ I am very grateful to Montserrat Batlloiri for pointing me to Par (1923) and for discussing the Modern Catalan data with me.
indefinite’ use (Alonso-Ovalle & Menéndez-Benito 2003, 2010, Jayez & Tovena 2006, Aloni & Port 2010), whereby they indicate that “the speaker is unable to give any further information about who or what satisfies her existential claim” (Alonso-Ovalle & Menéndez-Benito 2003:1). This epistemic import seems to resemble the ‘unknown’ component introduced by Haspelmath (1997) in his typology and also attributed to Latin aliquis, as seen above: in Section 4 I will discuss the extent to which these notions in fact correspond.

The epistemic use of the continuations of aliquis is impossible in Italian and French (5.b). In this function, Italian and French have developed a new item, qualche / quelque (5.c) from an original free-choice function (Foulet 1919, Jayez and Tovena 2011, Zamparelli 2007).

(a) María se casó con algún estudiante del departamento de lingüística: # en concreto con Pedro

‘María married a student of the linguistics department, # namely Pedro’

Italian:
(b) *Maria si è sposata con alcuno studente del dipartimento

‘Maria married a student of the linguistics department’

Italian:
(c) Qualche idiota ha dimenticato di spegnere la luce

‘Some idiot has forgotten to switch off the light’

Earlier stages of Italian and French show that alcuno and aucun had a broader distribution (cf. Stark 2006, Ingham 2011) and occurred with a non-negative meaning in non-veridical contexts, like in modalized sentences and in the conditional protasis.

Thus, it seems that the outcomes of aliquis enjoy a great versatility during the earliest stages of the Romance languages, and undergo a successive reduction in their contexts of use, which is more severe in Italian and French than in Spanish and Portuguese. In Ibero-Romance two tendencies seem to be at work: on the one hand, the epistemic function of the singular is very vital; on the other hand, the NPI uses are subject to the stricter constraint of negative inversion.

Table (1) summarizes the distribution of the continuations of aliquis in the various contemporary Romance languages.

Table (1): Distribution of aliquis and continuations from Latin to Romance

<table>
<thead>
<tr>
<th>LANGUAGE</th>
<th>Classical Latin</th>
<th>Romance</th>
</tr>
</thead>
<tbody>
<tr>
<td>FORM</td>
<td>aliquis</td>
<td>Sp algún, alguien</td>
</tr>
<tr>
<td>FUNCTION</td>
<td>‘SPECIFIC UNKNOWN’</td>
<td>SINGULAR</td>
</tr>
<tr>
<td></td>
<td>EPISTEMIC INDEFINITE</td>
<td>NEGATIVE POLARITY ITEM</td>
</tr>
<tr>
<td></td>
<td>Sp, Pt (It qualche, Fr quelque)</td>
<td>Sp, Pt, Ita, Fr</td>
</tr>
</tbody>
</table>
3. The diachronic challenge

In this section I discuss in which sense the Romance continuations of Latin *aliquis*, as well as the development of *aliquis* in later stages of Latin, pose a challenge to diachronic theories of semantic change. I will first present Haspelmath’s (1997) semantic map for indefinites and the diachronic generalizations in encodes (Section 3.1). Then I will discuss Haspelmath’s notion of specificity and its implementation in the recent literature on Latin indefinites, showing the synchronic and diachronic limits of an analysis of *aliquis* solely in terms of specificity (Section 3.2).

3.1 A semantic map for indefinites

The extension into non-specific functions of forms continuing a specific indefinite is rarely observed. Haspelmath (1997:153-154) mentions only the uncommon case of ‘dunno’ indefinites, i.e. indefinites originating from expression meaning ‘I don’t know who / what’ (e.g. French *je ne sais quelle tristesse* ‘I don’t know which sadness’ = ‘some kind of sadness’); he treats them as specific unknown indefinites, which can diachronically extend into non-specific functions (questions, conditionals).

Haspelmath notes that this is the only kind of development running against his generalization that diachronic shifts affecting indefinites operate, through semantic weakening, unidirectionally from right to left on his semantic map. In the case of *aliquis*, however, the observed development is not only left-to-right, but also uncommonly wide-ranging, reaching the opposite pole of the semantic space, as becomes clearer from the schematic representation of Haspelmath’s map given in Figure (1):

Figure (1): schematic representation of Haspelmath’s (1997) semantic map for indefinites

<table>
<thead>
<tr>
<th>specific</th>
<th>specific</th>
<th>irrealis</th>
</tr>
</thead>
<tbody>
<tr>
<td>known</td>
<td>unknown</td>
<td>non-specific</td>
</tr>
<tr>
<td>question</td>
<td>indirect negation</td>
<td>direct negation</td>
</tr>
<tr>
<td>conditional</td>
<td>comparative</td>
<td>free choice</td>
</tr>
</tbody>
</table>

Haspelmath’s (1997) map shows the distribution of indefinites according to functions or uses, that is contexts of use. An indefinite series can be found –and actually is typically found– in more than one function. Adjacent functions are related in a systematic and semantically motivated way.

Interestingly, Haspelmath’s map is also designed to make predictions with respect to the directionality of the diachronic developments affecting indefinites: “where markers gradually acquire new functions, they will first be extended to those functions that are adjacent to the original functions on the map, and only later to functions that are further away” (Haspelmath 1997:63). Some ‘cycles’, i.e. recurrent developments, emerge clearly from Haspelmath’s survey (e.g. the ‘free choice cycle’, the ‘negative cycle’), as well as precise directional tendencies (e.g. free choice to NPI but not vice versa). Such diachronic conclusions have not gone unchallenged (see for instance the debate on the directionality of change among negative polarity elements, Jäger 2010, Biberauer & Roberts 2010), and certainly a more fine-grained analysis of contexts and meaning effects is needed in order to reach safer conclusions with respect to the mechanisms involved in the grammaticalization processes (weakening vs.

---

5 Haspelmath (1997:154) proposes that the unidirectionality of grammaticalization can be preserved also in this case once one considers a further meaning dimension, emphasis (with respect to the speaker’s lack of knowledge): the change with ‘dunno’ indefinites would also be an instance of weakening, from more to less emphatic, since non-specific functions are by default ‘unknown’.
strengthening, loss vs. gain of features). Nonetheless, Haspelmath’s map has a clear appeal for formal theories of the syntax-semantics interface.

First, it is in line with the Matthewson-Kratzer conjecture on cross-linguistic variation in indefinites (cf. Kratzer & Shimoyama 2002): indefinites would distribute across a uniform semantic space language after language, and cross-linguistic variation would amount to morpho-syntactic properties of lexical items, hopefully reducible to a restrictive format.

Furthermore, from a diachronic point of view, the map singles out recurrent paths of change and suggests principled explanations for diachronic clines, based on observed synchronic correlations and theoretical models thereof, as for instance the continuum of downward-entailing contexts (cf. Jäger 2010, Hoeksema 2010), and the continuum of so-called ‘referentially deficient indefinites’ (cf. Giannakidou 1998, Giannakidou & Quer 2010), which correspond to the central and right-hand functions in Haspelmath’s map.

The Romance data seen in Section 2 suggest that, in the course of their histories, the Romance languages witnessed sensible fluctuations in the behavior of the polarity-sensitive continuations of aliquis. It is well known that licensing contexts for polarity items may vary over time, and it has been proposed that such diachronic variability corresponds to the synchronic fine-grained typology of polarity sensitivity (cf. Ramat 1998 on Italian veruno, Hoeksema’s 2010 analysis of Dutch enig). Martins (2000) offers a comparative diachronic account of some Romance polarity items in terms of a system of syntactic features encoding interface-relevant semantic information on the lexical item. She documents the existence of a diachronic trajectory from weak to strong polarity items in Romance. This amounts to a restriction of licensing polar contexts, i.e. to a reduction in polar versatility, which Martins model as reduction in featural underspecification.

In this sense, then, the behavior of the continuations of aliquis during the historical stages of Romance is not surprising: once an element is polarity-sensitive, it is diachronically plausible that it undergoes diachronic processes affecting its licensing contexts. What is not expected, though, is that the source of such polar sensitivity be a specific indefinite. The change from ‘positive’ to ‘negative’ is usually attested with generic elements (cf. French personne ‘anyone, no one’ < Late Latin PERSONAM ‘person’, Catalan cap < Latin CAPUT / Late Latin CAPUM ‘head’), which can also function as minimizers (cf. French rien < Latin REM ‘thing’). It is less plausible, semantically, that a specific indefinite ends up being licensed by negation, as is the case with e.g. Italian alcuno. I believe that, in fact, the diachronic picture becomes much clearer once the ‘unknown’ component in the meaning of Latin aliquis, and thus its specificity, is re-evaluated in light of recent research on similar indefinites in contemporary languages, like Spanish algún, French quelque and Italian qualche. I will present some insights from this research in Section 4. Before that, I shortly turn to the notion of specificity adopted in Haspelmath (1997) and in subsequent work on Latin, in order to better set the terms of the debate.

---

6 Note that also the frequent change from NPI to negative indefinite runs against Haspelmath’s generalization about directionality.

7 Martins (2000) discusses only cases where the licensing contexts become more restricted. This change would comply with the Subset Principle, since it would be motivated in an acquisitional perspective by assuming that less salient contexts of licensing become not robust enough to be successfully acquired by new generations of speakers (Martins 2000:192). These previously available contexts become less salient in language use mainly due to “competition between negative and positive indefinites in non-negative modal contexts” (Martins 2000:215). The history of aliquis would represent a case where, during a first stage of the development, licensing contexts are actually broadened, from a subset of weak polarity contexts to a greater number thereof, and eventually to strong polarity contexts. It is possible that in this kind of diachronic process, which does not comply with the Subset Principle, further factors are at work, like changes affecting paradigmatic complementarity among indefinites (e.g. in Latin between aliquis and quis) and removing morphological blocking (cf. Pereltsvaig 2004 and Penka 2011:76, 212-227).


3.2 Specificity and *aliquis*

In an articleless language like Latin, indefinite markers represent the only means of overtly marking the indefinite interpretation of a nominal phrase. Specificity is most often left unmarked (6.a). When specificity is morphologically realized in a nominal phrase, Latin lexically distinguishes whether the identity of a specific referent introduced into the discourse is known (*quidam*, 6.b) or unknown (*aliquis*, 6.c) to the speaker.

(6) Classical Latin (Petr. Sat. 29.1)

(a) *non longe ab ostiarii cella canis ingens,*

not far from porter:GEN.SG lodge:ABL.SG dog:NOM.SG huge:NOM.SG

chain:ABL.SG tie:PTCP.NOM.SG in pariete

erat pictus

‘Not far from the porter’s lodge a huge dog, secured to a chain, was painted on the wall’

Classical Latin (Petr. Sat. 12.3)

(b) *nec diu moratus rusticus quidam*

and.not long delay:PTCP.NOM.SG countryman:NOM.SG a.certain:NOM.SG

familiaris oculus meis

familiar:NOM.SG eye:ABL.PL my:ABL.PL

cum muliereula comite propius accessit

with girl:ABL.SG accompanying:ABL.SG closer come:IND.PRF.3SG.ACT

‘In a little while a countryman, whom I knew by sight, came up with a girl’

Classical Latin (Cic. S. Rosc. 76)

(c) *litteras, credo, misit*

letter:ACC.PL suppose:IND.PRS.1SG.ACT send:IND.PRF.3SG.ACT

*alicui sicario*

some:DAT.SG killer:DAT.SG

‘He sent a message, I suppose, to some killer’

Haspelmath (1997:253-255) considers the ‘specific unknown’ function of *aliquis* as primary; as mentioned above, a specific unknown indefinite can appear in ‘prototipical realis’ contexts (affirmative declarative sentences with a perfective past or ongoing present predicate). Haspelmath remarks that *aliquis* is also found in irrealis (modalized) contexts, questions and conditionals, but attributes its occurrence in the two latter contexts to a sort of morphological suppletion with respect to the more common bare indefinite *quis* (*aliquis* would substitute for *quis* where *quis* cannot satisfy its enclisis requirements). In the literature on Latin indefinites, this classification is not uncontroversial. If, for instance, Bertocchi et al. (2010:25) state that “*aliquis* is generally, though not always, specific”, Orlandini (1983:233), basing on pragmatic criteria, classifies *aliquis* as non-specific. The difference is due to the notion of specificity adopted in the various works. As mentioned, Orlandini (1983) assumes a pragmatic definition of specificity, based on the speaker’s epistemic state: a specific indefinite expresses the speaker’s certainty with respect to the referent’s identity (unique identifiability).

---

8 I use the term ‘marker’ in order (i) to encompass adnominal and pronominal uses, and (ii) to remain vague as to whether in the various cases we are dealing with a determiner, a quantifier, or an adjectival modifier (each of which would have a syntactically distinct position within the nominal phrase).

9 In fact, the presence of *aliquis* in conditionals seems to be governed by more subtle semantic factors, cf. Bertocchi et al. (2010), Bortolussi (2010).
Under this approach to specificity, the speaker’s epistemic state substantially corresponds to the state yielding the referential interpretation of a definite description; the difference consists in the epistemic state of the hearer, who is familiar with the referent in the case of a definite noun phrase, whereas does not know the referent’s identity in the case of a specific indefinite.

Haspelmath (1997:38) adopts the following definition: “an expression is specific if the speaker presupposes the existence and unique identifiability of its referent”. 10 According to Haspelmath, it makes sense to apply the ‘known’ / ‘unknown’ distinction only in the case of specific indefinites, since only specific indefinites presuppose unique identifiability, whereas with non-specific ones the referents “are not identifiable in principle” (Haspelmath 1997:45, 133). In the case of specific unknown indefinites, the (possibility of) referent identifiability is presupposed, but the speaker is not able to carry out the identification (Haspelmath 1997:45). The German indefinites prefixed with irgend- can fulfill the ‘specific unknown’ function, as well as the Greek ka-series and the Russian -to series, to cite some frequently discussed indefinites; the English some-series, when used specifically, is ambiguous between a ‘known’ and an ‘unknown’ reading.

Bertocchi et al. (2010:24-28), as well as Bortolussi (2010), take Haspelmath’s definition as their point of departure, and consider presupposition of existence and unique identifiability as the two basic features constituting specificity. They distinguish, however, between those cases where presupposition of existence and unique identifiability go hand in hand (either they are both present or they are both absent) and cases where instead there is no unique identifiability, but the existence of a referent is nonetheless presupposed. With aliquis (7.a), there is always presupposition of existence but there is variation as to whether, according to context of occurrence, it carries or not a commitment to unique identifiability. When the referent is not uniquely identifiable, aliquis is considered non-specific. Unlike aliquis, the bare indefinite quis (7.b), occurring in non-veridical contexts, is neutral with respect to the presupposition of existence, and the NPI quisquam (7.c) is negatively specified with respect to it.

(7) Classical Latin (Cic. Att. 13.15):
(a) expectabam aliquem meorum
  ‘I was expecting one of my messengers’

Classical Latin (Cic. De Or. 1.102):
(b) si quis quid quaereret
  ‘if anyone had anything to ask’

Classical Latin (Caes. B. C. 3.87.6):
(c) nec uero ex reliquis fuit quisquam
  ‘and in fact there was no one among the others who hesitated to swear’

From the discussion in Bertocchi et al. (2010), the system that I schematically present in (8) emerges:

10 In his review of Haspelmath’s book, Dahl (1999) discusses Haspelmath’s notion of specificity and his use of the term ‘presupposition’.
(8) Classification of Latin indefinites in Bertocchi et al. (2010)
   
   (a) *quidam* [+ presupposition of existence]; [+ unique identifiability]
   (b) *aliquis* [+ presupposition of existence]; [± unique identifiability]
   (c) *quis* [∅ presupposition of existence]; [– unique identifiability]
   (d) *quisquam* [– presupposition of existence]; [– unique identifiability]

This system tries to capture the intuition, shared by all the traditional analyses, that the indefinites in (8) are organized along a (non-)referentiality scale, with *quidam* being the referentially strongest term, *aliquis* occupying an intermediate position, and *quis* and *quisquam* as the weakest terms. The set of features that are used, however, confronts us with a number of problems. The two most relevant ones, which I will discuss below, concern (i) the adopted notion of presupposition and (ii) the empirical adequacy of the proposed feature system. These problems are not limited to the approach to Latin sketched above, of course: numerous analyses have been proposed for specificity, and none uncontroversially accounts for all aspects that have been connected to this notion. Farkas (1994) argues that the term ‘specificity’ encompasses at least three different classes of phenomena (epistemic, scopal and partitive specificity). Von Heusinger (2011:1027-1028) distinguishes at least seven different types of specificity. If “[a] prototypical specific indefinite is assumed to have wide scope, a referential reading, an existential presupposition, and to indicate discourse prominence” (von Heusinger 2007:273), many non-canonical cases displaying only a subset of these properties are known from the literature, and their analysis is still matter of debate.

As for (i), a presuppositional treatment of specific indefinites has been proposed and successively defended in the theoretical literature only for a very restricted class of indefinites, and, crucially, it has been tied to a particular syntactic position (subject of individual level predicates / landing site for scrambled indefinites) or discourse status (topic), cf. von Fintel (1998), Endriss (2009:138-151), von Heusinger (2011:1051-1052) for discussion. That is, the kind of presupposition involved would not be lexically encoded as a property of the indefinite in the lexicon, but rather derived as a contextual effect, i.e. from the informational structure and / or from the interaction with other operators. Now, a first observation on Latin *aliquis*, pending further corpus work to substantiate it, is that it very rarely appears as subject of an individual-level predicate in a declarative sentence: in all of Caesar I found only one such example, given in (9).

(9) Classical Latin (Caes. B. C. 1.2):

   \[\text{dixerat} \quad \text{*aliquis*} \quad \text{leniorem} \quad \text{sententiam}\]

   say:IND.PLUPRF.3SG.ACT someone:ACC.SG milder:ACC.SG opinion:ACC.SG

   ‘someone had expressed a milder opinion’

Besides the rarity of such examples, in the sentence in (9) what could be interpreted as presupposed (in the sense of being part of the common ground) is the contextually established restrictor set (i.e. the members of the assembly on which Caesar is reporting). We would thus have a partitive effect, or contextual domain restriction, rather than existential import. It is therefore not clear whether the meaning effects of Latin *aliquis* that have been analysed as presuppositional are in fact brought about by its lexical specification or by the interaction with its contexts of use.

A similar reasoning applies to the uniqueness presupposition: if it were brought about by the lexical item itself, it is difficult to find an account for the fact that *aliquis* can oscillate between triggering and not triggering it.

Turning to (ii), the feature system seen in (8) does not account for all the contexts of occurrence and the possible interpretations of *aliquis*, as in fact acknowledged by Bertocchi et al. (2010). First of all, the quite frequent scalar uses of *aliquis* (Bertocchi et al. 2010:45-48)
are left out of the picture. In these uses, *aliquis* expresses either indifference with respect to the possible identification, and may carry a derogatory effect (cf. 10.a), 11 or the existence of an undetermined degree (especially with mass nouns), with variable meaning effects according to context (cf. 10.b).

(10) Classical Latin (Petr. Sat. 36):
(a) *suspicatus ad aliquam urbanitatem*

\begin{align*}
\text{suspect:PTCP.NOM.SG} & \quad \text{to some:ACC.SG} & \text{joke:ACC.SG} \\
\text{totiens iteratam vocem pertinere} & \\
\text{so often repeat:PTCP.ACC.SG} & \quad \text{word:ACC.SG} & \quad \text{pertain:INF.PRS.ACT}
\end{align*}

‘suspecting that this word repeated so often was part of some kind of joke’

(b) *omnia quae putas aliquam spem*

\begin{align*}
\text{everything:ACC.PL} & \quad \text{that:ACC.PL} & \quad \text{think:SUBJ.PRS.2SG.ACT} & \quad \text{some:ACC.SG} & \quad \text{hope:ACC.SG} \\
\text{mihi posse adferre mutandarum rerum} & \quad \text{can:INF.PRS.ACT} & \quad \text{bring:INF.PRS.ACT} & \quad \text{change:GER.GEN.PL} & \quad \text{thing:GEN.PL}
\end{align*}

‘everything that you may think could bring me some hope of a change’

Moreover, the definition of specificity as presupposition of existence and unique identifiability does not account for those cases where *aliquis* resists scoping under other operators even if the referent is not identifiable. In the literature on Latin, assumptions on the presuppositional nature of *aliquis* have been adduced to account for its properties in terms of scope: *aliquis* is able to take wide scope with respect to some structurally superior operators. This is what is known in the literature as ‘scopal specificity’ (cf. e.g. Farkas 1994, von Heusinger 2011). Scopal specificity, in its strictest interpretation, refers to the ability of some quantifiers to take exceptionally wide scope over scope islands that usually trap quantifiers in their scopal domain, like e.g. conditional clauses (cf. a classical example in 11, cited after von Heusinger 2011:1027):

(11) (a) *If a friend of mine from Texas had died in the fire, I would have inherited a fortune.*

\begin{align*}
\text{(possible: there is a specific friend of mine and if he had died...)} & \\
\text{(b) If each friend of mine from Texas had died in the fire, I would have inherited a fortune.} & \\
\text{(not possible: for each of my friends, if one of them had died...)}
\end{align*}

In a broader interpretation of scopal specificity, also cases of ambiguity like (12), taken by Haspelmath (1997:37) to represent the core of specificity, have been interpreted as scopal effects, due to the interaction between the quantifier and the operator (opaque contexts with modal verbs / propositional attitudes / distributive contexts):

(12) *Nobuko wants to marry a native speaker of Ainu*

Specific: Nobuko has a specific native speaker of Ainu in mind, and she wants to marry him

Non-specific: Nobuko will choose her husband within the set of those individuals who have Ainu as native language

\[11\] In such cases, according to Bertocchi et al. (2010:46-48), *aliquis* behaves “next to a free choice pronoun”.
Note, however, that a similar kind of ambiguity can arise also in non-opaque contexts, as in (13): this is what Farkas (1994) and Von Heusinger (2011) label ‘epistemic specificity’:

(13) A student in the syntax class cheated on the final exam
    Specific: ‘she was sitting right in front’
    Non-specific: ‘but I don’t know who’

Clearly, in these cases, what is necessary in order to have a specific interpretation is ‘having a particular referent in mind’, that is unique identifiability. Therefore, as in Bertocchi et al. (2010), and unlike Haspelmath (1997), when the referent is ‘unknown’ the interpretation is (epistemically) non-specific.

Now, there are contexts where, despite the referent being unknown (epistemic non-specificity), aliqua is able to take wide scope with respect to other operators (scopal specificity). An environment where this can be safely assessed is negation: aliqua systematically scopes over negation, thus yielding a scopally specific reading, although the identity of the referent remains unknown. The example in (14) can be paraphrased as follows: ‘you use the verb carere [to lack] when there is something that you don’t have (whatever it may be) and you are aware that you do not have it’.

(14) Classical Latin (Cic. Tusc. 1.88):
    dicitur enim alio modo etiam ‘carere’
    say:IND.PRS.3SG.MP in.fact other:ABL.SG way:ABL.SG also lack:INF.PRS.ACT
    cum aliqua non habeas
    when something:ACC.SG not have:SUBJ.PRS.2SG.ACT
    et non habeas te sentias
    and not have:INF.PRS.ACT you:ACC.SG perceive:SUBJ.PRS.2SG.ACT
    ‘so carere [to lack] is used in a different sense when there is something that you don’t have and you are aware that you don’t have it’

These mismatches between epistemic and scopal specificity are not uncommon cross-linguistically: unique identifiability does not always bring about wide scope (hence the possibility of intermediate scope for specific indefinites, cf. von Heusinger 2011:1035-1046), and wide scope indefinites do not always express referent identifiability (Jayez & Tovena 2006, Alonso-Ovalle & Menéndez-Beníto 2010, Romero 2010, Ebert et al. 2013). For the latter type, which has become known as the class of ‘epistemic indefinites’, alternative analyses have been recently proposed which do not resort to the notion of specificity in order to account for wide-scope readings. These approaches focus on the interplay between the semantic contribution of epistemic indefinites and the variable pragmatic effects they trigger according to the context; they capitalize on the semantic source of the ‘lack of speaker’s knowledge’ to account for the context-dependence shown by epistemic indefinites.

In Section 4 I present the ongoing debate on epistemic indefinites. In Section 5 I propose that Latin aliqua also belongs to the class of epistemic indefinites.

---

12 Szabolcsi (2004:421-422) argues that specificity has to be distinguished from (simple) positive polarity. That is, not all expressions that are incompatible with a narrow reading under negation have to be specific. They might be featurally characterized in the lexicon as Positive Polarity Items, thus restricting their special scope requirements to clausemate negation, whereas specific expressions take exceptionally wide scope with respect to a broader inventory of operators. In the case of aliqua, here I am only considering the behavior with respect to negation for its relevance in a Romance comparative perspective; however, it is well known (cf. Bertocchi et al. 2010:37-41) that in the Classical use aliqua is also avoided in some types of antecedents of conditionals or can take exceptional scope in this environment.

13 Note that here the term ‘epistemic’ is employed—somewhat deceivingly—to indicate those indefinites that signal ‘lack of speaker’s knowledge’. I will say more on terminology in Section 5.
4. Epistemic indefinites

Epistemic uses arise when a specialized indefinite marker conveys information about the speaker’s state of knowledge with respect to the referent of the indefinite phrase. In fact, even the plain indefinite article can trigger pragmatic inferences about the speaker’s epistemic state: example (15), containing a neutral indefinite article a, can have different interpretations according to the conversational context it is uttered in: a doctor the speaker has in mind (specificity), a doctor whose identity is not known to the speaker (ignorance), or any doctor (indifference / free choice).

(15) You may marry a doctor

The specialized forms known as ‘epistemic indefinites’ systematically bring about inferences related to the speaker’s epistemic state. Moreover, they display variability in the meaning effects associated with them according to the operators they interact with.

Before I discuss this aspect further, a short terminological note is in order. The term ‘epistemic’ has been employed by Jayez & Tovena (2006) to indicate, in a broad sense, all those indefinites that are sensitive to the parameter ‘knowledge of the speaker’, that is, both indefinites requiring the speaker to know the identity of the referent (Fr. un certain ‘a certain’) or not to know it (Fr. un N quelconque, ‘any’). The correspondence with Haspelmath’s ‘known / unknown’ classification breaks down when we consider that for Haspelmath speaker’s knowledge is a relevant parameter only in the case of specific indefinites, while for the rest of the system it is simply non-applicable. On the contrary, Jayez & Tovena (2006) distinguish between different varieties of free-choicedness and keep identification separate from specificity (which they intend as wide scope).

In other work, e.g. Alonso-Ovalle & Menéndez-Benito (2003) and Aloni & Port (2010), the term ‘epistemic’ has been restricted to indefinites indicating lack of speaker’s knowledge, such as e.g. Spanish algún, French quelque, Italian qualche, Romanian vreun, German irgendein. This usage has become predominant and I follow it here. Note, however, that, as observed above, every indefinite (in fact every reference-establishing expression) contributes information about the speaker’s epistemic state, and is thus sensitive to the knowledge of the speaker. So in my opinion one should rather say that epistemic indefinites are polarized with respect to this parameter, being oriented either towards the ‘known’ or ‘unknown’ endpoint.

The two main questions guiding recent research on epistemic indefinites concern the nature of their knowledge-sensitivity and the way to model the contextual variability observed with the ‘epistemic effect’. Here I will focus on Alonso-Ovalle & Menéndez-Benito’s (2003, 2010) work on Spanish algún, which elaborates on Kratzer & Shimoyama’s (2002) treatment of German irgendein and derives the epistemic effect as a conversational implicature resulting from the constraints that epistemic indefinites impose on their quantificational domain. This approach is not without problems, as discussed e.g. by Aloni & Port (2010), Lauer (ms), cf. also the overview provided in Alonso-Ovalle & Menéndez-Benito (2013). However, it is very promising from a historical perspective, since, as we will see below, it suggests a scale along which indefinites would be organized, depending on the domain shift they impose. This scale could be argued to have diachronic implications, in terms of directionality of change.

4.1 The ‘anti-singleton’ constraint

Alonso-Ovalle & Menéndez-Benito (2003, 2010) define Spanish algún as an ‘anti-singleton’ indefinite, whose domain of quantification cannot be narrowed down to an individual or singleton. A sentence like (16) is uttered when more than one part of the house are an epistemic possibility for the speaker, and would be infelicitous when the part where Juan in fact is has been already established in the common ground. It is possible, though, to exclude some possibilities: (16) is felicitous even if it has been established that some rooms are
excluded as epistemic alternatives.

(16) Spanish (Alonso-Ovalle & Menéndez-Benito 2010:8-9):

*Juan puede estar en alguna parte de la casa*

‘Juan may be in a part of the house’

The fact that some alternatives can already be discarded shows that the domain is not maximally widened. This latter feature represents the difference with respect to the quantificationally stronger free-choice effect, where all members of the domain, even the most marginal ones, are an option (e.g. Spanish *cualquiera*, Italian *qualunque*), cf. Kadmon & Landman (1993). This is the reason why Alonso-Ovalle & Menéndez-Benito (2008) speak of ‘minimal domain widening’. In Alonso-Ovalle & Menéndez-Benito (2010:17-20) the ‘anti-singleton’ constraint is modeled as a constraint (a presupposition) on the possible values of a subset selection function.

Shifts in the quantificational domain have been adduced also to account for specificity: according to Schwarzschild’s (2002) pragmatic approach, specificity would be the effect of extreme domain narrowing on the existential quantifier at the contextual level, resulting in a ‘singleton indefinite’. This approach has far-reaching implications with respect to scope-taking abilities of specific indefinites, which cannot be commented on here. Let us just note that, if a version of domain narrowing is adopted, it is possible to recognize a scale of operations on quantificational domains, encoded in different classes of indefinites: extreme domain narrowing > minimal domain widening > domain widening, as schematically represented in Figure (2).

**Figure (2): continuum of operations on quantificational domains**

On the basis of the ‘minimal domain widening’ feature, the contextual variability displayed by epistemic indefinites is accounted for in pragmatic terms, as the effect of Gricean reasoning in the presence of various operators. In Section 4.2 I will illustrate the proposed mechanism for the four main contexts that have been discussed in the literature, and in Section 5 I will discuss how Latin *aliquis* behaves in such contexts.

4.2 Contextual variability

Aloni & Port (2010) identify four main uses or functions of epistemic indefinites and survey their cross-linguistic distribution. They designate with ‘use / function’ a meaning-context pair, and define the four basic functions of epistemic indefinites as given in Table (2) and exemplified with German *irgendein* in (17). Note that the ‘specific modal variation’ function corresponds to the ‘specific unknown’ function in Haspelmath (1997).

---

14 Aloni & Port (2010) consider the epistemic effect (which they call ‘modal variation’) a conventionalized implicature, and account for it, within a dynamic semantics framework, in terms of shifts in the conceptual covers (naming, ostension, description) used to identify the referent.

15 Aloni & Port (2010:2) say that they follow Haspelmath (1997) in their definition of ‘function’. However,
Table (2): functions (meaning-context pairs) of epistemic indefinites in Aloni & Port (2010)

<table>
<thead>
<tr>
<th>abbreviation</th>
<th>function</th>
<th>context</th>
<th>meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) spMV</td>
<td>specific modal variation</td>
<td>episodic sentences</td>
<td>ignorance</td>
</tr>
<tr>
<td>(ii) epiMV</td>
<td>epistemic modal variation</td>
<td>under epistemic modal</td>
<td>ignorance</td>
</tr>
<tr>
<td>(iii) deoFC</td>
<td>deontic free choice</td>
<td>under deontic modal</td>
<td>indifference</td>
</tr>
<tr>
<td>(iv) NPI</td>
<td>negative polarity</td>
<td>downward-entailing</td>
<td>plain existential</td>
</tr>
</tbody>
</table>

(17) German (from Aloni & Port 2010):

(a) **Irgendein Student hat angerufen (# nämlich Peter)**
    ‘Some student called (#namely Peter)’
    (spMV)

(b) **Maria muss irgendeinen Arzt geheiratet haben**
    ‘Maria must have married some doctor’
    (epiMV)

(c) **Maria muss irgendeinen Arzt heiraten**
    ‘Maria must marry some doctor or other’
    (deoFC)

(d) **Niemand hat irgendeine Frage beantwortet**
    ‘Nobody answered any question’
    (NPI)

As discussed by Kratzer & Shimoyama (2002), German *irgendein* brings about an ignorance effect in positive declarative contexts (17.a), as well as under the scope of epistemic modals (17.b); under deontic modals, instead, it brings about a free choice flavor (17.c); when occurring in negative contexts (17.d) no special effect is recognizable, and the indefinite behaves like a plain existential.

An indefinite in a given language qualifies for a function if *both* of the following conditions are fulfilled (cf. Aloni & Port 2010:2): the use of the indefinite is grammatical in the context and it gives rise to the meaning specified by the function. That is, besides the case where an indefinite is simply ungrammatical in a given context, it is also possible that a given indefinite is indeed grammatical, but does not have the meaning specified by the function. Aloni & Port (2010:2) provide the example of English *some*: *some* is grammatical in the scope of a root modal like deontic *may*, but does not have the free choice reading expected for the deoFC function.

Note, moreover, that the functions singled out by Aloni & Port (2010) do not exhaust the possible readings of an indefinite in the modal contexts, since they are defined as scope under the modal. However, at least some epistemic indefinites systematically allow also for a wide scope reading: (17.c) is in fact ambiguous and allows for wide scope of the indefinite over the deontic modal. Under the wide scope interpretation of epistemic indefinites, the conveyed meaning is ignorance anchored to the speaker, and not to the attitude holder or the subject of the modal verb (cf. Alonso Ovalle & Menéndez Benito 2010:2-3, Lauer ms).

Comparative research has also uncovered further dimensions of variation: the epistemic effect conveys total ignorance with some determiners (e.g. Italian *un N qualsiasi*, whereas it is compatible with partial ignorance or requires partial ignorance with others (e.g. respectively Spanish *algún* vs. Romanian *vreun*, cf. Fălaşu 2010).

Table (3), adapted from Aloni & Port (2010:4), shows the distribution of some epistemic indefinites cross-linguistically.16

---

16 Aloni & Port (2010:4) arrange the functions in the order given in Table (3) to mirror the expectations arising in their theoretical account. Note that this way they preserve contiguity among same-polarity values for any
Various accounts have been given for the cross-linguistic variability apparent from Table (3), and no one is conclusive (cf. Alonso-Ovalle & Menéndez-Benito 2013 for a survey). The extent of domain widening is certainly one factor: *irgendein* has been analyzed by Kratzer & Shimoyama (2002) as bringing about maximal domain widening, whereas for *algún* and (*un*) *qualche* the weaker anti-singleton constraint would be operative. However, as noted by Aloni & Port (2010), approaches based on domain widening do not account for the fact that epistemic indefinites are sensitive to the (deontic or epistemic) modal base. Therefore at this stage the comparative survey in Table (3) can only be taken as an indication of the kind of cross-linguistically attested combinations among uses/functions. In what follows I will focus on Spanish *algún* to discuss how the observed meaning effects come about via pragmatic reasoning; singular *algún* yields an ignorance effect with both epistemic and deontic modals, that is, it is not sensitive to the modal base, but only to the modal’s quantificational force, which will allow us to set aside the problem mentioned above. In Section 5 I will proceed to add Latin *aliquis* to the picture seen in Table (3).

The hypothesis developed in the analyses capitalizing on domain widening is that conditions on the domain of quantification (domain widening or anti-singleton constraint) observed with epistemic indefinites cause the indefinite nominal phrase to interact with other operators in the clause (e.g. modal verbs, negation), and to ultimately become dependent on (some of) them for its licensing. The idea goes back to Kratzer & Shimoyama’s (2002) Hamblin semantics for indefinites. Domain-widened indefinites induce a set of alternatives, which originate as individual alternatives and then grow into propositional alternatives; this way they are ‘visible’ to modals and can be distributed over the set of accessible worlds introduced by the modals.

As in Kadmon & Landman’s (1993) original proposal, domain widening must happen for a reason: in downward-entailing contexts, it yields a stronger reading, since domain widened expressions are maximally informative in scale-reversal contexts. In non-downward-entailing contexts, domain widening can in principle be used for two other reasons: to avoid making a false claim in view of the available evidence, and to avoid unwarranted exclusion of alternatives, i.e. exhaustivity inferences.

In the case of Spanish *algún*, under necessity modals, as in (18), for every accessible world, there has to be some alternative (Juan is in the bedroom, Juan is in the living room, Juan is the bathroom, etc.) that is true in such world. The anti-singleton constraint ensures that the domain is not restricted to a singleton, i.e. that the room is not the same in all alternatives.

\[ Juan \text{ tiene que estar en alguna habitación de la casa} \]

‘Juan has to be in some room of the house’

According to Alonso-Ovalle & Menéndez-Benito (2010), uttering (18) triggers a quantity implicature, i.e. competition with alternative domain restrictions where the set is reduced to a singleton, thus yielding a stronger statement. The cooperative hearer will conclude that these competitors must be false and that the speaker wanted to avoid making a false claim, and this will derive the ignorance effect as an implicature, cf. (19).

(19) Implicature with *algún* and necessity modals in Alonso-Ovalle & Menéndez-Benito (2010:20, their example 59)

(a) Assertion: \( \square (\text{Juan in bedroom} \lor \text{Juan in living room} \lor \text{Juan in bathroom}) \)

(b) Implicature: \( \neg \square (\text{Juan in bedroom}) \land \neg \square (\text{Juan in living room}) \land \neg \square (\text{Juan in bathroom}) \)

Also cases where no overt modal is present are treated as implicitly modalized in Kratzer & Shimoyama (2002) and Alonso-Ovalle & Menéndez-Benito (2003, 2010): a widest-scope covert assertoric operator (with universal quantifying force) ranging over the speaker’s epistemic alternatives is assumed in order to derive the ignorance effect in sentences like (20), already seen as (5.a). The pragmatic reasoning is parallel to what proposed for overt necessity modals.

(20) Spanish (Alonso-Ovalle & Menéndez-Benito 2010:2):

(a) \( \text{María se casó con algún estudiante del departamento de lingüística} \)

‘María married a student of the linguistics department’

With possibility modals, as in (16) above, the pragmatic reasoning based on avoidance of a false claim does not work, since the speaker holds at least one pragmatic competitor to be true. Alonso-Ovalle & Menéndez-Benito propose that in this case anti-exhaustivity drives the inference: the hearer justifies the anti-singleton component as a signal not to prematurely exclude viable alternatives. The assertion is interpreted as meaning that there is at least one accessible world where Juan is in a room of the house, and it is pragmatically enriched as to mean that the speaker does not know which (otherwise he would have used an exhaustivity-triggering assertion): this results again in the ignorance epistemic effect.

Alonso-Ovalle & Menéndez-Benito (2010) do not discuss negation, since they only consider pre nominal *algún*, which cannot scope under negation. \(^{17}\) However, as seen in Section 2, *algún* with nominal inversion behaves like a NPI, and it is in fact treated as such by Aloni & Port (2010), as seen in Table (3). Of course, a unified treatment is much desirable, and it is actually called for by the observation that domain widening is a widespread strategy to strengthen negative assertions and that it is often the case that elements inducing domain widening function as NPIs in downward-entailing contexts and as free-choice items otherwise, English *any* being the textbook example (Kadmon & Landman 1993, cf. also Haspelmath 1997:152 for the diachronic connection). German *irgendein* is also one such case, although its behavior with negation is complex and cannot be fully described here (cf. Kratzer & Shimoyama 2002 and Penka 2011:204-227 for discussion). A clear-cut case of NPI behavior is (21.a), where *irgendein* co-occurs with a negative indefinite:

\(^{17}\) They discuss denial in arguing against a treatment of the epistemic effect as a presupposition: use of *algún* under *no es verdad que* ‘it is not true that’ shows that the effect does not project, as a presupposition would instead do (Alonso-Ovalle & Menéndez-Benito 2010:10-13). An objection to this line of reasoning is discussed in Lauer ms.
(21) German (Penka 2011:206-207):
(a) Niemand hat irgendein Auto gesehen
‘Nobody has seen any car’
(b) Niemand hat ein Auto gesehen
‘No one saw a car’

In (21.a) no epistemic effect arises because computing the implicature under negation would be vacuous: the pragmatic reasoning only yields statements that are already implied by the assertion (cf. Kratzer & Shimoyama 2002, Lauer ms). The use of irgendein in this context may just yield a more emphatic reading than (21.b) when irgendein is stressed, and forces narrow scope (whereas wide scope of the existential is possible with ein). Penka (2011:218-221) proposes that, in fact, irgendein can be used in the scope of an antimorphic operator only if the global meaning is richer (contributing emphasis or denial) than what would arise with a negative indefinite. This kind of paradigmatic competition, which Penka treats as morphological blocking, would explain the fact that irgendein cannot occur with the sentential negator nicht in the intended NPI meaning, whereas its NPI use is possible with emphatic adverbials like keinesfalls ‘in no case’. This connects with what we observed in Section 2 about nominal inversion and the Romance continuations of aliquis: the emphasis added by the operation of nominal inversion allows for the use of epistemic indefinites under negation, despite their epistemic component being vacuous, and despite there being negative indefinites (e.g. Spanish ningún, Italian nessuno) as paradigmatic competitors.

5. Latin aliquis

In the previous section we have seen how modeling the constraint that algún imposes on its quantificational domain in terms of minimal domain widening can derive the ignorance component as a pragmatic effect. The wide scope uses, as well as the apparently unembedded ones, are in fact quantified over by a covert assertoric operator, and the epistemic component is anchored to the speaker. In intensional contexts narrow scope readings can arise, where algún imposes the anti-singleton constraint on the worlds in the quantification domain of the modal. This analysis accounts for contextual variability (in fact, predicts it) better than an analysis based on specificity: epistemic indefinites have the possibility of being interpreted specifically (having wide scope over an intensional operator) while still contributing the ignorance component, or they can be interpreted as dependent from the intensional operator.

This analysis could also be exploited in a diachronic perspective: by introducing (minimal) domain widening, epistemic indefinites could represent a ‘bridging context’ towards other functions, like the free choice or the NPI function. At the same time, the cross-linguistic survey in Aloni & Port (2010) shows that often epistemic indefinites only display a subset of the uses seen in Table (3), hinting at a possible diachronic processes (loss / gain of semantic features) affecting the inventory of functions.

In this section I present data from Latin that support the classification of aliquis as an epistemic indefinite. I will discuss, in turn, data from modalized contexts and data from negative environments. We will see that aliquis behaves similarly to Spanish algún and Italian (un) qualche with respect to modals and propositional attitudes. In negative contexts, while

---

18 In Kratzer & Shimoyama’s (2002) Hamblin semantics framework, the impossibility for irgendein to occur under sentential negation is interpreted as an intervention effect. Kratzer & Shimoyama (2002:24) note that irgendein under negation is rescued by stress, as in (i):

(i) ich hab’ nicht irgendwas gelesen
   *‘I didn’t read anything’
   ‘I didn’t read just ANYthing’

Penka (2011:206) interprets this reading as metalinguistic negation of the free-choice implicature.
the Classical language does not show signs of NPI uses, evidence of a change in this respect can be already found in Late Latin, especially since the 4th century CE.

5.1 *Aliquis* in modalized contexts

I have mentioned that one of the arguments for analyzing *aliquis* as a specific indefinite is that it can occur in ‘prototypical realis’ contexts, as subject of affirmative declarative sentences with a perfective past or ongoing present predicate. We have seen, however, that these instances are quite rare (cf. Section 3.2). Consider again the example in (1), repeated as (22):

(22) Late Latin (Vulg. Luc. 8.46):

\[
\text{tetigit me *aliquis*, nam ego}
\]

\[
\text{touch:IND.PRF.3SG.ACT me:ACC.SG someone:NO}\text{M.SG for I:NO}\text{M.SG}
\]

\[
\text{novi virtutem de me exisse.}
\]

\[
\text{know:IND.PRF.1SG.ACT power:ACC.SG from me:ACC.SG go.out:INF.PRF.ACT}
\]

‘Someone did touch Me, for I was aware that power had gone out of Me’

Here in fact the assertion has to be interpreted as implicitly modalized: ‘I believe / I am certain that someone touched me’, since Jesus infers that someone touched him from indirect evidence. That means that this context cannot count as a ‘prototypical realis’ one. In fact, the use of *aliquis* in modalized contexts is very frequent. Traina & Bertotti (1965:187) note that *aliquis* is often used to talk about future states of affairs. Some examples with future tense are given in (23). 19

(23) Early Latin (Plaut. Bacch. 638a):

(a) \[
\text{deus respiciet nos *aliquis*}
\]

\[
\text{god: NOM.SG protect:IND.FUT.3SG.ACT us:ACC.PL some: NOM.SG}
\]

‘some god will protect us’ (# namely Juppiter)

(b) \[
\text{ostium pultabo atque intus evocabo aliquem foras}
\]

\[
\text{door:ACC.SG knock.at:IND.FUT.1SG.ACT and inside call:IND.FUT.1SG.ACT someone:ACC.SG outside}
\]

‘I will knock at the door and I will call someone outside from the house’

Classical Latin (Petr. Sat. 74.2):

(c) \[
\text{nam aut incendium oportet fiat,}
\]

\[
\text{for either fire: NOM.SG be.necessary:IND.PRS.3SG.ACT become:SUBJ.PRS.3SG.ACT}
\]

\[
\text{aut *aliquis* in vicinia animam abiciet}
\]

\[
\text{or someone:NO}\text{M.SG in neighborhood:ABL.SG soul:ACC.SG give.up:IND.FUT.3SG.ACT}
\]

‘Either there must be a fire, or someone close by is just going to give up the ghost’

In (23.c) Trimalchio is commenting, upon hearing a rooster crow, that it must mean bad luck, and makes a prediction on possible outcomes: clearly he is not thinking of a specific referent here. Epistemic possibility, compatible with the meaning of *aliquis*, can also be conveyed by indicative present tense, as for instance in (24), where the adverbs *forsitan* ‘perhaps’ and *utique* ‘be it as it may’ express the speaker’s epistemic attitude. 20

---

19 For the modal use of the future tense in Latin see Magni (2010:263-264).

20 Note that *forsitan* ‘perhaps’ is usually found with the subjunctive in Classical prose, at least until Cicero (1st cent. BCE); the use in (24) belongs to the colloquial style.
(24) Classical Latin (Petr. Sat. 115):

\[
\begin{align*}
\textit{hunc & forsitan, proclamo,} & \quad \textit{in aliqua & parte & terrarum} \\
\textit{this:ACC.SG & perhaps & say:IND.PRS.1SG.ACT} & \quad \textit{in some:ABL.SG & part:ABL.SG & land:GEN.PL} \\
\textit{secura & expectat & uxor, & forsitan} & \\
\textit{unsuspecting: NOM.SG & wait:IND.PRS.3SG.ACT & wife:NOM.SG & perhaps} \\
\textit{ignarus & tempestatis & filius, & aut pater;} & \\
\textit{unaware: NOM.SG & storm:GEN.SG & son:Nom.SG & or father:NOM.SG} \\
\textit{utique & reliquit & aliquem,} & \\
\textit{anyway & leave:IND.PRF.3SG.ACT & someone:ACC.SG} \\
\textit{cui & proficiscens & osculum & dedit} & \\
\textit{who:DAT.SG & leave:PTCP.NOM.SG & kiss:ACC.SG & give:IND.PRF.3SG.ACT} \\
\end{align*}
\]

[context: the speaker is looking at the body of a man dead at sea] ‘I say: perhaps in some part of world an unsuspecting wife is waiting for him, perhaps a son who is not aware of the storm, or a father; at any rate, he left someone, whom he kissed in leaving’

The use of *aliquis* in implicitly modalized contexts is very richly exemplified in Latin texts, and seems to be one of the primary functions of *aliquis*.

Turning to overt modals, Latin can express epistemic and deontic modality by means of subjunctive forms or by means of a set of modal verbs (cf. Magni 2010:206-212). Modal verbs for possibility are the verb *possum* ‘can’ or the impersonal *licet* ‘it is permitted’ / ‘it is possible’. However, in the Classical texts I analyzed, these verbs most often express dynamic (ability) rather than epistemic or deontic modality. While *aliquis* is frequently found in combination with these modals, I could not find many clear examples of epistemic or deontic readings. One instance where the deontic possibility reading is plausible is given in (25): here the modal *possum* ‘can’ is embedded in the complement clause of the verb *concedo* ‘permit’:

(25) Classical Latin (Cic. Verr. 2.1.32):

\[
\begin{align*}
\textit{vos & quaeso & datehoc} & \\
\textit{you:ACC.PL & beg:IND.PRS.1SG.ACT} & \quad \textit{give:IMP.PRS.2PL.ACT} & \quad \textit{this:ACC.SG} \\
\textit{et & concedite & pudori & meo,} & \\
\textit{and allow:IMP.PRS.2PL.ACT & modesty:DAT.SG} & \quad \textit{my:DAT.SG} \\
\textit{ut & aliquam & partem & de & istius & impudentia} & \\
\textit{that & some:ACC.SG & part:ACC.SG & of & this:GEN.SG & shamelessness:ABL.SG} \\
\textit{reticere & possim} & \\
\textit{keep.silent:INF.PRS.ACT} & \quad \textit{can:SUBJ.PRS.1SG.ACT} \\
\end{align*}
\]

‘Please, do permit this, and grant to my modesty that it may be allowed to refrain from describing some of his shamelessness’

The part of Verres’ misdeed that Cicero wants to pass over in silence is not further specified, although clearly some alternatives can already be discarded. The effect is one of modal variation, dependent from the possibility modal.

The evidence with overt modals is more abundant in the case of necessity, both for epistemic and deontic modality. Overt modal verbs expressing such notions in Latin are *debeo* ‘must’, the impersonal *oportet* ‘it is necessary’ and the phrase *necesse est* ‘it is necessary’. The examples in (26) show that, both with epistemic (26.a) and with deontic (26.b-c) modal base, the meaning contribution of *aliquis* is ignorance, and the modal variation effect

---

21 Epistemic possibility can be expressed by the potential subjunctive: *aliquis* combines with the present or the perfect potential subjunctive in expression like *aliquis dicat* ‘someone may say’ (e.g. Ter. And. 639). Since these forms are quite rare, a wider corpus study is needed to assess the extent to which *aliquis* appears in this context.
observed with Spanish *algún* obtains in Latin as well. In (26.c) the indefinite *aliquid* has narrow scope also with respect to the universal quantifier *omnia*.

(26) Classical Latin (Cic. S. Rosc. 105):
(a) *necesse est aliquem dixisse*

necessary be:IND.PRS.3SG.ACT some:ACC.SG say:INF.PRF.ACT

*municipem aut vicinum*

citizen:ACC.SG or neighbor:ACC.SG

[context: when something like that happens, judges, you usually say:] ‘some fellow citizen or neighbor must have told him’

Classical Latin (Col. 12.52):
(b) *sed si vasta sunt oliveta,*

but if vast: NOM.PL be: IND.PRS.3PL.ACT olive-groves: ACC.PL

*necesse est aliqua pars eorum*

necessary be: IND.PRS.3SG.ACT some: NOM.SG part: NOM.SG they: GEN.PL

*maturum fructui reservetur*

mature: DAT.SG fruit: DAT.SG reserve: SUBJ.PRS.3SG.MP

‘but if the olive-groves are very big, it is necessary that some part be reserved to the harvest’

Classical Latin (Quint. Inst. 3.6.52):
(c) *omnia enim ista referri ad aliquid*

all: ACC.PL indeed this: ACC.PL refer: INF.PRS.MP to something: ACC.SG

*necesse est*

necessary be: IND.PRS.3SG.ACT

[context: Quintilian discusses the rhetorical organization of a juridical cause, which at a certain stage has to discuss questions like ‘if someone has the right to accuse’, ‘if he should do it now’, ‘if he should do it this way’] ‘because it is necessary that all of these questions be referred to something’

Deontic modality can be expressed in Latin also by means of a periphrastic construction combining the gerundive with a form of the verb ‘to be’: in (27) it is not the case that there is something specific that has to be contributed to the community; *aliquid* is embedded under the modal and also in this case contributes ignorance with respect to the identification of the referent.

(27) Classical Latin (Cic. Off. 1.52):

*quare et his utendum est et semper*

therefore and this: ABL.PL adopt: GER.NOM.SG be: IND.PRS.3SG.ACT and always

*aliquid ad communem utilitatem afferendum*

something: NOM.SG to common: ACC.SG interest: ACC.SG contribute: GER.NOM.SG

‘Therefore, one should adopt these principles and always contribute something to the common interest’

We can conclude that *aliquis* under deontic modals does not have the free choice import that has been observed e.g. for German *irgendein*. This supports also for Classical Latin *aliquis* an analysis in terms of minimal domain widening: this indefinite signals, in the same way as Spanish *algún*, that its domain cannot be reduced to a singleton set.
5.2 Aliquis and negation

As discussed in Section 3.2 and seen in example (14), _aliquis_ systematically takes wide scope with respect to negation in Classical Latin. Existential quantification under the scope of the sentential negator is expressed by dedicated NPI-series, most often the _quisquam_ (pronoun) / _ullus_ (adjective) series, as in (28.a-b) (cf. also 7.c), and the archaic -piam series. Latin, being a ‘double negation’ language, also has ‘self-licensing’ negative indefinites (e.g. _nemo_ ‘nobody’, _nihil_ ‘nothing’) which express negative meaning without co-occurring with adverbial negation, cf. (28.c).

(28) Classical Latin (Caes. B. C. 3.69):
(a) _neque quisquam omnino consisteret_
   and.not any:NOM.SG at.all stop:SUBJ.IMPF.3SG.ACT
   ‘and no one at all stopped’

(b) _sine timore ullo_
   without fear:ABL.SG any:ABL.SG
   ‘without any fear’

(c) _interit tamen nemo_
   die:IND.PRS.3SG.ACT however nobody.NOM.SG
   ‘however nobody died’

If _aliquis_ co-occurs with negation, either it is interpreted outside its scope (cf. 14), or the negation is interpreted as metalinguistic (cf. 29.a, where negation is brought about by the negative verb _nolo_ ‘not-want’). In case of double negation, _aliquis_ is possible, as expected given the overall positive polarity of the context (29.b). In these uses, _aliquis_ typically expresses a relevant quantity, similarly to the examples seen in (10).

(29) Early Latin (Plaut. Curc. 319-320):
(a) PH. _Iam edes aliquid._
   now eat:IND.FUT.2SG.ACT something:ACC.SG
   ‘Now you shall eat something’

   CV. _Nolo hercle aliquid_
   not.want:IND.PRS.1SG.ACT for.Hercules something:ACC.SG
   ‘for Hercules, I do not want _something_: I prefer _certain_ to _something_’

   _certum quam aliquid mauolo._
   certain:ACC.SG than something:ACC.SG prefer:IND.PRS.1SG.ACT
   ‘for Hercules, I do not want _something_: I prefer _certain_ to _something_’

(b) _non sine aliquo divino numine_
   not without some:ABL.SG divine:ABL.SG providence:ABL.SG
   ‘not without some divine providence’

More rarely _aliquis_ can co-occur with the preposition _sine_ ‘without’ in a downward-entailing context, as in (30.a). The interpretational difference with respect to the NPI _ullus_ in (30.b) consists in the fact that _aliquis_, also in this context, expresses the existence of a relevant quantity, especially in combination with abstract mass nouns, while the NPI is a maximal domain widener (cf. Traina & Bertotti 1965:189).
Latin *aliquis* as an epistemic indefinite

(30) Classical Latin (Caes. B. C. 73):
   (a) *sine aliquo vulnere*
       without some:ABL.SG damage:ABL.SG
       ‘without relevant losses’
   
   Classical Latin (Cic. Att. 9.13.4):
   (b) *sine una spe*
       without any:ABL.SG hope:ABL.SG
       ‘with no hope’

A similar contrast has been observed in other downward-entailing contexts, like the restriction of a universal quantifier: most often the negative polarity *quisquam / ulla*-series is found (cf. 31.a); when *aliquis* is used, as in (31.b), it does not refer to a scalar endpoint (Devine & Stephens 2013:375-376). Thus, also in this case no exhaustivity effect is connected with it.

(31) Classical Latin (Cic. Verr. 2.3.28):
   (a) *omnes... qui ullam agri glebam possiderent*
       ‘All those who owned any clod of land’
   
   Classical Latin (Cic. Parad. 40):
   (b) *omnes qui aliquid scire videntur*
       all:NOM.PL who:NOM.PL something:ACC.SG know:INF.PRS.ACT seem:IND.PRS.3PL
       ‘Everyone who seems to know something’

Coming back to negation, there is only one systematic case where *aliquis* is interpreted in the scope of negation: this happens with the negative complementizer *ne*. *Ne* functions as negation in prohibitions and in some non-veridical environments (especially non-embedded uses of the subjunctive in e.g. the expression of desires, suppositions, concessions; cf. Ernout & Thomas 1951:195); moreover, it introduces negative purpose clauses (cf. 32.a), as well as the complement of propositional attitude verbs expressing fear (cf. 32.b), wish, commands.

(32) Classical Latin (Cic. Tusc.1.88):
   (a) *illud excutiendum est, ut sciat quid sit ‘carere’, ne relinquatur aliquid erroris in verbo*
       that:NOM.SG investigate:GER.NOM.SG be:IND.PRS.3SG.ACT
       briefly worry:IND.PRF.1PL.ACT lest ‘carere’ (to lack):INF.PRS
       we have to investigate this, in order to know what ‘carere’ (to lack) is, lest any / some misunderstanding be left in the interpretation of the word

   Classical Latin (Petr. Sat. 30):
   (b) *sine dubio paulisper trepidavimus, ne contra praeceptum aliquis nostrum limen transiret*
       without doubt:ABL.SG briefly worry:IND.PRF.1PL.ACT lest against instruction
       someone:NOM.SG we:GEN.PL threshold:ACC.SG walk.past:SUBJ.IMPF.3SG.ACT
       ‘Undeniably for a short while we were worried that some of us could walk past the threshold against the instructions’
Orlandini (2001:40-41), Bortolussi (2010:21-22), Devine e Stephens (2013:374-375) all agree that the negation introduced by *ne* has special semantic and pragmatic properties that can account for the behavior of *aliquis* without losing the generalization with respect to the ‘normal’ clausemate negation. According to Bortolussi (2010:21), *ne* takes widest scope over the entire proposition in virtue of its position as a complementizer. Devine & Stephens (2013:374) treat negative purpose clauses as an instance of extracausal negation, similar to denial in that it ‘applies to the whole proposition’. Bortolussi (2010:21) shows that with denials *aliquis* is possible and preserves the speaker-ignorance component even in positive contexts, cf. (33):

(33) Classical Latin (Ov. Rem. Am. 281):

\[
\begin{align*}
\textit{non aliquis} & \quad \textit{socios} & \quad \textit{rursus ad arma vocat} \\
\text{not someone:ACC.SG} & \quad \text{ally:ACC.PL} & \quad \text{back to arm:ACC.PL call:IND.PRS.3SG.ACT}
\end{align*}
\]

‘it is not true / the case that there is someone who is calling the allies back to arms’

Scholars, however, disagree as to the exact interpretation of *aliquis* in examples like (32.a). According to Orlandini (2001:40), *aliquid* behaves here like a negative polarity item, and reinforces the negation in expressing the endpoint of a scale; she translates it as “*la moindre equivoque*”, i.e. ‘any misunderstanding whatsoever’. Devine & Stephens (2013:374), on the other hand, treat *aliquid* in (32.a) as a plain existential and translate “So as to avoid some residual error in the use of the word”, i.e. so as to avoid a situation in which there is some residual error. Also in view of examples like (32.b), where no extreme domain widening is observable, I tend to agree with Devine & Stephens that it is not necessary to analyse *aliquis* as a NPI in contexts like (32.a); it is nonetheless possible that these contexts might have created ambiguity at an early stage, allowing in fact for both interpretations and favoring reanalysis.

Signs of a reanalysis leading *aliquis* to behave like a NPI indefinite under the scope of sentential negation can be traced to the Late Latin stage. Particularly interesting data can be found in texts between the late 3rd and the early 5th century CE, an age that also witnesses the first steps in the development of Negative Concord (cf. Molinelli 1988, 1989). Some examples for the expansion of *aliquis* to negative contexts within the scope of negation are already found in the Vulgate translation of the Bible (4th cent. CE), cf. (34):

(34) Late Latin (Vulg. Luc. 11.36):

\[
\begin{align*}
\textit{si ergo} & \quad \textit{corpus} & \quad \textit{tuum} & \quad \textit{totum} & \quad \textit{lucidum} & \quad \textit{fuerit} \\
\text{if therefore body:Nom.SG} & \quad \text{your:Nom.SG} & \quad \text{all:Nom.SG} & \quad \text{bright:Nom.SG be:Subj.Prf.3SG} \\
\textit{non habens} & \quad \textit{aliquam} & \quad \textit{partem} & \quad \textit{tenebrarum} \\
\text{not have:PTCP.Nom.SG} & \quad \text{any:Acc.SG} & \quad \text{part:Acc.SG} & \quad \text{darkness:Gen.PL} \\
\textit{erit lucidum} & \quad \textit{totum} \\
\text{be:Ind.Fut.3SG.ACT} & \quad \text{bright:Nom.SG} & \quad \text{all:Nom.SG}
\end{align*}
\]

‘If therefore your whole body is full of light, with no dark part in it, it will be wholly illumined’

This use becomes more frequent in later texts. Bertocchi et al. (2010:44-45, 93) note that Gregory of Tours (6th cent. CE) has many cases of co-occurrence of *aliquis* and *quisquam*, where no semantic difference between the two indefinites can be detected. An interesting example from the 6th century CE, where the context makes clear that all the previously listed alternatives have to be excluded is presented in (35):
Late Latin (Anthem. 76):

\( si \ q\ i\ s \ crudos \ lactes \ vult \ bibe, \)
if anyone:NOM.SG raw:ACC.PL milk:ACC.PL want:IND.PRS.3SG.ACT drink:INF.PRS.ACT
mel \ habeant \ admixtum \ vel \ vinum
honey:ACC.SG have:SUBJ.PRS.3PL.ACT mixed:PTCP.ACC.SG or \ wine:ACC.SG
aut medus; \ et si non fuerit \ aliquid \ de istis
or mead:ACC.SG and if not be:SUBJ.PRF.3SG.ACT any:NOM.SG of this:ABL.PL
pocius, \ sale \ mittatur \ modicum
beverage:ABL.PL salt:NOM.SG add:SUBJ.PRS.3SG.MP little:NOM.SG
‘if one wants to drink raw milk, they should mix honey, or wine, or mead; and should there not be any of these beverages, add a little salt’

The development of \textit{alquis} into a NPI must have been completed, at least in some geographical areas, by the 4\textsuperscript{th}-5\textsuperscript{th} cent. CE. Molinelli (1989:622-623), in discussing the development of Negative Concord, cites interesting examples from Augustine, where he comments on Bible translations and indicates what would be the correct Latin usage. Interestingly for us, he reacts to Negative Concord structures, which he deems unacceptable, with alternative correct formulations where, in fact, a non-Classical use of \textit{alquis} as NPI can be observed, cf. (36):

Late Latin (Aug. Loc. In Hept. 2, de Ex 54.239):

\begin{align*}
\text{non est} & \text{ relictum} \quad \text{viride} \quad \text{nihil} \\
\text{not be:IND.PRS.3SG.ACT} & \text{leave:PTCP.NOM.SG} \quad \text{green:NOM.SG nothing:NOM.SG} \\
in \text{lignis} & \text{dicendum} \quad \text{fuit} \\
on \text{wood:DAT.PL} & \text{say:GER.NOM.SG} \quad \text{be:IND.PRF.3SG.ACT} \\
\text{more} & \text{locutionis} \quad \text{nostrae:} \\
\text{custom:Aobl.SG} & \text{language:GEN.SG our:GEN.SG} \\
\text{non est} & \text{relictum} \quad \text{viride} \quad \text{aliquid} \quad \text{in lignis} \\
\text{not be:IND.PRS.3SG.ACT} & \text{leave:PTCP.NOM.SG} \quad \text{green:NOM.SG any:NOM.SG on wood:Abl} \\
\text{‘there isn’t anything [nihil] green left on the trees} & \text{should rather be according to our way of speaking: ‘there isn’t anything [aliquid] green left on the trees’}
\end{align*}

5.3 Summary and outlook

From the data discussed in Section 5.1 and 5.2 we can draw some conclusions, which will have to be tested by further extensive corpus work. With respect to the four functions / uses discussed in a comparative perspective by Aloni & Port (2010), \textit{alquis} behaves similarly to the Romance epistemic indefinites included in Table (3), in that it does not allow for free-choice uses under deontic modals, unlike German \textit{irgendein} or English \textit{any}: the meaning contribution is always ignorance. This suggests that an anti-singleton constraint, rather than maximal domain widening, is imposed on the quantificational domain of \textit{alquis}. An analysis of \textit{alquis} solely in terms of specificity cannot account for the cases where it takes narrow scope with respect to the modal operators. However, \textit{alquis} can also appear in unembedded positive contexts: here \textit{alquis} patterns with \textit{algún} and \textit{un qualche} in Table (3), and unlike Romanian \textit{vreun}. Many of these contexts seem to be implicitly modalized, thus suggesting the presence of a covert assertoric operator, as in the treatments of epistemic indefinites discussed in Section 4.2. Finally, with respect to negation, Classical Latin \textit{alquis} patterns like Italian \textit{un qualche} in not having NPI uses, whereas in Late Latin it starts behaving like Italian \textit{alcuno} and Spanish (post-nominal) \textit{algún}. 
Table (4) summarizes the results adding *aliquis* to the cross-linguistic picture in Table (3):

<table>
<thead>
<tr>
<th>Latin</th>
<th>spMV</th>
<th>epiMV</th>
<th>NPI</th>
<th>deoFC</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>aliquis</em></td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>

In the discussion I just touched in passing on the very frequent cases where *aliquis* receives a scalar interpretation, expressing the existence of an undetermined degree, with variable meanings according to context, which cannot be reconciled with a classification as specific indefinite (cf. Section 3.2). These uses have not been discussed in connection to epistemic indefinites in the literature, but Bertocchi et al. (2010:46-48, 55-56, 63) have connected them to factors like scale orientation and scale type (quality vs. quantity). In some cases, according to Bertocchi et al. (2010), *aliquis* behaves similarly to a free-choice pronoun. Future research will have to address the issue whether also these uses can be accounted for by assuming an anti-singleton constraint (rather than a requirement of maximal widening, as in free choice indefinites) on the quantificational domain of *aliquis*.

6. Conclusion

In this paper I have discussed the synchronic and diachronic problems that arise once Latin *aliquis* is analysed as a specific indefinite, and I have proposed to consider it an epistemic indefinite. *Aliquis* manifests context-variability effects similar to those observed in the literature for Spanish *algún*, Italian *qualc*he, German *irgendein* among others. I proposed that epistemic interpretations may constitute a bridging context, an intermediate step in the extension of *aliquis* to further environments, comprising negation. The key change is the development of domain widening, starting with epistemic use (minimal domain widening according to Alonso-Ovalle & Menéndez-Benito 2008). Extension to negative contexts, which were incompatible with *aliquis* in Classical Latin, can be observed starting from the 4th century CE. The developments in Romance follow the same direction in all languages, but reach different stages. Future research will have to establish how these differences in the language-specific histories are connected to other co-occurring changes (especially in the domain of negation) and to dynamics governed by the paradigmatic pressure of other indefinites. The general impression is that Spanish, in displaying epistemic uses of *algún*, represents a more conservative stage than Italian and French, which lost these uses (attested at early stages) and only show continuations of *aliquis* in downward-entailing contexts.

7. References

Biberauer, Theresa & Ian Roberts 2010. Comments on Jäger’s ‘Anything is nothing is something’: On the diachrony of polarity types of indefinites. Natural Language and Linguistic Theory, 28.4, 823-836.
663-678.


XXX Congresso Internazionale della Società di Linguistica Italiana. Roma: Bulzoni, 397-409
Discourse Effects of Differential Object Marking in Romanian *

Klaus von Heusinger & Sofiana Chiriacescu

klaus.vonheusinger<at>uni-koeln.de
sofiana.chiriacescu<at>uni-koeln.de

1. Differential Object Marking in Romanian - Observations and Research Issues

Pe-marking in Romanian is an instance of differential object marking (DOM), i.e. the marking of the direct object under certain conditions. It is commonly assumed that pe-marking (generally accompanied by clitic doubling) is obligatory with (modified) definite human postverbal noun phrases, as in (1) and with all other forms that are higher on the Referentiality Scale, while it is optional with indefinite noun phrases (2) (Niculescu 1965, Pană-Dindelegan 1997, von Heusinger & Onea 2008, Stark & Sora 2008, Ciovârmache & Avram 2013).

(1) (a) Doctorul il examinează pe băiatul bolnav.
Doctor.DEF looks at PE boy.DEF sick
‘The doctor examines the sick boy.’
(b) #Doctorul examinează băiatul bolnav.
Doctor.DEF looks at boy.DEF sick
‘The doctor examines the sick boy.’

(2) (a) Toţi bărbaţii -o iubesc pe o femeie.
All men CL love PE a woman
‘All men love a woman.’ (specific – wide scope)
(b) Toţi bărbaţii iubesc o femeie.
All men love a woman
‘All men love a woman.’ (specific / wide scope – non-specific / narrow scope)

The contrast in (2) is generally associated with specificity. While the pe-marked indefinite direct object in (2a) has only a (scopally) specific interpretation in the sense that there is one woman such that all men love her, the unmarked indefinite in (2b) allows for a specific or non-specific reading (Farkas 1978, Dobrovin-Sorin 1994). (It is more accurate to say that pe-marking with indefinites excludes a non-specific reading since unmarked indefinite direct objects are compatible with a specific reading as well.)

There is an interesting exception in the otherwise quite robust distribution of pe-marking with definite noun phrases. Modified human definite direct objects obligatorily receive pe-marking, as in (1a), but the co-occurrence of pe and the definite article is blocked if the noun phrase is not (further) modified. This blocking effect on pe-marking derives from an independent syntactic rule of Romanian which holds for most prepositions, e.g. the

---

* We are grateful to the audience of the workshop Theoretical implications of the syntax/semantics interface in Romance held in November 2012 in Cologne for constructive comments and considerable assistance. In particular we like to thank Susann Fischer for constructive comments. Our research was supported by the German Science Foundation by a grant to the project C2: Case and referential context, as part of the Collaborative Research Center 732 Incremental Specification in Context at the University of Stuttgart. The second author is grateful for the support of the Transilvania University of Brasov, Romania. A slightly revised version of this article will be published in Revue Roumaine de Linguistique.
The Discourse Structuring Potential of Differential Object Marking

preposition *la ‘to’ in (3A), but also for *pe in its case-marking function, as in (3a). There are two alternative constructions for the ungrammatical sentence (3a), namely (3b), in which the definite article is present and the *pe-marker (and the clitic) are absent, and (3c), in which the *pe-marker (and the clitic) precedes the noun phrase in the absence of the definite article, yielding a definite reading.

(3) A: Un băiat merge la doctor.
   a boy goes to doctor.
(a) *Doctorul îl examinează pe băiatul.
   Doctor.DEF CL examines PE boy.DEF
(b) Doctorul examinează băiatul.
   Doctor.DEF examines boy.DEF
(c) Doctorul îl examinează *pe băiat.
   Doctor.DEF CL examines PE boy

‘A boy goes to the doctor. The doctor examines the boy.’

We can summarize our observations made so far: The alternation of *pe-marking with postverbal human direct objects is not restricted to indefinite NPs, but also applies to definite (unmodified) NPs. This observation raises the following main questions:

(i) The function of *pe for indefinites is generally associated with the notion of specificity. Can we assign an analogous function to *pe-marked definites as for example signaling the contrast between a referential and an attributive reading in the sense of Donnellan (1966)?

(ii) Even if we can find a parallel feature for definites, the contrast between (3b) and (3c) cannot be accounted for in terms of scope or specificity. Thus we have to find out whether there are other features that are associated with *pe-marking.

(iii) A more general question regarding *pe-marking and DOM in other languages as well is whether the marker has a genuine (lexical) function, or whether the expressed function is derived from the particular construction and some additional inferences.

In this paper we argue that: a) *pe-marking has the same or a very similar function for definites and indefinites in contexts with operators; b) *pe-marking expresses specificity (or more exact: incompatibility with non-specificity) for indefinites, and referentiality for definites; c) specificity cannot account for contrasts in transparent sentences, i.e. simple declaratives with no operators; d) in such contexts, *pe-marking expresses the discourse pragmatic function of Discourse Structuring Potential (DSP) for both, definites as well as indefinites. DSP is measured in (i) a higher likelihood of subsequent mention (Givón 1983, Ariel 2001, Arnold 2010) and (ii) a higher topic shift potential (Givón 1983). Following the literature (Prince 1981, Ionin 2006) we assume that specificity and DSP are closely related and that both functions are lexically encoded in *pe (similar to the lexical contribution of articles, specific adjectives like a certain, or discourse particles like then, therefore, etc.).

The paper is structured as follows: In section 2, we provide a brief overview on the conditions of *pe-marking in Romanian and discuss the alternation between *pe-marking and the use of the definite article for unmodified definite direct objects, in particular. In section 3, we introduce the discourse-pragmatic concept of Discourse Structuring Potential, and in section 4, we report the findings of two sentence-continuation experiments that tested the discourse behavior of indefinite and definite noun phrases. In section 5, we discuss these findings and present a general perspective for *pe-marking in Romanian and DOM in general.
2. Pe-marking and the Referentiality Scale

There is some consensus in the literature that the most important synchronic conditions triggering DOM in Romanian are animacy, definiteness, specificity and topicality (Farkas 1978, Dobrovie-Sorin 1994, Cornilescu 2001, von Heusinger & Onea 2008, Kamp & Bende-Farkas (submitted), among others). We focus on definiteness and specificity, which are mapped to the Referentiality Scale (Aissen 2003 among others), as illustrated in Table 1. We restrict our investigation to human direct objects in postverbal position.

Table (1): *Pe*-marking of postverbal human direct objects in Romanian depending on the Referentiality Scale

<table>
<thead>
<tr>
<th>pe-marking for human postverbal direct objects</th>
<th>pers. pron.</th>
<th>&gt; PN</th>
<th>&gt; def. NP</th>
<th>&gt; spec. indef NP</th>
<th>&gt; non-spec. indef NP</th>
</tr>
</thead>
<tbody>
<tr>
<td>modified</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+/-</td>
<td>–</td>
</tr>
<tr>
<td>unmodified</td>
<td>+</td>
<td>+</td>
<td>+/-</td>
<td>+/-</td>
<td>–</td>
</tr>
</tbody>
</table>

Full personal pronouns referring to animate entities are always marked with *pe* and doubled by a clitic in synchronic Romanian¹, as (4) shows. Proper names referring to humans are always *pe*-marked, as in (5). Modified human definite NPs in direct object position are generally *pe*-marked, like in (6a) while the form without *pe* is rather marginal, as in (6b).

(4) Maria îl asculta pe el.
Mary CL listens PE he
‘Mary listens to him.’

(5) Am vazut-o pe Maria.
Aux. seen-CL PE Mary
‘I have seen Mary.’

(6) A: Un băiat merge la doctor.
a boy goes to doctor.
(a) Doctorul îl examinează pe băiatul bolnav.
Doctor.DEF CL examines PE boy.DEF sick
‘The doctor examines the sick boy.’
(b) #Doctorul examinează băiatul bolnav.
Doctor.DEF examines boy.DEF sick
‘A boy goes to the doctor. The doctor examines the sick boy.’

2.1 Indefinite NPs

*Pe*-marking of indefinite human direct objects is optional and the literature (Farkas 1978, Dobrovin-Sorin 1994) assumes that specificity is the main triggering parameter for *pe*-marking. Following Farkas (1994) and von Heusinger (2011) we can distinguish between different kinds of specificity. We restrict the discussion of specificity to scopal specificity, as in (7), specificity in opaque contexts (*referential specificity*) as in (8), and epistemic specificity in transparent contexts as in (9). Scopal specificity with extensional operators and referential specificity with intensional operators triggers *pe*-marking. While the sentence (7a) is ambiguous between a specific (or wide scope) reading and a non-specific (or narrow scope)

¹ DOM in Romanian is generally accompanied by clitic doubling, i.e. the occurrence of a co-indexed weak pronoun. A doubling clitic is optional, obligatory or blocked, depending on semantic features of the head noun and further syntactic constraints. In this paper we will address the phenomenon of DOM in Romanian as a whole, thus, in the following sections, will not make an explicit distinction between clitic doubling and *pe*-marking (but see Dobrovie-Sorin 1994, Gramatica Limbii Române 2005).
reading, the non-specific reading in (7b) is ruled out due to the presence of *pe* (Dobrovie-Sorin, 1994). The same variation in readings between specific and non-specific readings is maintained for constructions with intensional operators, like in (8). To be more exact: non-specificity blocks the appearance of *pe* (see also Tigau 2012 for Romanian; Lenoetti 2004, Lopez 2012 for Spanish).

(7) scopal specificity
(a) Toţi bărbaţii iubesc o femeie.
All men love a woman
‘All men love a woman.’ (specific / non-specific)
(b) Toţi bărbaţii -o iubesc pe o femeie.
All men CL love PE a woman
‘All men love a woman.’ (specific)

(8) specificity in opaque contexts (referential specificity)
(a) Ion caută o secretară.
John looks for a secretary
‘John looks for a secretary.’ (specific / non-specific)
(b) Ion o caută pe o secretară.
John CL looks for PE a secretary
‘John looks for a secretary.’ (specific)

Thus, (8a) could have an interpretation where John is looking for a particular secretary (the specific reading) as well as one where any secretary will do (the non-specific reading), while (8b) only allows the specific interpretation.

It seems that we cannot extend this contrast to epistemic specificity, as illustrated in (9). Epistemic specificity is understood as the knowledge of the speaker or of some other salient agent about the identity of the referent. If *pe* showed a contrast with respect to epistemic specificity, we would expect the referent associated to the indefinite *pe un prieten* (‘pe a friend’) in (9b) to be known by the speaker or some other salient agent, while the sentence (9a) would allow both an epistemic specific and an epistemic non-specific reading.

(9) epistemic specificity
(a) Petru a vizitat un prieten.
Petru has visited a friend
‘Petru visited a friend.’
(b) Petru l-a vizitat pe un prieten.
Petru CL has visited PE a friend
‘Petru visited a friend.’

Clear judgments for these contexts are difficult to get. It seems that both sentences are compatible with a continuation like: (i) *I do not know the friend*, or (ii) *I do know the friend*. It thus seems that the contrast in (9) is due to a different feature that cannot be thoroughly captured by epistemic specificity. This provides evidence that the contrast is associated with a different kind of feature. We assume that the formal alternation between the two forms can be associated with the weaker discourse-pragmatic property Discourse Structuring Potential. We will elaborate upon this property in section 3 and 4.
2.2 Definite unmodified NPs

Definite modified direct objects are always *pe*-marked. However, unmodified definite direct objects also show an optional *pe*-marking due to a grammatical rule, which blocks the co-occurrence of the enclitic definite article with *pe* as in (10a). In (10), the discourse referent is given, definite, referential, specific, accessible etc. and can be picked up by the definite noun phrase with the enclitic article in (10b) or by *pe* and the “bare” noun (and clitic doubling) in (10c). Informants do not have clear intuitions about the contrast between the readings of (10b) and (10c) and there are not many similar instances in corpora, where such an option is available.

(10) Context: Un băiat merge la doctor. (A boy goes to the doctor.)
   (a) *Doctorul îl examinează pe băiatul.
       Doctor.DEF CL examines PE boy.DEF
   (b) Doctorul examinează băiatul.
       Doctor.DEF examines boy.DEF
   (c) Doctorul îl examinează pe băiat.
       Doctor.DEF CL examines PE boy

The alternation between *pe*-marking and the lack of the definite article vs. the lack of *pe*-marking and the definite article is further restricted by various blocking factors. Here, we provide the possessive dative as one example for illustration (see for more examples Chiriacescu 2007, von Heusinger & Onea 2008, von Heusinger & Chiriacescu 2011). At sentence level, *pe*-marking is ruled out whenever the definite article is modified by a possessive preverbal (11a) or postverbal dative (11b), even in cases where the NP is further modified by an adjective (see for a similar blocking effect on DOM in Spanish by possessive datives, Bruggè & Brugger 1996, 29-30).

(11) (a) Maria îşi înţelege (*pe) buna prietenă.
       Maria DAT understands PE good.DEF friend
       ‘Maria understands her good friend.’
   (b) Înţelegându-şi (*pe) frumoasa soţie a făcut […]
       understanding-DAT PE beautiful.DEF wife has made
       ‘Understanding his beautiful wife, he made […]’

In the following we will test whether the typical semantic-pragmatic parameters of definite noun phrases, i.e. (i) anaphoricity, (ii) uniqueness vs. familiarity, (iii) scope, (iv) referential readings in opaque contexts and (v) referential vs. attributive readings in transparent contexts, influence the use of *pe*-marking.

An anaphoric definite noun phrase is linked to its coreferent antecedent, as in (12). In a small survey we found various examples of the form in (12) where we replaced one form for the other, but we did not find any significant difference in meaning between the alternate forms.
(12) Ion a cunoscut un politician şi un scriitor.
    John has met a politician and a writer
    ‘John met a politician and a writer.’
    (a) A doua zi a văzut politicianul la televizor.
        the next day has seen politician.DEF on TV
    (b) A doua zi i-a văzut pe politician la televizor.
        the next day CL has seen PE politician on TV
    ‘He saw the politician on TV next day.’

The following context allows us to make an interesting distinction with respect to (situational)
salience or familiarity on the one side and descriptive uniqueness on the other. Both sentences
(13a) and (13b) are felicitous in the given context, however, they give rise to different
interpretations on the side of the hearer, i.e. whether he can see the escaping prisoner or not.
In the pe-marked version (13a), the hearer is not only informed of the existence of the escaped
prisoner, but he is also instructed to locate the referent in the immediate situation of utterance.
If the policeman within the prison utteres sentence (13a), then the policeman outside the prison
must see the runaway. Otherwise, he would ask a wh-question to find out which prisoner
escaped, where he went, etc. On the other hand, if the first policeman utters sentence (13b)
with the definite article and without pe, the hearer does not have to see the runaway to
understand what happened and how he should react. He would have to look where (and who)
the (unique) runaway actually is.

(13) Context: A policeman is standing guard outside a prison, which is surrounded by a
twenty-foot wall. Suddenly, he hears the voice of a colleague policeman from the other
side:
    (a) Prinde -l pe fugar.
        catch CL PE runaway
    (b) Prinde fugarul.
        catch runaway.DEF
    ‘Catch the runaway.’

Definite noun phrases generally show wide scope, but we can construct cases where definite
noun phrases receive narrow scope with respect to a universal quantifier, as in (14). Here we
find an interesting contrast: In (14a), the phrase proprietarul (‘the owner’) could be
interpreted as ‘Each renter salutes his/her owner’, even if the noun phrase is further modified
by the adjective bogat (‘rich’). However, the pe-marked definite in (14b) clearly underlines
the fact that the mentioned owner is the same for each renter. In both readings the definite
description expresses a uniqueness presupposition, but with respect to different domains. The
definite article allows for local domains created by the universal quantifier, while the pe-
marked version only allows for wide scope, similar to demonstrative expression and to the
scopal specificity of indefinite noun phrases discussed above.

(14) (a) Toţii chiriaşii salută proprietarul bogat.
        all renters salute owner.DEF rich
        ‘All renters salute the rich owner.’
    (b) Toţii chiriaşii îl salută pe proprietar / pe proprietarul bogat.
        all renters CL salute PE owner / PE owner.DEF rich
        ‘All renters salute the owner / the rich owner.’

Definites in opaque contexts show different readings - they can have a de re or a de dicto
reading (Quine 1956). We cannot go into the details of the analysis of these constructions at
this point, but a simplistic interpretation would suggest that the contrast could be reconstructed with scope. For the *de re* reading, the noun phrase takes scope over the intensional operator *seek*, while in the *de dicto* reading, the noun phrase takes narrow scope. Example (15b) with the *pe*-marked definite expresses a *de re* reading, i.e. there is an administrator such that Peter is looking for him or her. The property reading is excluded with *pe* in such contexts, but not in others (see Cornilescu 2013 for a discussion of objects of reflexives and middles). The form *administratorul* without *pe* and with the definite article has a *de re*- and *de dicto* reading. The latter one refers to a situation in which Peter is seeking an administrator, whoever s/he might be.

(15) (a) Petru caută administratorul.
    Petru seeks administrator

(b) Petru il caută pe administrator
    Petru CL seek PE administrator.

‘Peter seeks for the administrator.’

The classical contrast of definite noun phrases in a transparent sentence, i.e. in a simple declarative sentence without further operators, is one between a *referential* and an *attributive* reading (Donnellan 1966). In the referential reading, the speaker has a particular individual in mind, while in the attributive reading, any individual that matches the descriptive content qualifies as a good referent. This contrast is very similar to the (epistemic) specific vs. non-specific readings of indefinites (Partee 1970):

(16) (a) La inaugurarea aeroportului Braşov, cetăţenii vor invita primarul.
    at inauguration.DEF airport.DEF Braşov citizen.DEF will invite mayor.DEF

(b) La inaugurarea aeroportului Braşov, cetăţenii il vor invita pe primar.
    at inauguration.DEF airport.DEF Braşov citizen.DEF CL will invite PE mayor

‘At the inauguration of the Brasov airport, the citizens will invite the mayor.’

Both (16a) and (16b) are felicitous in this context, but with different readings. (16a) refers to whoever person might occupy the mayor position at the time the airport will be opened (a particular function). In (16b), the *pe*-marked noun phrase is not tight to the function its associated referent designates, but to the individual that occupies this position, say Mr. Jones. Note that only functional nouns like *mayor, president, murderer* etc. show a clear contrast between a referential and an attributive reading. This observation does not hold for sortal concepts like *boy* as in example (3), for which another explanation is needed.

2.3 Comparing conditions for DOM with indefinites and definites

We can now compare the referential functions expressed by *pe*-marking of indefinite direct objects with those referential functions expressed by *pe*-marking of definite direct objects. While the alternation is well documented and discussed for indefinites, there is not very much work on the alternation with definites. In Table 2, we summarized the observations of the last two subsections. Both *pe*-marked and unmarked definites can be anaphoric, but only *pe*-marked definites refer to referents that are directly perceivable, while unmarked ones are uniquely identified by their descriptive content. These observations hold for definite noun phrases only. *Pe*-marked indefinites and definites show wide scope, while unmarked noun phrases can have wide or narrow scope (at least in the case of the indefinites). *Pe*-marked indefinites and definites show referential or wide scope (or *de re*) readings in opaque contexts, and they tend to get a referential reading in transparent contexts. Again, this analysis seems more appropriate for definites than for indefinites.
### Table (2): Referential properties expressed by *pe* for indefinites and definites

<table>
<thead>
<tr>
<th></th>
<th>anaphoric</th>
<th>immediate situation</th>
<th>scope</th>
<th>opaque context</th>
<th>transparent context</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>pe + indef. N</strong></td>
<td>n.a.</td>
<td>n.a.</td>
<td>wide</td>
<td>wide</td>
<td>spec/non-spec</td>
</tr>
<tr>
<td><strong>Ø + indef. N</strong></td>
<td>n.a.</td>
<td>n.a.</td>
<td>narrow/wide</td>
<td>narrow/wide</td>
<td>spec/non-spec</td>
</tr>
<tr>
<td><strong>pe + N-Ø</strong></td>
<td>anaphoric</td>
<td>familiar / visible</td>
<td>wide</td>
<td>wide</td>
<td>referential</td>
</tr>
<tr>
<td><strong>Ø + N+def.</strong></td>
<td>anaphoric</td>
<td>unique</td>
<td>narrow</td>
<td>narrow</td>
<td>attributive</td>
</tr>
</tbody>
</table>

The overview presented above includes transparent contexts in which we find definite and indefinite direct objects with and without *pe*-marking. While informants report that there are differences in readings, we cannot find one clear referential property that determines this contrast. We therefore assume that the alternation between the *pe*-marked and the unmarked form is due to the discourse function of the (in)definite noun phrase. Definite and indefinite noun phrases do have a “cataphoric” force or “forward looking function” and thus express a certain (additional) level of activation of the referent they are associated with (see von Heusinger 2007), in a way explained in the next section.

### 3. Forward looking function and Discourse Structuring Potential

A body of linguistic and psycholinguistic research has investigated various factors that influence the comprehension and production of different types of referring expressions (Givón 1983, Gundel, Hedberg & Zacharski 1993, Kehler et al. 2008, Arnold 2010). The majority of these studies focused on anaphora resolution, as it is commonly assumed that reduced referring expressions correlate with highly accessible or prominent entities. Furthermore, to determine the accessibility of a referent, researchers have generally employed a backward-looking perspective, determining the factors that license the usage of a particular type of referring expression at a particular stage in the discourse. In other words, given a certain type of referring expression (e.g. a pronoun), the factors that license its use were investigated.

In contrast to personal pronouns, which refer to previously mentioned and focussed entities, definite noun phrases display different kinds of *forward-looking* referential properties: First, they can be used for discourse-new entities or for entities introduced by a bridging or inference relation as well (Hawkins 1978, Vieira & Poesio 2000). Second, both types of definite noun phrases (familiar ones and first mentioned ones) change the accessibility or the activation of the associated discourse referents (von Heusinger 2003, 2007). Indefinite noun phrases are more often related to a “forward looking function” as their main function is to introduce a new discourse item that can be used as an antecedent for subsequent anaphoric terms (Karttunen 1976, Heim 1982, Kamp 1981 / 2013). In this study we extend this *forward-looking* perspective and test the effects of production-driven biases licensed by *pe*-marked and unmarked indefinite and definite direct objects in Romanian. We investigate the Discourse Structuring Potential (DSP), which can be best measured by two textual characteristics of their associated referents that pertain to the following discourse (Chiriacescu & von Heusinger 2010, Chiriacescu 2011, Deichsel & von Heusinger 2011).

The first metric, referential persistence, reflects the likelihood that a particular referent will be picked again in the ensuing discourse (Givón 1983, Kehler et al. 2008). The second metric for DSP, topic-shift potential, is defined in terms of the likelihood that a referent will be mentioned in grammatical subject position. We focus on the subject position because different linguistic and psycholinguistic studies (e.g. Crawley & Stevenson 1990) have shown that referents mentioned in the syntactic subject position are more salient or accessible in a given
discourse than referents mentioned in other syntactic positions (e.g. as direct or indirect direct objects). For the sake of simplicity, the first instance in which a direct object referent becomes the grammatical subject in a matrix clause is treated as an instance of topic shift. Despite being mentioned in a rather non-preferential grammatical position (i.e. as a direct object), we expect pe-marked definite and indefinite noun phrases to show higher values for both metrics than their non-pe-marked counterparts.

4. Web-based experiment for prominence with pe-marking

To investigate whether the presence of the pe-marker boosts the prominence or salience of the referents associated with the direct objects realized as definite unmodified noun phrases, we used the metrics for discourse prominence developed for the experiment with indefinite noun phrases (cf. Chiriacescu & von Heusinger 2010). More precisely, we analyzed whether pe-marked definite noun phrases are (i) referentially persistent in the subsequent discourse (i.e. whether the referent headed by pe is likely to be continued), and (ii) more susceptible to shift the topic (i.e. in the sense of Givón 1983, Ariel 2001, among others) of the current discourse.² Let us now consider how we predict participants’ responses to pattern with respect to the two different metrics we tested. First, in light of the findings from the pe-marking experiment with indefinite noun phrases (Chiriacescu & von Heusinger 2010) and other experimental investigations, which showed that accessible / salient referents are more likely to be subsequently mentioned (Givón 1983, Gernsbacher & Shroyer 1989, Arnold 1998, among others), we predict that referents headed by pe will be referentially more persistent in the ensuing discourse, compared to referents marked with the simple definite article. Second, given the observation that important or salient referents tend to be mentioned in topic position (which in English generally corresponds to the grammatical subject position, e.g. Ariel 2001, Arnold 1998), we predict that in comparison to their unmarked counterparts, pe-marked direct objects will (i) be mentioned more often in the subsequent text, and will (ii) become the new topic in the following discourse.

4.1 Method

Participants

Twenty native speakers of Romanian participated in the experiment on pe-marking with indefinite noun phrases and other twenty native speakers of Romanian participated in the experiment on pe-marking with definite noun phrases. They received no incentive for taking part in the survey. It took about twenty minutes to complete an experiment.

Materials

The methodology used in this experiment was an open-ended sentence-continuation task. Participants were presented target items consisting of mini-discourses, as in table 4 and 5. Their task was to read the given story fragments and add five logical and natural-sounding sentence continuations for each of them. The first two sentences of each test item set the context of the story, and contained individual references to two characters. The first character was the clearly established topic of the mini-discourse, as it was mentioned in subject position at least once and was the referent the story was about. In the last sentence of each mini-discourse, the critical referent was introduced as an indefinite or definite noun phrase in direct

² Independently of these two textual characteristics, we considered the type of referring expression used to pick up the referent of the direct objects. We will not discuss the findings of this metric in this paper, but note that the likelihood of subsequent mention does not point in the same direction as the likelihood of being realized with a pronoun (see also the discussions in Kehler et al. 2008, Kaiser 2010, Chiriacescu & von Heusinger 2010). For the purposes of this paper, just note that pronominalization does not reflect discourse prominence as defined in this paper.
object position. We only manipulated the morphological realization of the target referents, which resulted in two conditions for each experiment, i.e. one condition in which *pe* heads an indefinite unmodified noun phrase (in Exp1) and a definite unmodified noun phrase (in Exp2) (see the left columns of Tables 3 and 4 below), and another condition in which the same direct objects are not *pe*-marked, i.e. they are headed by the simple indefinite article in Exp1 or followed by the enclitic definite article in Exp2 (see the right columns of Tables 3 and 4).

**Table (3): Sample experimental items from Exp1 on *pe*-marked indefinite NPs**

<table>
<thead>
<tr>
<th><em>pe</em>-condition</th>
<th>non-<em>pe</em>-condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aseară a fost extraordinar de cald. Pentru că nu mai rezista în casă, Graur să-a hotărt să iasă în oraș. Pe drum l-a văzut <em>pe un copil</em> intrând intr-un magazin. “It was extraordinarily warm outside yesterday evening. Because it was unbearable for him to stay home anymore, Graur decided to go downtown. On his way there he saw a <em>child</em> entering a store”.</td>
<td>Aseară a fost extraordinar de cald. Pentru că nu mai rezista în casă, Graur să-a hotărt să iasă în oraș. Pe drum a văzut <em>un copil</em> intrând intr-un magazin. “It was extraordinarily warm outside yesterday evening. Because it was unbearable for him to stay home anymore, Graur decided to go downtown. On his way there he saw a <em>child</em> entering a store”.</td>
</tr>
</tbody>
</table>

**Table (4): Sample experimental items from Exp2 on *pe*-marked definite NPs**

<table>
<thead>
<tr>
<th><em>pe</em>-condition</th>
<th>non-<em>pe</em>-condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>La petrecerea de aseara, Andrei a cunoscut un politician și un cântăreț de renume. Astăzi l-a întâlnit <em>pe politician</em> în piață. “At yesterday evening’s party, Andrew met a politician and a famous singer. Today he met a <em>politician</em> at the market.”</td>
<td>La petrecerea de aseara, Andrei a cunoscut un politician și un cântăreț de renume. Astăzi a întâlnit <em>politicianul</em> în piață. “At yesterday evening’s party, Andrew met a politician and a famous singer. Today he met a <em>politician</em> at the market.”</td>
</tr>
</tbody>
</table>

**Procedure and data analysis**

The first five main clauses (including subordinate ones, if there were any) of each continuation story provided by the participants were analysed. Two independent judges coded for two aspects of the direct objects: (i) their referential persistence and (ii) their topic shift potential. We coded 10 continuations for each condition of the two experiments (i.e. 10 responses for *pe*-marking with indefinite NPs and another 10 responses for non-*pe*-marking with the simple indefinite article in Exp1. In Exp2 we coded 10 responses for *pe*-marking with definite NPs and 10 responses for non-*pe*-marking). The global topic of the first and second sentence (e.g. *Graur and Andrei* in the test items in Table 3 and Table 4) received Subscript 1. Subscript 2 was used for the noun phrase whose form was manipulated in the critical sentence (e.g. *pe un copil* or *pe copilul* (‘*pe a child*’ in Exp1 or ‘*pe the child*’ in Exp2) in the left columns of Tables 4 and 5. Example (17) represents an example response for test item 3 for the *pe*-condition, and Table 5 illustrates the coding methods used.
(17) Example responses and coding methods from the story continuation experiment

La petrecerea de aseara, Andrei1 a cunoscut **un politician**2 și **un cântăreț de renume**3. Astăzi (pro)1 l-a întâlnit pe **politician**2 în piață.

‘At yesterday evening’s party, Andrew1 met a politician2 and a famous singer3. Today he1 met the politician2 at the market.’

| S1: | (pro)1 stia ca asta este sansa lui1. |
|     | ‘He1 knew that that’s his1 chance.’ |
| S2: | **Politicianul**2 era un pic grizonat, slabut, cu accent bariton. |
|     | ‘The politician2 had some greyish hair, was thin with baritone voice.’ |
| S3: | **Andrei**1 s-a dus spre el2 si (pro)1 i-a cerut ajutorul sa (pro)1 alegea un pepene bun. |
|     | ‘Andrei1 went towards him2 and he1 asked (him2) for help to choose a tasty water melon.’ |
| S4: | **Politicianul**2 s-a intros si (pro)2 i-a raspuns cu un aer distrat. |
|     | ‘The politician2 turned around and (pro)2 responded him1 in a distracted voice.’ |
| S5: | Il2 chema don Giuseppe si (pro)2 era inginer zootehnist de meserie. |
|     | ‘His2 name was don Giuseppe and he2 was a zootechnician engineer.’ |

Table (5): Coding methods for the test item presented in (17)

<table>
<thead>
<tr>
<th>Coding methods</th>
<th>First referent</th>
<th>Target referent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Subject)</td>
<td>(Object)</td>
</tr>
<tr>
<td>Anaphoric forms and grammatical function</td>
<td>refer per item / S</td>
<td>refer per item / S</td>
</tr>
<tr>
<td>S1</td>
<td>[pro1] (pron1)</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>[Sub1] (IO1)</td>
<td>2</td>
</tr>
<tr>
<td>S2</td>
<td>[def NP2]</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>[Sub2]</td>
<td>2</td>
</tr>
<tr>
<td>S3</td>
<td>[PN1, pron2] [pro1, CL2] (pro1)</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>[Sub1, PP2] [Sub1, IO2] (Sub1)</td>
<td>5</td>
</tr>
<tr>
<td>S4</td>
<td>[def NP2] [pro2, CL1]</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>[Sub2] [Sub2, IO1]</td>
<td>6</td>
</tr>
<tr>
<td>S5</td>
<td>[CL2] [pro2, DO2] [Sub2]</td>
<td>0</td>
</tr>
</tbody>
</table>

Referential persistence is measured by referents mentioned per sentence (referent / S) and the sum of all items up to S5 (i.e. a cumulative measure). Comparing the sums indicates at what stage in the discourse we have more anaphoric expressions referring to one referent compared to another. In the example (17) above, the referent of the target item in the pe-condition (i.e. the referent of the politician) exceeds in persistence the referent of the first referent (i.e. Andrei) in the last continuation sentence (S5). Furthermore, we verified in what sentence the target referent (i.e. the politician) becomes the subject and topic of a main clause. In example (17), this happens in sentence continuation 2 (S2).

4.2 Results

40 participants provided continuations for the initial story fragments. The results from the two metrics, referential persistence and topic shift potential reflect the discourse status of the stories’ referents. In the following sections we discuss the findings of the two textual characteristics in detail.

4.2.1 Referential persistence

The first textual characteristic investigated was referential persistence. Figure 1 displays the mean values for referential persistence of all referents of the test items in Exp1 with indefinite noun phrases. For the pe-condition, we notice a strong likelihood of the referent to me
mentioned in the following discourse. On the contrary, the direct object referents in the non-
*pe*-marked condition are picked up in the subsequent discourse less often.

Figure (1): Referential persistence of subject and object referents in Exp1 on indefinite direct objects

![Mean values (cumulative) for referential persistence of subjects and objects](image)

The same holds for the second experiment (Exp2) on definite noun phrases. Figure 2 shows
that the referents of the *pe*-marked direct objects realized as definite noun phrases are more
frequently re-mentioned than the unmarked ones. The predictions concerning this metric are
confirmed, as the *pe*-marked referents were picked up more often in the subsequent discourse
than the referents of the unmarked direct objects.

Figure (2): Referential persistence of subject and object referents in Exp2 on definite direct objects

![Mean values (cumulative) for referential persistence of subjects and objects](image)

In sum, participants preferred a continuation story that evolved around the referent of the
subject, thus taking it up more often, unless the direct object referent was *pe*-marked. In such
a case, the referent of the *pe*-marked referent becomes a better competitor for the subject
referent in terms of referential persistence.

### 4.2.2 Topic shift potential

The second textual characteristic investigated was the topic shift potential of direct object
referents. Recall that each mention of a direct object in grammatical subject position was
counted as an instance of topic shift. The counts for the topic shift potential are cumulative.
The findings condensed in Figure 3 and 4 reveal several patterns. First, the referent of the *pe-
marked direct object displays a stronger preference to become a subject in the continuation
sentences (S1-S5) than the referent of the non-*pe*-marked direct object referent. Second, while
almost all participants mentioned the referent of the *pe*-marked direct object sooner or later as
a subject in the continuation text, the unmarked direct object became a subject in less than
25% of cases. Third, Figure 5 shows that the referent of the unmarked direct object was never
picked up in subject position in the first two continuation sentences (S1 and S2) provided by
the participants. On the contrary, the referent of the *pe-*marked direct object was picked up in the first two continuation sentences, even though the rate was not high.

Figure (3): Topic shift potential of referents in both conditions for Exp1 on indefinite direct objects

![Graph](image1)

Figure (4): Topic shift potential of referents in both conditions for Exp2 on definite objects

![Graph](image2)

The findings concerning the topic shift potential of direct objects confirmed the initial predictions, as the referents of the *pe-*marked direct objects displayed a higher expectancy to be mentioned again as topics in a main clause (i.e. in subject position) in comparison to the unmarked ones. The observations hold for both Exp 1 with indefinite noun phrases, and Exp2 with definite noun phrases.

4.3 Discussion

The findings with respect to the Discourse Structuring Potential of direct objects realized as definite noun phrases parallel those reported in Chiriacescu & von Heusinger (2010) about the discourse behaviour of indefinite noun phrases in direct object position. Up to the last continuation sentences (S5), the *pe-*marked referents (i) exceeded their unmarked counterparts in referential persistence (76% vs. 24%) and (ii) became the topic of the discourse more often than the non-*pe-*marked referents (in 80% vs. 15% of the cases).

The referential persistence and topic shift underline the privileged status of the *pe-*marked referents (whether expressed by an indefinite NP or by a definite unmodified NP) and thus confirmed Predictions 1 and 2 (cf. Chiriacescu 2011 on similar effects of indefinite- *this* in English and indefinite- *son* in German; Deichsel & von Heusinger 2011 and Deichsel 2013 for indefinite *dies* (*this*) in German). The likelihood of a referent to be mentioned in the subsequent discourse is not a reflex of the high activation of that referent, but rather represents a mechanism employed by the speaker to link the hearer’s attention to an entity which will be further elaborated upon. Psycholinguistic research (Levy 2008) has convincingly shown that statistical regularities are observed at different levels of linguistic output. It seems that hearers identify frequency patterns in order to predict what is likely to
occur in the following context. The referential persistence of the \textit{pe}-marked indefinite descriptions analysed here shows that language users make use of such regularities at the discourse level as well.

5. Conclusion - The function of \textit{pe} in Romanian

In light of the findings of the experiment presented in section 4, we argue that \textit{pe} signals the Discourse Structuring Potential of the referents it precedes. More concretely, it was shown that \textit{pe}-marking has the same or a very similar function for definites and indefinites in neutral (transparent) contexts, as illustrated in (3) and (9). Such referents were shown to be more recurrent in the following discourse and to be more prone to shift the topic of the discourse. These observations add an additional dimension to the analysis of DOM: besides referential (specificity), lexical (animacy), information structural and backward-looking discourse properties, DOM indicates forward-looking discourse properties (i.e. a \textit{Discourse Structuring Potential}) as well. These properties are not pragmatic, but a “built in” or semantic, similar to specific adjectives like \textit{a certain}, or discourse particles like \textit{then}, \textit{therefore}, etc.

6. References


Ionin, Tania. 2006. This is definitely specific: specificity and definiteness in article systems. *Natural Language Semantics* 14(2): 175-234.


1. Introduction

This paper aims at putting together a series of facts – all of them pertaining to the domain of DP interpretation- that have, to my knowledge, never been connected before, and seem to point towards a significant generalization on the role of contrast. Such facts involve different kinds of expressions, from bare plurals, indefinite DPs and interrogatives to strong pronouns and definite descriptions. In a nutshell, the idea I would like to put forward is the following one. None of the expression types I just mentioned is inherently contrastive in their basic semantics. However, their use in certain environments is typically associated with a contrastive reading. It seems, thus, that contrast appears as a pragmatic effect triggered by the interaction between the linguistic meaning of the nominal expression and certain features of the context: more precisely, my claim is that contrast operates as a last resort interpretive mechanism that is able to “rescue” a nominal placed in a “hostile” environment, i.e. an environment where the nominal does not fit in adequately for some reason. Contrast is thus triggered as an inferential solution for some interpretive mismatches. It is convenient to make clear from the beginning that, instead of dealing with constructions that encode contrast, I intend to concentrate on constructions where contrast is pragmatically inferred as a component of the interpretation assigned to a DP. The central feature of my account lies in taking advantage of an adequate division of labor between the grammatical system (syntax and semantics) and the cognitive principles that are responsible for our inferential abilities (pragmatics). Most of the data I will discuss will come from Spanish and other Romance languages.

An analysis of this cluster of phenomena should ideally shed some light both on the semantics / pragmatics interface in the interpretation of NPs / DPs, and on the nature of contrast and related notions like D-linking and Contrastive Topic. Needless to say, my perspective will also raise questions that I will not be able to answer in a fully satisfying way, but hopefully such questions will provide us with some stimuli for extending the research. I try to offer some sketchy ideas on these issues in section 4. The rest of the paper is organized as follows. Section 2 is devoted to contrast in indefinite expressions: bare nouns receive a special attention, but indefinite determiners and interrogatives are considered as well. Section 3 is a brief sketch on contrast in definite DPs, in particular definite descriptions and strong pronouns. Finally, in section 4 a couple of speculative ideas on contrast in interpretation close the paper.

2. Indefinites

2.1. Bare plurals as topics

Bare nominals and indefinites provide us with valuable data on the role of contrast in the acceptability of NPs and DPs. Let’s start the discussion with a review of some basic facts in the grammar of bare nouns in Romance, with special attention to bare plurals in Spanish.

---

1 This section essentially reproduces ideas developed in Leonetti (forthcoming).
A well known constraint on the distribution of bare nouns in Spanish is the Naked Noun Constraint (NNC). In its original formulation (Suñer 1982: 209), the constraint states that “an unmodified common noun in preverbal position cannot be the surface subject of a sentence under conditions of normal stress and intonation”. This rightly excludes examples like (1) and (2), usually considered ungrammatical sentences in Spanish:

(1) *Niños jugaban en la calle.
   Children play.PST.3PL in the street
   ‘Children were playing in the street.’

(2) *Turistas llegaron a la ciudad.
    Tourists arrive.PST.3PL to the city
    ‘Tourists arrived in the city.’

Suñer’s formulation of the NNC mentions some factors that are related to information structure, namely “conditions of normal stress and intonation”. This is important for the generalization to hold, since a special intonation contour that marks the preverbal subject as contrastive Focus (cf. (3a)) or as a dislocated Topic (cf. (3b)) can make sentences like (1) and (2) acceptable:

(3) (a) TURISTAS llegaron a la ciudad.
    tourists-FOC arrive.PST.3PL to the city

(b) Turistas, llegaron (pero no demasiados).
    tourists-TOP arrive.PST.3PL (but not too many)

A natural account of the contrast between (1)-(2) and (3) should be based on two assumptions. The first one is uncontroversial: the “preverbal position” occupied by the bare plural is not the same in (1)-(2) –the canonical subject position (Spec, IP in the generative tradition)- and (3) –a peripheral position, and a different one in (3a) and (3b), independently of the particular view of the left sentential periphery that one might choose. The second one may be controversial, but in my opinion gives rise to the most promising account of the constraint on bare nouns in preverbal position: it amounts to assuming, following Suñer’s original insight, that the NNC is related to information structure. More precisely: bare nouns have to be licensed by semantic incorporation (in any of its versions, cf. Dobrovie-Sorin & Laca 2003, Cohen & Erteschik-Shir 2002, Espinal 2009), but such operation is constrained by information structure, and banned from a predominantly topical position like Spec, IP in null-subject languages. The NNC arises as a result of the incompatibility of semantic incorporation (needed for non-referential expressions like bare nouns) and the topical nature of the preverbal subject position in null-subject languages. In Suñer’s (1982:225) terms, “Naked nouns never appear as subjects in preverbal position because these nouns cannot be interpreted as the theme of the sentence”. If this is correct, we have an account of the ungrammaticality of (1)-(2). But why are the examples in (3) perfectly acceptable? Apparently, focalization and dislocation have the power of repairing the anomaly derived from the incompatibility between bare nouns and the preverbal position. However, things are actually a bit more complicated. Focalization and dislocation cannot give acceptable results unless bare plurals are independently licensed as subjects in postverbal position, i.e. unless semantic incorporation can successfully take place in a postverbal position (cf. Laca 1996:259-260). The examples in (4) show that bare plurals cannot make good subjects of transitive verbs –at least, those that represent Individual-Level predicates- even in postverbal position, probably because semantic incorporation cannot apply

---

to external arguments in a straightforward manner, and in that case operations of focalization or dislocation cannot rescue them.

(4) (a) *Tienen orcas una gran aleta dorsal.

have.3PL orcas a big findorsal

‘Killer whales have a big dorsal fin.’

(b) *ORCAS tienen una gran aleta dorsal.

(c) *Orcas, tienen una gran aleta dorsal.

What focalization and dislocation tell us with respect to bare plurals is not that they can provide licensing mechanisms for them, but rather that both operations are associated with a meaning ingredient that makes preverbal bare plurals fully acceptable once they are moved away from the basic position where semantic incorporation licenses them. Such ingredient is contrast.

Contrast is obviously present in contrastive focalization, and I will have nothing special to say about this issue. Focalized bare plurals cannot be interpreted as topics, so they fall out of the scope of the NNC and are correctly predicted to be grammatical. Contrast is also present in left dislocations, as (3b) shows, and in this case a more detailed discussion is worth, since at first sight we face a serious counterexample for the NNC: bare plurals should be excluded when they are placed in topic position, but when they are dislocated they are explicitly marked as sentential topics and nonetheless are perfectly acceptable. The central questions are two: 1. why are dislocated bare plurals acceptable, in spite of being interpreted as topics, while they are excluded as preverbal subjects (in null subject languages)? 2. does contrast play some role in the licensing of dislocated bare plurals?

I addressed such questions in Leonetti (forthcoming), and here I will simply take up some ideas from that paper. The different behavior of bare plurals as preverbal subjects and dislocated topics is a result of the different properties of two kinds of topics. The crucial fact is that left dislocation does not impose the same interpretive requirements on nominals than Spec,IP does. Preverbal subjects in Spanish are unmarked, default topics, in sentences that express categorical judgments. I assume that Spec,IP is not endowed with a [+topic] feature. Topicality is not encoded in the preverbal subject position in null-subject languages. It is rather inferred by default, unless the context favors a thetic–all-focus–interpretation for the sentence. Left dislocation, on the other hand, marks a constituent as an external, marked topic. Dislocated phrases are linked to a sentence-internal anaphoric element—a clitic, or a null resumptive—. Dislocated bare plurals, in particular, are licensed in the internal position, where semantic incorporation is possible. Notice that this still leaves the central question without an answer—why bare plurals are possible topics only when detached, and not in the preverbal subject position? My claim is that the asymmetry is related to another fact already pointed out in Suñer (1982:231): all left-dislocated bare nouns must be contrastive. As Suñer puts it (1982: 236), “left-dislocated naked nouns are thematic but contrastive…naked nouns cannot be non-contrastive themes”. In fact, the dislocated bare plural in (3b), turistas ‘tourists’ is interpreted as a contrastive topic: it typically evokes a set of contextual alternatives (in (3b), < tourists, workers, immigrants, pilgrims…>, but other sets might be activated in different contexts). Moreover, it has been observed (cf. Brunetti 2009) that Romance left dislocation is quite often contrastive, probably as a consequence of its marked status with respect to the unmarked topical nature of the subject in SV(O) sentences. In Spec,IP, on the contrary, contrastive readings are not available for bare nouns, under normal conditions.

A way to put together all these facts could be the following one. Left dislocation encodes an instruction to interpret the dislocated phrase as the address under which the information carried by the sentence is entered. This is what the classical notion of aboutness amounts to. The instruction has to be obligatorily satisfied in the interpretive process. Definite, referential,
familiar DPs are prototypical topics, and thus expressions that make optimal addresses for information update. On the other hand, indefinite DPs and all kinds of non-referential nominals—for instance, bare nouns—are non-prototypical topics. My claim is that indefinites and bare nouns can in fact act as sentence topics, only if some kind of inferential adjustment is triggered, as a part of the interpretive process, that solves the mismatch between the instruction associated to dislocation and that of the linguistic expression: as a result of this reinterpretation mechanism, the expression is assigned a reading that allows the hearer to comply with the instruction. I suggest that the contextual effect triggered by the need to satisfy the instruction associated with external, marked topics is contrast. The prediction is, thus, that all indefinite and non-referential topics will be contrastive (see section 2.3 for additional data).

From this point of view, contrast is a significant factor in obtaining a relevant interpretation for bare nominals used as sentence topics. Contrast makes bare plurals become acceptable topics. In left dislocation, it is triggered by the need to comply with a specific interpretive procedure. In preverbal subjects, on the contrary, nothing forces a contrastive reading, as there is no specific instruction to follow. Moreover, semantic incorporation is banned. The consequence is that bare nominals as subjects cannot be interpreted as topics. Now we have an answer for the previous question about external and internal topics, and such answer is based on the role of contrast as a sort of last resort interpretive mechanism. It is quite clear that at this point a new set of relevant questions enters the picture: how is contrast able to rescue the acceptability of certain nominal expressions? Why is it contrast, instead of some other notion, that is involved in the distribution of bare nouns and indefinites?

Despite their importance, I prefer to delay the discussion of these points until the final section, when more empirical evidence for the role of contrast will be available and a wider perspective on the facts will be within reach. For the moment, I will simply add some more data concerning contrast and acceptability of bare plurals in Spanish (in section 2.2), and a brief excursion on the relevance of contrast to conclude this section.

2.2 Bare plurals: additional evidence for the role of contrast

Until now, I have limited the discussion on bare nouns to two positions—preverbal subject and dislocated topic. But there is additional evidence for the active role of contrast in the interpretation of bare nouns. In this section I focus on three issues that should reinforce the empirical support for the generalization put forward in the previous section: marked objects, postverbal subjects, and modifiers inside the NP.

2.2.1 Marked objects

A well known feature of Spanish syntax is Differential Object Marking (DOM): animate (and mostly specific) direct objects must be marked with the preposition a. Less familiar, and not entirely systematic, is the fact that unmodified bare plurals tend to be excluded as marked objects (cf. Brugè & Brugger 1996, Leonetti 2004, López 2012: 52-53), as shown in the examples in (5). In unmarked objects, without a, bare plurals are acceptable (if the verb allows semantic incorporation).

(5) (a) #En el poblado vi a pescadores. (cf. Vi pescadores.)

In the village see.PST.1SG to fishermen

‘In the village I saw fishermen.’

I cannot offer a full discussion of this claim here, but the generalization seems to be correct, both for bare nouns and for indefinite DPs. I don’t know of any indefinite topic that has a non-contrastive interpretation. See Erteschik-Shir (1997), Arregi (2003), Kiss & Gyuris (2003), and Giurgea & Remberger (2011) for related, though not identical, proposals.
(b) #Estoy contratando a traductores. (cf. Contrato traductores.)
Be.PRS.1SG hiring to translators
‘I am hiring translators.’

The constraint imposed by marked objects on bare plurals seems to be equivalent to the NNC for preverbal subjects. In fact, the same factors, i.e. focalization and modification, are able to rescue the acceptability of bare plurals in both contexts. The examples in (6) show the effects of (in situ) contrastive focus in DOM contexts, and those in (7) illustrate the effects of restrictive modifiers.

(6) (a) En el poblado vi a PESCADORES, no a turistas extranjeros.
In the village see.PST.1SG to fishermen-FOC, not to tourists foreign
‘In the village I saw FISHERMEN, not foreign tourists.’
(b) Yo contrato a TRADUCTORES, no a REDACTORES.
I hire.PRS.1SG to translators-FOC, not to editors-FOC
‘I hire TRANSLATORS, not EDITORS.’

(7) (a) En el poblado vi a pescadores ya mayores.
In the village see.PST.1SG to fishermen already old
‘In the village I saw old fishermen.’
(b) Yo contrato a traductores con experiencia.
I hire-PRS.1SG to translators with experience
‘I hire experienced translators.’

The data suggest that bare plurals are excluded as marked objects for the same reason that they are excluded as preverbal subjects. Topicality could be invoked as the semantic feature that is responsible for the limited distribution of bare plurals in DOM contexts, if the preposition a is analyzed as a sort of secondary topic marker (as in Leonetti 2004), but this is actually not crucial for my argumentation, and it is not an issue that I can deal with here. From a purely descriptive point of view, the essential fact is that both preverbal subjects and marked objects are hostile environments for bare plurals, and in both cases contrast is the mechanism that is able to make them acceptable, either through focalization or through restrictive modification (I am assuming that modifiers trigger contrast inside NPs/DPs, since they evoke a set of contextual alternatives that contrast with the explicit information in the nominal). This demonstrates that the difference in acceptability between (1)-(2) and (3) is not an isolated phenomenon concerning only subjects, but rather an instance of a general interpretive mechanism.

López (2012:52) accounts for the limited distribution of bare plurals assuming that their Case requirements must be satisfied by incorporation and that they cannot move, neither to Spec,IP as subjects, nor to the functional slot where marked objects—analyzed as scrambled objects—must rise. As for the effects of focalization and modification, he briefly mentions the proposal that “the addition of a modifier or of contrastive focus turns the nominal phrase into a bigger type of constituent that can be selected by K” (2012:53), i.e. the two factors turn the nominal into a phrase that does not need to incorporate to satisfy its Case requirements and thus behaves like a true argument. I agree with the basic intuition about focus and modification, but we still do not fully understand why they allow speakers to override the incompatibility of DOM with bare nouns. In López’s proposal, focalization and modification are apparently able to change the syntactic status of bare nouns; however, there is no clear evidence supporting this interpretation of the facts. I prefer to assume that the key factor for the acceptability of bare nominals in (6) and (7) is the presence of contrast, as I have been arguing for here: if acceptability increases, it is due to this component of interpretation.
Moreover, I am reluctant to accept a purely syntactic view of the constraint in DOM contexts, given that some counterexamples can be found even in the recent *Nueva Gramática de la Lengua Española* (RAE 2009:§15.11g-h, §34.9), with unmodified bare plurals as marked objects:

(8) (a) … así que al final contraté a albañiles.
    ‘… so I ended up hiring bricklayers.’

If examples like the one in (8) are considered acceptable, it may well be that the constraint is much softer than a genuine syntactic restriction, and probably we need a thorough and detailed revision of the alleged incompatibility of DOM and bare plurals—something that falls outside the limits of this paper. However, RAE (2009:§15.11g) crucially emphasizes that *albañiles* ‘bricklayers’ in (8) receives a “type reading”, i.e. a reading that highlights the kind of persons that were hired, and not the plurality of individuals. I would claim that a type reading is simply the result of focalizing the bare plural, thus activating a set of contextual alternatives where *bricklayers* contrasts with, say, *plumbers, painters and electricians*. Again, contrast is required to obtain an optimal interpretation of the bare plural in a syntactic context that otherwise would exclude it. This points towards a semantic-pragmatic account of the problem and disfavors a purely syntactic one.

In any case, the proposal in López (2012) contains some valuable insights, like the analysis of Spanish DOM as a case of object scrambling. Scrambled objects—in languages that resort to such syntactic device—behave in many respects like marked objects in DOM languages. Interestingly, object scrambling in Germanic languages provides us with further evidence for the effects of contrast on bare nouns. Kallulli (2005:32) points out that existential bare plurals cannot scramble in German and Dutch, unless they are contrastively focused, as shown in (9).

(9) German:

(a) Anna hat nicht Zeitungen gelesen.
    ‘Anna hasn’t read newspapers.’

(b) *Anna hat Zeitungen nicht gelesen.

(c) Anna hat ZEITUNGEN nicht gelesen.

Kallulli (2005:36) mentions similar effects with bare singulars in Albanian and Greek: when bare singulars occur as subjects of unergative and transitive predicates, they must necessarily be focused, as in (10)-(11).

(10) Albanian:

    GJARPËR e kafshoi An-ën.  (Albanian)
    snake-FOC her bit An-the
    ‘It was a snake that bit Anna.’

---

4 Scrambling and contrastive readings are not necessarily related in the same way in all languages, and not all kinds of object scrambling trigger the same interpretive effects. However, all I need to support my proposal is that certain instances of object scrambling show properties that are parallel to those exhibited by preverbal subjects in null subject languages. Detailed analyses of scrambling and contrast in different languages can be found in Neeleman & Vermeulen (eds.) (2012).
It was a snake that had bitten Costas.’

Scrambling is not involved this time, but there is a clear connection with all previous data, since the constraint affects bare nouns as external arguments. The common property of marked objects and external arguments is that they represent prominent positions that tend to obstruct and block semantic incorporation. Such obstacles for incorporation are typically circumvented by means of contrast, both in preverbal subjects / external arguments and in marked / scrambled objects, as the next section will try to make clear.

2.2.2 Postverbal subjects in Spanish
The distribution of bare plurals as postverbal subjects in Spanish provides us with interesting clues on the factors that trigger contrast as a repair mechanism. It is widely assumed that bare plurals can occur as postverbal subjects of unaccusative verbs with acceptable results. This is expected, if semantic incorporation can usually take place from internal argument – complement- positions. As for subjects of transitive and unergative verbs –external arguments-, the classical stance is that bare plurals are excluded, unless certain special conditions hold. I will concentrate on one of such conditions: word order as an indication of focus structure. The relevant examples are reproduced in (12), where the verb is transitive (aprobar ‘pass’).

(12) (a) Han aprobado el examen (hasta) [extranjeros].
    have.PRS.3PL passed the exam till foreigners.
    ‘Even some foreigners passed the exam.’
(b) #Han aprobadó extranjeros el examen.
    Han.PRS.3PL passed foreigners the exam
(c) Han aprobadó EXTRANJEROS, el examen.
    Han.PRS.3PL passed FOREIGNERS-FOC the exam
    ‘FOREIGNERS passed the exam.’

The sentences in which the subject occurs in final position (a, c) are perfectly acceptable. Example (b) displays a VSO order, and is slightly degraded. The difference is related to focus structure: (a) and (c) force narrow focus on the postverbal subject and make a set of contextual alternatives associated with it easily available, while (b) has a wide focus interpretation –typical in VSO sentences- with no contrast on the subject. Thus, it is the availability of contrast in (a) and (c) that makes the difference for the bare plural. And it is as relevant for preverbal subjects as for postverbal subjects, if they are external arguments. Though I started the discussion with the classical NNC, now we face a pervasive phenomenon that affects bare nouns in different positions and had already been correctly perceived in Suñer (1982).

2.2.3 Bare nouns and contrast
Up to this point, I have reviewed situations in which bare plurals need an extra factor to optimize their interpretation, this factor being contrast. Such situations share a significant feature. They involve positions that either impose special requirements on bare nominals – dislocated topics, marked / scrambled objects- or correspond to external arguments, thus hindering semantic incorporation: in a few words, they represent prominent syntactic positions that are in a sense “hostile” to bare nominals. The striking fact is that contrast is relevant for the acceptability of bare plurals precisely in these positions, and not elsewhere –
although contrast may independently occur as a component of interpretation due to contextual factors. In the cases under discussion, contrast seems to appear as a last-resort inferential solution for an interpretative mismatch. The pattern is always the same: a bare plural is placed in a grammatical context where, for some reason, it does not fit in adequately, and far from giving rise to ungrammaticality, it is supplemented by a contrastive component that makes the construction acceptable. If this view is correct, the insertion of contrast to adjust the interpretation of a nominal expression is akin to other well-known interpretive operations, such as coercion and accommodation, that can “rescue” otherwise deviant strings (cf. Leonetti and Escandell-Vidal 2011, de Swart 2011). Such operations cannot solve purely grammatical mismatches—in that case, the result is ungrammaticality—but they are able to provide acceptable readings for strings that include a semantic mismatch. A reasonable hypothesis about the role of contrast in the interpretation of NPs/DPs is that it should be treated as one particular case of a more general procedure for mismatch resolution that includes coercion and accommodation as well.

In what follows, I intend to gather more evidence supporting contrast as a mismatch resolution device for nominals. The big question is under what conditions, and to what extent, contrast can “rescue” deviant strings. For the moment, it may be useful to enlarge our data collection, and show that the idea is relevant not only for bare plurals, but for indefinite DPs too.

2.3 Indefinites as topics

2.3.1 Left dislocated indefinite DPs

As already pointed out in §2.1, the prediction is that all topical indefinites are in some sense contrastive. With dislocated indefinite DPs, I believe that the prediction is clearly borne out—other topical indefinites are discussed in §2.3.2. The following examples illustrate both the case of specific interpretations and the case of non-specific interpretations (often discarded or ignored in the literature, but quite common at least in Romance Left Dislocation):

(13) Italian (Rizzi 2005):
    Un libro, l’ho letto. Italian (Rizzi 2005)
    ‘A book, I’ve read.’

(14) Catalan (Valduví 2002):
    Un gelat, me’l menjaria amb molt de gust. Catalan (Valduví 2002)
    ‘An ice-cream, I would eat with pleasure.’

(15) Spanish (Leonetti 2011):
    Dos países de África, sí (los) habíamos visitado.
    ‘(At least) two countries in Africa, we had visited.’

Some comments on (13) can give a precise idea of the role of contrast. The example in (13) is ambiguous between a specific and a non-specific reading. In the specific reading, the speaker is asserting that (s)he read a particular book that (s)he is presumably able to identify. Such book is the entity the sentence is about. This is considered as the most natural reading—and sometimes the only one—in the literature on indefinites and topicality. It is commonly assumed that topicality constrains the range of interpretations that an indefinite DP may have: more precisely, it constrains the choice of a value for the discourse referent that indefinites introduce, and specificity arises as an effect of such constraint. But a non-specific reading is
possible as well. On the non-specific reading, the sentence can be paraphrased as ‘It is true that I read at least one book’: the speaker communicates no intention to refer to any specific book, and the hearer is not directed towards the insertion of a new file card in the discourse model. The indefinite DP does not introduce a discourse referent. It is intriguing how a non-specific indefinite can survive in a syntactic position that forces its interpretation as a link, i.e. as an adequate address for information update. The reason why indefinites often make poor topics is well known: they do not encode an instruction to identify an accessible referent. With indefinites, then, a new address has to be built, and the question is how this is accomplished in the case of non-specifics.

I assume that the conflict between the instruction associated to links and the semantics of indefinite DPs must be solved by pragmatically inferring an interpretation of the indefinite that is able to satisfy the requirement imposed by the topic. The two readings of un libro ‘a book’ in (13) represent two different ways to solve the conflict.

For the specific reading to be obtained, the hearer has to suppose that there is a certain property that the speaker is not making explicit and is restrictive enough to reduce the search for a referent to a particular individual. The set of books where the referent belongs may have been contextually specified: in this case, the information required to connect the topic to the previous discourse has already been provided – i.e. the set of books is available for the hearer too, and a partitive reading is naturally inferred; if the set has not been previously mentioned, the hearer tries to accommodate it, with the same result. The indefinite defines a partition on a given set, and a new address for information update is built, in accordance with the requirements of dislocation. The result is a partitive specific interpretation of the indefinite DP, where contrast distinguishes the intended referent from the rest of the books in the contextual set. Notice that contrast is necessarily involved in the specification of the intended reading.

As for the non-specific reading, its felicity conditions are stricter. This is not surprising, given that non-specific indefinites are less amenable to be used as topics. In the non-specific interpretation, (13) would hardly be used ‘out of the blue’ or in a neutral context: it would typically count as a refusal of a previous contextual assumption like ‘You haven’t read any book’ (thus, it would represent a sort of echoic use). Notice that the non-specific topic is again associated with a contrastive reading (‘one, but possibly no more’, or ‘one book, but possibly nothing else’), by which one item is contrasted with other possible items inside a series or a group. The same happens in the two remaining examples in (14) and (15), where the topical indefinites are preferably non-specific. In the Catalan sentence in (14), un gelat ‘an ice-cream’ introduces a set of contextual alternatives, roughly corresponding to the set of possibilities for the dessert: {ice-cream, yoghurt, crème caramel, apple pie…}. A contrastive reading is made salient, by which the speaker communicates his/her willingness to have an ice-cream, and excludes other cases like eating some yoghurt, eating some crème caramel, eating some apple pie… In the Spanish example in (15), the topic is dos países de África ‘two countries in Africa’, and the contrast set could be something like {no country in Africa, one country in Africa…}; as suggested by the gloss, the salient interpretation is ‘It is true that we had visited at least two countries in Africa’.

I believe the evidence is enough to reach two conclusions. The first one is that, as already observed in Umbach (2004), building on a previous proposal attributed to Regina Eckardt, it is the kind of contrast sets that are available in the context that determine the salience of specific and non-specific readings. Partitive specificity results from the contrast between one element –or more- in a set and the remaining members of such set. Non-specific readings, on the other hand, result from other contrast sets where the alternatives are triggered by focus on the determiner (as in {no book, one book, some books…}) or on the common noun (as in {ice-cream, yoghurt, crème caramel, apple pie…}). It all depends on the process of identifying embedded focus structures inside the topical constituents of the sentence that Erteschik-Shir
(1997) defined as ‘subordinate update’. This simply reinforces the assumption that contrast plays a major role in the interpretation of topical indefinites. The second conclusion is that dislocated topics in (13)-(15) clearly belong to the class of contrastive topics (cf. Kiss and Gyuris 2003, Krifka 2007, Giurgea and Remberger 2011, Torregrossa 2012). The typical features of contrastive topics are salient in the examples:

- the most characteristic one is the activation of contextual alternatives in the interpretation, already commented. The refutative value of utterances with non-specific indefinite topics –i.e. their restrictive condition of use in discourse- is an effect of choosing one alternative and discarding the rest of alternatives in a contextual domain.

- a second typical feature is an incompleteness or uncertainty implicature: the speaker indicates that (s)he is unable or unwilling to support more informative readings of the utterance. In (13), for instance, the speaker claims that at least one book has been read, but the exact quantity of books that were read is not specified. In (15), it is said that two African countries had been visited, but the speaker is not explicit about the total number of countries that (s)he actually visited. According to Giurgea and Remberger (2011), the indefinite topic corresponds to the weakest claim in a series: the speaker asserts that this claim is certainly true or highly probable, contrasting it with stronger claims for which the truth value is unknown.

It may be worth to add a brief clarification concerning the status of contrastive topics. Some authors believe that contrastive topics have to be formally distinguished from aboutness topics (cf. Frascarelli and Hinterhölzl 2008 for a syntactic approach). I assume a different perspective, in which aboutness topics and contrastive topics are rather viewed as two ways in which a dislocated topic can fit in the context. The syntax and semantics of dislocation remain the same, but the overall interpretation is inferentially elaborated along different paths depending on the nature of the topical constituent (definite or indefinite, specific or non-specific) and the context. The specification of contrastive relations is a part of the inferential process by which an optimal interpretation is accessed by the hearer. If this is true, the distinction aboutness / contrastive is entirely dependent on the context (and on intonation, an issue I cannot deal with here). Notice that if the two kinds of topic are not kept separated, then the widely held assumption that topicality is associated to specificity (in indefinites) cannot be maintained: it works for aboutness topics only.

The correlation between non-specific indefinite topics and contrastiveness is, thus, well established. My conclusion is that indefinite dislocated topics are always contrastive in some sense. Contrast is here forced as a sort of last resort mechanism that operates in the search for an optimal interpretation, in particular in the search for a reading that allows the hearer to take an indefinite expression as an acceptable topic (i.e. an acceptable address for information update). The address for information update that the topic must provide is, in the case of non-specific indefinites, harder to determine than with specific indefinites and with definite DPs: reconstructing it is an inferential operation that requires combining the indefinite DP with different pieces of background information (in (10), for instance, the address would be something like ‘As for reading books,...’). This is why the conditions of use are stricter for non-specifics. In any case, contrast is involved in making indefinites acceptable dislocated topics. To sum up, the survival of weak, non-specific readings in indefinite topics is dependent on contrast. It seems clear that the facts cannot be unrelated to what happens with bare plurals: the role of contrast is the same, i.e. providing interpretive material that allows a nominal expression to fit in a grammatical context that tends to block its occurrence.

---

5 Parallel facts in Hindi are discussed in Dayal (2003).
2.3.2 Spanish unos and preverbal subjects

A look at indefinite DPs as preverbal subjects in Spanish can give us additional evidence for the role of contrast in interpretation. As I already mentioned, preverbal subjects in null subject languages like Spanish tend to be interpreted as sentence topics, though this is not an obligatory requirement. An indefinite preverbal subject typically receives a specific reading if it is topical. In such case, the specific reading may involve contrast (when it is partitive) or not. More interesting is the case of indefinite subjects that for some reason are not amenable to a specific reading. The plural indefinite unos ‘some’ provides us with relevant data. As Laca and Tasmowski – De Ryck (1996:123) and Gutiérrez-Rexach (2010), among many others, point out, unos receives a characteristically weak interpretation and it is usually incompatible with an Individual-Level predicate. In fact, (16a), with an Individual-Level predicate like ser inteligente ‘to be intelligent’ is slightly deviant (at least, out of context). However, if by means of a continuation with otros ‘others’, as in (16b), a contrastive interpretation is forced, with two groups of boys compared to each other, the use of unos becomes perfectly acceptable:

(16) (a) #Unos niños son inteligentes.
    some boys be.3PL intelligent
    ‘Some boys are intelligent’

(b) Unos niños son inteligentes, otros no tanto.
    some boys be.3PL intelligent others no so much
    ‘Some boys are intelligent, others not so much’

In (16b) the hearer is compelled to assign a partitive reading to the subject unos niños. Contrast appears as a result of the partition established on the set of boys. Partitivity is in fact the variety of specific interpretation that is typically linked to topics. The relevant fact is that the acceptability of a weak indefinite as topic is again dependent on contrast.

Generic indefinites are a case of indefinite topical subject where the context blocks specific readings and partitivity is, thus, irrelevant. Nevertheless, generic indefinites too are associated to contrast, as Erteschik-Shir (1997: 121) suggests. The idea is that generic indefinites, like a beaver in (17), are licensed by contrast inside a restrictive set available from the context or inferable (for instance, {beaver, racoon, bear…}), and such restrictive set is the entity that counts as topic (it is obtained from background material). Thus, even if a non-specific indefinite DP is not an adequate address for information update per se, contrast licenses its use as a topic in generic contexts:

(17) A beaver builds dams.

In Erteschik-Shir’s view, given that an indefinite is not an optimal topic, contrast inside a contextually accessible set is inferred to force an interpretation that solves the clash between the semantics of the indefinite DP and the condition imposed by the predicate (or the generic context). This is in line with all the data previously presented. When used as topics, indefinite DPs and bare plurals undergo the same kind of interpretive process. Specifically, every time we place in a topic position some expression that can hardly count as an aboutness topic (not only non-specific indefinites and bare nominals, but also predicative adjectives or infinitives), a contrastive reading is obtained.

---

6 A different proposal can be found in Cohen (2008).
2.4 Interrogatives: D-linking

Interrogatives are to some extent similar to indefinites (in fact, they are the same thing in a number of languages). So it would be not surprising that contrast could be relevant to their interpretation under the same conditions that hold for indefinites. The parallelism I intend to focus on concerns interrogative subordinate clauses. It is well known that indirect interrogatives are weak islands (wh-islands): they ban the extraction of adjunct wh-phrases, as the contrasting status of the two stereotypical examples in (18a) and (18b) shows.

(18) (a) What do you wonder [how John could fix _]?  
(b) *How do you wonder [what John could fix _]?

In (18a) an argument –the direct object- is extracted, whereas in (18b) it is an adjunct that is extracted (how), with the result that acceptability clearly decreases. The literature on wh-islands soon pointed out that certain factors can affect the acceptability of wh-extractions, and the most widely discussed factor is Discourse-linking (D-linking). D-linked wh-phrases are usually defined as nominals that define a partition on a contextually given and presupposed set of entities (i.e. a set that is already in the mind of the speaker or has already been established in discourse). Extraction of argument wh-phrases from wh-islands, according to the literature, gives more acceptable results when the wh-phrases are D-linked (typically with which N phrases instead of bare interrogatives like who or what); with non-D-linked wh-phrases, extraction results in slightly deviant sentences. The Italian examples in (19a-b) illustrate the contrast (cf. Cinque 1990, Cruschina 2011: ch. 4); (19c) shows that another kind of extraction, namely focalization of arguments, is possible in Italian across weak islands.

(19) (a) ??A chi ti chiedi quanti soldi hai dato?  
   to whom REFLECT ask.2SG how.much money have.2SG given  
   ‘To whom are you wondering how much money you gave?’

(b) A quale dei tuoi figli ti chiedi quanti soldi hai dato?  
   to which of your sons REFLECT ask.2SG how.much money have.2SG given  
   ‘To which of your children are you wondering how much money you gave?’

(c) A GIANNI mi chiedo quanti soldi hai dato.  
   to John-FOC REFLECT ask.1SG how.much money have.2SG given

Notice that D-linking implies a contrast among contextual alternatives that are mutually manifest to the speaker and the hearer. In (19b), the possible values for the answer belong to a restricted set that is contextually specified. This gives us the clue to recognize a pattern that should be already familiar by now: if the wh-phrase is non-D-linked and there is no contrast inside a given set, the acceptability of extraction is degraded, but if the wh-phrase is D-linked and contrast is activated, then the survival of the phrase in a hostile environment like a weak island is guaranteed. One could reasonably claim that contrast acts here again as a mismatch resolution mechanism. As for focalization in (19c), it is again contrast the interpretive property that D-linked wh-phrases share with fronted foci. Then, it is expected that focalization is as acceptable as extraction of D-linked wh-phrases. A syntactic parallelism between the two categories is defended in Cruschina (2011: ch. 4).

Contrast is encoded in certain D-linked phrases (in case they show a partitive structure, as in (19b)), but can also be forced as an inferred feature of the interpretation in others, such as in bare wh-operators that receive a certain prosodic prominence in a syntactic context where their acceptability is expected to be unclear, as in (20). Here contrast is only one among different possible contextual enrichments of prosodic prominence:
¿A QUIÉN no sabes cuánto dinero le has dado?
‘To whom don’t you know how much money you gave?’

The overall picture shows now a clear parallelism with what we already noticed in the behavior of bare plurals and indefinites. D-linking happens to be just a particular case of a more general phenomenon involving contrast. At least two important questions emerge at this point. One is why D-linking is able to affect acceptability levels in weak islands. The other is what D-linking can tell us about the nature of wh-islands.

As for the first question, there is no clear answer in the generative tradition, for all I know. I would be prone to accept an account in terms of processing, along the lines of Hofmeister and Sag (2010): the idea is that D-linked wh-phrases contribute to narrow down the list of focus alternatives that have to be considered in answering a question, thus reducing computational effort and improving acceptability. This is in accordance with the effects already observed for contrast in bare plurals and indefinites: contrast offers some kind of reward that makes worth the processing of a phrase in a context where it does fit in adequately.

As for the second question, the fact that D-linking improves acceptability in weak islands suggests that weak islands do not behave as purely syntactic constraints. Were they purely syntactic restrictions, their effects should be stronger and more systematic. If semantic and discourse factors like D-linking are able to dissolve such effects, it seems natural to think of wh-islands as resulting from some semantic mismatch or from a combination of factors that conspire to increase processing difficulty. This, in principle, favors a processing view of islands constraints, against a strictly grammatical view. The solutions usually invoked in formal syntax, such as assuming that D-linked and non-D-linked wh-phrases are syntactically distinct from each other (because of the presence of [+topic] features or different internal configurations), are an attempt to recast the problem in syntactic terms, but in my view there is no serious evidence supporting them. I would rather assume that a contrastive reading is the main property distinguishing D-linked DPs from non-D-linked ones. Contrast seems to be unable to ‘rescue’ extraction from strong islands (cf. coordinated structures, complex DPs). This might be an indication that strong islands are of a different nature and call for a different explanation.

Thus, the whole issue looks like a particular case of the more general phenomenon I am describing here: contrast is an interpretive device that is able to provide acceptable readings of DPs that occur in syntactic contexts where they would not fit in adequately otherwise. Wh-phrases may survive in certain hostile environments –i.e. when they are extracted from weak islands– if contrast appears, and this should be related to the role of contrast in the interpretation of indefinite DPs and bare nouns.

3. Definite DPs

3.1. Definites in existential contexts

This section briefly presents some extensions of the previous ideas in the field of definiteness. The first issue I want to examine is the Definiteness Effect (DE), in particular in French existential impersonal constructions, where the subject must be obligatorily postverbal and indefinite. The data are taken from the recent discussion of the problem in Zimmermann (2012). The ban against definite DPs is quite clear in examples like (21).

(21) *Il arrive les/ ces/ ses filles. (cf. Il arrive des filles)
it arrive.3SG the these his/her girls
As it happens in many DE contexts in other languages, there are counterexamples to the constraint. Such counterexamples occur in allegedly ‘highly restrictive’ conditions that French grammars try to specify accurately. Some examples appear in (22):

(22) (a) Il reste / manque les filles.
     it remain.3SG / be-missing.3SG the girls
     ‘There remain the girls. / The girls are missing.’

     (b) Il a dormi ici Jean, Paul et François.
     it have.3SG slept here Jean Paul and François
     ‘Jean, Paul and François have slept here.’

     (c) Il est arrivé l’électricien.
     it be.3SG arrived the electrician
     ‘The electrician arrived.’

The problem is that grammars usually offer a list of contexts and factors, but not a true account of the availability of counterexamples. Zimmermann (2012) proposes an alternative unified view based on information structure, where features such as definiteness, anaphoricity or referentiality are no longer the central pieces in the explanation. According to Zimmermann, French existential impersonal constructions impose the condition that the postverbal DP must be focused. This requirement is trivially satisfied by indefinite DPs either as instances of informational focus or as instances of contrastive focus. Not surprisingly, definite DPs have to obey a stricter condition: they must be instances of narrow focus / contrastive focus; otherwise they are excluded. Their referents are implicitly or explicitly contrasted with other entities in a given contextual set. The examples in (22) satisfy this condition. And the classical examples of violations of the DE in English existential constructions (so-called “list” or “enumerative” readings of definite DPs) seem to obey it as well, as shown in (23) (from Abbott 1993).

(23) (a) – Is there anything to eat?
     – Well, there’s the leftover chicken from last night.

     (b) – I guess we’ve called everybody.
     – No, there’s still Mary and John.

The facts surely deserve a more detailed discussion, but I believe that Zimmermann’s approach is correct. For my purpose, it is enough to observe that a) definites obey a stronger constraint than indefinites, since existential constructions are “hostile” environments for definite nominals, and b) the constraint involves narrow focus. My proposal is that narrow focus becomes relevant just because it triggers contrast among contextual alternatives, and contrast allows definite DPs to “survive” in an existential context. Contrast, then, seems to have beneficial effects both for the interpretation of definite DPs and for the interpretation of indefinite DPs. I am aware that this proposal has to face at least one serious obstacle: one could rightly wonder why this mechanism is active in French (and maybe English), but not in other languages. Spanish existential sentences, for instance, do not admit counterexamples to the DE under the same conditions as French existential impersonal sentences. At this point I would simply recall that the DE is a complex phenomenon, and we cannot expect that DE contexts display exactly the same properties in every language. Additional factors could override the role of contrast under certain conditions. This remains an open question.
3.2. Strong pronouns in null-subject languages

In Romance linguistics it is common to describe strong pronouns (more precisely, overt subject pronouns), in opposition to weak / null pronouns, as special forms that are mostly used for emphasis and contrast. Strong pronouns are obligatory when they are in focus, and when the antecedent contrasts with another referent in the domain of discourse. A recent proposal about Romance overt subject pronouns is put forward in Mayol (2010): in an attempt to offer a precise analysis of the contrastive import of these pronouns, the author claims that non-focal contrastive overt subject pronouns are Contrastive Topic markers. As such, they trigger topic alternatives and convey an implicature of uncertainty by which the speaker ignores whether other alternatives are true or not. The Catalan example in (24), adapted from Mayol (2010), gives an intuitive grasp of the proposal.

(24) –¿Qué voldran per sopar?
   What want.FUT.3PL for dinner
   ‘What will you have for dinner?’
A: –Bé, doncs jo vull pollastre.
   well then I want.1SG chicken
B: –Doncs, jo vull sopa.
   then I want.1SG soup
A: –Well, I’ll have chicken.
B: -Well, I will have soup.’

In A’s response, the pronoun jo ‘I’ introduces topic alternatives ({‘B will have soup’, ‘C will have soup’...}), with the implicature that the truth of the alternatives is not known by the speaker. The answers in (24) count as partial answers to the question. Mayol rightly points out that answering with a null pronoun instead of jo would be perfectly grammatical, but it would not introduce the alternatives and would be understood as a complete answer. Leaving aside the details of Mayol’s hypothesis, as well as the fact that other non-pronominal expressions (for instance, proper names) could be contrastive topics too in the same context, I just assume that a contrastive component is typically associated with the interpretation of overt subject pronouns, and that the notion of contrastive topic is as useful in this case as it was for the analysis of indefinites in section 2.3. However, the question that attracts my interest is this: why are overt subject pronouns typically contrastive? An answer like ‘Because their semantics includes a [+contrast] feature’ would not be particularly illuminating. Moreover, it is surely false, since strong pronouns are not always contrastive (for instance, when they occur in contexts where they are the only possible choice, null pronouns being excluded for some grammatical reason). Contrast in the interpretation of overt pronouns must rather appear as a result of choosing to use the pronoun in a certain context. But how?

It seems to me quite natural to apply to strong pronouns the same ideas on contrast that I applied to bare plurals and indefinites. In order to do this, it is essential to take into account that strong pronouns in Null-Subject languages compete with null / weak pronouns in the expression of subject arguments. When two alternative forms are in competition (i.e. a weak pronoun vs a strong pronoun, or an unstressed pronoun vs. a stressed pronoun), they usually convey different interpretations. In this case, the weak form is optimally interpreted as a continuing topic, whereas the other one represents a shifted / contrastive topic, as we can observe in (25).

(25) La presidenta empujó a la secretaria y Ø / ella la insultó.
   the president push.PST.3SG to the secretary and she her insult.PST.3SG
   ‘The president pushed the secretary and she insulted her.’

If Ø is chosen, the interpretation is that the null pronoun refers to the discourse topic (the subject of the first sentence, la presidenta). If ella is chosen, the insertion of the overt pronoun means that something other than what is usually conveyed by a null pronoun has to
be expressed, and the interpretation is that there is a topic shift, and a new topic is introduced: the overt pronoun takes la secretaria as its antecedent, and the contrast established between the two possible antecedents becomes a component of its interpretation. Contrastive readings arise not because contrast is encoded in the linguistic meaning of pronouns, but because there must be some justification for choosing to use an overt pronoun instead of a less costly item like a null pronoun, and contrast is a natural pay-off for such choice, given that referent identification is at stake and contrast is usually subordinated to the determination of reference. Thus, contrast is a mechanism operating in the inferential phase of interpretation, and in particular in establishing the reference of a nominal expression. The parallelism with contrast in bare plurals and indefinites could seem to be obscured by the fact that in the case of pronouns there are no “hostile environments”. However, the situation is quite similar. Being in competition with a more economical item is equivalent to being placed in a “hostile environment”, in the sense that some reason for using the costlier form must be found in order to obtain a relevant reading.

To sum up, strong pronouns receive contrastive readings in positions where they alternate with weak / null pronouns, as a result of the competition between two group of items. Strong pronouns are marked forms: they are more complex than weak / null pronouns. The use of a marked element must be rewarded with a marked reading, and contrast – between referents, or between situations – appears in the use of strong pronouns as an extra component of interpretation (cf. de Hoop 2003, Kaiser 2010).

4. Speculations / Consequences

In the preceding sections I have been pursuing two goals. One is building a unified account of a series of facts that, to my knowledge, had never been gathered under a single perspective. Such series of facts includes ways to circumvent a classical constraint on bare plurals in Spanish, the behavior of indefinite DPs as topics (both in dislocation and as preverbal subjects), certain counterexamples to the Definiteness Effect in French impersonal existential, and the interpretation of overt subject pronouns in Null-Subject languages. If my proposals are on the right track, this should at least provide us with some insights on the interpretation of NPs/DPs.

The second goal is more theory-oriented, and more ambitious too. My aim is to investigate how mismatch resolution works in interpretive tasks (see Leonetti and Escandell-Vidal 2011 for some ideas). I have been trying to describe a recurring pattern: contrast is pragmatically inferred as a way of solving some kind of interpretive mismatch, either a mismatch involving a DP and its position, or the competition between two alternative expressions in the same position. This interpretive pattern presupposes that a number of linguistic mismatches are in fact solved by speakers and hearers in their interaction, that such process is to a certain extent systematic, and that interpretation is the result of combining decoding and inference. We already have strong evidence that these assumptions are essentially correct. But a look at the pattern of occurrence of contrast in nominals still raises several questions. Here I want to focus on two questions that can lead us to explore crucial aspects of the problem:

1. Why is the resolution mechanism linked to contrast?
2. How is the resolution process constrained?

At this point, the answers have to be rather speculative.

1. Why contrast? Contrast provides the interpreter with a reward for adjusting the interpretation of a constituent in a context where “something more” is required to access an adequate reading. I would like to cast the idea in Relevance-theoretic terms (cf. Sperber and Wilson 1986), exploiting the assumption that processing effort must be counterbalanced by contextual effects in utterance interpretation. The data I examined represent examples of inferential processes in building the basic explication of an utterance – i.e. the proposition that
the utterance explicitly communicates; reference determination is always a contribution to the specification of propositional explicatures. In Leonetti forthcoming I suggest for bare plurals that a marked use of a bare nominal in a context where semantic incorporation is not straightforward forces the hearer to wonder what is “behind” such choice, and contrastivity is a natural way out because it introduces a set of contextual alternatives that have to be considered in interpreting and allow the hearer to figure out the consequences of enlarging the context that way. The intuition can perfectly work for other kinds of nominals. Considering a range of contextual alternatives is a way of ensuring that some contextual effects will be obtained, as contrast establishes a connection with previous discourse and a basis for inferentially expanding the context. Once the available set of alternatives satisfies the need for rewarding benefits created by a costly operation, optimal relevance is achieved. In very simple terms, contrast provides a direct way of enriching interpretation. This could explain why it is so pervasive. It remains to be clarified whether contrast is relevant too in the interpretation of non-nominal categories.

2. What are the limits of the resolution process? It seems clear that contrast cannot repair whatever kind of mismatches. Intuitively, only “soft” restrictions affecting NPs/DPs can be neutralized by resorting to a contrastive reading. The Definiteness Effect and the Naked Noun Constraint are among them. This means that they can no longer be treated as syntactic constraints. If they were, they would behave as “strong” constraints—the ones that cannot be violated and give rise to ungrammaticalities. Thus, this provides evidence for considering them as constraints based on semantic and discourse principles, rather than on syntactic rules. This is one of the positive consequences of investigating mismatch resolution: it allows us to determine which mismatches can be repaired and, as a consequence, what is the nature of the corresponding constraints. Obviously, there is much to be gained for a better understanding of the syntax—semantics interface and the grammar—pragmatics distinction.

An in-depth study should give us more precise indications about the conditions under which mismatch resolution takes place. The task exceeds the limits of this paper. A model for such study could be the ample literature devoted to accommodation phenomena and the discourse conditions for accommodation of presuppositions. Accommodation is, in fact, another case of mismatch resolution. Ultimately, a detailed investigation should lead to a better understanding of the notion contrast.

5. References

Espinal, Mª Teresa 2009. Bare nominals in Catalan and Spanish. Their structure and meaning. Lingua 120, 984-1009.
On contrastive readings in the interpretation of NPs/DPs


Giurgea, Ion & Eva Maria Remberger 2011. Indefinites as contrastive topics. Paper presented at the 44th Annual Meeting of the Societas Linguistica Europaea, Universidad de La Rioja (Logroño, Spain), September 2011.


1. Introduction

Since Kuteva (1998; 2001) the proximative (‘be about to’) as an imminent gram and the avertive (‘was on the verge of taking place, but did not take place’) expressing imminence, past time reference and counterfactuality are considered to be two distinct but closely related grams. Given the functional overlap of proximative and avertive when the proximative is used in past contexts, the separation of the two grams is a known problem (cf. Kuteva 2001). Furthermore, Bellosta von Colbe (2001a; b) gives a first overview about the periphrases in Romance languages by using the related terms de conatu and frustrierte Imminenz (‘frustrative’). In spite of their common functional relationship I will discuss that a closer look into the pragmatics-semantics interface allows to distinguish between de conatu, frustrative, proximative and avertive using pragmatic tests like the cancellability test (cf. Schwellenbach 2009).

Regarding their diachronic relationship, there has been no consensus on the direction of their grammaticalization path: Despite the grammaticalization path observed by Kuteva (1998; 2001) in Bulgarian which gives rise to an avertive-to-proximative development via generalization (i.e. loss of specificity with respect to pastness and counterfactuality), Bellosta von Colbe (2001a; b) proposes the alternative route proximative-to-ave rtive (i.e. addition of specificities by the conventionalization of inferences) for Romance languages.

The aim of the present article is threefold: Firstly, I will determine the difference between the two categories in Romance languages and their related terms ‘frustrative’ and imperfectum de conatu. Secondly, I will show that the proximative periphrases in Romance languages are characterized by an imperfective morphology which allows the cancellability of counterfactuality by reason of semantic underspecification. In contrast, the avertive periphrases are characterized by a perfective morphology while their expression of counterfactuality should not be cancellable. Finally, the grammaticalization path proximative-to-avertive will be discussed by empirical evidence from Ibero-Romance languages.

2. Distinguishing avertive, proximative, frustrative and imperfectum de conatu

According to Kuteva (1998; 2001; 2009) the avertive is a ‘semantically elaborate grammatical category’ situated at the intersection of aspectuality, temporality and modality denoting imminence, past time reference and counterfactuality. In contrast, the proximative denotes a ‘semantically straightforward grammatical category’ being a purely aspec tal gram involving imminence. Given the underspecification with regard to the temporal and the modal domain, the proximative can be used in present, past and future tense “with (crucially) no implication that the situation did not actually occur” (Kuteva 2009:19). I formalize the difference between avertive and proximative as illustrated in table 1:

---

* This paper is based on my talk given at the “VI NEREUS International Workshop on theoretical implications of the syntax/ semantics interface in Romance”. I am grateful to the organizers for inviting me to participate and to contribute to this volume. Special thanks go also to Liane Teger, Thomas Siwonia and Laura Zester.
Regarding the Romance languages, the avertive meaning (i.e. ‘was on the verge of taking place but did not take place’) has long been treated under the terms of ‘frustrative’ and *imperfectum de conatu*. As shown in a previous work (cf. Schwellenbach 2009) the well-known pragmatic tests can be used in order to distinguish between the feature counterfactual as part of the conventional semantic meaning as in the case of the avertive or if it is merely part of a context-dependent pragmatic meaning as in the case of the proximative. While proximative and *imperfectum de conatu* merely implicate the counterfactual as can be verified by cancellability, the avertive is specified to express the three features [+imminence], [+past time reference] and [+counterfactual].

In the case of the *imperfectum de conatu* the functional overlap with the avertive is due to a conversational implicature of the imperfect tense (cf. Schwellenbach 2009). As the imperfect morpheme is underspecified with respect to the avertive features (i.e. imminence, counterfactuality and past time reference), the avertive meaning has to be induced by the context, as in the typical ‘schema of incidence’ (Pollak 1960) shown in (1). The context dependency of the implicature is illustrated in (2).

(1) **Spanish:**

Salía cuando llegó su madre. (+> No salió.)

‘He was about to leave when his mother arrived.’ (+> He didn’t leave.)

(2) **Spanish:**

Salía cuando llegó su madre, pero no la vio. (+> Salió.)

‘He was leaving when his mother arrived, but he didn’t see her.’ (+> He left.)

In contrast, the term ‘frustrative’ (Cartagena & Gauger 1989; Wandruszka 1969) is used as an unspecific hypernym with reference to any construction which may express the avertive meaning irrespective of being implied or implicated and therefore regardless of dealing with a proximative, an avertive or an *imperfectum de conatu*. In most cases the term is used in order to refer to a proximative with the implicature of ¬p, as in (3) and (4).

(3) **Spanish:**

Estaba para salir cuando sonó el teléfono. (+> No salió.)

‘He was about to leave, when the phone rang.’ (+> He didn’t leave.)

(4) **Catalan:**

Estava a punt de marxar quan va sonar el telèfon. (+> No salió.)

‘He was about to leave, when the phone rang.’

<table>
<thead>
<tr>
<th>Aspectuality</th>
<th>Avertive</th>
<th>Proximative</th>
</tr>
</thead>
<tbody>
<tr>
<td>λp(x)[IMM(R,^p(x))]</td>
<td>λp(x)[IMM(R,^p(x))]</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Temporality</th>
<th>Avertive</th>
<th>Proximative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pp(x)</td>
<td>Pp(x) v p(x) v Fp(x)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Modality</th>
<th>Avertive</th>
<th>Proximative</th>
</tr>
</thead>
<tbody>
<tr>
<td>¬p</td>
<td>◊p v ◊¬p (¬□¬p)</td>
<td></td>
</tr>
</tbody>
</table>
As in the case of the *imperfectum de conatu*, the counterfactuality of the proximative is easily cancellable:

(5) Spanish:
    Estaba para ganar la carrera y, de hecho, la ganó.
    ‘He was about to win the race, and, in fact, he won it.’

In the case of the avertive, the polar component ¬p has become part of the conventional meaning, as in (6), and is therefore not cancellable, as shown in (7).

(6) French:
    J’ai failli tomber.
    ‘I almost fell.’

(7) French:
    #J’ai failli tomber et je suis tombé.
    ‘I almost fell and I fell.’

Therefore, the ‘frustrative’ is an underspecified term referring to the functional overlap of proximative, avertive and *imperfectum de conatu* which differ in their semantic-pragmatic representation, as illustrated in table (2).

Table (2): The functional overlap of frustrative, proximative, avertive and *imperfectum de conatu*

<table>
<thead>
<tr>
<th></th>
<th>Examples</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>F R U S T R A T I V E</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>imperfectum de conatu</em></td>
<td>Salía cuando llegó su madre.</td>
<td>Pp(x) ↠ λPλx[Imm(R,¬p(x))]</td>
</tr>
<tr>
<td></td>
<td>Estaba para ganar la carrera.</td>
<td>λPλx[Imm(R,¬p(x))]</td>
</tr>
<tr>
<td>Proximative</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avertive</td>
<td>J’ai failli tomber.</td>
<td>λPλx[Imm(R,¬p(x))]</td>
</tr>
</tbody>
</table>

Regarding the lexical constructions, there is no consensus whether the polar component is part of the inherent meaning or if it is supported contextually as it will be discussed in the next chapter.
3. Avertive and proximative in Romance languages

3.1 Lexical constructions

To analyze the meaning of the lexical constructions (almost VP, fr. presque VP, sp. casi VP, por poco VP etc.), the ‘conjunctive analysis’ of proximal and polar component (Ducrot 1972; Sevi 1998; Horn 2002, Schwenter 2002) can be applied. According to Sevi (1998), Horn (2002) and Schwenter (2002) the meaning of ‘Alonso almost won the race’ (8) can be analyzed by the conjunction of the proximal component asserting that ‘He was about to win the race’ and the polar component expressing that ‘he didn’t win it’.

(8) Spanish:
Alonso casi gana la carrera.
Alonso almost win.PRS.3SG the race.
He was about to win the race. (Proximal Component)
He didn’t win it. (Polar Component)
‘Alonso almost won the race’

In order to answer the question whether we are dealing with a proximative or an avertive meaning it is necessary to reconstruct briefly the long-standing debate in pragmatics. It is discussed whether the polar component is a use-conditional or a truth-conditional part of the meaning of ‘almost VP’ (cf. Ducrot 1972; Sadock 1981; Sevi 1998; Horn 1991; Horn 2002; Schwenter 2002). Thus, answering the question whether the polar component is either a generalized conversational implicature (GCI) as argued by Sadock (1978; 1981) and Ziegeler (2000; 2006), a semantic presupposition as suggested by Ducrot (1972) and Martin (2005), an entailment as pointed out by Horn (2002) and Schwenter (2002), or a conventional implicature (cf. Jayez & Tovena 2008) is crucial in order to verify if we are dealing with a proximative or an avertive construction. The main problem of the debate is that ¬p is not cancellable, as shown in (9), but it can be felicitously reinforced without redundancy as in (10):

(9) Spanish:
#Alonso casi gana la carrera y la ganó.
Alonso almost win.PRS.3SG the race, and it win.PST.3SG
#‘Alonso almost won the race, and he won it.’

(10) Spanish:
Alonso casi gana la carrera, pero no la ganó.
Alonso almost win.PRS.3SG the race, but it win.PST.3SG
‘Alonso almost won the race, but he didn’t.’

Nevertheless, as Horn (1991) has shown reinforceability on its own is not convincing enough to stand for an implicature, because logical inferences may also be felicitously reinforced without redundancy, as in the case of the following presuppositions:

(11) “It is unfortunate that our results vitiate one of the few apparently reliable tests for distinguishing conversational implicata from logical inferences and semantic entailments, but vitiate it they do.”

(Horn 1991:336)
(12) I don’t know why I love you, but I do.  
(Horn 1991:322)

Furthermore, Horn (1991) discusses that the information in the second conjunct (i.e. in the adversative conjunct) is not redundant but based on a rhetorical contrast between ‘almost VP’ and ‘not p’ participating in oppositely-oriented argumentative scales within the theory of argumentation established by Anscombe & Ducrot (1976). Therefore, the polar component has been described as entailed but “somehow backgrounded” (Sevi 1998) or “assertorically inert” (Horn 2002). Moreover, Jayez & Tovena (2008) argue for a conventional implicature\(^1\) which can be considered part of the class of entailments. As illustrated in the following figure taken from Potts (2005:23), meanings can be divided into context dependent meanings and entailments.

**Figure (1): A Meaning Graph (Potts 2005:23)**

- **meanings**
  - **context dependent**
  - **entailments**
    - **at-issue entailments**
      - not invariably speaker oriented, vary under holes, plugs
    - **conversational implicatures**
      - Not conventional, not speaker oriented, not backgrounded
    - **conversationally-triggered presuppositions**
      - not speaker oriented, backgrounded
    - **conventional presuppositions**
      - not speaker oriented, backgrounded
    - **CIs**

In the case of the proximative, the polar component is characterized by context dependency, whereas the polar component of the avertive is part of the class of entailments. The semantic-pragmatic representation of the polar component (i.e. entailment or context dependent meaning) will be addressed in the next chapter in order to discuss the functional overlap between proximative and avertive periphrases in Romance languages.

### 3.2 Verbal periphrases

In Romance languages proximative and avertive are expressed by verbal periphrases featuring the Latin auxiliaries *stare* ‘to be’, *habere* ‘to have’, *fallere* (*<* fallere*) ‘to fail’, *ire/ andare* ‘to go’ (cf. Kuteva 1998; Bellota von Colbe 2001a; 2001b; Heine & Kuteva 2002). I suggest that the proximative periphrases in Romance languages exhibit an imperfective morphology characterized by underspecification with regard to the realization of the events, whereas the avertive periphrases in Romance languages exhibit a perfective morphology being specified with regard to the realization or non-realization of events. Even though avertive and proximative share the common aspectual feature [imminence], solely the avertive is specified to express the

---

\(^1\) The term of the ‘conventional implicature’ is sensitive given that the contradiction in terms is to be found in the fact that an implicature which becomes conventional no longer is an implicature but part of the conventional meaning. For further discussion on the criticism of the conventional implicature see Bach (1999) and Potts (2005).
counterfactuality (i.e. the non-realization of the event). The polar component of the proximative is merely implicated and therefore easily cancellable due to the underspecification of the proximative.

Furthermore, it will be distinguished between highly grammaticalized and less grammaticalized avertive periphrases in Romance. The auxiliary of the former is characterized by the development of restriction to perfective morphology, as in the case of the periphrasis in French which is today restricted to the compound past tenses.

(13) French:

J’ai failli tomber.
have.PRS.1SG fail.PTCP fall.INF
‘I almost fell.’

In contrast, the avertive in the Ibero-Romance languages is expressed by the less grammaticalized periphrases which share the grammaticalization source (i.e. the auxiliary) with the proximative or with another gram. I suggest that the Ibero-Romance\(^2\) periphrases \textit{stare} + PP/PREP + infinitive in present and imperfect tense express the proximative and in perfective tenses express the avertive (cf. Schwellenbach 2009).

\textbf{3.2.1 The case of the \textit{stare}-periphrases}

\textbf{3.2.1.1 Proximative}

I will focus on the three Spanish periphrases \textit{estar por} + infinitive, \textit{estar para} + infinitive and \textit{estar a punto de} + infinitive, but the analysis is adaptive to the equivalent periphrases in Catalan (\textit{estar a punt de} + infinitive, \textit{estar per} + infinitive), Galician (\textit{estar para} + infinitive) and Portuguese (\textit{estar para} + infinitive, \textit{estar por} + infinitive and \textit{estar a ponte de} + infinitive).\(^3\)

In present and in imperfect tense they express the proximative, i.e. the situation described by the main verb is about to occur (i.e. the realization or non-realization of the event is not specified).

(14) Spanish:

Está para llover.
be.PRS.3SG PREP rain.INF
‘It is about to rain.’

(15) Spanish:

La carrera está por comenzar.
the race be.PRS.3SG PREP begin.INF
‘The race is about to begin.’

\(^2\) The Italian periphrasis \textit{stare per} + infinitive seems to be more conservative developing only the proximative and conserving the periphrasis \textit{essere per} + infinitive. For this reason the case of Italian has to be presented in a separate work.

\(^3\) The problems concerning the semantic and syntactic status of the imminential periphrases as expressions of lexical or syntactic aspectual categories are shown in Laca (2002).
(16) Spanish:
El tren está a punto de salir.
the train be.PRS.3SG on the point of depart.INF
‘The train is about to depart.’

In the case of animate subjects, the proximative meaning (i.e. be about to) is inferred from the modal meaning of intention (i.e. be in favor of), as in (17) and (18).

(17) Spanish:
Estoy por irme.
be.PRS.1SG PREP go-away.INF
Modal reading: ‘I am in favor of going away.’
Aspectual reading: ‘I am about to go away.’

(18) Catalan:
Estic per marxar.
be.PRS.1SG PREP go-away.INF
Modal reading: ‘I am in favor of going away.’
Aspectual reading: ‘I am about to go away.’

Within the scope of negation and combined with animacy, estar por + infinitive and estar para + infinitive are restricted to express this modal reading.

(19) Spanish:
No estoy para salir de marcha.
not be.PRS.3SG PREP go-out.INF
Modal reading: ‘I am not in favor of going out.’

(20) Spanish:
No estoy por ir a la fiesta.
not be.PRS.3SG PREP go.INF to the party
Modal reading: ‘I am not in favor of going to the party.’

(21) Spanish:
No está a punto de salir.
not be.PRS.3SG on the point of leave.INF
Modal reading: ‘He is not in favor of leaving.’
Aspectual reading: ‘He is not about to leave.’

While the proximative with estar por + infinitive and estar para + infinitive seems to be restricted to affirmation, it is possible to enforce the negation of imminence in the case of the periphrasis estar a punto de + infinitive.

(22) Spanish:
La guerra no estaba a punto de terminar.
the war not be.IMPF.3SG on the point of finish.INF
Aspectual reading: ‘The war was not on the point of finishing’
This is due to the fact that the imminence needs not to be inferred from a modal meaning but is lexically expressed by the prepositional phrase *a punto de*.

On the contrary, if the event within the scope of imminence is negated, the periphrasis may only take the modal reading ‘I am in favor of not going to the party.’

(23) Spanish:

\[
\begin{align*}
\text{Estoy} & \text{ por no ir a la fiesta.} \\
\text{be.PRS.1SG} & \text{ PREP not go.INF to the party.}
\end{align*}
\]

Modal reading: ‘I am in favor of not going to the party.’

Given the underspecification of the proximative, the functional overlap with the avertive is due to an implicature of counterfactuality. This can be shown by cancellability, as in (24), and reinforceability, as in (25).

(24) Spanish:

\[
\begin{align*}
\text{Juan estaba a punto de meter un gol, y, de hecho, lo metió.} \\
\text{John be.IMPF.3SG on the point of score.INF a goal and in fact it score.PST.3SG}
\end{align*}
\]

‘John was about to score a goal, and, in fact, he scored.’

(25) Spanish:

\[
\begin{align*}
\text{Juan estaba a punto de meter un gol, pero no lo metió.} \\
\text{John be.IMPF.3SG on the point of score.INF a goal, but not it score.PST.3SG}
\end{align*}
\]

‘John was about to score a goal, but he didn’t score.’

As I suggested in Schwellenbach (2009) this implicature of counterfactuality may be described as a ‘counterfactual implicature’ in terms of Ziegeler (2000; 2006). According to Ziegeler (2000) modality can be ordered on a pragmatic Horn-scale (S, W) in which the modalized utterance is understood as a weaker (W) version (i.e. less informative and therefore weaker in its truth value) of the non-modalized utterance (S). Asserting an utterance which is weaker in its truth value (W), as in the case of the underspecified periphrases in imperfect tense, may Q-implicate the negation of the stronger predication (¬S). In order to point out the interaction of Q-based and R-based implicatures in the grammaticalization of *almost* Ziegeler (2000) establishes the following model:

Table (3): Pragmatic developments in the grammaticalization of proximatives (Ziegeler 2000:1762)

| Stage I: | proximity to X (prediction of X via R-based implicatures) |
| Stage II: | proximity to X, but Y, therefore not-X (prediction of X is explicitly cancelled, inducing Q-based implicatures) |
| Stage III: | proximity to X (but Y, therefore not-X) (cancellation is omissible, and implicit) |
| Stage IV: | proximity to X ('not X' generalised as part of the meaning) |

4 As a reminder: Horn’s (1984) hearer-based Q-Principle (‘Make your contribution sufficient. Say as much as you can.’) is based on Grice’s (1975) maxims of Quantity, whereas his speaker-based R-Principle (‘Make your contribution necessary. Say no more than you must’) is based on the Gricean maxims of Quantity, Relation and Manner. “In the case of Q-based implicite, the assertion of ‘W’ Q-implicates ¬S. […] In the case of R-based implicata, the assertion of ‘W’ R-implicates S.” (Horn 1984: 20-21)
The stages I-III will be adopted in order to analyze the implicature of the STARE-periphrases in Romance languages. The imminential periphrases may R-implicate the prediction of the event, which corresponds to the first stage in the model of Ziegeler (2000). At stage II the R-based implicature of prediction is cancelled by an adversative clause inducing a Q-based implicature which derives the inference of counterfactuality, as shown in (26).

La carretera estaba por acabar, pero nadie trabajaba.
the race be.IMPF.3SG PREP finish.INF but nobody work.IMPF.3SG
Se habían esfumado
REFL.3PL have.IMPF.3PL disappear.PTCP
‘The highway was about to finish, but nobody was working. They disappeared.’

La carretera estaba por acabar (‘the highway was about to finish’) R-implicates the prediction of finishing. The information that nadie trabajaba (‘nobody was working’) in the adversative conjunct may cancel the R-based implicature triggering simultaneously the Q-based inference that ‘the highway didn’t finish’. Similarly, in the schema of incidence shown in (3) and (4) the R-based implicature of the proximative is cancelled by the sudden interruption due to the occurrence of another event which induces the Q-based implicature. In contrast, stage III appears when the explicit cancellation of the R-based implicature may be omissible, as in (27).

(27) Spanish (El País, 05/12/2007, “La solidaridad de los marinos de Santa Pola galardonada”, Spain):
tres niños pequeños, que estaban a punto de ahogarse
three children small that be.IMPF.3PL on the point of drown.INF
‘Three small children who were on the point of drowning.’

To sum up, the STARE-proximative appears with an imperfective morphology and may functionally overlap with the avertive. This functional overlap is induced by the ‘counterfactual implicature’ (Ziegeler 2000) of the STARE-proximative. In contrast, the STARE-periphrases with perfective morphology express the avertive, as will be shown in the next chapter:

3.2.1.2 Avertive

Accepting that the STARE-proximative in Romance is restricted to imperfective morphology, I will show that the STARE-periphrases used in tenses of perfective aspect express the avertive due to the fact that the implicature of counterfactuality has become conventionalized.

Firstly, the polar component is not cancellable in the case of the Spanish estar a punto de + infinitive (cf. Carrasco Gutiérrez 2006) and is at least difficult to cancel in the case of estar por + infinitive and estar para + infinitive, as shown in the following examples (28)-(33).5

5 As Portner (2010:226) has shown for English modals with past tense morphology in combination with the perfect, the interpretation of counterfactuality as a conversational implicature (cf. Condoravdi 2002) is problematic due to the uncannellability: ?At that point, he might have won the game, and in fact he did.
(28) Spanish:
Estuvo para ganar la carrera (#y, de hecho, la ganó).
be.PST.3SG PREP win the race.INF (#and in fact it win.PST.3SG)
Proximal Component: He was about to win the race.
Polar Component: He didn’t win the race.
‘He almost won the race.’ (#and, in fact, he won it.)

(29) Spanish:
Estuve por caerme (#y, de hecho, me caí).
be.PST.1SG PREP fall.INF-REFL.1SG (#and in fact me fall.PST.1SG)
Proximal Component: I was about to fall.
Polar Component: I didn’t fall.
‘I almost fell.’ (#and, in fact, I fell.)

(30) Portuguese:
Esteve para morrer (#e, de fato, morreu).
be.PST.3SG PREP die.INF (#and in fact die.PST.3SG)
Proximal Component: He was about to die.
Polar Component: He didn’t die.
‘He almost died.’ (#and, in fact, he died.)

(31) Galician:
Estiven para marchar (#e, de feito, marchei).
be.PST.1SG PREP go-away.INF (#and in fact go-away.PST.1SG)
Proximal Component: I was about to go away.
Polar Component: I didn’t go away.
‘I almost went away.’ (#and, in fact, I went away.)

(32) Catalan:
Vaig estar a punt de comprar-lo (#i, de fet, el vaig comprar.).
go.PRS.1SG be.INF on the point of buy.INF-it (#and in fact it go.PRS.1SG buy.INF)
Proximal Component: I was about to buy it.
Polar Component: I didn’t buy it.
‘I almost bought it.’ (#and, in fact, I bought it.)

(33) Catalan:
Vaig estar per dir-li la veritat (#i, de fet, li vaig dir.).
go.PRS.1SG be.INF PREP say.INF-him the truth. (#and in fact him go.PRS.1SG say.INF)
Proximal Component: I was about to tell him the truth.
Polar Component: I didn’t tell him the truth.
‘I almost told him the truth.’ (#and, in fact, I told him the truth.)

Indeed, as in the case of the lexical avertive constructions the reinforceability of the polar component without inducing redundancy is possible.
(34) Spanish:
Estuvo para ganar la carrera, pero no la ganó.
Proximal Component: He was about to win the race.
Polar Component: He didn’t win the race.
‘He almost won the race, but he didn’t.’

"Ya me dirás con quién" estuve a punto de decirle yo.
Proximal Component: He was almost about to leave.
Polar Component: He didn’t leave.
‘When I was already almost about to leave, I heard the voice of the inspector who came running.’

Nevertheless, according to Horn (1991) the reinforceability is not a distinctive criterion between pragmatic and logical inferences, as illustrated above. Furthermore, the reinforceability of the asserted proximal component is equally possible without inducing redundancy. In (36) the proximative periphrasis is within the scope of the lexical avertive construction casi VP without being redundant. Furthermore, the reinforceability of the asserted proximal component in (37) is twofold.

Cuando ya casi estaba a punto de partir,
Proximal Component: He was about to leave.
Polar Component: He didn’t leave.
‘When I was already almost about to leave, I heard the voice of the inspector who came running.’

(37) Spanish (El Correo, 15/08/2008, “Sesenta de cintruénigo; [...]”, España, LEXIS NEXIS):
Ilumbe casi casi se llenó. Faltaría más.
Proximal Component: He almost filled. It would be missing more.
‘The Illumbe Bullring almost filled. It would be missing more.’

Secondly, contrary to the proximative periphrases with imperfective morphology, the avertive periphrases with perfective morphology are semantically equivalent to the lexical avertive construction casi VP (cf. Gómez Torrego 1999:3376). Therefore, the periphrasis given in example (35) can be replaced by the lexical avertive construction, as in (38):

(38) Spanish:
"Ya me dirás con quién" casi le dije yo.
Proximal Component: He almost said to him.
Polar Component: He didn’t say.
‘You’ll surely tell me with whom”, I almost said to him.’
Thirdly, contrary to the forced negation of the proximative periphrases, the negative particle within the scope of the avertive reverses the polarity in “was on the verge of not taking place, but did take place.”

(39) Spanish (Miguel de Cervantes Saavedra, 1582, Don Quijote de la Mancha, CORDE):

Don Quijote estuvo por no creer a Don Antonio.

Proximal Component: Don Quijote was about to not believe him.

Polar Component: Don Quijote believed him.

‘Don Quijote almost didn’t believe Don Antonio.’

Context: Admirado quedó Don Quijote de la virtud y propiedad de la cabeza, y estuvo por no creer a Don Antonio; pero, por ver cuán poco tiempo había para hacer la experiencia, no quiso decirle otra cosa sino que le agradecía el haberle descubierto tan gran secreto.

(CORDE: Cervantes Saavedra, Miguel de, 1582, Don Quijote de la Mancha, Parte II, Capítulo LXII, “Que trata de la aventura de la cabeza encantada, con otras niñerías que no pueden dejar de contarse”)

Translation: Don Quixote marveled at the power and properties of the head, and almost didn’t believe Don Antonio. But when he saw how little time he had to wait to find out, he decided not to say anything else except to thank him for telling him such a great secret.

(Cervantes Project, Texas A&M University, Don Quixote. Part I (1605) and Part II (1615). English translation by John Ormsby, London, 1885)

(40) Spanish (Gustavo Sainz, 1965, Gazapo, Mexico CDE):

¡Estuve a punto de no regresar!

Proximal Component: He was about to not return.

Polar Component: He returned.

‘I almost didn’t return.’

(41) European Portuguese (CETEMPúblico, LINGUATECA):

E o concerto que esteve para não se realizar --

Proximal Component: The concert was about to not take place.

Polar Component: The concert took place.

‘The concert that almost didn’t take place.’

(42) Galician (La Voz de Galicia, 1995, VG1995-04-20/2, CORGA):

Pouco despois de iniciarse as conversas, que estiveron a punto de non celebrarse, o ministro de Exteriores, Andrei Kózirev, dixo que […]

Proximal Component: The conversations, which almost didn’t take place, the foreign minister, Andrei Kózirev, said that […]
The avertive in French shows the same scope interaction with the negative particle. As in the case of the STARE-avertive, the avertive has scope over the negative particle so that the proposition becomes true.

(43) French (Raymond Queneau, *Journaux 1914-1965*, 1996, p. 741, FRANTEXT) :

Il a failli ne pas voir Napoléon passer.

He have.PR.SG fail.PTCP not see.INF Napoleon pass.INF

‘He almost didn’t see Napoleon.’

In contrast, the two polar components of lexical and periphrastic avertive may be combined without reversing the polarity of the proposition, as shown in (44) and (45). They do not cancel each other due to the “assertorically inert” entailment (Horn 2002) of the lexical constructions suggested above (cf. chapter 3.1).


Casi estuvo a punto de enfadar al público

almost be.PST.SG on the point of anger to-the public

Proximal Component: He was about to anger the public.
Polar Component: He didn’t anger the public.

‘He almost angered the public.’

(45) French (Ouest-France, mardi 7 juin 2011, La folle semaine d'Erwan Benech, France, LEXIS NEXIS) :

Caroline Cayeux a presque failli tomber de sa chaise

Caroline Cayeux have.PR.SG almost fail.PTCP fall.INF of her chair

Proximal Component: Caroline Cayeux was about to fall off her chair.
Polar Component: Caroline Cayeux didn’t fall off her chair.

‘Caroline Cayeux almost fell off her chair.’

Therefore, the STARE-avertive is characterized by perfective morphology. The polar component is not cancellable and is compatible with the negation of the infinitive reversing the polarity of the proposition in ‘was on the verge of not taking place but did take place’. In the next chapter it will be discussed how and when the polar component of the STARE-avertive has become conventionalized.

### 3.2.1.3 The diachronic relationship between proximative and avertive

There has been no consensus on the diachronic relationship between proximative and avertive in Romance languages. Despite the grammaticalization path observed by Kuteva (1998; 2001) in Bulgarian, which gives rise to an avertive-to-proximative development via generalization (i.e. loss of specificity with respect to past time reference and counterfactuality), Bellosta von Colbe (2001a; b) proposed the opposite route proximative-to-avertive for Romance languages which is comparable to the route proximative-to-counterfactuality observed by König (1993) and Heine (1994) in African languages. Nevertheless, given the different grammaticalization sources (i.e. *fallire, stare, habere, ire* etc.) in Romance languages, it is necessary to distinguish between distinct grammaticalization paths. As is to be observed in the grammaticalization path of the avertive in French developed out of *fallire* there is neither a necessitated avertive-to-proximative development nor a necessitated proximative-to-avertive development in Romance. On the one hand, the French avertive grammaticalize into the avertive in the 16th century without requiring a
detour through the proximative (cf. Gougenheim 1929). The French avertive will not be considered in greater detail, but this is due to the fact that the polar component of the lexeme *fallire* is grammaticalized earlier than the proximal component (i.e. imminence). On the other hand, it seems unlikely that the avertive in French will lose the polar component and grammaticalize into the proximative.

Nonetheless, the route proximative-to-avertive is indeed not a necessitated but a possible emergence of the avertive as will be shown for the STARE-periphrases in Ibero-Romance languages. I suggest that in the case of the STARE-periphrases in Ibero-Romance the avertive emerges from the proximative by the conventionalization of the inference of counterfactuality. Based on corpus data given from the *Corpus del Español* and the *Corpus do Português* I will verify the four stages of conventionalization of the counterfactual implicature defined by the model of Ziegeler (2000).

The proximative in Latin is expressed by the future active participle plus *esse* (‘to be’) as illustrated in (46).

(46) Latin (Revelation 12:4):

> et draco stetit ante mulierem quae erat paritura ut cum peperisset filium eius devoraret
>
> ‘and the dragon stood in front of the woman who was about to give birth, so that he might devour her child the moment it was born’

When the Latin proximative is provided with the infinitive perfect active of *esse* (i.e. *fuisse*), the analytic form expresses counterfactuality (cf. Fleishman 1982).

(47) Latin (Livius, *Ab urbe condita II*, Ch. 1; Part.4):

> Quid enim futurum fuit si [...]  
>
> ‘What would have happened if [...]’

Thus, the counterfactuality arises when the Latin proximative is combined with the perfective aspect. Nevertheless, the analytic form did not evolve the avertive. The avertive in Latin needs to be expressed lexically, as in (48).

(48) Latin:

> Paene dixi.  
>
> almost say.PST.1SG  
>
> ‘I almost said.’

As Palermo (2004) has shown for the Italian periphrases, *essere per* + infinitive arises earlier than *stare per* + infinitive and the frequency of *essere per* + infinitive is significantly higher than *stare per* + infinitive up to the 17th century. Contrary to the conservative Italian, the Ibero-Romance languages completely replace the *SER*-periphrases with the STARE-periphrases. As in the case of *essere per* + infinitive, *ser por* + infinitive and *ser para* + infinitive may express the modal meaning of volition. Combined with animacy and, further, within the scope of negation, the modal volitional meaning of *estar para* + infinitive presented above in (19) is already expressed by *ser para* + infinitive in the 13th century.
Sarah Schwellenbach

(49) Old Spanish (Anónimo, 1250, Bocados de oro, Spain, CDE):

\[ \text{no es para recibir ningún consejo} \]

not be.PRS.3SG PREP receive.INF any advice

‘He is not in favor of receiving any advice’

Looking at the first stage in the model of Ziegeler (2000), in Ibero-Romance languages there is also a rivalry between *ser* por + infinitive (as shown in 50, 52) and *estar* por + infinitive (as shown in 51) until the 15th century expressing prediction of X via an R-based implicature in a paradigm which is still dominated by venitive verbs as Spanish *venir* ‘to come’ and Portuguese *vir* ‘to come’. The fulfillment of the action is expected by the speaker.

(50) Old Spanish (Alfonso X el sabio, Siete Partidas, 1252–1284, CDE):

\[ \text{prophetar las cosas que son por venir.} \]

Predict the things that be.PRS.3PL PREP come.INF

‘predicting the things that are about to come.’

(51) Old Spanish (Alfonso X el sabio, Siete partidas, 1252–1284, CDE):

\[ \text{querer tomar el poder de dios para saber las cosas que} \]

want.INF take.INF the power of God to know.INF the things that

\[ \text{estan por venir.} \]

be.PRS.3PL PREP come.INF

‘wanting to take the power of God in order to know the things that are about to come.’

(52) Old Portuguese (Livro de vita Christi, 1446, CDP):

\[ \text{mas livra-nos de mal, de todo mal, […] Ou: de todo mal, passado,} \]

but deliver-us from evil from all evil Or from all evil past

\[ \text{presente e que he por vir.} \]

present and what be.PRS.3SG PREP come.INF

‘deliver us from evil, all evil […]. Or: from all evils, past, present and what is to come.’

Until the 15th century *ser* por + infinitive is increasingly replaced by the proximative *estar* por + infinitive. In the 14th century stage II is already lexically expressed by *estar en punto de* + infinitive (‘be on the point of doing, but […]’) cancelling the R-based implicature by an adversative clause and inducing a Q-based implicature of counterfactuality, as shown in (53).

(53) Old Spanish (Pedro Alfonso de Barcelos, anonymous translation, Crónica de 1344, CDE):

\[ \text{muchas veces estoujeron en punto de vencer la batalla /} \]

many times be.PST.3PL on the point of come.INF the battle but the

\[ \text{Romanos eran muy pocos & los otros eran muchos} \]

Romans be.IMPF.3PL very few and the others be.IMPF.3PL many

‘many times they were about to win the battle, but the Romans were very few in number and the others were many.’

From the 15th century the Spanish and Portuguese proximative *estar* por + infinitive and *estar para* + infinitive appears increasingly with the Q-based implicature of counterfactuality (stage II and III), as in the ‘schema of incidence’ shown in (54).
(54) Old Portuguese (Crónica de D. Fernando, 1431-1443, CDP):

Estando para partir, chegou Fernam d’Andrade com carta del Rey
be.GER PREP leave.INF arrive.PST.3SG Fernam d’Andrade with letter from-the king
‘He was about to leave, when Fernam d’Andrade arrived with a letter from the king’

As of 16th century the polar component of estar por + infinitive and estar para + infinitive no longer needs to be contextually reinforced and becomes conventionalized, so it becomes an avertive, as shown in the following examples. The counterfactuality expressed by the avertive periphrases in the indicative mood is accepted by the speaker as being equivalent with the expression of counterfactuality induced by the subjunctive, as indicated by the enumeration given in example (55).

(55) Early Modern Spanish (Juan Boscán, Poesías, 1514–1542, CDE):

¡Cuántas vezes estuvo por tornarse, cuántas veces quisiera
How many times be.PST.IND.3SG PREP return.INF how many times want.IMPF.SUBJ.3SG
hallar estorvo, y cuántas no quisiera ser nacido
find.INF obstacle and how many not want.IMPF.SUBJ.3SG be.INF born.PTCP
‘How many times did he almost return, how many times he has wanted to find the obstacle, and how many times he would not have been born.’

(56) Early Modern Spanish (Garcilaso de la Vega, El Inca, La Florida, 1578, CDE):

le dio una gran cuchillada en el pescuezo de que estuvo
him give.PST.3SG a big stab-with-a-knife in the neck of that be.PST.3SG
para morir
PREP die.INF
‘he gave him a big stab with a knife in the neck of which he almost died.’

(57) Early Modern Portuguese (Fernão Mendes Pinto, Peregrinação, 1603, CDP):

despois de bem açoutado, & pingado com huas torcidas de azeite, de que
after of good flog.PTCP and drip.PTCP with some drops of olive-oil of that
esteue para morrer,
be.PST.3SG PREP die.INF
‘After a good flogging and dripping with some drops of olive oil of that he almost died’

According to the data given from the Corpus del Español (CDE) and the Corpus do Português (CDP) the avertive expressed by the verbal periphrases estar por + infinitive and estar para + infinitive appears from the 16th century by the conventionalization of the implicature of counterfactuality. Therefore, the route proximative-to-avertive could be verified for the stare-periphrases in Ibero-Romance languages.

4. Conclusion

It has been shown that a closer look into the syntax-semantics interface allows to distinguish between de conatu, frustrative, proximative and avertive. The avertive differs from the proximative, frustrative and imperfectum de conatu by entailing counterfactuality, as can be verified by the cancellability test. Due to their underspecification, proximative and imperfectum de conatu conversationally implicate counterfactuality and functionally overlap with the avertive. In the case of imperfectum de conatu and proximative, the implicature of counterfactuality is
easily cancellable. Regarding the romance languages, the term ‘frustrative’ is used as an underspecified hypernym which may refer to a proximative, avertive or an imperfectum de conatu.

Given such underspecification, the proximative periphrases exhibit an imperfective morphology whereas the avertive periphrases exhibit a perfective morphology characterized by specification in regard to the non-realization of the event. In the case of French the development of restriction to perfective morphology stands for a highly grammaticalized avertive periphrasis. In Ibero-Romance languages the avertive shares the grammaticalization source with the proximative as in the case of the STARE-periphrases. The STARE-proximative in imperfect tense may implicate counterfactuality via a Q-based implicature and may therefore functionally overlap with the avertive. This Q-based implicature has been identified as the ‘counterfactual implicature’ defined by the model of Ziegeler (2000).

Regarding the diachronic relationship, the route avertive-to-proximative could not be confirmed for Romance languages. However, the avertive does not necessarily need to evolve from the proximative as can be seen in the case of the French avertive which grammaticalizes into the avertive in the 16th century without requiring a detour through the proximative. In the case of the STARE-periphrases in Ibero-Romance languages, the proximative grammaticalizes into the avertive by the conventionalization of the inference of counterfactuality. The four stages of Ziegeler (2000) could be verified for the STARE-periphrases in Ibero-Romance languages.

5. References

Anscombe, Jean-Claude & Oswald Ducrot 1976. L’Argumentation dans la langue. Langages 10 (42), 5-27.
On the interpretation of preverbal subjects in Spanish in contexts of narrow information focus

Melanie Uth
melanie.uth@uni-koeln.de

1. Introduction

As concerns the realization of narrow information focus in Spanish, syntactic approaches such as Zubizarreta (1998), Costa (2001), Büring & Gutiérrez Bravo (2001), Gutiérrez-Bravo (2008) generally assume that the nuclear accent is necessarily assigned to the final accented syllable of the sentence, so that non-contrastively focused subjects have to be placed in sentence-final position in order to be marked prosodically (cf. 1b. vs. 1a.). This phenomenon is generally conceived of as being due to a syntactic rule that is sensitive to prosodic constraints and that is labeled ‘Regla P’ or ‘p-movement’ by Zubizarreta (1998, 1999), cf. the generalization under (2).

(1) ¿Quién te regaló la botella de vino?
   (a) *[María] me regaló la botella de vino.
   (b) Me regaló la botella de vino [María].
   (Zubizarreta 1998: 125f)

(2) We may say that the reordering in (1b.) is due to prosodic reasons. More precisely, the constituents are reordered in order to put the focused constituent in the position which the neutral nuclear accent is assigned to in the clause (...). We refer to this mechanism by means of the designation Regla P.

(Podemos decir que el reordenamiento esquematizado en [1b. (MU)] está motivado por razones prosódicas. Más precisamente, el reordenamiento de constituyentes tiene lugar para dejar la constituyente foco ... en la posición donde cae el acento nuclear neutro dentro de la cláusula (...). Nos referimos a tal mecanismo con el término de Regla P., Zubizarreta 1999: 423f)

Contrary to that, the results of many phonetic experiments suggest that the nuclear accent may very well be assigned in situ to e.g. preverbal subjects in contexts of narrow information focus:

„We use the term narrow focus in contrast with broad focus, but do not distinguish between types of narrow focus (e.g. contrastive and non-contrastive). With regards to the word order and intonation facts considered here, in both Castilian Spanish and Neapolitan Italian there appears to be no difference between different types of narrow focus.” (Face & D’Imperio 2005: 285)

Whereas the majority of the relevant phonetic analyses rely on read pairs of Wh-questions and answers, which necessarily display a fixed word order, Gabriel (2007, 2010) pursues semi-spontaneous elicitation experiments that consist of two steps. First, short picture stories are shown to the informants in order to introduce the relevant referents and the scene the informants are supposed to be asked for. Afterwards, the same pictures are shown to the informants once again, but this time accompanied by (written) Wh-questions asking for the
subject or object referents introduced before. The informants are requested to imagine that the questions are asked by interlocutors that don’t know each other and that are unfamiliar with the questions that were asked before. One of the main results of these experiments is that, in line with the data of the phonetically oriented approaches, “[c]onstructions with transitive verbs exhibit a strong tendency towards the pre-verbal placement of a [non-contrastively (MU)] focused subject when the object is realized as a full nominal DP constituent” (Gabriel 2010: 189), a result that is in diametrical opposition to the grammaticality judgments of Zubizarreta (1998, 1999) and other syntactic approaches.

In this paper, we argue that methodological considerations give reason to readdress the controversy concerning the syntactic realization of non-contrastively focused subject constituents in Spanish from a pragmatic perspective. More precisely, we would like to give consideration to the hypothesis that the preverbal subjects in contexts of narrow information focus found for Spanish by Gabriel (2007) might be instances of what we will call exhaustive focus in what follows, the (eventual) exhaustive marking being due to certain pragmatic strategies similar to the ones that ‘rescue’ bare subjects in sentences such as Trataron de salvarlo médicos #(famosos) (‘Famous doctors tried to save him’) or TURISTAS llegaron a la ciudad (‘Tourists arrived to the city’) according to Leonetti (2012, in press). Such data lead us to consider the question of the syntactic realization of non-contrastively focused subject (and object) constituents in Spanish as ultimately unresolved up to this day, so that we see an urgent need to discuss elicitation methods that might help to further clarify the issue.

The outline of the article is as follows. In section 2, we retrace the long-standing debate concerning the syntactic realization of non-contrastively focused subject constituents in Spanish and we present the experimental design of Gabriel (2007, 2010) in more detail. In section 3, we draw attention to the exhaustive nature of the focus type generally called ‘contrastive focus’ by Zubizarreta (1998, 1999) and others, and we discuss the methodology employed by Gabriel (2007, 2010) against the background of pragmatic accounts on exhaustivity, assuming that exhaustivity is to be traced back to pragmatic inferences guiding the interpretation of cleft sentences (e.g. Horn 1981, Lambrecht 1994, Dufter 2009, Drenhaus et al. 2011), just as well as the interpretation of fronted objects (Skopeteas & Fanselow 2011). In section 4, we present an experimental design which we designed in order to mitigate the effects of pragmatic inferences of the above kind, and we discuss the results of an elicitation experiment we recently carried out in Quintana Roo, Mexico, on the basis of this material. In section 5, we give a summary of the main conclusions and an outlook for future research.

2. Non-contrastively focused subjects in Spanish: The debate

The syntactic literature on focus marking in (standard) Spanish generally distinguishes between two different kinds of focus, i.e. information focus and contrastive focus, the former corresponding to the new information of a sentence, that is not yet part of the common ground of the speaker and the hearer, whereas the latter is generally understood as “a quantification-like operation which involves exhaustive identification on a set of entities” (Gutiérrez Bravo 2008: 164), hereby capturing the fact that the various kinds of the so-called ‘contrastive focus’ (corrective, affirmative etc.) are all based on the exhaustive marking and/or interpretation of the corresponding constituents (cf. ibid.: 164f as well as Kiss 1998, Zubizarreta & Vergnaud 2005, Adli 2011: 114, and the references cited therein).¹ According to most of the syntactic approaches, the two focus types are realized by fundamentally different grammatical means in Spanish. As Zubizarreta (1998, 1999), Costa (2001), Domínguez (2004), Gutiérrez Bravo

---

¹ There is certain terminological confusion as concerns the denomination of the different focus types. For the time being, we refer to the relevant focus types as ‘information focus’ and ‘contrastive focus’, respectively, following Zubizarreta & Vergnaud’s (2005) terminology. However, in section 3 we refine our terminology based on semantic considerations following Umbach (2004), Büring (2006), Krifka (2007) and others.
(2006) and others argue, in standard Spanish, the non-contrastively focused constituent is obligatorily placed in the sentence final position in order to coincide with the ‘neutral nuclear stress’ (“acento nuclear neutro”, Zubizarreta 1999: 4229, cf. 3b. and c.), which obligatory falls on the final constituent (3a).2

(3) (a) Nuclear Stress Rule, NSR, romance version:
In Spanish, nuclear stress falls on the rightmost accented word within the Intonational Phrase.
(b) Focus Prominence Rule, FPR:
The focused constituent must contain the intonational nucleus of the Intonational Phrase, where the intonational nucleus is identified as the syllable that bears the main phrasal prominence.
(c) ¿Quién compró los discos?
Los discos, los compró [una muchacha]. (ibid.)

By contrast, phrasal prominence to signal contrastive focus is, in general, not generated by means of the NSR, but is attributed to the so-called Contrastive/Emphatic Stress (“acento nuclear enfático”, ibid.), which “may fall on any accentable morpheme” in the sentence.3

(4) (a) Emphatic/Contrastive Stress Rule, E/CSR:
The emphatic stress may fall on any accentable morpheme.
(b) El gato de botas ROJAS se comió un ratón, y no el de botas AZULES.
(Zubizarreta 1999: 4230)

Contrary to that, the results of many phonetic experiments suggest that the (neutral) nuclear accent may very well be assigned in situ to e.g. preverbal subjects in contexts of narrow information focus (cf. e.g. Face 2001, Face & d’Imperio 2005, Cabrera Abreu & García Lecumberri 2003, Estebas-Vilaplana & Prieto 2010, Hualde 2005). Whereas the majority of these analyses rely on read pairs of Wh-questions and answers, which necessarily display a fixed word order, and which are hence ultimately irrelevant for our discussion, the results of Gabriel’s (2007) semi-spontaneous elicitation experiments equally argue for the possibility to displace the nuclear accent in order to signal narrow information focus in Spanish. These experiments consist of two steps. First, short picture stories are shown to the informants in order to introduce the relevant referents and the scene the informants are supposed to be asked for (cf. table 1, Ia.-IIb.). Afterwards, the same pictures are shown to the informants once again, but this time accompanied by written Wh-questions asking for the subject or object referents introduced before (cf. e.g. table 1, Ic.-IId.). The informants are requested to imagine that the questions are asked by interlocutors that don’t know each other and that are entirely unfamiliar with the questions that were asked before. The participants were asked to answer by means of complete sentences, and in order to motivate this request, they were told that that the data shall afterwards be used to train learners of Spanish as a second language.

2(3a., b.) and (4a.) are our own translations of the Spanish generalizations in Zubizarreta (1999: 4229f.), partly adapted on the basis of Adli (2011: 116, 119).
3With this, Zubizarreta (1998, 1999) argues that ‘nothing prevents the contrastive focus to be identified by the neutral nuclear accent’, ibid.: 4229, our translations) via p-movement in cases in which “the NSR also applies”, assigning “main prominence to the [contrastively focused constituent (MU)]” (Zubizarreta 1998: 134), the problem of this hypothesis being that it finally comes down to assuming true optionality in the grammar. The issue is not relevant for our purposes, however.
Table (1): Experimental design of Gabriel (2007) with exemplary choice of 4 stimuli/Wh-questions

<table>
<thead>
<tr>
<th>(Ia.)</th>
<th>(Ilia.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>María compra el diario en el kiosco.</td>
<td>Blancanieves secuestra a Tarzán...</td>
</tr>
<tr>
<td>‘María buys the newspaper at the kiosk.’</td>
<td>‘Snow White kidnaps Tarzan...’</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(Ib.)</th>
<th>(Ilb.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Después se lo da a su hermano.</td>
<td>... y se lo entrega a los siete enanitos.</td>
</tr>
<tr>
<td>‘Shortly after, she gives it to her brother.’</td>
<td>‘... and hands him over to the Seven Dwarfs.’</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(Ic.)</th>
<th>(Ilc.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>¿Quién compra el diario en el kiosco?</td>
<td>¿Quién secuestra a Tarzán?</td>
</tr>
<tr>
<td>‘Who buys the newspaper at the kiosk?’</td>
<td>‘Who kidnaps Tarzan?’</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(Id.)</th>
<th>(Ilid.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>¿Quién da el diario a su hermano?</td>
<td>¿Quién entrega a Tarzán a los 7 enanitos?</td>
</tr>
<tr>
<td>‘Who gives the newspaper to her brother?’</td>
<td>‘Who hands over Tarzan to the Seven Dwarfs?’</td>
</tr>
</tbody>
</table>

The 18 Hispanic informants that were interrogated in the realm of the experiment are from different regions of Spain (14 participants) and Latin America (4 participants) and are mostly graduate students or part of the university stuff. Table 2 presents the results with respect to the syntactic realization of the non-contrastively focused subject constituents pertaining to the stimuli (Ia.) and (Ib.), showing that the informants almost exclusively make usage of constructions with preverbal subjects, whereas there is not any evidence in favor of $p$-movement in the data.4

4 The answers to stimuli (Ic.) and (Id.) are not presented by Gabriel (2007). Double-object constructions (answers to e.g. ¿Qué le da María a su hermano?) or answers involving a direct object and a goal constituent (answers to ¿A quién le entrega Blancanieves a los siete enanitos?) would normally also be relevant for the present debate. However, for reasons of clarity, we restrict the discussion to subject constituents in this paper.
Table (2): Realization of subjects with respect to stimuli (Id., IIc., table 1) according to Gabriel (2007)

<table>
<thead>
<tr>
<th>Syntactic construction</th>
<th>Example</th>
<th>share</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preverbal, nom. object</td>
<td>f[Maria]f le da el diario a su hermano.</td>
<td>26</td>
<td>72%</td>
</tr>
<tr>
<td>Preverbal, pron. object</td>
<td>f[Maria]f se lo da.</td>
<td>1</td>
<td>3%</td>
</tr>
<tr>
<td>Cleft</td>
<td>f[Maria]f es quien le da el diario a su hermano.</td>
<td>6</td>
<td>17%</td>
</tr>
<tr>
<td>P-movement</td>
<td>Le da el diario a su hermano f[Maria]f.</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>Final</td>
<td>Se lo da f[Maria]f.</td>
<td>3</td>
<td>8%</td>
</tr>
<tr>
<td>Table</td>
<td></td>
<td>36</td>
<td>100%</td>
</tr>
</tbody>
</table>

From this, Gabriel (2007) concludes that Zubizarreta’s (1998, 1999) far-reaching generalization concerning p-movement in Spanish has to be relativized in view of the empirical results, since, in sentences/constructions with a direct object DP, the participants of Gabriel’s (2007) study do not feel any need to place the non-contrastively focused subject constituent in sentence final position in the context of narrow information focus.

3. Contrast, exhaustivity and pragmatic inferencing: The problem

From the above it is evident that the debate concerning the positioning of non-contrastively focused subjects in Spanish highly depends on Zubizarreta’s (1998, 1999) categorical distinction between narrow information focus, considered as new, i.e. non-presupposed information of an utterance, on the one hand, and ‘contrastive focus’, conceived of “a quantification-like operation which involves exhaustive identification on a set of entities” (Gutiérrez Bravo 2008: 164). However, the utility of the notion of ‘contrast’ in order to disentangle different types of focus is likewise highly debated in the relevant literature. Thus, whereas Zubizarreta’s (1998, 1999) distinction is based on Chomsky (1971), who argues in favor of a categorical difference between ‘ordinary’ syntactic focus and expressive, i.e. contrastive focus, Bolinger (1961, 1972) claims that any focus is ultimately contrastive, suggesting that “[a]s the alternatives are narrowed down, we get closer to what we think of as contrastive accent” (Bolinger 1961: 87). Indeed, many influential semantic accounts on focus part from the assumption that “[f]ocus generally indicates the presence of alternatives that are relevant for the interpretation of linguistic expressions” (Krifka 2007: 18, based on Rooth 1985, 1992). Moreover, it is generally assumed that a speaker focusing a constituent aims at relating it to the Common Ground (CG) by establishing a set of alternatives containing an element that is (semantically) given in the CG (cf. e.g. Büring 2006: 148). Against the background of this general semantic principle, the fact that the ‘focusing mechanism’ may be used in order to achieve different communicative goals (pragmatic uses of focus, Krifka 2007: 21-25) accounts for the different subtypes of focus. One of the subtypes is e.g. ‘exhaustive focus’, which is employed in order to signal that “the focus denotation is the only one that leads to a true proposition, or rather more general: that the focus denotation is the logically strongest one that does so” (ibid.: 33). However, as we already alluded to above, the syntactic approaches generally also relate the so-called ‘contrastive focus’ (Chafe 1976) or ‘identificational focus’ (Kiss 1998) to exhaustive interpretation. For example, similarly to that proposes Gutiérrez Bravo (2008) for ‘contrastive focus’ in Spanish, Kiss (1998) suggests for the Hungarian ‘identificational focus’ construction that the corresponding focused constituents refer to “[the exhaustive (MU)] subset of the set of contextually or situationally

---

5 Roughly speaking, the Common Ground can be equated with the “information that is mutually known to be shared and continuously modified in communication”, (Krifka 2007: 15).
6 Chafe (1976) proposes a similar albeit distinct differentiation assuming that contrastive-ness is characterized by the fact that the number of candidates [of the candidate set (MU)] is limited and the sentence corrects an explicit or implicit assumption made by the hearer, whereas sentences supplying new information allude to an unlimited set of possibilities, cf. Krifka (2007: 32f) for a more detailed discussion.
given elements for which the predicate phrase can potentially hold” (ibid.: 245), whereas (narrow) informational focus does not imply exhaustive identification of a/several candidate/s of a given set but equals to the non-presupposed information in a sentence. Evidently, based on the notion of exhaustivity, the dualistic view on the focus category/ies is easy to reconcile with the semantic accounts in the way proposed by Krifka (2007). However, note that this reconception of the focus dichotomy requires a reconsideration of the debate concerning the preverbal non-contrastively focused subject constituents in the above mentioned elicitation experiments, since there are several reasons to suspect that the informants who opted for the prosodic marking of a preverbal subject constituent in the context of questions such as ¿Quién compró el periódico? reacted this way in order to signal that “the focus denotation is the only one (...) or (...) the logically strongest one that leads to a true proposition” (cf. above).

This supposition becomes even more probable against the background of recent work on exhaustivity suggesting that the exhaustive interpretation of clausal constituents is very often due to pragmatic inferences, which highly depend on the assumptions of the speaker concerning the information “the questioner is interested in” (Schulz & von Rooij 2006: 7). With this, it is important to note that Schulz & von Rooij (2006), just as Groenendijk & Stoeckhof (1984) and others, are dealing with a phenomenon that may be subsumed under the label implicit exhaustivity, whereas we are concerned with the (possible) explicit marking, or explicit triggers, of exhaustivity in structurally marked answers to Wh-questions, which are rather comparable to English it-clefts (cf. e.g. Horn 1981, Lambrecht 2001, Huber 2006), to fronted object constructions of different languages (Skopeteas & Fanselow 2011), or to the Hungarian preverbal focus position (Kiss 1998), phenomena that we will relate to explicit exhaustivity in what follows.

While there are indeed several semantic analyses of e.g. English it-clefts and the Hungarian focus position (cf. especially Lambrecht 2001, Kiss 1998), many scholars nowadays agree on the hypothesis that both constructions are to be analyzed in terms of (non-truth-functional) pragmatics, or CG management, respectively. Among the many analyses of cleft constructions in different languages, Dufter (2009) cites abundant evidence in favor of a pragmatic account on the exhaustivity of English it-clefts and its Romance counterparts. For example, he points out that exhaustive modifiers such as only occur as modifiers of cleft constructions, albeit it is normally not possible to combine several expressions with exhaustive semantics (cf. 5a.). Another piece of evidence comes from clefts with universal quantifiers, since universal quantifiers seem to be generally incompatible with exhaustive modifiers (5b.).

(5) (a) It’s only on that basis that a balanced approach can be found.
   ?Only exclusively the wife.
 (b) In this case, it is everyone who is being discriminated against.
   ?Only everyone

Dufter (2009: 96f)

Against the background of such data, the author convincingly argues that the exhaustivity of cleft constructions should best be analyzed as a generalized conversational implicature in the sense of Levinson (2000) and Horn (1981), which may be reinforced (cf. 5a.) or cancelled (cf. e.g. It is also out of respect for your electorate that (...) I do not wish to set myself up as a judge, Dufter 2009: 96), depending on the demands of the communicative setting, or the CG management, respectively.

A further piece of evidence in favor of the pragmatic nature of the ‘exhaustivity restriction’ of it-clefts comes from a recent investigation by Drenhaus et al. (2011), showing for German

---

7 For details concerning the distinction between CG content and CG management cf. Krifka (2007: 16f).
that the violation of exhaustivity is much less acceptable in sentences containing an exhaustive modifier (*Only Maria can play the piano and, besides, Luise and Jana*, ibid.: 326f) than in cleft constructions (*It is Maria who plays the piano and, besides, Luise and Jana*, ibid.). The additional ERP-study conducted with the same contrastive stimuli suggests that the neurological processing of the exhaustivity violation in *it*-clefts differs categorically and qualitatively from the one of the exhaustivity violation in the *only*-sentences. More precisely, according to the authors, the ERP-effect observed with the former violation can be interpreted as N400 effect, which is generally held to be an indicator of semantic/pragmatic integration but “cannot be taken as an indicator of truth-value violations” (Drenhaus et al. 2011: 324, based on results from Fischler et al. 1983), whereas the violation of exhaustivity in *only*-sentences resulted in a centro-parietal positivity at 600-800ms which can be interpreted, according to Drenhaus et al. (2011), either “as an indicator of general reanalysis” (ibid.: 334) or as “an indication for the actualization and updating of a mental model” (ibid.).

Similarly, in recent work on the Hungarian preverbal Focus Position (Hungarian FP, in what follows), it is convincingly shown that the exhaustivity generally held to be a semantic characteristic of this construction (cf. e.g. Kiss 1998, Szabolcsi 1981, 1994), is also best to be analyzed as a pragmatic phenomenon. For example, Wedgwood (2009) shows that a question introduced by the pluralized Wh-phrase kik (who) may be restrictively answered by means of a restrictive modifier (6a.), whereas it is not possible to explicitly restrict the focused candidate(s) by just localizing the corresponding NP/proper name in the Hungarian FP (6b.).

(6)

Kik hívták fel Emilt?
‘Who called Emil?’
(a) Csak Anna hívt fel Emilt.
‘Only Anna called Emil.’
(b) # Anna hívt fel Emilt
‘Anna called Emil.’

Wedgwood (2009: 104)

A similar albeit more elaborated comparison of restrictive modifiers and the Hungarian FP is accomplished by Onea & Beaver (2011). In the first part of the paper, the authors reject the semantic analyses of the Hungarian FP in a way very similar to what we saw above with respect to Dufter’s discussion of cleft constructions. To cite but one example, Onea & Beaver (2011) point to the fact that constituents in the Hungarian FP may co-occur with the restrictive modifier csak, as evidenced by (7).

(7)

Péter CSÁK MARIT (*meg-*)csőkolta meg.
Peter only Mary kissed
‘Peter kissed only Mary.’

Onea & Beaver (2011: 346)

Secondly, the researchers report on a series of perception experiments they conducted in order to show that “immediately pre-verbal focus in Hungarian is (...) clearly less exhaustive than sentences containing exclusives like only, which we take to be clearly conventionally exhaustive” (ibid.: 348). For example, the informants had to judge the acceptability of the replies under (8a.-c.) which were given to the several focus constructions in table 3, which in turn served as more or less exhaustive descriptions of various visual stimuli triggering more or less important exhaustivity violations (cf. ibid.). With this, the authors part from the hypothesis that the exhaustive modification in (1a., table 3) leads to a contradiction with respect to the pictorial stimulus, so that in this case, the informants would most probably select answer (8c.). Contrary to that, the ‘default sentences’ are “true but possibly
pragmatically misleading” (ibid.: 351) triggering the answers (8a.) or (8b.). However, the “sentences in the second condition (…) are false or misleading depending on whether the exhaustiveness feature is part of the truth conditional content or not” (ibid.).

Table (3): Experimental design of Onea & Beaver (2011: 350-355, experiment 1)

<table>
<thead>
<tr>
<th>(I.) Exhaustive modifier:</th>
<th>(II.) Hungarian FP:</th>
<th>(III.) ‘Default sentences’:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Csak MARCI fogott meg egy lepkét. Only Marci caught PRT a butterfly ‘Only Marci caught a butterfly’</td>
<td>MARCI fogott meg egy lepkét. Marci caught PRT a butterfly ‘Marci caught a butterfly’</td>
<td>Marci meg-fogott egy lepkét. Marci PRT caught a butterfly ‘Marci caught a butterfly’</td>
</tr>
</tbody>
</table>

(8) (a) Yes, and Peter caught a butterfly too.  
(b) Yes, but Peter caught a butterfly too.  
(c) No, Peter caught a butterfly too.

The main result of this experiment is that the constructions involving the Hungarian FP are significantly less likely to be contradicted for ‘untrue exhaustivity’ than the only-sentences (i.e. significantly fewer share of ‘No,’-answers in the condition II.). On the other hand, the informants preferred the ‘Yes, but’- answer over the ‘Yes, and’-reply in this context (i.e. II, the Hungarian FP construction), whereas they predominantly opted for the ‘Yes, and’-reply in the context of the ‘default sentences’ (III.). Against the background of these data, the authors conclude that, first of all, “the exhaustiveness effect associated with immediately pre-verbal focus in Hungarian is not semantic” (ibid.: 358), and secondly, that there is nevertheless a strong “pragmatic tendency to exhaustify pre-verbal focus in Hungarian” (ibid.).

Coming back to our discussion concerning the debate on preverbal non-constrastively focused subjects in Gabriel’s (2007) elicitation data, the most important insight from the above discussion is that there seems to be a considerable pragmatic flexibility in the realization and interpretation of exhaustive meaning components or pragmatic inferences, respectively, the exact realization or interpretation being to a large extent dependent on CG management, i.e. on the way the speakers present the information that is to be added to the Common Ground. Furthermore, it has become evident that the CG management in turn highly depends on the communicative setting and/or the speaker’s assumptions concerning the relevant presuppositions, i.e. the CG content (cf. again Schulz & von Rooij 2006).

With this in mind, we would like to point to the fact that the communicative settings given to the informants of the above mentioned elicitation experiments are rather vague, so that, in the context of the corresponding Wh-questions, the participants might feel motivated to just confirm what they saw when having been confronted with the relevant picture story, thereby opting for prosodic marking of preverbal subjects in order to signal that “the focus denotation is the logically strongest one that [that leads to a true proposition]” (cf. again Krifka 2007: 33), i.e. in order to pragmatically strengthen the (clues for) exhaustive interpretation. Moreover, the fact that the referents are introduced by means of proper names might further detract the informants from appropriately reenacting the intended communicative settings since there is an important pragmatic mismatch between the familiarity with the referents suggested by the usage of proper names and the underspecified nature of the (unknown?) questioners and the undefined communicative setting. These methodological considerations
lead us to consider the question of the syntactic realization of non-contrastively focused subject constituents in Spanish as ultimately unresolved up to this day. In the next section, we will present an experimental design which we developed in order to mitigate the effects of pragmatic inferences of the above kind, and we discuss the results of a corresponding elicitation experiment we recently carried out in Quintana Roo, Mexico, on the basis of this material.

4. Disentangling explicit exhaustivity and non-contrastive narrow focus: The proposal

Resuming the above discussion concerning the syntactic and prosodic realization of non-contrastively focused subject constituents in Spanish, it is to be recalled, first of all, that the grammaticality judgments suggesting obligatory p-movement are diametrically opposed to the elicitation data in favor of the possibility to prosodically mark the focused subject in preverbal position, the latter construction being traditionally held as pertinent to ‘contrastive focus’ only. However, we drew attention to the fact that the preverbal realization of the non-contrastively focused subjects might be motivated by the fact that the speakers did not properly reenact the intended communicative setting, but felt the need to signal exhaustivity due to pragmatic reasons, a supposition that becomes even more plausible in view of the abundant evidence suggesting that exhaustive marking and/or interpretation of focused constituents is a matter of CG management, meaning that it is highly dependent on the (speaker’s assumptions concerning the) intended communicative setting.

In order to empirically clarify our methodological suppositions, we recently designed a threefold elicitation material based on the elicitation design of Gabriel (2007) and we conducted a corresponding elicitation study with 15 speakers in Quintana Roo, Mexico. The speakers are of a similar age as the ones surveyed by Gabriel (2007) and have a similar educational background. In one third of the queries, conducted on the basis of 5 informants, we tightly followed the design of Gabriel (2007). In the second part, we introduced the relevant referent by means of indefinite noun phrases instead of proper names, and in the last part, we modified the pictorial stimuli in order to be more precise as concerns the communicative settings. Crucially, we designed the elicitation material of the last part aiming at tightly replicate the contexts that motivated Zubizarreta’s (1998, 1999) formulation of the generalization concerning p-movement in Spanish. Furthermore, against the background of the above discussion, we had to be sure to base our query on well-defined communicative settings. For these purposes, we elaborated slightly more complex pictorial stimuli, in order to create contexts that do not bear the risk of a simple confirmation or reaffirmation of the information given to the informants in the introduction (cf. table 4, Ia.-Ig. as well as table 5, IIa.-IId.). In the next step, the informants were shown the same pictures once again, but this time accompanied by questions asked by the friends and neighbors of the ‘main characters’, which were equally introduced before by means of the corresponding introductory slides. In tables 4 and 5, the stimuli used for focus elicitation are exemplified by means of the ones used to elicitate non-contrastively focused subject constituents (cf. table 4, Ih.-Ii. as well as table 5, Ile.-Ilf.). The participants were told that they will see the pictures once again, but this time accompanied by speech balloons that they shall complete by giving (contextually) appropriate answers. As in the context of Gabriel’s (2007) experiments, the informants of the study were asked to answer by means of complete sentences and to behave as natural as possible.

Many thanks to Cristoph Gabriel for leaving to us the original slides of the enquiry of 2007 as well as to Henrike Rödiger for designing the further stimuli.
On the interpretation of preverbal subjects in Spanish in contexts of narrow information focus

Table (4): Experimental design of elicitation study, condition 3, story 2 & stimuli for subject focus

<table>
<thead>
<tr>
<th>(Ia.)</th>
<th>(Ib.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ella es Aruma Hernández Casas. ‘This is Aruma Hernández Casas.’</td>
<td>Aruma compra un periódico en una tienda, y una conocida con su hija la están mirando desde lejos. ‘Aruma buys a newspaper in a kiosk, and a riend with her daughter are looking at her from a distance.’</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(Ic.)</th>
<th>(Id.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Después, Aruma encuentra a su amigo Don Hernando y dan un paseo en un parque cercano. ‘Afterwards, Aruma encounters her friend Don Hernando. They go for a walk in a near public park.’</td>
<td>En el parque, encuentran a unos amigos y se paran un rato a platicar con ellos. ‘In the park, they encounter some friends and they stay a bit in order to have a conversation.’</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(Ie.)</th>
<th>(If.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Luego, Aruma acompaña a Don Hernando a casa y le da el periódico. Sus vecinos los están mirando desde su jardín. ‘Aruma walks home Don Hernando and she gives him the newspaper, Don Hernando’s neighbors being in their backyard looking at them.’</td>
<td>Don Hernando entra a su casa, prepara la cena y espera a su hermana, leyendo el periódico. ‘Don Hernando enters his home, prepares the dinner and waits for her sister, reading the newspaper.’</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(Ig.)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Por fin, llega su hermana y empiezan a cenar. ‘Finally his sister gets home and they have dinner.’</td>
<td></td>
</tr>
</tbody>
</table>
(Ih.) Stimulus for subject focus 1:

Who bought the newspaper in the kiosk?

(Ii.) Stimulus for subject focus 2:

Who gave the newspaper to you?

Table (5): Experimental design of elicitation study, condition 3, story 1 & stimuli for subject focus

(Iia.) Blancanieves secuestra a Tarzán, y dos enanitos los están mirando de detrás de una roca.
‘Snow White kidnaps Tarzan, two dwarfs hiding behind a rock and looking at them.’

(Iib.) Después, Blancanieves entrega a Tarzán al pueblo de los Siete Enanitos, …
‘Afterwards, Snow White brings Tarzan to the village of the Seven Dwarfs, …’

(Iic.) … lo encierra en un huacal y vuelve a salir para colectar hongos para la cena.
‘She locks him up in a cage and leaves the village in order to collect mushrooms for the dinner.’

(Ild.) Luego, regresan los cinco enanitos restantes a su pueblo. Están todo asustados por la enorme criatura en el huacal.
‘Shortly after, the remaining five dwarfs come back to the village being totally scared by the huge creature sitting in the cage.’

(Ile.) Stimulus for subject focus 1:

Who is kidnapping Tarzan?!

(Ilf.) Stimulus for subject focus 2:

Who brought Tarzan to our village?!
As concerns the results of our study, we would like to stress, first of all, that the informants generally behaved in a much more natural way in the third condition than in the two conditions replicated in analogy to Gabriel (2007). That is to say, the utterances obtained in the realm of condition 3 are characterized by their great naturalness as concerns both, prosody (no astonishment, no boredom, no listing intonation) just as well as the verbalization of respective contents (flexibility, creativity and idiosyncratic behavior), so that we can be quite sure that the participants fully engaged in the ‘game’ and faithfully reenacted the intended communicative setting.

Moreover, table 8 lists the realization of the subject constituents in the utterances pertaining to the stimuli (Ih.) and (II.) of table 4, as well as (IIf.) and (IIIf.) of table 5, which are to be compared to the answers that are given to the corresponding stimuli of condition 1 (‘Aruma compra un diario... y después se lo da a un amigo’, ‘Blancanieves secuestra a Tarzán... y se lo entrega a los siete enanitos,’ => ¿Quién...? x 4) and condition 2 (‘Una mujer compra un diario en una tienda... y después se lo da a un amigo’, ‘Una princesa atrapa a un bandido... y se lo entrega a un grupo de enanitos.’ => ¿Quién...? x 4), which are listed in tables 6 and 7, respectively.

Table (6): Realization of subjects of stimuli for non-contrastive focus on the subject, condition 1

<table>
<thead>
<tr>
<th>Syntactic construction</th>
<th>Example</th>
<th>share</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preverbal, nom. object</td>
<td>f[Aruma]f compró el diario en la tienda.</td>
<td>12</td>
<td>60%</td>
</tr>
<tr>
<td>Preverbal, pron. object</td>
<td>f[Blancanieves]f lo secuestra.</td>
<td>1</td>
<td>5%</td>
</tr>
<tr>
<td>Cleft</td>
<td>f[Aruma]f es quien le da el diario al hermano.</td>
<td>6</td>
<td>30%</td>
</tr>
<tr>
<td>P-movement</td>
<td>Le da el diario a su hermano f[Aruma]f.</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>Others final</td>
<td>Se lo da f[Aruma]f.</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>Others non-final</td>
<td>El diario lo compra f[una mujer llamada Aruma]f - en una tienda de abarrotes.</td>
<td>1</td>
<td>5%</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>20</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table (7): Realization of subjects of stimuli for non-contrastive focus on the subject, condition 2

<table>
<thead>
<tr>
<th>Syntactic construction</th>
<th>Example</th>
<th>share</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preverbal, nom. object</td>
<td>f[La mujer]f le está dando el diario a un amigo.</td>
<td>18</td>
<td>90%</td>
</tr>
<tr>
<td>Preverbal, pron. object</td>
<td>f[Una Princesa]f lo secuestra.</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>Cleft</td>
<td>f[una princesa]f está atrapando a un bandido.</td>
<td>2</td>
<td>10%</td>
</tr>
<tr>
<td>P-movement</td>
<td>Le da el diario a su amigo f[una mujer]f.</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>Others final</td>
<td>Se lo da f[una mujer]f.</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>20</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table (8): Realization of subjects of stimuli for non-contrastive focus on the subject, condition 3

<table>
<thead>
<tr>
<th>Syntactic construction</th>
<th>Example</th>
<th>share</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preverbal, nom. object</td>
<td>f[Aruma]f compró el diario y me lo entregó.</td>
<td>4</td>
<td>20%</td>
</tr>
<tr>
<td>Preverbal, pron. object</td>
<td>f[Blancanieves]f nos lo trajo.</td>
<td>1</td>
<td>5%</td>
</tr>
<tr>
<td>Cleft</td>
<td>Fue f[Blancanieves]f quien lo entregó aquí.</td>
<td>3</td>
<td>15%</td>
</tr>
<tr>
<td>P-movement</td>
<td>Le da el diario a su hermano f[Aruma]f.</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>Others non-final</td>
<td>Me lo dio f[Aruma]f en la mañana. // Vino f[Blancanieves]f corriendo y se nos entregó.</td>
<td>2</td>
<td>10%</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>20</td>
<td>100%</td>
</tr>
</tbody>
</table>
Comparing the data presented in tables 6 to 8, note, first of all, that there is not any important difference between the shares of condition 1 and 2, meaning that the type of referring expressions used to introduce the referents (i.e. proper names versus indefinite descriptions) did not seem to play any role as concerns the realization of the subject constituents in the answers corresponding to the relevant stimuli or questions, respectively. Still, there are important differences between the shares of condition 3 compared to the ones of conditions 1 and 2 (cf. tables 6 and 7), and to the data of Gabriel (2007) (cf. table 2). One of the most evident differences is that there are much more subject constituents in final position in condition 3 of our elicitation material (50%) than in the conditions 1 and 2 (0%, apart from clefts) or in Gabriel’s (2007) data (8%). Accordingly, the informants realize much fewer preverbal subjects in condition 3 (25%) than in conditions 1 and 2 or in Gabriel’s data (65%, 90% and 75%, respectively). Nevertheless, there are two further important aspects concerning the data of the third condition, which ultimately strengthen Gabriel’s (2007) position. First of all, there is not any unequivocal data in favor of structures that might be traced back to what Zubizarreta (1998, 1999) calls ‘prosodic movement’, whereas there is, secondly, unmistakable evidence in favor of the possibility to localize non-contrastively and/or non-exhaustively focused subject constituents in preverbal position in Spanish (cf. the stimuli in tables 4 and 5 and the corresponding shares in table 8, as well as examples such as ¿Quién te dió el diario? – [Aruma] compró el diario y me lo entregó. ‘Aruma bought the newspaper and gave it to me.’). As such, the results may be taken as important additional evidence in favor of the possibility of preverbal subjects in contexts of narrow information focus in Spanish.

5. Conclusions and future research

For the time being, we may draw the following three conclusions from our preliminary investigation. First of all, the communicative setting is highly important in the realm of elicitation experiments related to information structure. Secondly, the sentence final position seems to be by far the most common position for non-contrastively focused subject constituents in Spanish in constructions without full nominal object constituents, whereas the abundant preverbal subjects in the conditions 1 and 2 are very probably to be traced back to the fact that the speakers felt the need to signal exhaustivity for the (pragmatic) reasons mentioned in section 3. As concerns the comparison of our data with the results of Gabriel (2007, 2010), note that the fact that “focused subjects are preferably realized in clause-final position” in constructions “involving the pronominalization of the object is equally emphasized by Gabriel (2007, 2010: 195), already. On the other hand, our data suggest, again along the lines of Gabriel (2007), that it is an entirely grammatical, albeit marginal option for native speakers of Spanish to localize non-contrastively and/or non-exhaustively focused subjects in preverbal position (cf. again the shares presented in table 8).

However, it should be noted that our results may at best be conceived of as a preliminary approach to the (methodological) issues related to narrow focus realization in Spanish. For example, one immediate desideratum of the present investigation is the prosodic analysis of the relevant subject constituents, since the realization of the different (sub)types of focus is a matter of syntax and prosody in Spanish. In this context, it would be particularly interesting to investigate and/or discuss in detail the prosodic realization of the preverbal subjects in the third condition in view of the question if they are prosodically marked by a nuclear accent or not. Secondly, it is evident that our experiment can only be judged as a pilot study due to the low number of participants. There is an urgent need to conduct follow-up experiments in order to receive well-founded data on the basis of a considerably larger amount of informants. This seems to be especially important in view of the long-standing dispute concerning the possibility of preverbal subjects and/or p-movement in the realm of non-contrastively focused subjects in Spanish. In this respect, it is furthermore to be noted that our data do not virtually
contain any evidence as concerns the grammaticality of $p$-movement, since the fact that the informants omitted this strategy does not mean that it is not part of the grammar of their native language. For this reason, it is very important to design and conduct appropriate perception experiments (e.g. again along the lines of Gabriel 2007) in order to investigate the acceptability of constructions involving $p$-movement. Finally, another drawback of the present investigation is that we cannot entirely exclude the possibility that any preverbal subject constituent is (explicitly) marked for exhaustivity in our elicitation experiments, e.g. for (methodological) reasons that are unknown so far.\footnote{Ultimately, this reasoning even extends to the postverbal subjects figuring in the corresponding environments.} Therefore, it would be reasonable to conduct a perception experiment along the lines of Onea & Beaver (2011) in order to explicitly test the several syntactic constructions produced by the informants of our pilot study with respect to possible exhaustive meaning components or pragmatic inferences, respectively. For this purpose, we need to develop pictorial stimuli displaying at least two referents each which are carrying out one and the same action (such as e.g. buying a newspaper etc.). Furthermore, the informants should be asked to rate the appropriateness of the ‘Yes, and…’-‘Yes, but…’- and ‘No…’- answers as described by Onea & Beaver (2011: 351) with respect to stimuli descriptions with (i) prosodically marked preverbal subjects (Aruma compra el periódico), (ii) $p$-movement (Compra el periódico Aruma), (iii) only/sólo-sentences (Sólo Aruma compra el periódico) and (iv) ‘default sentences’ (Aruma compra el periódico, cf. Onea & Beaver, ibid. on the latter category). Given this setup, if e.g. focused subjects in preverbal position are indeed marked for exhaustivity in Spanish, we would expect the results of condition (i) to pattern similarly to the only-sentences, or to trigger at least the ‘yes, but…’-answer signaling pragmatic incompatibility, whereas if preverbal focused subjects are not exhaustive in Spanish, we would expect the results of the very same condition to pattern more similarly to the default sentences.

Nevertheless, despite of the provisional nature of our investigation, the data that resulted from our elicitation experiments may be conceived of as important additional evidence in favor of (i) the possibility of preverbal subjects in contexts of narrow information focus in Spanish and (ii) the marginality of the word order which is presented in the introduction in examples (1b.) and (3c.), and which is traced back to $p$-movement by Zubizarreta (1998, 1999) and others. Furthermore, as already emphasized in a similar context by Gabriel (2007, 2010), our data equally suggest that the sentence final position is nevertheless the preferred position for non-exhaustively focused subject constituents in sentences without nominal direct object in Spanish. Finally, it has become evident that the exact nature of the communicative setting is in general highly important in the realm of elicitation experiments related to information structure.

6. References


Leonetti, Manuel. in press. Information structure and the distribution of Spanish bare plurals. In Johannes Kabatek & Albert Wall (eds.), *New Perspectives on Bare Noun Phrases in Romance*. Amsterdam: John Benjamins.


Rooth, Mats E. 1985. *Association with focus*: University of Massachusetts at Amherst.


## Addresses of the Contributors

<table>
<thead>
<tr>
<th>Martin Becker</th>
<th>Sofiana Chiriacescu</th>
</tr>
</thead>
<tbody>
<tr>
<td>Universität zu Köln</td>
<td>Universität zu Köln</td>
</tr>
<tr>
<td>Romanisches Seminar</td>
<td>Institut für deutsche Sprache und</td>
</tr>
<tr>
<td>Albertus-Magnus-Platz</td>
<td>Literatur I Sprachwissenschaften</td>
</tr>
<tr>
<td>50923 Köln</td>
<td>Albertus-Magnus-Platz</td>
</tr>
<tr>
<td>Germany</td>
<td>50923 Köln</td>
</tr>
<tr>
<td></td>
<td>Germany</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sonia Cyrino</th>
<th>María Teresa Espinal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Department of Linguistics</td>
<td>Centre de Lingüística Teórica</td>
</tr>
<tr>
<td>IEL – Language Studies Institute</td>
<td>Departament de Filología Catalana</td>
</tr>
<tr>
<td>Unicamp – State University of Campinas</td>
<td>Universitat Autònoma de Barcelona</td>
</tr>
<tr>
<td>Rua Sérgio Buarque de Holanda, 571</td>
<td>08193 Bellaterra</td>
</tr>
<tr>
<td>Campinas, SP – 13084-971</td>
<td>Spain</td>
</tr>
<tr>
<td>Brazil</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Susann Fischer</th>
<th>Chiara Gianollo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Universität Hamburg</td>
<td>Universität zu Köln</td>
</tr>
<tr>
<td>Institut für Romanistik</td>
<td>Institut für deutsche Sprache und</td>
</tr>
<tr>
<td>Von-Melle-Park 6</td>
<td>Literatur I Sprachwissenschaften</td>
</tr>
<tr>
<td>20146 Hamburg</td>
<td>Albertus-Magnus-Platz</td>
</tr>
<tr>
<td>Germany</td>
<td>50923 Köln</td>
</tr>
<tr>
<td></td>
<td>Germany</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Klaus von Heusinger</th>
<th>Manuel Leonetti</th>
</tr>
</thead>
<tbody>
<tr>
<td>Universität zu Köln</td>
<td>Catedrático de Lengua Española</td>
</tr>
<tr>
<td>Institut für deutsche Sprache und Literatur I</td>
<td>Departamento de Filología</td>
</tr>
<tr>
<td>Sprachwissenschaften</td>
<td>Facultad de Filosofía y Letras</td>
</tr>
<tr>
<td>Albertus-Magnus-Platz</td>
<td>Universidad de Alcalá</td>
</tr>
<tr>
<td>50923 Köln</td>
<td>c/ Trinidad, 5</td>
</tr>
<tr>
<td>Deutschland</td>
<td>28801 Alcalá de Henares (Madrid)</td>
</tr>
<tr>
<td></td>
<td>Spain</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sarah Schwellenbach</th>
<th>Melanie Uth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Universität zu Köln</td>
<td>Universität zu Köln</td>
</tr>
<tr>
<td>Romanisches Seminar</td>
<td>Romanisches Seminar</td>
</tr>
<tr>
<td>Albertus-Magnus-Platz</td>
<td>Albertus-Magnus-Platz</td>
</tr>
<tr>
<td>50923 Köln</td>
<td>50923 Köln</td>
</tr>
<tr>
<td>Germany</td>
<td>Germany</td>
</tr>
</tbody>
</table>
Die Arbeitspapiere können über Frau Simeoni bezogen werden. Teilweise sind sie auch elektronisch verfügbar und können unter folgender Adresse abgerufen werden:

http://ling.uni-konstanz.de/pages/publ/arbeitspapiere.html

Unter dieser Adresse findet sich auch eine vollständige Liste der seit 1988 erscheinenden Arbeitspapiere. Diejenigen Hefte, die in der unten stehenden Liste mit einem Stern * versehen sind, sind vergriffen.

Die elektronisch verfügbaren Arbeitspapiere können auch über das Konstanzer Online-Publikations-System (KOPS) der Bibliothek der Universität Konstanz unter folgender Adresse bezogen werden:

http://www.ub.uni-konstanz.de/kops/schriftenreihen_ebene2.php?sr_id=1&la=de

75. Susanne Günthner: From Subordination to Coordination? Verb-second position in German causal and concessive constructions. Februar 1996. DM 5,-.
77. Zvi Penner: From Empty to Doubly-Filled Complementizers. A Case Study in the Acquisition of Subordination in Bernese Swiss German. Oktober 1996. DM 10,-.
79.* Klaus von Heusinger und Urs Egli (Hrsg.): Proceedings of the Konstanz Workshop „Reference and Anaphorical Relations“. November 1996. DM 13,-.
83. Björn Wiemer: Displaced speech - systematic account and acquisitional background (illustrated by Polish and German). Oktober 1996/ Februar 1997. DM 7,-.
88.* Christoph Schwarze & Aditi Lahiri: Einführung in die französische Phonologie. April 1998. DM 10,-.

105. Christopher Schwarze: Aspetti semanticci della formazione delle parole, April 2001, DM 2,--.
111. Matthias Weisgerber (ed.): Proceedings of the Conference “SuB7 – Sinn und Bedeutung”; September 2003, € 5,--.
113. Marc-Olivier Hinzelin & Georg A. Kaiser: Das neutrale Pronomen ello im dominikanischen Spanisch und die Nullsubjekteigenschaft. Anhang: Korpus und Bibliographie zu ello und unpersönlichen Konstruktionen (mit Subjekt) im Spanischen; Oktober 2004, € 1,--.
115. Judith Meinschaefer: The prosodic domain of Italianironcamentois not the clitic group; Juni 2005, € 1,--.
121. Luigia Garrapa: Vowel Elision in two Varieties of Spoken Italian is constrained by Morphology; April 2007.
126. Christoph Schwarz: Les “adjectifs participes”: des lexemes construits en morphologie; Januar 2012, € 1,00.