

Arbeitspapier Nr. 113

**PROCEEDINGS OF THE WORKSHOP
“SEMANTIC AND SYNTACTIC ASPECTS OF
SPECIFICITY IN ROMANCE LANGUAGES”**

**Klaus von Heusinger &
Georg A. Kaiser (eds.)**

Fachbereich Sprachwissenschaft der Universität Konstanz

Arbeitspapier Nr. 113

**PROCEEDINGS OF THE WORKSHOP “SEMANTIC AND
SYNTACTIC ASPECTS OF SPECIFICITY IN ROMANCE
LANGUAGES”**

Klaus von Heusinger & Georg A. Kaiser

Fachbereich Sprachwissenschaft
Universität Konstanz
D-78457 Konstanz
Germany

Konstanz
April 2003

Schutzgebühr EUR 3,50

Fachbereich Sprachwissenschaft der Universität Konstanz
Sekretariat des Fachbereichs Sprachwissenschaft, Frau Ursula Haase,
D 185, D- 78457 Konstanz, Tel. 07531/ 88-2465

Participants of the Workshop

“SEMANTIC AND SYNTACTIC ASPECTS OF SPECIFICITY IN ROMANCE LANGUAGES”



von Heusinger Leonetti Kaiser Jacob Reich Stark
Gutiérrez Bende-Farkas Dirdal Fischer Parodi

Table of contents

Klaus von Heusinger & Georg A. Kaiser

Preface

Susann Fischer

Overridden Specificity 3

Rodrigo Gutiérrez-Bravo

Prominence Scales and Unmarked Word Order in Spanish..... 19

Klaus von Heusinger & Georg A. Kaiser

Animacy, Specificity, and Definiteness in Spanish..... 41

Manuel Leonetti

Specificity and Object Marking: the Case of Spanish a 67

Teresa Parodi

*Clitic Doubling and Clitic-Left Dislocation in Spanish and Greek as Native
and as L2 Grammars*..... 103

Uli Reich

Specifically Brazilian..... 119

Elisabeth Stark

Countability and Specificity in Old Italian..... 133

Preface

The semantic-pragmatic category *specificity* is a very recent category – it was introduced in the late sixties. Since then it has been used in many descriptive grammars for describing a great variety of grammatical phenomena, such as different articles, cases markings (e.g., differentiated object marking or DOM), the modality in relative clauses, word order and information structure, clitic doubling, to list only a few. Thus specificity seems to be a category that is realized in very different grammatical areas. Romance languages provide an interesting field for investigating specificity since they show some common realizations of specificity (such as the interaction of modality in the relative clause with the specificity of the head noun), but they also show language particular grammatical reflexes of specificity, such as Spanish *a* or Romanian *pe*, or clitic doubling in various Romance languages. For more information on specificity and other referential categories see the internet presentation of the *Research Network for Referential Categories in Spanish and other Romance Languages* or *Nereus* in short:

<http://ling.uni-konstanz.de/Nereus/>

The volume is a collection of papers given at the workshop *Semantic and syntactic aspects of specificity in Romance languages* at the Universität Konstanz, 10.-12. October 2002. During this workshop experts in Romance languages discussed different semantic and syntactic aspects of specificity. The presentations and discussions provided us with a more detailed picture of the linguistic realization of specificity in Romance languages, but also about a more detailed linguistic model of specificity in general. Two papers of the workshop are not included in this proceedings: “*Possession*”, *Noun Combining, and Specificity* by Daniel Jacob, and *Specific Indefinites in Romanian and the Marker PE* by Agnes Bende-Farkas. The workshop is also documented on the following web page:

<http://ling.uni-konstanz.de/pages/home/heusinger/konf-proj/02SpecRom/index-e.html>

We would like to thank all contributors and participants for their interest and committed engagement in the workshop. The quality of their contributions, and willingness to share ideas turned the workshop into a lively and inspiring event. By agreeing to put their contributions into writing, they moreover gave us the opportunity to present the workshop topics to a wider audience. We would like to thank the Center for Junior Research Fellows of the University of Konstanz for funding the workshop and the Universitätsgesellschaft Konstanz for financial support for the extra-academic program. Special thanks go to Katharina Ruchti for preparing the manuscript and to Christian Gassner for providing the web-presentation.

Konstanz, April 2003

Klaus von Heusinger
Georg A. Kaiser

Overridden Specificity

Susann Fischer

sfisher@ifla.uni-stuttgart.de

1. Introduction

As is very well known, Modern Romance clitics notoriously fill syntactic positions which are not available to other DPs and vice versa. This is illustrated in (1) with examples from Modern Catalan. (1c) shows that displacement of a full DP to a position between the subject and the finite verb is impossible, whereas it is allowed for clitics (1b). On the other hand, clitics are not allowed to stay within VP (1d).

- (1)
- a. La Mercè veu en Joan.
the Mercè sees the Joan
'Merçe sees Joan'
 - b. La Mercè **el** veu
the Mercè **him** sees
'Mercè sees him'
 - c. *La Mercè en Joan veu
 - d. *La Mercè veu **lo**

It has often been argued that what is syntactically encoded in these cliticization structures is the "familiar" interpretation of these constituents whose referent has to be identified with entities which have already been introduced in the discourse or are physically salient. In the literature, clitic movement has thus been related to "specificity effects" (cf. Sportiche 1996, Delfitto & Corver 1998). Here specificity is clearly related to the definition of "definiteness" in terms of the identity relation, found in Enç (1991), and also clearly overlaps with Diesing's (1992) claim, according to which familiar constituents have to be mapped outside VP.

In Old Romance, however, the pattern is not as clear-cut. Clitics are allowed to precede and follow the verb in matrix as well as in embedded sentences. This is illustrated in (2) with examples from Old Catalan. In (2a) and (2b) the clitic precedes the verb whereas in (2c), (2d) and (2e) the clitic follows the verb¹.

I would like to thank the audience of the workshop "Semantic and syntactic aspects of specificity in Romance languages" for inspiring comments and wonderful discussions throughout the workshop. This research was funded by the DFG grant DO544/1-1 "Non-structural Case".

¹ The texts of which data is presented here are: 13a) Lull, Ramon. Date of composition around 1288. *Llibre de Meravelles*. 13b) Desclot, Bernard. Date of composition between 1283 and 1288. *Crònica de Bernat Desclot*. 14a) Eiximenis, Francesc. Date of composition between 1373 and 1386. *Contes i Faules*. 14b) Metge, Bernard. Date of composition 1399. *Lo Somni*. The references follow the pattern used in my data-base (cf. Fischer 2002).

- (2) a. E lo cavaller **li** respòs e dix que ...
and the cavalier **him** answered and said that
'And the cavalier answered him and said that ...'
(13a/41)
- b. E ab la doctrina qui **li** donà son pare,
and with the catechism that **him** gave his father
'And with the catechism that his father gave to him,'
(13a/4)
- c. e donà-**la** per muler a l' emperador de Castela;
and gave.3sg-**her** as wife to the' emperor of Castille
'and he gave her in marriage to the emperor of Castille;'
(13b/1)
- d. ab los crestians que volien-**los** sobtar,
with the Christians that wanted.3pl-them surprise
'against the Christians who they wanted to surprise.'
(13b/122)

The following questions arise: How are post-verbal clitics in Old Catalan to be interpreted with respect to specificity? Has there been a change in how specificity is encoded in Old Catalan as opposed to Modern Catalan? Can we relate the change in clitic placement from Old to Modern Catalan to other changes in Catalan syntax?

Rather than a complete analysis, this paper marks a starting point of a project on word-order and textual information organization. I will discuss different ideas and will point out some connections that I think are of importance when looking at clitic placement and specificity. I suggest that there is a difference with respect to specificity between Old and Modern Catalan that might be connected to the development of the determiner system from Proto-Romance to Modern Romance. As long as the determiner system has not been fully grammaticalised the specificity of the clitics could be overridden by the clitics' being in the scope of an operator that changed sentence interpretation. More generally, I will argue that in Old Catalan word order contributes much to textual information organisation (cf. Fischer 2002), whereas in Modern Catalan a lot of the discourse organisation is done by the use of definite vs. indefinite determiners. As long as the determiner system was not fully developed to mark topic, foregrounding, backgrounding, focus, specificity etc. other means were needed in order to structure text discourse.

The paper is structured as follows: In Section 2 I give a brief overview of the clitic distribution in the Old Romance languages in general and discuss the so far proposed analyses and their problems. Section 3 shows that the analysis proposed by Fischer (2002) has advantages compared to the existing ones, one of which is the build in possibility to connect the difference in Old Romance clitic placement to the emergence of the d-system which will be looked at in section 4.

2 Old Romance clitic distribution in finite sentences

2.1 The Data

In all Old Romance languages clitics are allowed to precede and follow the verb. The following examples give an overview of the Old Romance languages. The (a) examples show clitics in a preverbal position, whereas the (b) examples show clitics in a postverbal position².

Old Spanish

- (4) a. Esto-**t** lidiare aqui antel Rey don alfonsso
 this-**you** dispute.1sg here before the King don alfonsso
 (Mio Cid, 3344)
- b. & fizo-**lo** traer preso
 and made **him** bring prisoner
 (Alfonso el Sabio-I.126r)

Old Portuguese

- (5) a. e pedio a Nosso Ssenhor quel**he** desse fruyto
 and asked to Our Senhor that-**him** let please
 (Vida de Eufrosina.:357)
- b. perguntou-**lhe** e disse-**lhe**....
 asked-**him** and said-**him** ...
 (Vida de Eufrosina.:359)

Old Italian

- (6) a. prevennero a una torre la quale **si**-chiamava la Torre de la Donzella
 came.3.pl to a tower the which **ref**-called.3pl the Tower of the Maiden
 (Tristano Riccardiano 560)
- b. E offerse-**gliene** due marchi di guadagno
 and offered.3sg-**him-it** two marks of interest
 (Novellino.26)

Old French

- (7) a. É Deus **m'en** deliurt de tutes anguisses.
 and God **me**'of-this liberates of all my fears
 (Li quatre livres des reis:54,XXVI,24,)
- b. au vallet et demande **li**
 in the valley? and ask **him**
 (Chrestien de Troyes:1454)

Old Romance clitics are traditionally seen to be related to the second position phenomena, i.e., they are assumed to obey the Wackernagel law (WL). Wackernagel's law refers to the generalization proposed by the 19th century philologist Jacob Wackernagel (1892) which states:

² Examples and translations of Old Spanish are taken from Fontana (1993: 19ff) and Old Italian from Wanner (1996: 542ff). The examples of Old French and Old Portuguese are taken from Kaiser (1992: 164ff) translations for French and Portuguese are mine and might not be totally correct.

- (8) WL:
 Inherently unstressed/unstressable words
 (a) are preferentially found in second position; and
 (b) specifically, they cannot stand in first position.

The WL has been one of the main subjects in linguistic descriptions in both synchrony and diachrony throughout its existence. In particular the Romance languages have seen numerous appeals to this principle of second position, mainly with regard to unstressed, object clitic pronouns of the medieval Romance languages. This is a curious development, since the predictions of Wackernagel's law are not very faithfully observed in some of the languages for which it was originally proposed.³ However, for the Romance languages it has been argued by different authors that they all developed a class of object pronouns which from proto-Romance until their medieval phases strictly obey the WL in its Romance disguise, i.e., the Tobler-Mussafia law (TM) (Wanner 1987; Rivero 1991; Kaiser 1992; Martins 1995 among many others):

- (9) TM: Unstressed object pronouns cannot stand in absolute initial position in the sentence.⁴

2.2 The different proposals to explain the clitic distribution in Old Romance

The TM transferred to generative theories holds the phonological component responsible for the linear order of the clitic verb sequence: either because syntax acts to satisfy the demands of the phonological component by invoking a last resort movement of the verb in front of the clitic (e.g. Rivero 1991, Cardinaletti & Roberts 1991), or because a phonological filter rules out all sequences where a clitic would end up in a forbidden position (e.g. Fontana 1993, 1997). Still others allow restricted movement operations to take place in the phonological component itself (e.g. Halpern 1995). In all these analyses Old Romance clitics are seen as enclitic elements which are forced to attach prosodically to a word on their left.

Flying in the face of any theory that considers clitics to be elements that due to their very prosody cannot begin a sentence are contexts in which clitics are found in first position: e.g. after a comma and/or after a prosodic break.

³ Especially relevant in this respect is Taylor's (1990) detailed discussion of the exceptions to the second position rule found in the Homeric Greek texts, since this is one of the prototypical languages for which the WL was proposed.

⁴ TM has not been formulated in exactly these concise terms, and in the further evolution of research it has come to cover a wider array of cases. TM here represents the case TM I of Mussafia (1888), i.e., absolute initial position. The above interpretation of TM is the one which is used in all recent analyses for the Romance languages (see Rivero 1986; Wanner 1987; Barbosa 1993 among others).

- (10) l' alberque e la viny que possehia, l' embargave algunes
 the' refuge and the vineyard that possessed.3sg **him'** hindered some
 vegades ...
 times
 'the refuge and the vineyard he possessed, hindered him sometimes ...'
 (13a/105,5)
- (11) a bona dona a son fill ajudar no podia, e, per la gran dolor
 the good women to her son help not could.3sg and for the great pain
 que havia, se meravellave com no moria.
 that had.3sg **ref** wondered how not died.3sg
 'The good woman could not help her son, and, she wondered why she didn't die
 of all the pain that she felt.'
 (13a/113,11)

In these sentences, there is only a prosodic break to the clitics' left to which it could attach, which casts at least some doubts on the claim that Old Romance object pronouns are enclitic elements that need a phonological host to their left.

Second, we run into great problems with all sentences in which clitics are positioned postverbally, even though the verb is preceded by a constituent, recall (2d) and consider (12) and (13).

- (12) E al maytí levaren-se ...
 and at the morning get-up.3pl-**ref**
 'And in the morning they got up ...'
 (13b/25,8)
- (13) E Déus donava-**ls** manera per què la caritat e devociò multiplicava
 and God gave-**them** way for that the charity and devotion multiplied
 'And God showed them how to multiply charity and devotion'
 (13a/188)

In these sentences there is no prosodic requirement for the verb to move in front of the clitic, since the clitic could already have attached to the adverb (12) or the subject in (13).

And, what about the sentences below: here we clearly see that the clitic leans to the right (14), (15), and to the left (16), (17).

- (14) Sènyor, dix lo fill jo.**m** meravell ...
 Lord said.3sg the boy **I.me** surprise.1sg
 'Lord, said the boy, I am surprised ...'
 (13a/64,24)
- (15) El rey se perdé, que no.l trobà hom, ne viu ne mort.
 the king ref. lost.3sg that not.**him** found.3sg man not alive not dead
 'The king wondered that he didn't find anyone neither alive nor dead.'
 (13b/14,1)

- (16) Lo compayó que yo avia amenat ab mi, **m'à** falit ...
 the companion that I have brought.1sg with me **me'**has failed.3sg ...
 The companion that I have brought with me has disappointed me;”
 (13b/54,22)
- (17) ... al qual dix que mal havie fet com l'avia gitat ... ,
 to.the whom said.3sg that bad had.3sg made how **it'**had.3sg turned
 ‘... to whom he said that it was wrong how he has changed ...’
 (13a/88,10)

From the discussion above we can conclude that Old Catalan clitics do not seem to be exclusively enclitic elements. If we consider sentences (14) to (17), it looks much more as if they are enclitic and proclitic depending on where a vowel is found.

Whereas the “prosodic approaches” all hold the enclitic character responsible for the clitic distribution in Old Romance, the syntactic approaches take different features into account, but still seek to explain why the object clitics never appear in initial position. Originally, it was the interpolation data that led a number of researchers to see Old Romance clitics as second position clitics in that these weak elements appear rather separated from the verb than not placed in second position. Consider the following examples from Old Spanish (18) and Old French (19).

- (18) assi como **les** dios auie prometido⁵
 so how **them** God had promised
 (Alfonso el Sabio-I.60)
- (19) Ke il **te** plus face aorer ...
 that he **you** more make.3sg adore.....
 (La vie de Saint Eustache. Paris: 1928)

Newer analyses see interpolation as being optional, but the variation of the clitic to either precede or follow the verb was often connected to the possibility of the clitic being separated from the verb, i.e., insofar as both characteristics are argued to depend on their categorial/lexical status as being phrases. The main reason to analyze medieval clitics as phrases was that they can be separated from the verb by the negation (e.g. Lema & Rivero 1989, Maaßen 1994). Under this assumption clitic movement is seen as phrasal movement to a scrambling position in front of the subject, and thus optional. However, Old Catalan contradicts these claims. Old Catalan complement clitics are never attested in a position other than adjacent to the verb, i.e., they are never separated from the verb by an adverb and they never appear to the left of the negation. Still, they precede and follow the finite verb in main as well as in embedded clauses (Fischer 2002).

Another view proposes that the Old Romance languages ought to be analyzed as verb second languages, in which the verb moves to C° in all matrix sentences (e.g. Benincà 1995, Manzini 1994). In order to explain the verb-clitic sequence in initial position, it is suggested that in those matrix-sentences in which SpecCP is empty, the verb moves on its own from I°

⁵ The Old Spanish example is taken from Fontana (1993) and the example for Old French is taken from (Kaiser 1992).

to C°, leaving the clitic behind. The evidence that the verb did move to C°, and not to another functional category, has always been seen in that postverbal clitics have not been attested in subordinate sentences. However, Old Catalan displays postverbal pronouns in subordinate sentences throughout the 13th and the 14th century.

- (20) E diu que lo primer respòs-**li** hòrreament e ab males paraules ..
 and said.3sg that the first answered-**him** horrified and with bad words
 ‘And he said that the first answered him horrified and with swear words ...’
 (14b/65,13)

Sentence (20) could be argued to display a bridge verb, as has been suggested by Adams (1987) for Old French. She convincingly argues that embedded V2 is possible in the complements to bridge verbs, and that the relationship between clauses in sentences of this type is paratactic, in that both clauses are main clauses. The class of bridge verbs of Old French is comparable to the class of those that in V2 Germanic languages typically allow complements with matrix properties, and like in German, in medieval French the complementizer is not present in these constructions, suggesting that these are cases of German-style embedded V2. In Old Catalan the complementizer is always present which seems to strengthen the fact of subordination, but for the sake of argumentation, I will assume the relationship between the two as being one of parataxis. However, consider sentence (21):

- (21) lo dit bon hom hac totes les vestedures pobres e mesquines que
 the said good man had.3sg all the clothing poor and shabby that
 la dita infantà portà-**li**.
 the said Infant carried.3sg **him**
 ‘that good man had all the poor and shabby clothing that the Infant had given to
 him.’
 (14b/34,19)

This sentence straightforwardly reveals that it would not be the right analysis to assume that the verb adjoins to C°. Even if the relative pronoun were in SpecCP, the verb-clitic sequence is separated from the relative pronoun by a subject, which needs to be in a Spec position below C°. Thus, I conclude that it is the correct assumption, for which I have argued also elsewhere (Fischer 2002), that there needs to be a further category between CP and TP.

3. Explaining Old Catalan data

3.1 A further proposal

The analysis to clitic placement that I proposed in order to explain Old Catalan clitic placement (Fischer 2002) is based on Chomsky’s (1993, 1995) Minimalist Program. In this framework, movement is triggered solely by the need to check the formal features associated with functional categories against the features of morphologically complex lexical units which raise to them. Feature checking is more or less reduced to feature deletion, whereby a “checked feature is marked invisible at the interface” (Chomsky 1995: 229). In order for a

derivation to converge, it must meet the condition of Full Interpretation. This principle states that no uninterpretable element can remain at the point where the derivation enters the semantic component. Lexical entries will contain two types of features: those relevant for PF and those relevant for LF. The types of features relevant only for LF, i.e., accessible to the computation, include formal features such as $[\pm N]$ and $[\pm V]$. Within this framework, certain formal features are strong, and force overt movement in spite of the principle of Procrastinate (overt movement being more costly/less economical than covert movement), while others are weak and therefore movement to check these features is procrastinated until LF. For Chomsky strong features are those that a derivation simply “cannot tolerate” and which trigger rules that eliminate them: “[strength] is associated with a pair of operations, one that introduces it into the derivation (actually, a combination of Select and Merge), and second that (quickly) eliminates it,” (Chomsky 1995: 233). Thus, all strong features must be erased by the checking operation in order for the derivation to converge.

In this analysis functional heads can only be adjoined to, not substituted into (although I provisionally maintain that their specifier is a substitution site). This means that clitics are always in an adjoined position, as Kayne (1991) proposes, and they never move to a category F by substitution. I also adopt Kayne’s fundamental insight that cliticization in Romance is always syntactic procliticization to a functional head: right adjunction is not an option made available by Universal Grammar. Finally I take Romance clitics as transitive D heads which subcategorize for either pro or a lexical complement (see, among others, Corver and Delfitto 1999; Rouveret 1999). Thus, Old Catalan object clitics are heads of pronominal DPs, satisfying the selectional requirement of the verb which governs them. Clitic placement can be seen as an instance of A-movement of a maximal projection followed by proper X° -movement (Sportiche 1989; Cardinaletti 1999). This formulation of clitic placement has proved to have many interesting consequences, and finds empirical support in Catalan where accusative clitic pronouns trigger past-participle agreement (22) throughout all centuries.

- (22) ... ab la força de la nació nostra **les** han superades;
 with the force of the nation ours **them** acc has.3pl surpassed.pl
 ‘... with our nations strength they surpassed them.’

(16a/84,15)

Under the assumption that agreement is licensed in a Spec-head configuration, the clitic heading the DP must have first undergone DP-movement into SpecvP, before subsequently moving further on as a head. Kayne (1991) argues that clitics left-adjoin to an Infl-type position. As for the motivation for this movement to an Infl-type position, an idea has been proposed that links clitic movement to the “specific” character of clitics (Sportiche 1996; Uriagereka 1995; Delfitto & Corver 1998; and Martins 1995). This kind of hypothesis has been proposed in different ways – not all of which conform with the Minimalist framework – depending on whether movement induced by specificity is considered to have a semantic or a morphological trigger and on what the relevant structural position with respect to specificity is (that being the position where clitics are hosted) – “Acc”, “Dat” in Sportiche (1996), “F” in Uriagereka (1995), and “ ” in Delfitto & Corver (1998). Elaborating on Diesing’s (1992) mapping hypothesis it has been argued that TP should be taken as the LF border separating

non-specific elements inside TP from specific elements scoped out of TP (Jonas and Bobaljik 1993; Martins 1995, and all the references therein). The specific character of clitics leads them to incorporate into AgrS, the functional head immediately outside the TP border. As for the assumption that TP is the border for specificity, the authors propose that it can be deduced if it is admitted that all predicates have a Davidsonian event argument; and that this is also true of nouns, even non-eventive ones (Higginbotham 1987). What specificity is is just an element taking wide scope with respect to the Davidsonian event operator in the sentence; in other terms, a DP is specific if and only if its event variable is not bound by the event operator, being instead bound by the discourse.

In the light of the above, I take clitics to be deficient elements that need to adjoin to a functional category outside VP (see also Rouveret 2000), namely T° . In doing so I conform with Diesing's (1992) claim that familiar constituents have to be mapped outside VP. To understand the full scenario: I assume that clitic movement involves two steps: the clitic moves as a DP up to SpecvP, from there the clitic moves on as a D head and – respecting shortest move – adjoins to the verb that is already situated in T° , thus becoming a proclitic. In this analysis the clitic has a fixed place in syntax and it is verb movement that influences clitic placement.

The question to ask in this context is: where does the verb move to produce the verb-clitic order? As the discussion in section 2 revealed, the verb cannot have moved to C° in all of these sentences. In my corpus of Old Catalan many sentences are attested that display a verb-clitic sequence even though this sequence is preceded by different constituents.⁶ These sentences do not fit in any of the prevailing theories. My data suggest that a further functional category above TP and below CP is needed, and since no clitic is ever preceded by the verb in a negated sentence, I make use of the old assumption that affirmation and negation are generated in the same position in the phrase-structure (Chomsky 1957). Building on Laka (1990)⁷, I take this position to be the functional category \bar{P} that hosts different sentence operators: negation, “emphatic” and “neutral” affirmation. With respect to Old Catalan clause structure I propose that different realizations for \bar{P} were available depending on what is expressed: negation vs. affirmation or emphasis.

⁶ See the following table for a summary of all findings

Fig. 1 *Clitic Placement in finite sentences*

Century	13th	14th	15th
Verb-clitic	163	169	140
XP-verb- clitic	76	56	20
Clitic-verb	32	41	75
XP-clitic-verb	590	470	546
all findings	861	736	782

⁷ The \bar{P} category has also been proposed for Portuguese by Martins (1995). However, the assumptions and consequences of her analysis are different from mine, since in her analysis the clitic and the verb may move to \bar{P} , depending on whether the verb or the clitic carries the strong feature. For an extensive discussion of the problems that arise with Martins approach see Fischer (2002).

- (23) a. \circ [-V] “neutral” affirmation
 b. \circ [+V] “emphatic” affirmation
 c. \circ [no] negation

Within a Minimalist framework \circ is like T an inflectional head with V-features, either strong or weak, that need to be checked. Thus if the V-feature is strong it must be checked in overt syntax, in the case of a weak V-feature checking needs to be delayed until LF.⁸ Under the proposed analysis the difference between the verb-clitic and clitic-verb sequences is a difference in sentence interpretation. Adopting this analysis we can account for the fact that Catalan clitics need to be analyzed as heads, never appear postverbally in negated sentences, that they also appear in a postverbal position in subordinate sentences, and that they lean to the right and to the left, that they are enclitic and proclitic⁹. The next section discusses in more detail the perceived difference in semantic interpretation between the verb-clitic vs. clitic-verb sequence.

3.2 The upgrade

The difference between the verb-clitic and clitic-verb sequences displays a difference in sentence interpretation. More specifically, I want to argue that in clitic-verb orders the clitics – as discussed before – move to a functional category outside VP, namely TP, because they are specific: bound in discourse, anchored to an expression used before in the text (cf. Heusinger 2002). Clitics cannot be focussed, they are background information, they refer to what has been mentioned before. However, with respect to the verb-clitic sequence, I suggest that the verb moved further on to \circ , deriving the order verb-clitic. In the scope of \circ [+V] the clitic’s specificity gets overridden, since it’s interpretation is now depending on the operator in \circ . The clitic is still part of the old information but the reversed order indicates that something unexpected or outstanding happens something that needs to be emphasized with respect to the development of the text, i.e., the weighting of textual information is different in verb-clitic sequences compared to clitic-verb sequences. Consider examples (24) and (25).

⁸ The analysis is cast within the framework of Chomsky (1993), where a distinction was made between strong and weak features, and nominal and verbal features both triggering displacement of heads and XPs. The analysis could be embedded within the framework of Chomsky (1998, 1999) along the following lines. \circ is like I \circ an inflectional head with a generalized EPP feature, which either triggers overt displacement of XPs or is satisfied via Agree. Unlike Chomsky (1998), however, I could assume that such features are not necessarily satisfied by Merge/Move XP. Merge/Move X \circ can also check them (see Alexiadou and Anagnostopoulou 1998 for a detailed argumentation). Additionally, I would like to mention that even though covert movement was not one of the options of Chomsky (1998, 2001), a lot of people argued in favor of covert movement (cf. Wilder 1997; Pesetsky 2000). Chomsky himself in “Beyond Explanatory Adequacy” (2001), “assumes this (covert movement S.F) to be correct, contrary to MI and DbP (2001: 32).

⁹ For an extensive discussion of the phonological part of this analysis which is cast within Distributed Morphology please refer to Fischer (2002).

- (24) I. jorn sdevench-se que I. juheu vench a aquell sant hom,
 one day happened.3sg-ref that a Jew came to that holy man,
 ‘And it came to pass that a Jew came to that holy man,’
 (Llull/78)
- (25) I. jorn se sdevench que lo ermità splugave son cilici,
 one day ref happened.3sg that the hermit put-on.3sg his penitential robe,
 ‘And it came to pass that the hermit put on his penitential robe,’
 (Llull/132)

The examples in (24) and (25) not only display the same structure in the matrix sentence but also the identical words. So, there is no syntactical or phonological difference which could motivate the verb once to precede and once to follow the clitic. Only if we look at the context in which the sentences appear in the text, the difference becomes obvious. The first sentence is part of a description of a man living as a hermit. This holy man lived on his own and avoided contact with people because he wanted to contemplate and understand God’s “Creation” within himself. But then, one day, a Jew came to the holy man in order to discuss and understand the trinity of God. The verb-clitic sequence here marks that the visit is something which stood out in the hermit’s daily routine, something which needs to be emphasized, because a hermit usually does not receive visitors. Sentence (25) in contrast describes an event that takes place every single day: the hermit puts on his robe like he does every morning. Additionally, this analysis provides a good explanation of why in the whole of my corpus no verb-clitic sequences are attested in negated sentences. Negative sentences are different from emphatical clauses, in that *se* is lexicalized, therefore *se* does not attract the verb to check off any strong feature, since the feature has already been checked off via merge¹⁰.

Taking this analysis as a working hypothesis, the next step is to discuss why word-order should contribute more to textual information structure in Old Catalan compared to Modern Catalan, and why we don’t find any postverbal clitics in Modern Catalan anymore.

¹⁰ As has been shown in Laka (1990), replies to yes/no questions crucially involve the *V* projection. We know that Latin did not have a word for “yes” (cf. Pinkster 1988 and all references in there). Instead expressions like *ita est, sane, vero* etc. are used next to the bare inflected verb. The minimal affirmative answer in Latin involves the bare inflected verb (ia), the negative answer involves the affix *non* together with the bare inflected verb (ib).

- | | | | |
|-----|---------------------------|----------|-------------|
| (i) | Legistine librum? | a. legi | b. non legi |
| | read.2sg book | read.1sg | no read.1sg |
| | “Have you read the book?” | “yes” | “no” |

The same effect is still observed in European Portuguese (see also Martins 1995a).

- | | | | |
|------|---|---------|--------|
| (ii) | Comeste o pãozinho que te trouxe? | a. comi | b. não |
| | eat.2sg the sandwich that you brought.1sg | eat.1sg | no |
| | “Did you eat the sandwich I brought you?” | “yes” | “no” |

European Portuguese obviously shows a conservative feature which is part of Latin, and which is also attested in Old and Modern Provençal (Jensen 1994). I assume that in the medieval period this feature was shared by more Romance languages. In my corpus I did not find minimal answers to yes/no questions, but as Catalan developed directly from Latin and was in constant contact with Provençal, it thus seems plausible that Old Catalan behaved similar to Latin in that the affirmative answer involved the bare inflected verb, i.e., the strong V-feature on *_* was checked by the verb.

4. Language change and the emergence of the d-system

The study of the emergence and development of the article-system is especially interesting in a language type with no articles, such as Classical Latin or Proto Germanic, to a language type with (fully) developed articles such as the Modern Romance and most Modern Germanic languages.

As for Romance, it is generally hold in traditional diachronic linguistics that the definite article develops from a demonstrative. Additionally, Vincent argues "... pretty uncontroversially it must be said, that the developments in Romance of the two categories of article and clitic pronoun, neither of which were attested in Latin, are linked ..." (Vincent 1997: 149). I would like to suggest to go one step further in proposing that the development of the article system is not only linked to the clitic pronoun system, but that only as long as the article system was not fully grammaticalised in order to mark topic, focus etc. postverbal clitics were available, i.e. word-order contributed much more to information organization in Old Romance than in Modern Romance.

However, as this paper only marks the starting point of a project I will mainly point out some connections that I think are of importance when looking at the emergence of the d-system in Romance and of information organization in texts.

Looking at the change from Old to Modern Catalan (and the other Romance languages) there are two changes that are eye-catching, and in my opinion indicate that my suggestion could be on the right track.

Going back in history and looking at Classical Latin we see that Latin didn't have articles in the modern sense, but only deictic/and or emphatic pronouns that could in different ways fulfil the function of giving overt expression to a textually or contextually highlighted item.

- (29) (a) *hic* this, or more precisely a marker of proximal deixis, i.e. "near the speaker"
 (b) *iste* a deictic associated with the second person, i.e. near the addressee
 (c) *ille* a marker of distal deixis, i.e. "distant from both speaker and the addressee"
 (d) *ipse* which marks emphasis or contrast

(s. Vincent 1997, Harris 1980)

Furthermore, in Classical Latin there are two classes of pronouns, corresponding to the distinction made in the traditional literature between weak and strong pronouns, but no clitics in the sense of Kayne (1975). In the development from Latin to Proto-Romance these deictic/emphatic pronouns started to disappear, and a class of items, etymologically derived from various forms of *ille* emerged that obey the diagnostics for cliticness identified in Kayne (1975). Whereas *ille* (along with the first and second person pronouns) is the only appropriate candidate for the change from weak pronoun to clitic, it is in competition with *ipse* as part of an emergent system of markers of the discourse status of arguments. Constituents in the preverbal field are marked by either *ille* or *ipse*, which attain the status 'articloid' during this period (cf. Vincent 1997: 165). He further argues that the next step consisted of two possibilities which both lead to the full article status for the reflexes of

ipse/illo. Either *illic* generalizes at the expense of *ipse*, or *ipse* generalizes at the expense of *illic* (cf. Vincent 1997: 166).

The early period of Romance literature is well documented in the studies by Selig (1992) and Stark (2001) and they both show that during this early period the marking of important referents was still facultative, i.e. nominal phrases could appear as bare singulars or bare plurals. Thus, there is a stage in the development from Latin to Romance where the clitic system is fully developed whereas the article system has not yet fully grammaticalized.

In the light of this I suggest that it is during this period that word-order took on some of the functions of the emphatic/deictic pronouns and which in the further development of the language was undertaken by the article system as soon as this was fully grammaticalized. I will now shortly compare the Old Romance languages with the Old Germanic languages in order to enforce the hypothesis proposed here.

Comparing the Old Germanic languages to the Old Romance languages a lot of similarities can be found two of which are important for the hypothesis proposed here: languages like Old High German, Old Norse, Old English and Middle English do not have a fully developed article system (cf. Philippi 1997, Leiss 2000), and exactly at the same time they display a reversed word-order in Stylistic Fronting constructions (Seefranz-Montag 1983, Fischer & Alexiadou 2001, Trips 2002, Fischer in press). Additionally, we find the same construction in Modern Icelandic and Faroese. Both display Stylistic Fronting and they do not possess an indefinite article. The examples below illustrate clear instances of Stylistic Fronting. The (a) examples in (30) and (31) represent canonical word-order, while those in (b) illustrate the order after SF has applied.

Icelandic

- (30) a. [Sá sem er fyrstur að skora mark] fær sérstökverðlaum
 he that is first to score goal gets special price
 ‘The first one to score a goal gets a special price’
- b. [Sá sem **fyrstur** er ___ að skora mark] fær sérstök verðlaum
 he that **first** is ___ to score goal gets special price
 (Jónsson 1991)

Old Catalan

- (31) a. Longament considerà lo hermità en la demanda que li hac
 long considered the hermit in the question that him had.3sg
 feta Felix.
 made Felix
 ‘For a long time the hermit considered the question that Felix had asked him.’
- b. com no li responia a la demanda [que **feta** li havie ___]
 how not him answered to the question [that **made** him had.3sg ___]
 (Fischer & Alexiadou 2001)

Old English

- (32) þatt **timmbredd** wass abufenn Godess arrke
 that built was above God's ark
 'that was built above God's ark'

(Trips 2002: 588)

With respect to Old Catalan¹¹ I have argued that Stylistic Fronting contributes to textual information organisation, in the same way as the reversed order in verb-clitic sequences contributes to textual information organisation (Fischer, to appear). It's understood that there is no way to claim that SF in Modern Icelandic contributes to textual information organisation. However, we should nevertheless reconsider the argument proposed here that word-order contributes to a greater extent to textual information organisation in the Old Germanic/Old Romance languages if compared to the Modern Germanic/Modern Romance languages. This seems rather plausible in the light of the fact that with respect to Catalan, postverbal clitics and Stylistic Fronting disappeared at exactly the same time, and the vanishing of these structures happened shortly after the article system is said to have fully been grammaticalized in the Romance languages (Vincent 1997).

5. References

- Adams, M. (1987). Parametric change: empty subjects in Old French. In: David Birdsong, (ed.). *Advances in Romance Linguistic*. Dordrecht: Foris Publications, 1-32.
- Alexiadou, A. and E. Anagnostopoulou. (1998). Parametrizing AGR: word-order, V-movement, EPP checking. *Natural Language and Linguistic Theory* 16: 491-539.
- Barbosa, P. (1993). On clitic placement in Old Romance and European Portuguese. In: K. Beals et.al. (eds.) *CLS 29: Papers from the 29th Regional Meeting of the Chicago Linguistic Society, Vol.1. (The Main Session)*. Chicago: CLS, 33-59.
- Benincà, P. (1995). Complement clitics in Medieval Romance: the Tobler-Mussafia Law. In: A. Battye and I. Roberts, (eds.). *Clause Structure and Language Change*. Oxford, 325-344.
- Cardinaletti, A. and I. Roberts (1991). Clause structure and x-second, MS University di Venezia.
- Chomsky, N. (1957). *Syntactic Structures*. The Hague.
- Chomsky, N. (1993). A minimalist program for linguistic theory. In: Kenneth Hale and Samuel Jay Keyser (eds.). *The View from Building 20*. Cambridge/Mass.: MIT Press, 1-52.
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge/Mass.: MIT Press
- Chomsky, N. (1999). Derivation by phase, In: M. Kenstowicz (ed.). *Ken Hale: A life in language*. Cambridge/Mass.: MIT Press (MIT Occasional Papers in Linguistics no. 18).
- Chomsky, N. (2000). Minimalist inquiries: the framework. In: R. Martin, D. Michaels, and Juan Uriagereka (eds.). *Step by Step*. Cambridge, Mass. MIT Press, 89-155. [also available as MIT Working Papers in Linguistics no.15, (1998), Department of Linguistics and Philosophy, MIT, Cambridge/Mass].
- Chomsky, N. (2001). Beyond explanatory adequacy. Ms. MIT.
- Delfitto, D. and N. Corver (1998). Feature primitives and the syntax of specificity. *Rivista di Linguistica*, 10.2, 281-334.
- Diesing, M. (1992). *Indefinites*. Cambridge/Mass.: MIT Press.
- Enç, M. (1991). The semantics of specificity. *Linguistic Inquiry* 22,1-25
- Fischer, S. (2002). *The Catalan Clitic System: a diachronic perspective on its syntax and phonology*. Berlin, New York: Mouton de Gruyter.

¹¹ Stylistic Fronting has also been shown to exist in Old French (Cardinaletti & Roberts 1991) and Old Spanish (Fontana 1993).

- Fischer, S. (in press). The diachronic relation between quirky subjects and stylistic fronting. In: K.V. Subbarao and Peri Bhaskarao (eds.). *Non-Nominative subjects*. Amsterdam: Benjamins.
- Fischer, S. (to appear). Stylistic fronting: A contribution to information structure. In: Arthur Stepanov, Gisbert Fanselow and Ralf Vogel (eds.): *Minimality Effects in Syntax*. Berlin: Mouton de Gruyter.
- Fischer, S. and A. Alexiadou. (2001). On stylistic fronting: Germanic vs. Romance. *Working Papers of Scandinavian Syntax* 68, 117-145
- Fontana, J. (1993). *Phrase Structure and the Syntax of Clitics in the History of Spanish*. Ph. D dissertation, University of Pennsylvania.
- Fontana, J. (1996). Phonology and syntax in the interpretation of the Tobler-Mussafia law. In: A. L. Halpern & A. M. Zwicky. *Approaching Second: Second Position Clitics and Related Phenomena*. Stanford, 41-84.
- Fontana, J. (1997). On the integration of second position phenomena. In: A. van Kemenade and N. Vincent (eds.). *Parameters of Morphosyntactic Change*, 207-250.
- Halpern, A. (1995). *On the Placement and Morphology of Clitics*. Stanford: Dissertations in linguistics CSLI publications.
- Harris, M. (1980). The marking of definiteness in Romance. In: Jacek Fisiak (ed.). *Historical Morphology*. The Hague: Mouton.
- von Heusinger, K. (2002). Specificity and definiteness in sentence and discourse structure. *Journal of Semantics* 19, 245-274.
- Higginbotham, J. (1987). Indefiniteness and predication. In: Eric Reuland and Alice G.B. ter Meulen (eds.). *The Representations of (In)definiteness*. Cambridge/Mass.: MIT Press, 43-70.
- Jensen, F. (1994). *Syntaxe de l'ancien occitan*. Tübingen: Niemeyer.
- Jonas, D. and J. Bobaljik. (1993). Specs for subjects: the role of TP in Icelandic. *MIT Working Papers in Linguistics* 18. Cambridge/Mass.: MIT Press
- Jónsson, J. G. (1991). Stylistic fronting in Icelandic. *Working Papers in Scandinavian Syntax* 48, 1-43.
- Kaiser, G. (1992). *Die klitischen Personalpronomina im Französischen und Portugiesischen: eine synchronische und diachronische Analyse*. Frankfurt am Main: Vervuert.
- Kayne, S. R. (1975). *French Syntax. The Transformational Cycle*. Cambridge/Mass.: MIT Press.
- Kayne, S. R. (1991). Romance clitics, verb movement, and PRO. *Linguistic Inquiry* 22, 647-686.
- Laka, I. M. (1990). *Negation in Syntax: On the Nature of Functional Categories and Projections*. MIT Working Papers.
- Leiss, E. (2000). *Artikel und Aspekt. Die grammatischen Muster von Definitheit*. Berlin, New York: De Gruyter.
- Lema, J. and M.L. Rivero (1989). Long head movement: ECP vs. HMC. *Cahier Linguistiques D'Ottawa* 18, 61-78.
- Maaßen, C. (1994). "Es war einmal ein Wackernagelpronomem ...". In: Brigitta Haftka (ed.). *Was determiniert Wortstellungsvariation? Studien zu einem Interaktionsfeld von Grammatik, Pragmatik und Sprachtypologie*. Opladen: Westdeutscher Verlag, 345-361.
- Manzini, M. R. (1994). Triggers for verb-second: Germanic and Romance. *The Linguistic Review* 11, 299-314.
- Martins, A.M. (1995). Clitic placement from Old to Modern European Portuguese. In: H. Andersen (ed.). *Historical Linguistics*. Amsterdam, 395-307.
- Pesetzky, D. (2000). *Phrasal Movement and Its Kin*. Cambridge: MIT Press.
- Philippi, J. (1997). The rise of the article in the Germanic languages. In: Ans van Kemenade and Nigel Vincent (eds.). *Parameters of Morphosyntactic Change*. Cambridge: Cambridge University Press, 62-93.
- Pinkster, H. (1988). *Lateinische Syntax und Semantik*. Tübingen: Francke.
- Rivero, M. L. (1986). Parameters in the typology of clitics in Romance and Old Spanish. *Language* 62, 774-809.
- Rivero, M.L. (1991). Clitic and NP climbing in Old Spanish. In: Héctor Campos and Fernando Martínez-Gil (eds.). *Current Studies in Spanish Linguistics*. Washington D.C.: Georgetown University Press, 241-282.
- Rouveret, A. (1999). Clitics, subjects and tense in European Portuguese. In: Henk van Riemsdijk (ed.). *Clitics in the Languages of Europe*. Berlin: Mouton, 639- 677.
- Rouveret, A. (2000). Clitics in romance. Talk presented at Humboldt-University in Berlin.
- Seefranz-Montag, A.(1983). *Syntaktische Funktionen und Wortstellungsveränderung: die Entwicklung „subjektloser“ Konstruktionen in einigen Sprachen*. München: Fink.
- Selig, M. (1992). *Die Entwicklung der Nominaldeterminanten im Spätlatein. Romanischer Sprachwandel und lateinische Schriftlichkeit*. Tübingen: Narr.
- Sportiche, D. (1996). Clitic constructions. In: Johan Rooryck and Laurie Zaring (eds.). *Phrase Structure and the Lexicon*. Dordrecht: Kluwer. 213 – 276.

- Stark, E. (2001). Indefiniteness and specificity in Old Italian texts. In: Klaus von Heusinger / Kerstin Schwabe (eds.). *Information Structure and the Referential Status of Linguistic Expressions*. ZAS Papers in Linguistics 23, 197-210.
- Tobler, A., (1875/1912). Besprechung von J. Le Coultre, De l'ordre des mots dans Chrétien de Troyes. *Vermischte Beiträge zur französischen Grammatik* 5, 395-414
- Trips, C. (2002). *From OV to VO in Early Middle English*. Phd. Dissertation. University of Stuttgart.
- Vincent, N. (1997). The emergence of the D-system in Romance. In: Ans van Kemenade and Nigel Vincent (eds.). *Parameters of Morphosyntactic Change*. Cambridge: Cambridge University Press, 149-169.
- Uriagereka, J. (1995). Aspects of the syntax of clitic placement in Western Romance. *Linguistic Inquiry*, 26, 79-123.
- Wackernagel, J. (1892). Über ein Gesetz der indogermanischen Wortstellung. *Indogermanische Forschungen* 1, 333-436.
- Wanner, D. (1987). *The Development of Romance Clitic Pronouns*. Berlin: Mouton de Gruyter
- Wanner, D. (1996). Second position clitics in Medieval Romance. In: Aaron L. Halpern and Arnold M. Zwicky (eds.). *Approaching Second: Second Position Clitics and Related Phenomena*. CSLI, 537- 578.
- Wilder, C. (1997). Phrasal movement in LF: re re readings, VP-ellipsis and binding. *NELS* 27, 1-15.

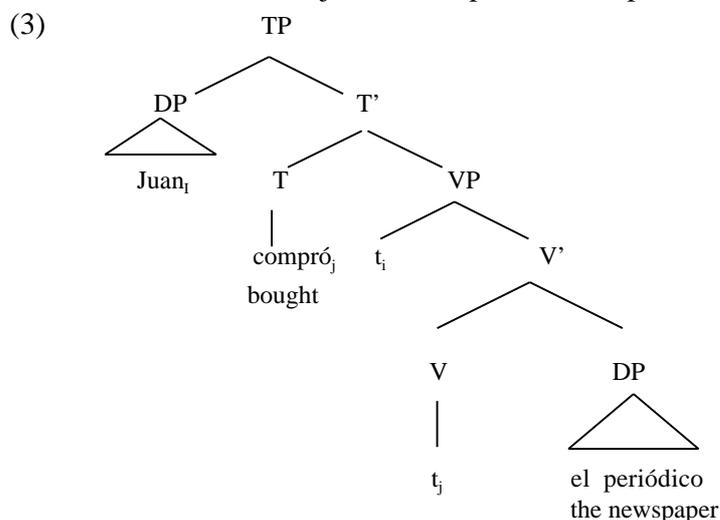
- (2) a. A Juan le gustan los chocolates.¹ **IOVS**
 to Juan CL-DAT like.3PL the chocolates
 'Juan likes chocolates.'
- b. Llegó tu hermano. **VS**
 arrivedyour brother
 'Your brother arrived.'

These data indicate that the word order of (1) is not the result of the subject moving to the preverbal position because of Case and/or agreement requirements of the subject (*contra* Costa 1998; 2001, and Fernández-Soriano 1999, for instance), since the subject does not occupy the preverbal position in these cases. They also point to the conclusion that the SVO order of (1) is not the result of the subject being a default topic, since again the subject does not occupy the preverbal topic position in the examples in (2).²

The proposal I will develop instead claims that clauses with different kinds of constituents in the preverbal position have different degrees of markedness depending on the semantic role of the preverbal constituent. I propose that this relative degree of markedness, in conjunction with the EPP, determines when and whether the preverbal position is occupied in the unmarked case.

1. The EPP and unmarked word order

First I will lay out my assumptions regarding the structure of SVO sentences in Spanish, which are fairly standard. Following Suñer (1994), I assume that the verb moves from V to T and movement of the subject from [Spec, V] to [Spec, T] in turn results in the order of (1).



¹ The abbreviations in the handout are as follows:

ACC	accusative	DAT	dative	PL	plural
Ag	Agent	EX	Experiencer	s	singular
CL	clitic	exp	expletive	Th	Theme

² See Gutiérrez-Bravo (2002) for other examples.

I also assume, following Suñer (1988) and Parodi (2002), that clitics in Spanish, including dative clitics such as the one in example (2a), are instances of morphological agreement. What needs to be determined first, though, is the factor responsible for the movement of the subject to [Spec, T]. Having seen that there is reason to think that this factor is not related to Case or agreement considerations, it is illustrating to bring into the discussion the analysis in Alexiadou & Anagnostopoulou (1998). These authors suggest that in null subject languages that show an SVO/VSO alternation (such as Greek and some varieties of Spanish) the EPP is satisfied via verb raising because in these languages verbal agreement morphology has the categorial status of a pronominal element. Given this, whether the word order observed is SVO, as in (4a), or VSO, as in (4b), depends essentially on pragmatic considerations.

- (4) GREEK
- a. O Petros pandrefitke tinIlektra. **SVO**
 Peter married Ilektra
- b. pandrefitke o Petros tinIlektra. **VSO**
 married Peter Ilektra
- (Alexiadou & Anagnostopoulou 1998:492)

Mexican Spanish (and other varieties of Spanish; see for instance Zagona 2002), provide strong evidence against this proposal, though. In this variety, transitive clauses where the preverbal position is empty are extremely marked, and downright ungrammatical in some cases. Compare (5) with (1).

- (5) MEXICAN SPANISH
- ??Compró Juan el periódico. **VSO**
 bought Juan the newspaper
 ‘Juan bought the newspaper.’

However, a crucial observation at this point is that VSO sentences can be “rescued” when a topic or a *wh*-operator occupies the same position as the subject in (1). This is shown in the examples in (6).

- (6) a. Ayer compró Juan el periódico. **Adv VSO**
 yesterday bought Juan the newspaper
 ‘Yesterday Juan bought the newspaper.’
- b. Por qué compró Juan el periódico? **Wh VSO**
 why bought Juan the newspaper
 ‘Whydid Juan buy the newspaper?’

My proposal is that, contrary to Alexiadou & Anagnostopoulou's analysis, the EPP is an active requirement in Spanish, and it is the factor responsible for the SVO order of (1).³ Furthermore, I propose that topics and *wh*-operators can "rescue" VSO sentences in Mexican Spanish because, just like the subject, they can also satisfy the EPP. In this I follow many contemporary definitions of the EPP that propose that it can be satisfied by constituents other than the grammatical subject (Collins 1997, Grimshaw 1997, Fernández-Soriano 1999, Chomsky 2000, *inter alia*). The similar behavior of (1) and (6) in contrast with (5) is thus explained under the assumption that fronted topics and *wh*-operators in Spanish also have [Spec, T] as their landing site (the Generalized TP analysis of Zubizarreta 1998: see Gutiérrez-Bravo 2000, 2002, for evidence).

The definitions of the EPP cited above allow for a definition of this requirement that can be independent of Case and agreement considerations (because it is not defined with respect to the subject grammatical relation). Accordingly, adopting the EPP in my analysis is consistent with our previous observation that Case and agreement do not appear to play a role in deriving the unmarked word order observed in Spanish. Specifically, from here on I adopt the definition of the EPP in (7).

(7) EPP

The specifier of the highest inflectional projection must be filled.

(Grimshaw 1997, Gutiérrez-Bravo 2000, 2002).

Observe now that the same pattern is observed with psych clauses in Mexican Spanish. Again, this suggests that the indirect object in (8a) satisfies the EPP in [Spec, T] in the same way that the subject does in (1) (see also Masullo 1993).

(8) a. A Juan le gustan los chocolates. **IOVS**
to Juan CL-DAT like.3PL the chocolates
'Juan likes chocolates.'

b. ??Le gustan a Juan los chocolates. **VIOS**
CL-DAT like.3PL to Juan the chocolates

At this point we can lay out an interim summary of the conclusions we can draw from the data above. First, the EPP seems to be an active requirement in at least some varieties of Spanish. Secondly, the EPP in Spanish can be satisfied by constituents other than the subject, and thirdly, the EPP is responsible in part for the word order observed in the unmarked case. However, a number of important issues remain. We need to explain why the EPP is satisfied by the indirect object and not by the subject in psych clauses like (8a). Perhaps more importantly, we also need to explain why the preverbal position is left empty in the case of unaccusative clauses (example 2b), in violation of the EPP. In the following sections, I develop an OT analysis that addresses these issues.

³ See Gutiérrez-Bravo (2000, 2002), and Goodall (2001) for further evidence that the EPP is an active requirement in Spanish.

2. Prominence scales and unmarked word order

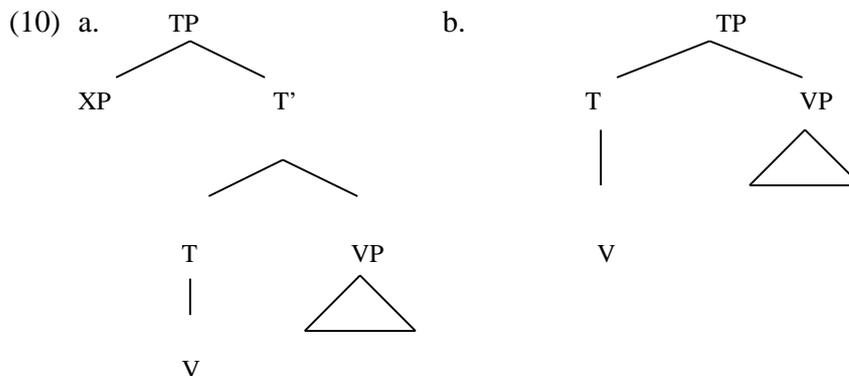
As a first step in the analysis, in order to disassociate the constituent that satisfies the EPP from any specific grammatical relation, I propose that it be referred to as the **Pole** of the clause.⁴ The Pole thus corresponds to the specifier of the highest inflectional head, independently of the function or grammatical relation of this specifier. There are many cases where the grammatical subject and the Pole of the clause overlap, as in SVO transitive sentences, but even in a language like English, where the correlation between subject and Pole is very strong there are instances where the Pole and the subject are syntactically distinct. For instance, my suggestion is that in English *there*-existential sentences, the expletive, which is the element that satisfies the EPP, corresponds to the Pole, whereas the post-verbal DP *some students* corresponds in turn to the grammatical subject as evidenced by the fact that it is this DP that shows agreement with the auxiliary (see Gutiérrez-Bravo 2002 for further discussion).

- (9) There are some students waiting outside the coffee shop.

In turn, in (8a), the indirect object corresponds to the Pole: it is the constituent that occupies [Spec, T] and so it satisfies the EPP, even when it is not the subject.

2.1 Economy of Structure and structural markedness

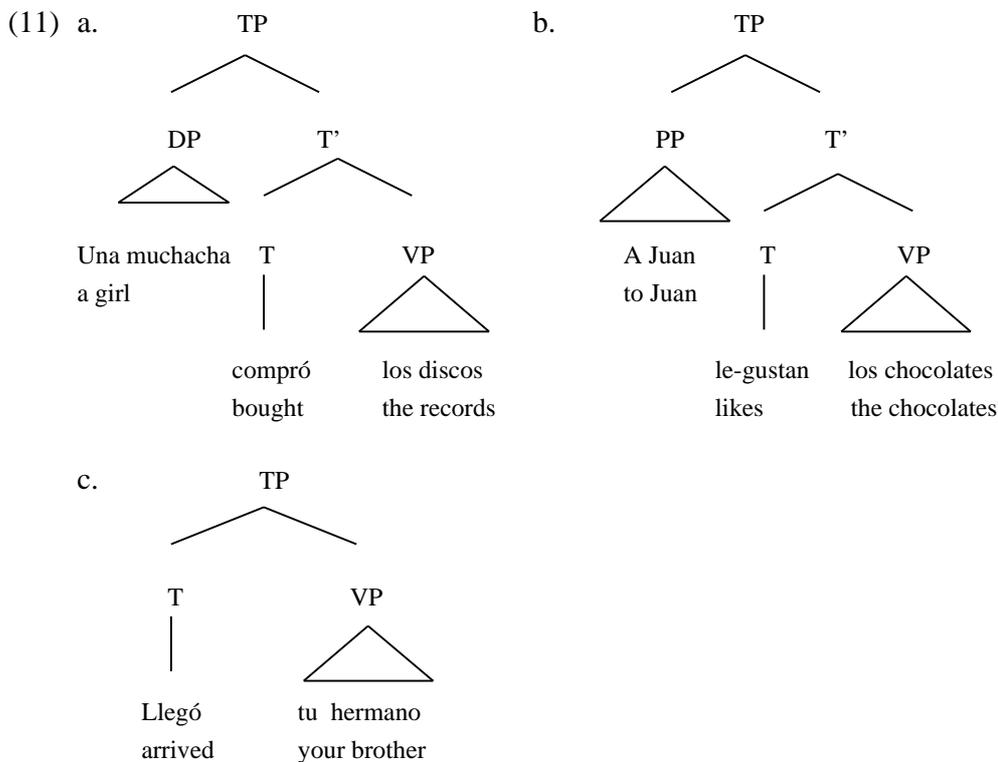
A second observation concerns the status of the Pole of the clause with respect to Economy of Structure (Chomsky 1993, 1995; Grimshaw 1993, 2001). Clauses with a Pole (i.e. (10a)) have an extra layer of structure, and so in terms of Economy of Structure these clauses are more marked than clauses without a Pole. Henceforth I refer to markedness in terms of Economy of Structure as *structural markedness*.



At this point, though, it is crucial to note that the EPP, which requires clauses to have a Pole, is in direct conflict with Economy of Structure. This is because the EPP favors the more marked structure in (10a), whereas Economy of Structure favors the less marked structure in (10b). My interpretation of the different word orders displayed by different classes of verbs in

⁴ In the remainder of this presentation I will concentrate on cases where the Pole corresponds to the constituent in [Spec, T]. Strictly speaking, however, the Pole corresponds to the specifier of the highest inflectional projection, which is mostly, but not always, [Spec, T]. See Gutiérrez-Bravo (2002) for details.

Spanish is that this conflict is resolved in favor of the EPP in transitive and psych clauses (as in ((11a-b)), but in favor of Economy of Structure in unaccusative clauses (as in (11c)).



In order to explain these differences, my proposal is that clauses with different constituents in the Pole position have different degrees of markedness, depending on the semantic role of the XP which functions as the Pole. In other words, clauses with a Pole are more marked than clauses without a Pole, but clauses with a Pole can still display different degrees of markedness when compared to one another. The degree of markedness of clauses with a Pole runs along the scale of semantic roles in (12) (see Larson 1988, Bakovic 1998, *inter alia*). Following this scale, a Pole that bears the *Agent* semantic role constitutes the least marked instance of a Pole, an *experiencer* constitutes the next least marked instance of a Pole, etc., and so on, until we reach XPs with a *reason* semantic role, which constitute the most marked Poles.

- (12) AGENT > EXPERIENCER > THEME > LOCATION > MANNER/TIME
>REASON

My suggestion is that in Spanish clauses where the Pole is an *agent* or an *experiencer* (11a-b) have a degree of structural markedness that is not enough to warrant a violation of the EPP. In other words, in these cases the EPP overrides structural markedness. However, clauses where the Pole is a *theme* (or lower in the scale in (12)) would have a degree of structural markedness high enough that it in this case it is preferable to violate the EPP, as in (11c). In

this case, structural markedness overrides the EPP). In what follows, I develop an Optimality Theoretic account of this state of affairs.

2.2 An OT analysis

Optimality Theory (Prince & Smolensky 1993), where the grammar of a language is a ranking of violable constraints, provides an ideal set of theoretical tools for accounting for the way in which the conflict between the EPP and structural markedness is resolved in Spanish. First, the discussion so far points to the conclusion that the EPP is not an all-or-nothing requirement in Spanish. This receives a straightforward explanation in an OT analysis where the EPP, as defined in (7), corresponds to a violable constraint. Secondly, the mechanisms exist in OT with which we can formalize the special structural status of the Pole and the sensitivity of the Pole position to the semantic role of the XP that occupies it.

With respect to the structural prominence of the Pole position, observe that the Pole is the most prominent constituent of the clause (it asymmetrically c-commands every other constituent in the clause), all else being equal. This can be represented by the structural prominence scale in (13) (Prince & Smolensky 1993).

- (13) Pole > Non-Pole⁵

Next, the relation between the Pole position and the scale of semantic roles in (12) can be formalized by Harmonic Alignment (Prince & Smolensky 1993; Aissen 1999a, 2000) of the structural scale in (13) with the Thematic Hierarchy in (12). Harmonic alignment of these two scales yields the Harmony scales in (14), which in turn translate into the subhierarchies of markedness constraints in (15).⁶

- (14) a. Pole/Agent > Pole/Experience > Pole/Theme > Pole/Location
> Pole/Manner-Time > Pole/Reason.
- b. Non-Pole/Reason > Non-Pole/Manner-Time > Non-Pole/Location
> Non-Pole/Theme > Non-Pole/Experiencer > Non-Pole/Agent
- (15) a. *Pole/Reason >>*Pole/Manner-Time>>*Pole/Location >>*Pole/Theme>>
*Pole/Experiencer >>*Pole/Agent
- b. *Non-Pole/Agent>>*Non-Pole/Experiencer >> *Non-Pole/Theme >>*Non-Pole/Location >>*Non-Pole/Manner-Time >>*Non-Pole/Reason

The harmony scale in (14a) expresses the fact that the least marked instance of a Pole is one filled by the *agent*, whereas a Pole filled by a *reason* adverb corresponds to the most marked

⁵ “Non-Pole” is a term that encapsulates every specifier and complement position in the lexical layer (VP or a VP shell). The constraints in (15b) which target these VP internal positions do not play a role in the analysis of the facts that follow.

⁶ See Prince & Smolensky (1993) for technical details.

instance. Inversely, the harmony scale in (14b) expresses the fact that a VP-internal *agent* corresponds to the most marked instance of a VP-internal argument or adjunct, whereas a *reason* adverb corresponds to the least marked instance of a VP-internal argument or adjunct. As mentioned before, these harmony scales are in turn translated into the hierarchies of markedness constraints in (15). The constraints in the hierarchy in (15a) target the relative markedness of a constituent with the specified semantic role when it appears as the Pole of the clause. In other words, *Pole/Agent is violated when the Pole ([Spec, T]) in what follows) corresponds to an *Agent*, *Pole/Experiencer is violated when the Pole corresponds to an *Experiencer*, etc. The most severe violation is incurred when the Pole is filled by a *reason* expression, whereas the least severe violation is incurred in when an *agent* functions as the Pole. Henceforth I will refer to this hierarchy as the Pole Hierarchy and it is this hierarchy that will crucially be resorted to in order to provide an analysis of unmarked word order in Spanish.⁷

Observe now that the hierarchy of constraints not only captures the intuition that clauses with different kinds of constituents in the Pole have different degrees of markedness, it further does so without making reference to grammatical relations. As previously discussed, this is a fundamental characteristic that must be taken into account in any account of word order in Spanish. Now that we have developed the necessary technical infrastructure, the unmarked word order facts can be accounted by embedding the EPP in the hierarchy in (15a), as in (16).

- (16) *Pole/Reason >> *Pole/Manner-Time >> *Pole/Location >> *Pole/Theme
>>EPP >> *Pole/Experiencer >> *Pole/Agent

We now need just one more constraint to complete the analysis. Spanish is a language that does not allow expletive insertion to satisfy the EPP. This indicates that the FULL-INTERPRETATION constraint in (17), which penalizes the presence of expletives in the output representation, stands at the top of the constraint ranking in (16).

- (17) FULL-INTERPRETATION
Parse lexical conceptual structure. Failed by expletives and auxiliary *do*.
(Grimshaw 1997)

Consider now how the ranking in (16) derives the different word orders in the examples in (1-2). The first case corresponds to an SVO transitive construction.

⁷ In contrast with this, the hierarchy in (15b) (which will not be a part of the analysis that follows) targets the relative markedness of a constituent with the specified semantic role when it surfaces in its VP-internal position. In this case, the most severe violation of the hierarchy in (15b) is incurred when an *agent* is found in a VP-internal position, whereas the least severe violation is incurred in when a *reason* expression occupies a VP-internal position. I also assume that adjoined positions (both in the lexical and in the functional layer of the clause) are never targeted by the constraints in either of these hierarchies, although this assumption plays no role in the analysis that follows.

- (18) Transitive clauses
 Unamuchacha compró los discos **SVO**
 a girl bought the records
 ‘A girl bought the records.’

(19) INPUT: <buy (x, y), x=a girl (Ag), y=the records (Th)>

	FULL-INT	*Pole/Theme	EPP	*Pole/Agent
☞ a. [_{TP} una muchacha compró [_{VP} los discos]] SVO				*
b. [_{TP} compró [_{VP} una muchacha los discos]] VSO			*!	
c. [_{TP} los discos los compró [_{VP} una muchacha]] OVS		*!		
d. [_{TP} <i>exp</i> compró [_{VP} una muchacha los discos]]	*!			

In the analysis of (18) in tableau (19) I assume, following Grimshaw (1997), that the input consists of a predicate-argument structure with no syntactic representation. This input is mapped into a number of candidates that are generated in accordance with X-Bar theory, but which display different word orders.⁸ Consider now the different losing candidates in (19) one by one. The VSO candidate (19b), just like example (5), leaves the Pole position empty, and thus loses to the SVO candidate (19a) because of its violation of the EPP constraint. Candidate (19c) satisfies EPP because the direct object occupies Spec-IP, the Pole position, but by doing so it incurs in a fatal violation of *Pole/Theme, since the semantic role of the constituent functioning as the Pole corresponds in this case to a *theme*. Finally, candidate (19d) avoids a violation of the EPP by insertion of an expletive (*exp*), but loses because of its violation of FULL-INTERPRETATION, as previously discussed.⁹ Candidate (19a), which incurs only in a violation of the low-ranked markedness constraint *Pole/Agent, thus emerges as the winner.

Now in the case of clauses with Psych predicates such as (20), the ranking in (16) ensures that it will not be the subject that will emerge in the preverbal position, but rather the IO, since given the Pole Hierarchy in (15) it is better to have an *experiencer* than a *theme* as the Pole of the clause, independently of the specific grammatical relations of the arguments involved.

⁸ I also assume that arguments outside the lexical projection of their predicate (VP in this case) must form a chain with a trace inside this lexical projection, which results in a violation of the STAY constraint (Grimshaw 1997), which penalizes the presence of traces in the output representation. However, the presence of traces left by the constituent that moves to [Spec, I] and the corresponding violations of STAY do not have any effect on the analysis that follows, and so they will be omitted from the tableaux that follow.

⁹ See also the analysis of Italian subject inversion in Grimshaw & Samek-Lodovici (1998).

- (20) Psych clauses
 A Juan le gustan los chocolates. **IOVS**
 to Juan CL-DAT like-3PLthe chocolates
 ‘Juan likes chocolates.’

The analysis is presented in tableau (21). Just as in the case of (19), the candidate that leaves the Pole position empty, candidate (21b), loses because of its violation of EPP, while (21d) loses because of its violation of undominated FULL-INTERPRETATION. The crucial comparison now is between the losing S-V-IO candidate (21a) and the winning IO-V-S candidate (21c). In contrast with what is observed in tableau (19), the subject-initial candidate (21a) now incurs a violation of the high-ranked *Pole/Theme constraint, because in this case the subject is a *theme*, and not an *agent*. This violation proves fatal when candidate (21a) is compared with candidate (21c), which instead violates the lower-ranked *Pole/Experiencer constraint by virtue of making the *experiencer* IO the Pole of the clause. The IO-V-S candidate (21c) thus emerges as the winner in this case.

- (21) INPUT: <like (x, y), x=chocolates (Th), y=Juan (Ex)>

	FULL-INT	*Pole/Theme	EPP	*Pole/Exper
a. [_{TP} los chocolates le gustan [_{VP} a Juan]] S V IO		*!		
b. [_{TP} le gustan [_{VP} los chocolates a Juan]] V IO S			*!	
☞ c. [_{TP} a Juan le gustan [_{VP} los chocolates]] IO V S				*
d. [_{TP} <i>exp</i> le gustan [_{VP} los chocolates a Juan]]	*!			

Lastly, consider unaccusative constructions like (22).

- (22) Unaccusative clauses
 Llegó tu hermano. **VS**
 arrived your brother
 ‘Your brother arrived.’

What is crucial in this case is that under the ranking in (16), it is better to leave the Pole position empty than to have a *theme* occupying it, because the degree of markedness of a clause with a *theme* as the Pole is higher than the degree of markedness of a clause that lacks a Pole altogether. Consequently, as opposed to the previous two cases, a verb-initial construction, candidate (b) in tableau (23) below, emerges as the output.

- (23) INPUT: <arrive (x), x=your brother (Th)>

	FULL-INT	*Pole/Theme	EPP
a. [_{TP} tu hermano llegó [_{VP}]] SV		*!	
☞ b. [_{TP} llegó [_{VP} tu hermano]] VS			*
c. [_{TP} <i>exp</i> llegó [_{VP} tu hermano]] exp-VS	*!		

The analysis presented in (18-23) represents the core of my proposal, and there are two crucial aspects related to it. First, by introducing the notion of Pole as; (a) the position in the clause where the EPP is satisfied, and; (b) a position sensitive to the semantic role of the XP that occupies it, it is possible to achieve this result without the need to invoke specific grammatical relations when determining which argument of the verb occupies the preverbal position. Secondly, by embedding the EPP constraint in the hierarchy in (16) we explain why the preverbal position needs to be occupied in some cases but not in others, a result that cannot be achieved in frameworks where satisfaction of the EPP is an all-or-nothing requirement.

It is further worth keeping in mind that the markedness constraints in the Pole Hierarchy are also violable constraints. This is important when we consider topicalization, where an XP that bears the *theme* semantic role, for instance, can occupy [Spec, T] and function as the Pole of the clause when it corresponds to a topic, as in the examples in (24). This is because the constraint that requires topics to be fronted to a left-peripheral position (the TOPICFIRST constraint of Costa 1998, 2001) outranks *Pole/Theme. The full analysis of topicalization is dealt with in Gutiérrez-Bravo (2002) to which I refer the reader for detailed discussion.

- (24) a. [Los discos]_{TOP} los compró una muchacha. **OVS**
 the records ACC-CL bought a girl
 ‘A GIRL bought the records.’
- b. [Tu hermano]_{TOP} llegó. **SV**
 your brother arrived
 ‘Your brother arrived.’

3. Other Cases

In this section I consider a number of other cases which are related to this proposal, and which I argue provide further support for the analysis developed above. The first case corresponds to different classes of psych verbs distinct from the class we have already analyzed. The second case corresponds to impersonal constructions.

3.1 Psych verbs

Spanish has different classes of psych predicates, roughly along the lines of those described for Italian in Belletti & Rizzi (1988). The verb *gustar* ‘like’, analyzed above, belongs to a class which is essentially equivalent to the *piacere* ‘please’ class of Belletti & Rizzi (1988).¹⁰ In a second class of Spanish psych verbs (roughly equivalent to the *temere* class of Belletti & Rizzi 1988) the *experiencer* is realized as the nominative subject and the *theme* as an accusative argument. An example with the verb *amar* ‘love’, is provided in (25).

¹⁰ Notice that the relevant word order facts for this class of psych verbs in Italian are the same as those observed in Spanish (Belletti & Rizzi 1988, Arnaiz 1998).

(25) UNMARKED WORD ORDER

Juan ama a María.	S	V	O
Juan loves ACC María	<EXP>		<THEME>
‘Juan loves María.’			

Observe that the unmarked word order facts for this class of verbs¹¹ are the ones predicted by my analysis. In terms of semantic roles, these predicates are just like *gustar* ‘like’, analyzed in (21), and so we expect them to behave in a similar way, independently of the specific grammatical relations that their arguments display.

A third class of psych verbs warrants a closer inspection. Belletti & Rizzi (1988) consider a third class in Italian, the *preoccupare* ‘worry’ class. In clauses with this class of verbs, the *experiencer* is realized as an accusative object and the nominative subject occupies the preverbal position, as shown in (26).

(26) ITALIAN (Belletti & Rizzi 1988:291)

Questo preoccupa Gianni.	SVO
this worries Gianni	

However, most of the Italian verbs in this class considered by Belletti & Rizzi (1988) belong to the *gustar*, ‘like’, class in Spanish, as shown in (27).

(27) A Juan le preocupan sus hijos.	IO	V	S
to Juan DAT-CL worry-3p his children	<EXP>		<THEME>
‘Juan is worried about his children.’			

Spanish does have some psych verbs with which the *experiencer* is realized as an accusative object in the post-verbal field and where the nominative subject appears in the preverbal position. Verbs of this kind include *atemorizar* ‘terrorize’, and *intimidar* ‘intimidate’. Examples are presented in (28).

(28) a. Juan atemoriza a sus hermanos.	SVO
Juan terrorizes ACC his siblings	
‘Juan terrorizes his siblings.’	
b. El gobierno intimida a los activistas.	
the government intimidates ACC the activists	
‘The government intimidates the activists.’	

My claim is that the word order observed in these cases is due to the fact that the subject is not really a *theme*, but rather is closer in its thematic properties to an *agent* (see also Pesetsky

¹¹ Other verbs in this class include *conocer* ‘know’, *admirar* ‘admire’, *amar* ‘love’, *odiar* ‘hate’, *despreciar* ‘loathe/despise’, etc.

1995).¹² The broad semantic role labels we have been using so far are not enough to clarify the state of affairs in (28), but Dowty's (1991) finer distinctions in terms of ProtoRole entailments do shed light on this issue. The subjects in (28), are perhaps not equivalent to transitive *agents*, but these verbs do necessarily entail their *volitional involvement in the event or state*, which is one of the entailments associated with the Agent Proto-Role in Dowty's analysis.¹³ This is in contrast with what is observed with respect to the object in (25) and to the subject in (27), where the psych verbs do not entail the *volitionality* of these arguments. The crucial observation is that the subjects in (28) already carry two of Dowty's five possible Proto-Agent entailments, *causing an event or change of state in another participant* and *existing independently of the event named by the verb*. Apparently, the additional *volitionality* entailment of the verbs in (28) is enough for their subjects to qualify as *agents*. The SVO order of (28) is thus consistent with my proposal.

Observe that Spanish has an additional class of psych verbs that provides support for this particular analysis. Verbs like *molestar* 'bother', *alterar* 'get on one's nerves', *conmover* 'be moved by', *asustar* 'frighten', show an alternation where the *experiencer* is realized as either an accusative DO or a dative IO. Crucially, this alternation is correlated to unmarked word order, as shown in the examples in (29).

- (29) UNMARKED WORD ORDER
- a. Los niños molestan a Juan. **SVO**
 the children bother ACC Juan
 'The children bother Juan.'
- b. A Juan le molestan los niños.¹⁴ **IOVS**
 to Juan DAT-CL bother the children
 '(The) children bother Juan.'

The relevant observation is that the meaning of these two examples is not the same. Example (29a) means that the children are intentionally engaged in doing things with the purpose of bothering *Juan*. In contrast, (29b) means that *Juan* finds the children bothersome, i.e., he is irritated by their mere presence or by the normal things that children usually do. This can be further shown by the fact that (29a) is compatible with certain adverbs like *a propósito* 'on purpose', whereas (29b) (with a subject-oriented reading) is not.

¹² Belletti & Rizzi (1988) in fact note that some verbs of the *preoccupare* class can have an interpretation where one of the arguments functions as an *agent*.

¹³ In this respect, it is relevant that Spanish *intimidar* in (28b) is not exactly equivalent to English 'to intimidate', since the former necessarily implies that purposeful actions are taken to bring about the intimidation. Accordingly, (28b) cannot mean that the activists are intimidated simply because of the presence or the amount of power of the government.

¹⁴ In some dialects of Spanish, human direct objects can be doubled by the clitic *le* (the so-called *leista* dialects). Mexican Spanish is not one of these dialects, so it is certain that the *experiencer* in (29b) is a dative indirect object, and not a clitic left-dislocated direct object.

- (30) a. Los niños molestan a Juan a propósito.
 the children bother ACC Juan to purpose
 ‘The children bother Juan on purpose.’
- b. ??A Juan le molestan los niños a propósito.
 to Juan DAT-CL bother the children to purpose
 (The children irritate Juan on purpose.)

These facts receive a straightforward explanation in terms of the analysis sketched above if the subject of (29a) has the *volitionality* entailment, which in turn the subject of (29b) lacks. Accordingly, the subject in (29a) qualifies as an *agent* and so it occupies the preverbal position and is compatible with adverbs like *a propósito* ‘on purpose’. In contrast, the subject of (29b) can be considered to be a *theme* and it appears in the post-verbal field in the unmarked case, as predicted by my analysis.

3.2 Impersonal constructions

Finally, I consider a number of active intransitive constructions, labeled *impersonal* constructions in the literature on Spanish, which have been recently analyzed in Fernández-Soriano (1999). These constructions are of two kinds. The first kind involves intransitive stative predicates and meteorological verbs that can appear with a preposed locative XP; some of these predicates (such as *faltar* ‘lack’) can also take a dative XP as their argument. The second class corresponds to intransitive eventive predicates (such as *suced*, *ocurrir* ‘happen’) that can take either a locative or a dative XP as an argument, which can also appear in the preverbal position. The two constructions are exemplified in (31) and (32), respectively.¹⁵

- (31) En Madrid nieva.
 in Madrid snows
 ‘It snows in Madrid.’
- (32) a. Aquí pasa algo.
 here happens something
 ‘Something’s going on here.’
- b. A Juan le pasa algo.
 to Juan DAT-CL happens something
 ‘Something’s going on with Juan.’

Fernández-Soriano provides considerable evidence that the locative and dative XPs in (31) and (32) are arguments of these predicates. She shows that they behave differently from adjuncts in a number of respects, such as extraction from coordinate clauses and interrogative

¹⁵ Examples (31) and (32a) are taken from Fernández-Soriano (1999).

inversion. She further shows that these arguments are base-generated in a position that is both structurally higher than the base position of the *theme* subject and the base position of the locative and dative (*goal*) arguments of other classes of predicates. These conclusions seem well-motivated and will not be contested here.

However, Fernández-Soriano goes on to propose that these locative and dative XPs are quirky subjects that appear in the preverbal subject position in the unmarked case. This would be problematic for my analysis for two reasons. First, in the proposal I have developed, *agents* and *experiencers* are predicted to appear in the preverbal Pole position in the unmarked case, but it is not immediately clear if the dative arguments of impersonal constructions can be grouped under either of these semantic roles. The second problem is that my analysis predicts that *locative* XPs should not appear in the Pole position in the unmarked case, because the constraint *Pole/Locative, which penalizes locative Poles, outranks the EPP (see the ranking in (16)). In relation to the analysis I have proposed so far, Fernández-Soriano's claims about the unmarked word order of these constructions require some clarification.

Consider first the predicates that take a dative argument, like *pasar* 'happen' in (32b). For these constructions I agree with Fernández-Soriano both in that the unmarked word order corresponds to the one in (32b), where the dative appears in the preverbal position, and in that it is fairly clear that these are not psych predicates in the general sense. However, this does not by itself exclude the possibility that the dative XP has the *experiencer* semantic role. In fact, Fernández-Soriano hints at this solution in two different ways. First, she notes that the dative XPs in these constructions do not behave like the *goal* arguments of ditransitive predicates with respect to secondary predication, interrogative inversion and other phenomena. Secondly, she notes that other predicates in this class express a psychological state in relation to some element (call it X) that is not realized as a nominal expression. Instead X is realized as a clause, as in (33a) or it has no overt syntactic expression whatsoever, as in (33b):

- (33) a. A Pedro le daba lo mismo [decírmelo o no].
 to Pedro DAT-CL gave the same to-tell-me-it or not
 'Pedro didn't care about telling me or not.'
- b. A Marta le fue bien en Buenos Aires.
 to Marta DAT-CL went well in Buenos Aires
 'Marta had a very good time in Buenos Aires.'

(Fernández-Soriano 1999: 123)¹⁶

Based on this observation, it is not unreasonable to suggest that the predicates in (33) simply correspond to a particular kind of psych predicates. In this class, the *theme* that corresponds to the origin/cause of the psychological state need not be realized as a nominal argument. Similar observations can be made with respect to (32b). In this case, the dative XP can be thought of as an *experiencer* that is affected by the state expressed by verb and the subject *theme*. Still, one probably has to acknowledge that the *experiencer* status of the datives in

¹⁶ The glosses are my own.

other verbs of this class is much less clear than in the examples in (32b) and (33). This is the case of the verb *faltar* ‘lack’, exemplified in (34):

- (34) A la silla le falta una pata.
 to the chair DAT-CL lacks a leg
 ‘The chair has a leg missing (Lit. “a leg is lacking from the chair”).’

The thematic characterization of the arguments of this verb does not seem to fit in the Thematic Hierarchy assumed up to now. The dative is clearly not a *experiencer*, but it is clearly not a *goal* either, since the verb expresses a state. Rather, *faltar* seems to establish something akin to a whole-part relation between the *dative* argument and the subject DP, where it is unclear if semantic roles as understood so far in this paper are of any relevance in defining this relation.¹⁷ However, it does not necessarily follow from this that the arguments of the verb must be equally prominent. In light of our discussion of semantic entailments in psych verbs, it is not unreasonable to think that the Pole may be sensitive to other prominence hierarchies, one of which could include *whole* > *part* as a part of the scale. This would explain why the dative appears in the preverbal position, since the whole would be more prominent than the part. Investigating this issue in detail, however, requires a more detailed analysis of the semantics of these predicates than is possible here.

The case of preposed locative XPs in (31) and (32a) is somewhat clearer. In this case there is considerable evidence that these preposed locative XPs are topics, and that, *contra* Fernández-Soriano’s analysis, they do not appear in the preverbal position in the unmarked case. Fernández-Soriano claims that the word order in (31) and (32a) is felicitous in an out-of-the-blue context, which corresponds to our diagnostic for unmarked word order. Although I do not share this judgement, it is possible that part of the problem has to do with the fact that (31) and (32a) correspond to the habitual present, which to my ear is easily amenable to an interpretation where the referent of the locative has already been introduced in the discourse and/or is readily identifiable by both speaker and hearer. In contrast, imagine a context where speaker A walks into a room and sees speaker B watching the news on television; he then utters the question in (35). In this context, the answers where the locative PP of meteorological verbs occupies the preverbal position are robustly infelicitous for all speakers of Mexican Spanish consulted:

¹⁷ This is not an idiosyncratic characteristic of Spanish, of course. For example, French can express some possessive relations by realizing the possessed entity as a nominative DP and the possessor as a dative, as is well known. See also Aissen’s (1999b) analysis of external possession in Tz’utujil (Mayan).

- (35) Qué pasa?
 what happens
 ‘What’s happening?’
- (36) a. Está nevando en Barcelona.
 it-is snowing in Barcelona
 ‘It’s snowing in Barcelona.’
- b. #En Barcelona está nevando.
 in Barcelona it-is snowing
- (37) a. Está lloviendo en Guadalajara.
 it-is raining in Guadalajara
 ‘It’s raining in Guadalajara.’
- b. #En Guadalajara está lloviendo.
 in Guadalajara it-is raining

As further evidence that fronted locatives in these constructions are functioning as topics, notice that for some speakers their acceptability is considerably downgraded in an out-of-the-blue context when they are indefinite, as shown in (38-39).¹⁸ Although this is not the case for all speakers, it is worth pointing out that, in contrast, preverbal *agents* and *experiencers* are not subject to this restriction, as shown in (40).

- (38) a. Aquí pasa algo.
 here happens something
 ‘Something’s going on here.’
- b. %En un salón pasa algo.
 in a classroom happens something
- (39) a. Aquí falta azúcar.
 here lacks sugar
 ‘There’s sugar missing here.’
- b. %En una mesa falta azúcar.
 in a table lacks sugar

¹⁸ Definiteness and specificity are variables that are not controlled for in Fernández-Soriano’s investigation. In most of her example sentences, the preverbal locatives are either deictic locative pronouns like ‘here’ or proper names corresponding to locations, whereas the post-verbal subjects are almost always indefinite or bare NPs. Given the nature of Spanish as a discourse-configurational language, it is important to control for these variables.

- (40) a. Una muchacha compró los discos.
 a girl bought the records
 ‘A girl bought the records.’
- b. A un reportero le gustó tanto tu discurso que
 to a journalist DAT-CL liked so-much your speech that
 lo quiere publicar.
ACC-CL he-wants to-publish
 ‘A journalist liked your speech so much that he wants to publish it.’

Consider also one of the crucial arguments presented by Fernández-Soriano in favor of her analysis of these locative XPs as quirky subjects. In her analysis, these locative XPs must receive quirky case in [Spec, T], which helps explain why they surface in this position in the unmarked case. In support of this proposal, Fernández-Soriano notes that the locative XPs are in complementary distribution with the dative XPs previously discussed, which are also quirky subjects. This is shown in (41)

- (41) a. Me falta café.
 to-me misses coffee
 ‘I am missing coffee.’
- b. Aquí falta café.
 here misses coffee
 ‘Coffee is missing here.’
- c. ??Aquí me falta café.
 here to-me misses coffee

(Fernández-Soriano 1999:121)

The argument is that if quirky case is discharged on the dative, then there is no possibility for the locative to receive quirky case in [Spec, T]. Example (41c), however, seems to me to be simply pragmatically anomalous, since one can readily construct other instances where the *locative* and the dative are not in complementary distribution.

- (42) a. En Barcelona nos pasó lo peor.
 in Barcelona to-us happened the worst
 ‘The worst happened to us in Barcelona.’
- b. En esta colonia nos faltan policías.
 in this neighborhood to-us lack policemen
 ‘We do not have enough policemen in this neighborhood.’

Fernández-Soriano claims that constructions like (42) are only possible when the locative is not an argument of the verb, in which case their interpretation is necessarily that of an adjunct. In other words, (41c) would have to be interpreted as something along the lines of

‘when I am here, I lack coffee’ (Fernández-Soriano 1999: 122). Although this is certainly a possible interpretation for (42a), (‘When we were in Barcelona, the worst happened to us’), example (42b) cannot have an interpretation along these lines (i.e. ‘When we are in this neighborhood, we don’t have enough policemen’). This points to the conclusion that the presence of the locative XP in [Spec, T] is not related to its status as either an argument or an adjunct, but rather to information structure considerations (i.e. topicalization).

Finally, a further argument can be found in the fact that fronted topics create islands for extraction, as is well known (see Rochemont 1989, Müller & Sternefeld 1993). The fronted locative XPs of impersonal verbs behave like topics in this respect, as shown in the examples in (43-44). Notice here that extraction is perfect when the locative XP appears in the post-verbal field. In contrast, preverbal subjects (Goodall 2001) and *experiencers* do not lead to such island effects, as shown in (45) (example (45a) is from Goodall 2001). This contrast holds both when the extracted *wh*-operator is D-Linked (as in (44a) and (45b)) and when it is not (as in (43a) and (45a)).

- (43) a. *Qué_i dices [que en Barcelona pasó t_i la semana pasada]?
 what you-say that in Barcelona happened the week past
- b. Qué_i dices [que pasó t_i en Barcelona la semana pasada]?
 what you-say that happened in Barcelona the week past
 ‘What did you say happened in Barcelona last week?’
- (44) a. ??Cuántos días_i dices [que en Barcelona nevó t_i la semana pasada]?
 how-many days you-say that in Barcelona snowed the week past
- b. Cuántos días_i dices [que nevó t_i en Barcelona la semana pasada]?
 how-many days you-say that snowed in Barcelona the week past
 For how many days did you say that it snowed in Barcelona last week?’
- (45) a. A quién crees [que Juan le dio el premio]?
 to whom you-think that Juan DAT-CL gave the prize
 ‘Who do you think that Juan gave the prize to?’
- b. Qué discurso_i dices [que a Juan le gustó tanto t_i [que lo quiere publicar]]?
 what speech say that to Juan DAT-CL liked so-much that ACC-CL wants to-publish
 ‘What speech did you say that Juan liked so much that he wants to publish it?’

Summing up, Spanish impersonal constructions clearly deserve more study, but their behavior does not appear to represent immediate counter-evidence against my proposal. On the one

hand, the dative arguments of impersonal verbs can be taken to be either *experiencers* or elements that are semantically more prominent than the subject DP along a scale different from the Thematic Hierarchy. On the other, the *locative* arguments of these verbs do not seem to surface in the Pole position in the unmarked case, and in those cases where they do appear in the preverbal field the evidence indicates that this is the result of topicalization.

4. Conclusions

In this paper I have developed an analysis where the unmarked word order in Spanish is determined by the semantic properties of the arguments of different classes of verbs, and not by their grammatical relation. I have argued that the relevant word order facts can be accounted for by appealing to the notion of the Pole as the constituent that occupies the highest inflectional specifier and satisfies the EPP, independently of its grammatical relation. I have then suggested that clauses with different constituents in the Pole position have different degrees of structural markedness. With respect to this I have suggested that Harmonic Alignment as part of an optimality-theoretic analysis can account for when and whether the highest inflectional specifier is filled in the unmarked case. The resulting analysis allows us to dispense altogether with the notion of a subject position in the inflectional layer of the clause in Spanish, and further provides an account for the cases where the highest specifier of the extended projection is occupied by a constituent other than the subject and for the cases where it is left empty. Finally, since semantic properties are shown to be the crucial factor that accounts for the relevant word order facts, this analysis opens the possibility that other semantic properties of DPs, such as specificity, may play a more important role than previously considered in determining word order in Spanish.

5. References

- Aissen, Judith. 1999a. "Markedness and subject choice in Optimality Theory", *Natural Language and Linguistic Theory* 17: 673-711.
- Aissen, Judith. 1999b. "External possessor and logical subject in Tz'utujil", in D. Payne & I. Barshi (eds.), *External Possession*. Amsterdam/Philadelphia: John Benjamins. 167-193.
- Aissen, Judith. 2000. "Differential object marking", University of California, Santa Cruz, ms.
- Alexiadou, A. y E. Anagnostopoulou. 1998. "Parametrizing AGR: word order, V-movement and EPP checking", *Natural Language and Linguistic Theory* 6:491-539.
- Arnaiz, Alfredo. 1998. "On the overview of the main word order characteristics of Romance", in A. Siewerska, J. Rijkhoff y D. Bakker (eds.), *Constituent Order in the Languages of Europe*. Berlin: Mouton de Gruyter. 47-73.
- Bakovic, Eric. 1998. "Optimality and inversion in Spanish", in P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis and D. Pesetsky (eds.), *Is the Best Good Enough?* Cambridge, Mass: MIT Press. 35-58.
- Belletiti, A. & L. Rizzi. 1988. "Psych verbs and theta theory", *Natural Language and Linguistic Theory* 6: 291-352.
- Chomsky, Noam. 1993. "A minimalist program for linguistic theory", in K. Hale & S. J. Keyser (eds.), *The view from Building 20*. Cambridge, Mass.: MIT Press. 1-52.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 2000. "Minimalist inquiries: The framework", in R. Martin, D. Michaels, & J. Uriagereka (eds.), *Step by Step: Essays on minimalist syntax in honor of Howard Lasnik*. Cambridge, Mass.: MIT Press.
- Collins, Christopher. 1997. *Local Economy*. Cambridge, Mass.: MIT Press.

- Contreras, Heles. 1976. *A Theory of Word Order with Special Reference to Spanish*, Amsterdam: North Holland.
- Costa, João. 1998. *Word Order Variation: A Constraint Based Approach*. Ph.D. dissertation, Leiden University.
- Costa, João. 2001. "The emergence of unmarked word order", in G. Legendre, J. Grimshaw y S. Vikner (eds.), *Optimality-Theoretic Syntax*. Cambridge, Mass: MIT Press. 171-204.
- Dowty, David. 1991. "Thematic proto-roles and argument selection", *Language* 67: 547-619.
- Fernández-Soriano, Olga. 1999. "Impersonal sentences in Spanish", *Syntax* 2: 101-140.
- Fant, Lars. 1984. *Estructura informativa en español: estudio sintáctico y entonativo*. Ph.D. diss., Uppsala University.
- Goodall, Grant. 2001. "The EPP in Spanish", in W. Davis y S. Dubinsky (eds.), *Objects and Other Subjects*. Dordrecht: Kluwer. 193-223.
- Grimshaw, Jane. 1993 "Minimal projection, heads and inversion.", Rutgers University, ms.
- Grimshaw, Jane. 1997. "Projection, heads and optimality", *Linguistic Inquiry* 28: 373-422.
- Grimshaw, Jane. 2001. "Economy of structure in OT", Rutgers University, ms.
- Grimshaw, J. & V. Samek-Lodovici. 1998. "Optimal subjects and subject universals", in: P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis and D. Pesetsky (eds.), *Is the Best Good Enough?* Cambridge, Mass: MIT Press. 193-219.
- Gutiérrez-Bravo, Rodrigo. 2000. "Subjects, left periphery and word order in Mexican Spanish". University of California, Santa Cruz, ms.
- Gutiérrez-Bravo, Rodrigo. 2002. *Structural Markedness and Syntactic Structure: A study of word order and the Left Periphery in Mexican Spanish*. Ph.D. dissertation, University of California, Santa Cruz.
- Larson, Richard. 1988. "On the double object construction", *Linguistic Inquiry* 19: 335-391.
- Masullo, Pascual J. 1993. "Two types of quirky subjects: Spanish versus Icelandic", in *Proceedings of NELS 23*. GLSA, University of Massachusetts, Amherst. 303-317.
- Müller, G. & W. Sternefeld. 1993. "Improper movement and unambiguous binding", *Linguistic Inquiry* 24: 461-507.
- Parodi, Teresa. 2002. "Clitic doubling and clitic-left dislocation in Spanish and Greek as native and as L2 grammars." Paper presented at *Semantic and Syntactic Aspects of Specificity in Romance Languages*, Konstanz, Konstanz University.
- Pestsky, David. 1995. *Zero syntax*. Cambridge, Mass.: MIT Press.
- Prince, A. & P. Smolensky. 1993. *Optimality Theory: Constraint interaction in generative grammar*. RuCCs Technical Report 2, Center for Cognitive Science, Rutgers University, Piscataway N.J. and Computer Science Department, University of Colorado, Boulder.
- Rochemont, Michael. 1989. "Topic islands and the subjacency parameter", *Canadian Journal of Linguistics* 34: 45-170.
- Suñer, Margarita. 1988. "The role of agreement in clitic-doubled constructions", *Natural Language and Linguistic Theory* 6: 391-434.
- Suñer, Margarita. 1994. "V-movement and the licensing of argumental *wh*-phrases in Spanish", *Natural Language and Linguistic Theory* 12: 335-372.
- Zagona, Karen. 2002. *The Syntax of Spanish*. Cambridge: Cambridge University Press.
- Zubizarreta, María Luisa. 1998. *Prosody, Focus and Word Order*. Cambridge, Mass.: MIT Press.

- (2) [-animate], [+definite], [+specific]:
- (2a) Vio **a las sierras.** (Puerto Rican Spanish)
 saw.past-3.sg the mountains
 ‘(S)he saw the mountains’
- (2b) Cosecharon **al maíz.** (Argentinian Spanish)
 harvest.past-3.pl the corn
 ‘They harvested the corn’

In Standard Spanish both sentences in (2) are ungrammatical since they violate the restriction that the particle *a* can only occur with a [+animate] direct object, cf. (1b). In American Spanish dialects, on the other hand, the particle can appear if the direct object is [+definite] and [+specific]. Definiteness is not a determining parameter for *a* in Standard Spanish, while animacy seems not to be a determining parameter for Spanish dialects allowing *a* with [-animate] objects. Specificity is an obligatory parameter in both variants, as summarized in table (3):

- (3) Parameters for determining the use of *a* with direct objects

Variant / category	animacy	definiteness	specificity
Standard Spanish	[+animate]	()	[+specific]
American Spanish	()	[+definite]	[+specific]

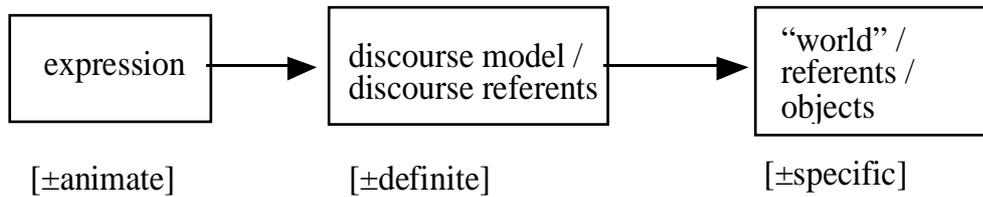
This very informal description must be developed in at least three directions: First, we have to give a clear definition, or at least a much better description, of our referential categories, in particular that of specificity. Second, we need more relevant data and a much deeper analysis of it in terms of the discussed referential categories. For example, we have to ask whether or not the use of *a* with [+definite] and [+specific] is only an exception or an acceptable variant. Third, we have to describe the interaction of the different categories: Is the interaction between animacy and definiteness similar or equal to that between specificity and definiteness? In this paper we intend to give first answers to these questions through an in depth analysis of the three referential categories involved, looking at the morphosyntactic reflexes of mood in relative clauses, the use of *a* with direct objects, and clitic doubling in Spanish.

In section 2, we present the three referential categories animacy, definiteness, and specificity, and discuss their particular values, and their representations as scales, hierarchies or polar features. In section 3, we account for three morphosyntactic phenomena in Spanish in terms of these referential categories: (i) mood selection in relative clauses; (ii) the conditions for the use of *a* with direct objects (“prepositional accusative”); and (iii) the restriction of clitic doubling in Standard and Río de La Plata Spanish, a variety of Spanish spoken in the area of Buenos Aires. It will be shown that the particular behaviour of clitic doubling in this variety of Spanish cannot be explained in terms of harmonic alignment of definiteness and specificity. In section 4, we therefore discuss different theoretical models of combining referential categories, such as subordination, cross-classification, and harmonic alignment. In Section 5, we give a short summary of our findings and directions for further research.

2. Referential categories

We use “referential categories” or “referential parameters” as a cover term for linguistic categories (or concepts) that are related to the semantic nature of an expression and its way [to] of referring. The class contains categories such as animacy, gender, number, specificity (or referentiality), genericity, definiteness and probably many more. We do not intend to define such a class or to discuss the nature of all the elements. Rather, we focus on three categories, namely animacy, definiteness, and specificity. Even in a very simple model of reference, as in (4), these categories occupy very different locations. Animacy is a lexical property of an expression, definiteness is a discourse pragmatic property, and specificity is a referential category.¹

(4) A simple model of reference:



2.1 Animacy

Simplified, we define animacy as a lexical feature of linguistic expressions that describes a certain property of the intended referent.² Animacy is often represented by a scale of different values, as in (5).

(5) Animacy Scale: human > animate > inanimate

As Silverstein (1976) and others have observed, animacy may determine certain morphosyntactic features in a language, such as grammatical hierarchies, number marking, or the lexical choice of question word. In English, German, French and Spanish, like in many other Indo-European languages, the choice of question word is determined by the animacy value for the noun phrase in the intended answer. The categorial cut is made between [+human] and [–human], as in (6):

(6) Animacy Scale and choice of question word

human	animate > inanimate
[+human]	[–human]
who / whom	what / what

¹ We do not want to take any position towards the question whether reference is a relation between the expression and its referent or whether it is a relation between a discourse representation and model theoretic objects, as in Discourse Representation Theory and as sketched in (4).

² One could also argue that it is a conceptual feature assigned to the referent by the speaker, which is reflected in the behavior of the linguistic expression associated with it. We do not want to discuss this subtle distinction or the question whether there is a distinction between natural and grammatical animacy.

- | | | |
|------|-----------------------------|-----------------------------|
| (6a) | Who have you seen? | What have you seen? |
| (6b) | Wen hast du gesehen? | Was hast du gesehen? |
| (6c) | Qui as-tu vu? | Qu' as-tu vu? |
| (6d) | ¿ A quién has visto? | ¿ Qué has visto? |

The categorial cut between two values of the scale is not always easy to make. Additionally, people may differ in the conception of certain objects with respect to these values. For example, a pet with a name is often regarded as part of the family, and therefore included in the [+human] rather than the [-human] category. There are many more such grey areas between categories that are otherwise clearly defined.

2.2 Definiteness

In a pre-theoretical definition, a definite singular expression unambiguously denotes or refers to one object, i.e. the object can be identified as the only one that is denoted by the expression. The fixed reference of a definite expression depends on as different parameters as the type of expression, semantic rules, and pragmatic strategies. There are several theories of definiteness, each of which focuses on a particular aspect of definiteness, a certain class of definite expressions, or one prominent use of them: (i) Russell's Theory of Definite Description is based on unique definite NPs; (ii) the anaphoric or familiarity theory of definiteness (Karttunen, 1976; Kamp, 1981; Heim, 1982) assumes that the anaphoric potential of definite NPs is the most characteristic; (iii) Löbner's (1985) theory of definiteness starts from the functional use; and (iv) the theory of definiteness as salience is based on the situational use of definite NPs (Lewis, 1979; Egli & von Heusinger, 1995). Again, we cannot present the whole discussion on this issue, but for a more comprehensive overview see Heim (1991) and von Heusinger (1997).

Here, we assume with Karttunen (1976), Kamp (1981), Heim (1982) and the dynamic tradition, that definiteness is a discourse-pragmatic property that indicates that the discourse referent associated with a definite expression can be identified with an already introduced discourse item. Thus, definiteness does not express the identifiability of the referent (in the world), a widespread view in descriptive grammars. Rather, definiteness expresses familiarity in a discourse structure. The discourse structure is understood as an intermediate structure between the linguistic expressions and their referents in "the world", as simplified in (4) above.

Besides the simple contrast between definite and indefinite, we also find different versions of "Definiteness Scales" The following is proposed by Aissen (2000:2), who refers to Croft (1988):

- (7) Definiteness Scale (Aissen, 2000):
 personal pronoun > proper noun > definite NP > indefinite NP

For certain discourse pragmatic functions, like anaphoric linkage, the Scale is divided by a categorial cut into two parts: The [+definite] expressions and the [-definite] expressions. A definite expression is accessible for anaphora, even if it is in the scope of a negation (or other operators), as illustrated by the contrast between (9) and (10).

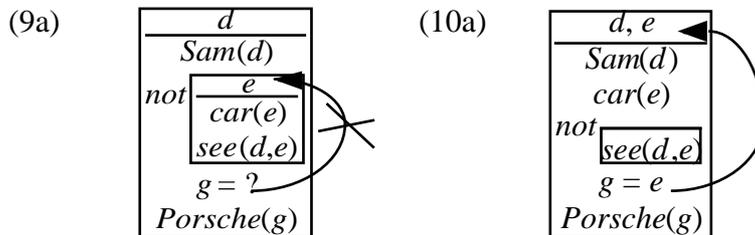
(8) Anaphoric Accessibility and the Definiteness Scale:

personal pronoun > proper noun > definite NP [+definite]	indefinite NP [−definite]
accessible even under negation	not accessible under negation

(9) Sam did not see a car. #It was a Porsche.

(10) Sam did not see the car. It was a Porsche.

Discourse Representation Theory (DRT) of Kamp (1981) and Kamp & Reyle (1993) explains the different accessibility structures of the two small text fragments (9) and (10) by assuming that the definite NP *the car* in (10) introduces a discourse referent *e* in the main box. This discourse referent is accessible for the discourse referent *g* associated with the pronoun *it* (cf. (10a)). In (9a) the indefinite introduces a discourse referent *e* only in the embedded box such that the pronoun discourse referent cannot access it.



To sum up, we assume that definiteness describes the discourse-pragmatic property of familiarity: definite expressions indicate that their associated discourse items can be identified with already introduced ones, while indefinite expressions indicate that their discourse referents are new.

2.3 Specificity

The concept of specificity was introduced in the late 60s by transferring the *de re-de dicto* distinction of definite NPs to indefinite NPs. The contrast is illustrated by example (11), which can be assigned two readings: the specific reading of *a monk* is motivated by the continuation (11a), while the non-specific reading can be continued with (11b) (see von Heusinger, 2002):

(11) Umberto Eco: "I desired to poison *a monk*."

(11a) He lived in the famous monastery Bobbio in the year 1347.

(11b) Therefore, Eco started to write a novel about a monastery.

Unlike animacy and definiteness, specificity is not assigned a scale, but rather a categorial distinction between [+specific] and [−specific], as in (12). Alternatively, one could also assume a two-part scale, as in (12'). However, it is not clear why [+specific] outrank [−specific].

- (12) Specificity Contrast [+specific] vs. [-specific]
 (12') Specificity Scale [+specific] > [-specific]

The literature on specificity is mainly concerned with specific indefinite NPs, which are grouped into different classes: (i) scopal specific indefinites, (ii) epistemic specific indefinites, (iii) partitive specific indefinites, (iv) intermediate specific indefinites or “relative specific indefinites” (see Farkas, 1995 for the first three classes and von Heusinger, 2002 for the last class).

2.3.1 Scopal specificity

Classically, the contrast between a specific and a non-specific reading of an indefinite is configurationally represented by scope interaction between the indefinite and some other operator, such as verbs of propositional attitude, negation or universal quantifiers as in (11) above and (13)-(14) below. In (14), the indefinite interacts with two operators such that we expect three readings, which the readers can easily work out by themselves.

- (13) Bill didn't see **a misprint**. (Karttunen 1976)
 (13a) There is a misprint which Bill didn't see.
 (13b) Bill saw no misprints.
 (14) Bill intends to visit **a museum** every day. (Karttunen 1976)

2.3.2 Epistemic specificity

There are examples that show the same (intuitive) contrast, but do not contain operators. For the specific reading of (15), we can continue with (15a), while the non-specific reading can be continued by (15b). This contrast is also often described as *referential vs. non-referential*. The specific indefinite refers to its referent directly, while the non-specific indefinite depends on the interpretation of other expressions in the context.

- (15) **A student** in Syntax 1 cheated on the exam. (Fodor & Sag 1982)
 (15a) His name is John.
 (15b) We are all trying to figure out who it was.

2.3.3 Partitive specificity

Milsark (1974) argues that indefinite NPs can either receive a weak (or existential) interpretation or a strong (or propositional) interpretation. In (16) the indefinite *some ghosts* receives a weak interpretation, but it gets a strong interpretation in (17), i.e. it presupposes that there are other groups of ghosts. The reading in (17) is generally called “partitive”.

- (16) There are **some ghosts** in this house.
 (17) **Some ghosts** live in the pantry; others live in the kitchen.

Enç (1991) claims, based on data from Turkish, that partitives denote an unknown element of a given set. Partitives always exhibit wide scope since the set from which they pick some elements

out is already mentioned. This means that partitives are complex expressions, formed by a choice from a definite set. We assume (contrary to Enç and others) that partitives comprise two independent referential functions: the first function ("the choice") can either be specific or non-specific, while the second function ("given set") must be definite. We therefore do not include them in the investigation of specific indefinites proper.

2.3.4 Relative specificity

The term "relative specific" or "intermediate specific indefinites" describes specific indefinites that depend on other expressions, and therefore show flexible scope behaviour. This observation was already made in early investigations of specificity. Contrary to Fodor & Sag (1982), Farkas (1981) shows with examples like (18) that indefinite NPs can have more than only a narrow scope non-specific reading (18a) and a wide scope specific reading (18c). They can also receive an "intermediate" scope reading (18b). According to this reading of (18), the indefinite *some condition proposed by Chomsky* has wide scope with respect to *three arguments* and narrow scope with respect to *each student*.

- (18) Each student has to come up with three arguments that show that **some condition proposed by Chomsky** is wrong. (Farkas, 1981)
- | | | |
|-------|---|--------------------|
| (18a) | each student > three argument > some condition | narrow scope |
| (18b) | each student > some condition > three argument | intermediate scope |
| (18c) | some condition > each student > three argument | wide scope |

Hintikka (1986) made a similar observation in his discussion of the expression *a certain*. In (19), he shows that the specific indefinite *a certain woman* can receive narrow scope with respect to the universal quantifier and still be specific: there is a specific woman for each man. Hintikka suggests that the specific indefinite NP is to be represented by a Skolem-function that assigns to each man the woman who is his mother. Once the reference for *man* is fixed (during the process of interpreting the universal quantifier), the reference for the specific indefinite is simultaneously fixed. In (19b), we informally mark this by indexing the indefinite NP with its anchor, here the variable for *man*.

- (19) According to Freud, every man unconsciously wants to marry **a certain woman** – his mother. (Hintikka, 1986)
- (19a) x [Man(x) Wants(x , marry(x , $f(x)$))]
with f : Skolem function from men onto their mothers
- (19b) x [Man(x) Wants(x , marry(x , [a woman] $_x$)

These observations motivate a revision of the pre-theoretical description of specificity in terms of obligatory wide scope or referential expression. It is shown that a specific indefinite NP need not depend on the speaker or the context of utterance; it can also depend on other linguistic entities, like the universal quantifier *each student* in (18b) or *every man* in (19).

2.3.5 A unified theory of specificity

Just like different aspects of definiteness have led to different theories of definiteness, the different aspects of specificity have invoked different theories of specificity. It is even controversial whether the different aspects belong to one and the same category or to different ones, or whether the different types of specificity are linked to each other by "family resemblance". We assume here that there is one underlying category of specificity. Our view is that the specific expression is linked or anchored to another expression (the anchor), and therefore its interpretation is independent of the direct linguistic context. The interpretation depends instead on the interpretation of the anchor expression. For a detailed account of this view, see von Stechow (2002).

3 Morphosyntactic reflexes of referential categories

(Indo-European) languages strongly differ with respect to the morphosyntactic expression of referential categories. In some languages definiteness is marked using articles, while specificity is indicated by more subtle morphosyntactic contrasts. Often, the referential behaviour can only be recovered by discourse-pragmatic strategies. Spanish, however, seems to be a good candidate for a language where the three referential categories, animacy, definiteness, and specificity have numerous morphosyntactic reflexes. In this section we only discuss (i) mood selection in relative clauses, (ii) the "prepositional accusative" with the particle *a*, and (iii) clitic doubling in Standard Spanish and Río de La Plata Spanish.

3.1 Mood in relative clauses

A well-known and often cited case where one of the referential categories discussed in this article, namely specificity, is expressed by morphosyntactic means in Spanish, are the relative clauses. In this kind of clauses, the mood form of the finite verb may change with respect to the specificity of the head of the relative clause. As illustrated in (20), the verb appears in the indicative form with a [+specific] relativized noun, while subjunctive is used when the noun receives a [–specific] interpretation (Rivero, 1975; Leonetti, 1999: 865; Pérez Saldanya, 1999: 3256):

- (20a) [–animate], [–definite], [+specific]:
 Busco **un libro** en el que se *analiza* el modo en las
 search-1.sg a book in which that REF analyse-IND the mood in the
 oraciones de relativo.
 clauses of relative

- (20b) [-animate], [-definite], [-specific]:
 Busco **un libro** en el que se *analice* el modo en las
 search-1.sg a book in which that REF analyse-SUB the mood in the
 oraciones de relativo.
 relative clauses
 ‘I am looking for a book in which the mood in relative clauses is analyzed’

This morphological mood distinction in Spanish is linked to the specificity contrast (12) in section 2.3. In (20a), the indicative forces a reading according to which there exists a book in the world representing the characteristics established in the relative clause. The subjunctive in (20b), on the other hand, does not imply the existence of such a book, and the head noun does not refer to any particular book, but expresses the property of being a book. The existence of such a book is not implied since it is embedded under *search*.³ This *de re-de dicto* distinction may be represented by a simple scope interaction between the indefinite and *search*, where (20a) corresponds to (21a) and (20b) to (21b) (Pérez Saldanya, 1999: 3259):

- (21a) [a book: x [x analyzes the mood in relative clauses & search x]]
 (21b) [search [a book: x & x analyzes the mood in relative clauses]]

Interestingly, this morphosyntactic contrast between a [+specific] and a [-specific] relativized noun is not restricted to *indefinite* nouns, but can also be observed with [+definite] nouns (Leonetti, 1999: 865):

- (22a) [-animate], [+definite], [+specific]:
 Busco **el libro** en el que se *analiza* el modo en las
 search-1.sg the book in which that REF analyse-IND the mood in the
 oraciones de relativo.
 clauses of relative
- (22b) [-animate], [+definite], [-specific]:
 Busco **el libro** en el que se *analice* el modo en las
 search-1.sg the book in which that REF analyse-SUB the mood in the
 oraciones de relativo.
 clauses of relative
 ‘I am looking for a book in which the mood in relative clauses is analyzed’

³ Note that “existence” does not mean that a noun must exist in the real world. As noted by Pérez Saldanya (1999: 3256), the specificity contrast can also be observed with NPs referring to a fictitious world:

- (ia) He soñado que quería visitar **una ciudad de Marte** que *estaba*
 have-1.sg dreamt that would-1.sg visit a town of Mars that was-IND
 habitada por alienígenas.
 inhabited by aliens.
- (ib) He soñado que quería visitar **una ciudad de Marte** que *estuviese*
 have-1.sg dreamt that would-1.sg visit a town of Mars that was-SUB
 habitada por alienígenas.
 inhabited by aliens.

However, according to Pérez Saldanya (1999: 3265), sentences like (22b) with [+definite, –specific] NPs are odd for some speakers. This seems to be due to the fact that [+definite] nouns in general presuppose a unique referent. Receiving a [–specific] interpretation, a [+definite] NP loses this existential presupposition and, therefore, does no longer refer to an object which is identifiable for the hearer, although it preserves its property to refer to an object in a unique way (Leonetti, 1980: 154). Given this ‘uniqueness condition’, the use of the subjunctive mood in relative clauses of *definite* nouns, indicating a non-specific head noun, is odd for some speakers. This oddness disappears when the sentence contains an element which allows us to deduce more easily the uniqueness of the [+definite] noun (cf. Leonetti, 1999: 865; Pérez Saldanya, 1999: 3265f):

- (23) [–animate], [+definite], [–specific]:
 Busco **el libro** en el que se analice **mejor** el modo
 search-1.sg the book in which that REF analyse-SUB better the mood
 en las oraciones de relativo.
 in the clauses of relative
 ‘I am looking for a book in which the mood in relative clauses is better analyzed.’

These examples show clearly that the indicative – subjunctive distinction strongly correlates with the specificity of the relativized noun. In other words, specificity plays a crucial role for the use of mood in relative clauses.⁴

3.2 The “prepositional” accusative

Standard Spanish generally marks [+animate] direct objects with the particle *a* independently of the definiteness of the object, as already shown in (1), repeated here as (24):

- (24a) [+animate], [±definite], [+specific]:
 Vi **a la / unamujer**. (Standard Spanish)
 see.past-1.sg the a woman

⁴ Note that the verbal mood does not always function as indicator of the specificity contrast between relativized nouns in Spanish. In some cases, the indicative does not prevent the [–specific] interpretation of the relativized noun. This may happen in cases, like those in (i), where the whole utterance has a generic character (Leonetti, 1999: 865; Pérez Saldanya, 1999: 3257):

- (ia) [+animate], [–definite], [–specific]:
Quien calla otorga.
 who is silent agrees-IND
 (ib) [+animate], [–definite], [–specific]:
 Óscar no se atrevería a dirigirse a **una chica** que no habla español.
 Oscar not REF venture-COND to address-REF to a girl who not speaks-IND Spanish

In other cases the mood can change independently of the specificity of the relativized noun, as it can be observed in comparative constructions like those in (ii), where the NP is used predicatively and therefore is neutral with respect to [±specific]:

- (ii) [+animate], [–definite], [±specific]:
 Se comporta como **una persona** que { oculta / oculte } algo.
 REFL behaves like a person who hides-IND hides-SUB something
 ‘(S)he conducts behaves like a person who hides something.’

- (24b) [+animate], [±definite], [+specific]:
 Vi (**a*) **la / una mujer.** (Standard Spanish)
 see.past-1.sg the a woman
 ‘I saw the / a woman.’

The direct object with this particle is traditionally called ‘prepositional accusative’ in order to distinguish it from the *indirect* (“dative”) object, which is always preceded by the preposition *a*. Some varieties of Spanish, especially American Spanish, allow the particle *a* to precede [–animate] direct objects, at least in certain contexts, as in (2), repeated as (25):⁵

- (25a) [–animate], [+definite], [+specific]:
 Vio *a* **las sierras.** (Puerto Rican Spanish)
 saw.past-3.sg the mountains
 ‘(S)he saw the mountains’
- (25b) [–animate], [+definite], [+specific]:
 Cosecharon **al maíz.** (Argentinian Spanish)
 harvest.past-3.pl the corn
 ‘They harvested the corn’

On the other hand, it can be observed that objects which are lexically characterized as [+animate] are used *without* the particle *a*. This is the case in examples like (26), provided by Brugè & Brugger (1996: 6):

- (26a) ... una fuente de vida nueva que purificaba **el hombre moral**
 a source of life new that purifies the man moral
 ‘... a source of new life which purifies the moral man’
- (26b) ?Las enfermedades y la guerra han exterminado **el hombre**
 the illnesses and the war have exterminated the man

The absence of *a* in these examples seems to be due to the fact that the direct object does not denote an individual person. It rather receives a “kind interpretation”, which can be, according to Brugè & Brugger (1996), associated with the feature [–animate]. What strongly supports this analysis is the observation that, if one asks for the object in a *wh*-question, one can use, besides *a quién* ‘who’, the *wh*-word *qué* ‘what’ (Brugè & Brugger, 1996: 7):

⁵ Note that in Standard Spanish some verbs require the particle *a* with direct objects, independently of the feature [±animate]. This often happens with verbs which normally have [+animate] direct objects, as for example *llamar* ‘to call’ or *matar* ‘to kill’, and then lexicalize the particle (Real Academia Española, 1973: 373; Bruyne, 2002: 309, fn.5):

- (i) [–animate], [+definite], [+specific]:
Llamar a la muerte.
 to-call the death
- (ii) [–animate], [+definite], [+specific]:
 Los griegos *mataron* entonces *a la poesía*
 the greeks kill.past-3.pl then the poetry

- (27a) –¿**Qué** purifica unfuente de vidanueva?
 what purifies a source of life new
 ‘What does a source of new life purify?’
 –El hombre moral
 the man moral
- (27b) –¿**Qué** han exterminado las enfermedades y la guerra?
 what have exterminated the illnesses and the war
 ‘What did the illnesses and the war exterminate?’
 –El hombre
 the man

As shown in (6) in section 2.1, the lexical choice of the *wh*-word in (27) indicates that a [–animate] object rather than a [+animate] one is intended in the answer. This lexical choice suggests that the object NPs in (26) to which the *wh*-words in (27) are referring may be interpreted as [–animate].

The choice of *a* with direct objects is not only determined by animacy, but also by specificity. This can be observed in clauses where a [+animate] object is relativized, and where specificity is reflected in the choice of mood for the predicate in the relative clause (see last section). In these cases, the particle *a* must precede [+specific] direct objects, but it is normally omitted when the object is [–specific] (Jaeggli, 1982: 56, fn.14; Brugè & Brugger, 1996: 31; Leonetti, this volume):

- (28a) [+animate], [+definite], [+specific]:
 Busco **a una cocinera que sabe** hablar inglés.
 search-1.sg a cook who knows-IND to-speak English
- (28b) [+animate], [±definite], [–specific]:
 Busco **una cocinera que sepa** hablar inglés.
 search-1.sg a cook who knows-SUB to-speak English
 ‘I am looking for a cook who can speak English’

What is important for our purpose is the observation, made by Leonetti (1999: 867), according to which *every* direct object used *without a* receives a weak or a non-specific interpretation. Thus, in existential constructions, which clearly favour a non-specific interpretation, the use of the marker *a* is not allowed. This is illustrated in (30) (see also Leonetti, this volume):

- (29) [+animate], [±definite], [–specific]:
 Había (*a) **unas / todas las mujeres en la plaza.**
 (there) was some / all the women in the place

The same observation can be made in constructions with multiple quantification. As illustrated in (30a), the lack of *a* implicates that the direct object has a narrow scope. It has narrower scope than the universally quantified subject *todos los encuestados* and receives a non-specific interpretation, represented in (30b) (Brugè & Brugger, 1996: 34f; Leonetti, 1999: 867):

- (30a) [+animate], [–definite], [–specific]:
 Todos los encuestados vieron **unapersona sospechosa**.
 all the interrogated persons saw a person suspicious
- (30b) x [interrogated persons (x) y [person(y) & saw (x,y)]]

To sum up, the use of *a* with direct objects is determined by two referential parameters, animacy and specificity, while the third referential parameter (definiteness) does not interact here. Only a [+animate] and [+specific] direct object is preceded by *a*, as summarized in table (31):

- (31) The use of *a* with direct objects: Animacy and Specificity

	[+specific]	[–specific]
[+animate]	<i>a</i>	–
[–animate]	–	–

We are aware that this table does not capture all conditions for the use of *a*. There are several additional factors which determine its use, most of them are extensively discussed by Leonetti (this volume). He notes, for instance, that in certain contexts [+animate] and [–specific] direct objects are introduced by *a*, an option which should be excluded according to table (31). According to Leonetti (1999: 866, this volume), in most of these cases the presence of *a* can be explained by the fact that the category animacy predominates the category specificity and that therefore the [+animate] feature may override the [±specific] feature:⁶

- (32a) [+animate], [–definite], [–specific]:
 Está buscando **a alguien**
 is-3.sg looking someone
 ‘(S)he is looking for someone.’
- (32b) [+animate], [–definite], [–specific]:
 Sería estupendo si contrataron **a un ayudante**
 would wonderful if contract-3.pl an assistant
 ‘It would be wonderful if they have contracted an assistant.’

Another interfering factor is the relation between the different grammatical functions *subject*, *indirect object* and *direct object*. It seems that the marking of [+animate] and [+specific] can depend on the values of these features in the noun phrases filling the other grammatical

⁶ As already noted in the preceding footnote, selection properties of the verb may determine the use of *a*. This is also the case when the object is [+animate] and [–specific] (Brugè & Brugger 1996, 45, Leonetti, this volume):

- (ia) [+animate], [–definite], [–specific]:
 Pepa quiere matar **a unpoliciá cualquiera**
 Pepa wants to-kill a policeman any
 ‘Pepa wants to kill any policeman.’
- (ia) [+animate], [–definite], [–specific]:
 Todas las niñas admiraban **a algúncantante**
 All thechildren admire.past-3.pl some singer
 ‘Every child admired some singer’

functions. The contrast between (33a) and (33b) lies in the animacy of the subject. In (33b) the subject is [-animate], while the object is marked with [+animate]. This violates the principle that the subject must not be lower on the animacy scale than the direct object.⁷

- (33a) [+animate], [±definite], [±specific]:
La diva conoce *a* muchos aficionados a la ópera.
 the diva knows many amateurs of the opera
- (33b) [-animate], [±definite], [±specific]:
 ***La ópera** conoce *a* muchos aficionados.
 the opera knows many amateurs
- (33c) [-animate], [±definite], [±specific]:
La ópera conoce muchos aficionados.
 the opera knows many amateurs

A similar observation can be made with respect to indirect objects. The omission of the particle *a* before the direct object is strongly preferred in double object constructions, even when the direct object bears the features [+animate] and [+specific]. According to the grammar of the Real Academia Española (1973: 374f), this happens in order to avoid ambiguity effects which emerge from the fact that the indirect object in Spanish is obligatorily marked by the preposition *a*:

- (34) [+animate], [±definite], [+specific]:
 (34a) ?Recomiende usted *a mi sobrino* al señor director
 (34b) Recomiende usted **mi sobrino** al señor director
 'Recommend my nephew to the director'

3.3 Clitic doubling

Clitic doubling is another domain in Spanish where the referential categories which we are discussing play a crucial role. Note that there is a crucial difference between clitic doubling with indirect objects and with direct objects. In the former case, clitic doubling is obligatory with pronouns and strongly preferred with proper nouns and [+definite] NPs in *all* varieties of Spanish (Jaeggli, 1982: 12; Suñer, 1988; Parodi, 1998; Fernández Soriano, 1999; Colantoni, 2002: 321):

- (35a) *Le* doy la carta **a él** (Standard Spanish)
 Cl-DAT give-1sg theletter to him
- (35b) *Le* doy la carta **a Juan** (Standard Spanish)
 Cl-DAT give-1sg theletter to Juan
- (35c) *Le* doy la carta **al vecino** (Standard Spanish)
 Cl-DAT give-1sg theletter to-the neighbour

⁷ An alternative explanation is that the [-animate] subject does not license a [+specific] direct object. This explanation is in line with the theory of specificity presented in section 2.3. According to this theory, specific NPs must be anchored to some [+animate] expressions.

With [–definite] indirect objects, clitic doubling is not obligatory, but it is the strongly preferred option (Parodi, 1998: 87):

- (36d) *Le doy la carta a un vecino* (Standard Spanish)
 CI-DAT give-1sg the letter to a neighbour

Clitic doubling with direct objects primarily depends on the category definiteness. However, it is not the simple contrast between [+definite] and [–definite] which determines the use of clitics in constructions with direct objects. Rather, the clitic doubling phenomena can be explained on the basis of Aissen's Definiteness Scale in (7), repeated here as (37):

- (37) Definiteness Scale (Aissen, 2000)
 personal pronoun > proper noun > definite NP > indefinite NP

Given this scale, we can observe that in Standard Spanish clitic doubling with direct objects is only possible with the leftmost elements, namely with personal pronouns. In this case, clitic doubling is obligatory (Jaeggli, 1982: 14; Parodi, 1998: 86; Fernández Soriano, 1999: 1248):

- (38a) *La veo a ella* (Standard Spanish)
 CI-ACC see-1.sg her
- (38b) **Veo a ella* (Standard Spanish)
 see-1.sg her
 'I see her'

With full NPs, on the other hand, clitic doubling is generally excluded in Standard Spanish (Parodi, 1998: 89; Fernández Soriano, 1999: 1249):

- (39a) ??*La veo a María* (Standard Spanish)
 CI-ACC see-1.sg Maria
 'I saw María'
- (39b) **La veo a la mujer* (Standard Spanish)
 CI-ACC see-1.sg the woman
 'I saw the woman'
- (39c) **La veo a una mujer* (Standard Spanish)
 CI-ACC see-1.sg a woman
 'I see a woman'

Interestingly, however, some dialects of Spanish, especially Río de Plata Spanish, allow clitic doubling when the direct object is a *full* NP. In these dialects clitic doubling is possible with proper nouns and with [+definite] NPs, as shown in (40a) and (40b). Clitic doubling with [–definite] direct object NPs, however, seems to be excluded, as illustrated in (30c) (Jaeggli, 1982: 19; Parodi, 1998; Fernández Soriano, 1999: 1251):

- (40a) *La* veo **a María** (Río de la Plata Spanish)
 Cl-ACC see-1.sg Maria
 ‘I see María’
- (40b) *La* veo **a la mujer** (Río de la Plata Spanish)
 Cl-ACC see-1.sg the woman
 ‘I see the woman’
- (40c) **La* veo **a una mujer** (Río de la Plata Spanish)
 Cl-ACC see-1.sg a woman
 ‘I see a woman’

Clitic doubling is determined by definiteness and grammatical role, as summarized in table (41). Río de La Plata Spanish differs from Standard Spanish in that it allows clitic doubling even for elements that are low on the Definiteness Scale. While this difference is not so obvious for indirect objects (the only difference is the optional clitic doubling for indefinite NPs in Standard Spanish), it is very explicit for direct objects. Río de La Plata Spanish allows clitic doubling with all but indefinite NPs, while Standard Spanish only allows it with personal pronouns.

(41) Clitic doubling in Spanish: Grammatical Function and the Definiteness Scale

		personal pronoun	proper noun	definite NP	indefinite NP
indirect object	Río de la Plata Spanish	+	+	+	+
	Standard Spanish	+	+	+	(+)
direct object	Río de la Plata Spanish	+	+	+	–
	Standard Spanish	+	(+)	–	–

3.4 Clitic doubling and the “prepositional accusative”

Clitic doubling and the particle *a* very often co-occur: With indirect objects they almost always come together, and in many cases they also appear together with direct objects. Therefore, in accordance with Kayne (1975) and Jaeggli (1982; 1986) it has been assumed that clitic doubling is licensed by the particle *a*, rather than attributed to the Definiteness Scale (see Kaiser, 1992 for a detailed discussion of this approach, often called 'Kayne's generalization'). He argues that in Río de la Plata Spanish *a* functions as a *dummy* case marker which is able to assign (abstract) case to the object NP, while the object clitic receives or "absorbs" the case directly from the verb. This assumption seems to be supported by the observation that in Río de Plata Spanish, as in many other dialects of American Spanish (cf. (2)), [–animate] direct objects may appear in combination with *a*, and that in this case clitic doubling is possible or even the strongly preferred option (Laca, 1987: 307; Suñer, 1988: 399; Fernández Soriano, 1999: 1251):

- (42a) [–animate], [+definite], [+specific]:
Lo vamos **a empujar al ómnibus** (Río de La Plata Spanish)
 Cl-ACC will.3.pl. push the bus
 ‘We will push the bus.’

- (42b) ¿Tú *la* friegas *a la* **cocina**? (Río de La Plata Spanish)
 you CI-ACC wipe the kitchen
 'Do you wipe the kitchen?':

However, this close connection between clitic doubling and the particle *a* would be surprising since we have seen that the use of the particle depends on specificity and animacy, while clitic doubling is determined by definiteness. We would therefore predict cases in which we find clitic doubling but not the particle *a* and vice versa. This prediction is born out by observations by Suñer (1988). She argues that *a* cannot function as a dummy case marker in Río de la Plata Spanish, since Río de la Plata Spanish also allows clitic doubling with [–animate] direct objects which are not preceded by the particle *a* (Suñer, 1988: 399):

- (43a) [–animate], [+definite], [+specific]:
 Yo *la* tenía prevista **esta** **muerte**.
 I her have previewed this death
- (43b) [–animate], [+definite], [+specific]:
 Yo *lo* voy a comprar **el** **diario** antes de subir.
 I him will buy the newspaper before to go upstairs

Given these examples and given the definiteness scale in (7)/(37), it seems that this scale is able to describe correctly the possibility of clitic doubling in Río de la Plata Spanish. In other words, definiteness – and *not* the presence or absence of *a* – is the relevant factor for clitic doubling in this dialect. The examples in (43) – and there are many more like these (see Parodi, 1998: 89) – show that clitic doubling and the use of the particle *a* with the direct object follow independent parameters: Clitic doubling is conditioned by one referential parameter, namely definiteness, while the use of the particle *a* is determined by two parameters: animacy and specificity.

3.5 Clitic doubling and more referential parameters

This is not the end of the story. Clitic doubling in Río de la Plata Spanish is conditioned by additional referential parameters. Recall that according to the literature it is claimed that clitic doubling is excluded with [–definite] direct objects in Río de la Plata Spanish (cf. (39c)). Suñer (1988) notes that under certain circumstances clitic doubling is possible with [–definite] direct objects in Río de la Plata Spanish. But, according to Suñer (1988: 396), this is only possible when the [–definite] object NP is [+specific] (cf. also Parodi, 1998: 88f):

- (44a) [+animate], [–definite], [+specific]:
 Diariamente, *la* escuchaba a **una** **mujer** **que** **cantaba** **tangos**.
 daily CI-ACC listen-3.sg a woman who sing-past-3.sg tangos
- (44b) [+animate], [–definite], [–specific]:
 La* busco **una **mujer** **que** **sepa** **inglés**.
 CI-ACC search-1.sg a woman who knows-SUB English

Given this observation, Suñer (1988: 397) argues that the "pertinent feature for doubling is [+specific] and not [+definite]". According to that, we have to modify our table (41) and replace [definite NP] and [indefinite NP] by [specific NP] and [non-specific NP] respectively, as in (45):

(45) Clitic doubling in Spanish: Definiteness Scale and Specificity

		personal pronoun	proper noun	specific NP	non-specific NP
direct	Río de la Plata Spanish	+	+	+	-
object	Standard Spanish	+	(+)	-	-

With such a table, we could argue that clitic doubling in Río de La Plata Spanish only depends on the referential parameter [\pm specific], assuming that personal pronouns and proper nouns are always [+specific]. If this were correct, we would still have to explain how such a change from [\pm definite] to [\pm specific] is possible and how we can integrate the Specificity Contrast into the Definiteness Scale.

This shift from [\pm definite] to [\pm specific] is also suggested by the analysis of further data from Río de La Plata Spanish. We start with the simplified categorization in (46), where clitic doubling only depends on [\pm specific].

(46) Clitic doubling in Río de La Plata Spanish: Specificity

[+specific]	[-specific]
+	-

Suñer (1988) argues that clitic doubling in Río de La Plata Spanish is only possible if the direct object is [+specific] and [+animate], while it is excluded with [+specific] [-animate] direct objects. This is shown by Suñer's (1988: 396) example in (47). Thus we can summarize Suñer's assumption in table (48):

- (47) [-animate], [+specific]:
 La* compramos (a) **esa novela.
 CL buy.past-1.pl that novel
 'We bought that novel'

(48) Clitic doubling in Río de La Plata Spanish: Specificity and Animacy (Suñer, 1988)

	[+specific]	[-specific]	(Suñer, 1988)
[+animate]	+	-	
[-animate]	- (cf. (47))	-	

Parodi (1998), on the other hand, argues that clitic doubling in Río de La Plata Spanish is only possible if the direct object is [+specific] and [+definite]. Her examples include (49), which she categorizes as [-definite] and [+specific]. Her view can be summarized by table (50):

- (49) [-definite], [+specific]:
 La* veo a **una **mujer**.
 ACC-CL see-1.sg a woman
 'I see a woman.'

- (50) Clitic doubling in Río de La Plata Spanish: Specificity and Definiteness

	[+specific]	[-specific]	(Parodi, 1998)
[+definite]	+	–	
[-definite]	– (cf. (49))	–	

The two authors agree that there are two interacting referential categories, one of which is specificity. However, they disagree about the other category. If we combine the two tables (48) and (50) into one table (51), the disagreement concerns only two cells, [+animate], [-definite], [+specific] and [-animate], [+definite], [+specific]:

- (51) Clitic doubling in Río de La Plata Spanish: Specificity, Definiteness, and Animacy

[+specific]	[+definite]	[-definite]	(Suñer, 1988 / Parodi, 1998)
[+animate]	+ / +	+ (cf. (44a)) / – (cf. (49))	
[-animate]	– (cf. 47) / + (cf. (52))	– / –	

Both authors give examples that confirm their view. Suñer shows (example (44a)) that [-definite] [+specific] objects are doubled if they are [+animate]. Parodi's example (49) seems to contradict this example, since it seems to show that clitic doubling with [+animate] objects is ungrammatical when they are [-definite] [+specific]. However, the object in (49) may be understood as [-specific] if there is no additional context. The context in (44a) makes the specific reading very prominent. Parodi, on the other hand, quotes (52) which shows that a [-animate] [+specific] object can be doubled if it is [+definite]. Again, this contradicts Suñer's example (47), where the demonstrative object is not doubled.

- (52) [-animate], [+definite] [+specific] (Parodi, 1998: 89)

La compro **la** **mesa**
 ACC-CL buy-1.sg the table
 'I buy the table'

We think that it must have become obvious from the discussion that there is no very clear picture of the conditions for clitic doubling in Río de La Plata Spanish, and that further research is necessary. The ambiguous situation may also indicate that we observe a process that is developing. However, if we try to fix the picture and describe a synchronic system, we feel forced to assume that clitic doubling depends on all three referential categories, as summarized in table (53). Even if we do not consider the two controversial cells, we still find a contrast between [+animate] and [+definite] vs. [-animate] and [-definite] for [+specific] objects.

(53) Clitic doubling in Río de La Plata Spanish: Specificity, Definiteness, and Animacy

[+specific]	[+definite]	[-definite]	[-specific]	[+definite]	[-definite]
[+animate]	+	(+)	[+animate]	–	–
[-animate]	(+)	–	[-animate]	–	–

4. Combining referential parameters

Referential parameters, such as animacy, definiteness, and specificity, determine different morphosyntactic contrasts in Spanish (as well as in other languages). We have seen in the last sections, that they do this in different combinations (or “conspiracies”). The mood in relative clauses is primarily determined by specificity. The use of the particle *a* preceding direct objects is controlled by specificity and animacy. In section 3.5, we argued that clitic doubling in Río de La Plata Spanish depends on all three referential categories. In the course of our presentation, we have presented different scales, tables and charts combining different parameters. In this section, we present some very preliminary ideas about possible ways of combining two or more referential categories. There are different ways in which the interaction of the three referential parameters are described: (i) as subordination of one parameter under another; (ii) as cross-classification of two or more parameters; and (iii) as harmonic alignment between a two-part scale and a multi-part scale.

4.1 Subordination

Specificity is often understood as secondary referential property of NPs that applies only to indefinite NPs and it is often included into Aissen’s Definiteness Scale (7)/(37), such as in Aissen (2000: 2):

- (54) Definiteness Scale (Aissen 2000: 2)
 personal pronoun > proper noun > definite NP > specific indefinite NP
 > non specific indefinite NP

According to this view, definite NPs are used if both the speaker and hearer can identify the referent, specific indefinite NPs if only the speaker can identify the referent, while non-specific indefinite NPs indicate that none of them can identify the referent:

(55) The “identifiability” criteria for definiteness and specificity

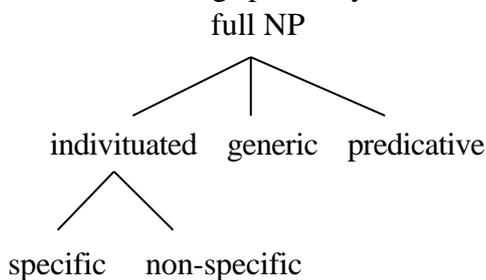
<i>identified by</i>	definite (+ specific)	indefinite specific	indefinite non-specific
speaker	+	+	–
hearer	+	–	–

However, this view is incorrect for theoretical as well as empirical reasons. The functional “explanation” or motivation for subordinating specificity under definiteness cannot be correct since the discussion of the last three decades has convincingly shown that definiteness cannot be explained in terms of “identifiability”. Definiteness is explained in terms of uniqueness,

anaphoric linkage and familiarity, functional concepts, or situational salience (see section 2.2). The empirical problems with such a table will be discussed in the next section.

There might be other cases, where subordination of one parameter under another is appropriate. For instance, if we assume a contrast of “ways of reference” as in (56) and the contrast between specific and non-specific as in (57), then we can combine the two contrasts such that the latter contrast only holds for the feature [+individuated] of the first contrast, as illustrated in the tree (58):

- (56) Ways of reference
individuated vs. generic vs. predicative
- (57) Specificity
specific - non-specific
- (58) Subordinating specificity under ways of reference



4.2 Cross classification

Many studies on grammatical contrasts that are triggered by referential parameters assume the subordination of specificity under indefinite full NPs. However, some studies that investigate the phenomena in more detail give good evidence that definiteness and specificity form, rather, a cross-classification as in (59), where the bold cell is the crucial one: [+definite], but [-specific]. The discussion in section 3.1 has shown that mood selection in relative clauses depends on specificity. A model like (54) or (55) would predict that all definite NPs are specific and therefore trigger indicative mood in relative clauses. However, evidence from Spanish, as in (60 a+b), shows that there are non-specific definite NPs, which therefore can trigger subjunctive mood.

- (59) Cross-classification of definiteness and specificity

	[+definite]	[-definite]
[+specific]	<i>la mujer que sabe inglés</i>	<i>una mujer que sabe inglés</i>
[-specific]	<i>la mujer que sepa inglés</i>	<i>una mujer que sepa inglés</i>

- (60a) [+animate], [+definite], [-specific]:
 Busco (a) **la mujer que sepa** inglés.
 search-1.sg. the woman who knows-SUBJ English
 ‘I look for the woman who knows / will know English.’
- (60b) [+animate], [+definite], [+specific]:
 Busco *a* **la mujer que sabe** inglés.
 search-1.sg. the woman who knows-IND English
 ‘I look for the woman who knows English.’

The discussion of clitic doubling in Río de La Plata Spanish suggests that we even have three parameters that cross-classify. Parodi (1998, 91) therefore proposes to extend Aissen’s Definiteness Scale (54) by the two additional parameters specificity and animacy, which results in the following scale (61).

- (61) Integrated Definiteness Scale (including animacy and specificity)

1.+2. pron.>	3.pron. >	full NP / animate > full NP / definite > full NP / specific >	full NP / inanimate full NP / indefinite full NP / non-specific
--------------	-----------	---	---

Lazard (1984, 283) proposes a very similar combined scale of definiteness and humanness (= animacy) for “actance variation” (i.e. differential object marking, or DOM) in various languages.

- (62) Combined scale of definiteness and humanness

1	2	3	4	5	6
1.+2. pron.	3. pron Proper Names	Definite Human	Indefinite Non-human	Mass	Generic

It is still unclear what it means to have a scale with parallel contrasts in certain cells, as in (61) and (62). In these cases we have only a partial order and would need additional context to decide a global order.

4.3 Harmonic Alignment

Often different scales align in such a way that their higher values and their lower values more easily combine than a high value with a low value etc. This can be illustrated by the alignment of the Relational Scale (63) and the Animacy Scale (64) (see Aissen, 2000: 6):

- (63) Relational Scale: Subject > Object
 (64) Animacy Scale: Human > Animate > Inanimate

It is not possible to combine the two scales into one, but harmonic alignment (borrowed from Optimality Theory – see details in Aissen, 2000) allows us to modify each of the two parts in the

two-part scale in (63) by each of the values of the multi-part scale (64). Thus, harmonic alignment of grammatical function with animacy forms two scales, one on subjects and one on objects. Each expresses the relative markedness of possible associations with the various degrees of animacy.

(65a) Subject/Human > Subject/Animate > Subject/Inanimate

(65b) Object/Inanimate > Object/Animate > Object/Human

The scale on subjects (65a) expresses that human subjects are less marked than animate ones, which, in turn, are less marked than inanimate ones. The scale (65b) on objects shows the reverse, namely that inanimate object are less marked than animate objects, and so on. The latter scale (65b) describes one of the parameters of “differential object marking” (or DOM) in many languages (Bossong, 1985). Spanish realizes this DOM by the prepositional accusative (see section 3.2) and by clitic doubling (see sections 3.3-3.5).

It seems that harmonic alignment is appropriate if we combine a two-part scale with a many-part scale that have a similar orientation or markedness. However, if we have two many-part scales or two scales without an intrinsic orientation (such as specificity), we cannot rely on harmonic alignment. Furthermore, if we combine more than two scales, we have to include cross-classification, as in 4.2.

This very brief discussion of combining referential parameters has shown that we must carefully examine the parameters and the empirical facts in a language before we can apply one of the discussed compositional rules.

5. Summary and direction for further research

In this article, we have described different morphosyntactic phenomena in Spanish in terms of the interaction of the referential categories animacy, definiteness, and specificity. Each of these categories refers to a different cognitive-semantic level: Animacy is a lexical or cognitive feature of NPs, definiteness is a discourse-pragmatic property of the discourse item representing the NP, and specificity expresses a semantic property that determines the referent in a particular way. Even though these three categories are associated with quite different semantic domains, their grammatical reflexes are closely interrelated, as can be shown for the mood in relative clauses, the prepositional accusative, and clitic doubling in Spanish. Our analysis has not only provided a uniform description in terms of the interaction of the discussed referential categories, it has also demonstrated data that allows us to evaluate different theoretical models of combining animacy, definiteness, and specificity. Still, further research is necessary in all three areas: We need more detailed investigations of relevant morphosyntactic contrasts, better theoretical models of underlying referential categories, and more discussion about the interaction of referential categories.

6. References

- Aissen, Judith 2000. *Differential Object Marking: Iconicity vs. Economy*. Ms. UCSC.
- Bossong, Georg 1985. *Empirische Universalienforschung. Differentielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Narr.
- Brugé, Laura & Gerhard Brugger 1996. On the Acusative *a* in Spanish. *Probus* 8, 1-51.
- Colantoni, Laura 2002. Clitic doubling, null objects and clitic climbing in the Spanish of Corrientes. In: J. Gutiérrez-Rexach (ed.). *Words to Discourse: Trends in Spanish Semantics and Pragmatics*. Oxford: Elsevier, 321-336.
- Croft, William 1988. Agreement vs. case marking and direct objects. In: M. Barlow & C.A. Ferguson (eds.). *Agreement in Natural Language. Approaches, Theories, Descriptions*. Stanford: Center for the Study of Language and Information, 159-179.
- Egli, Urs & von Heusinger, Klaus 1985. The Epsilon Operator and E-Type Pronouns. In: U. Egli et al. (eds.). *Lexical Knowledge in the Organization of Language*. Amsterdam: Benjamins, 121-141.
- Enç, Mürvet 1991. The Semantics of Specificity. *Linguistic Inquiry* 22, 1-25.
- Farkas, Donka [1981] 1985. *Intensional Description and the Romance Subjunctive Mood*. New York: Garland.
- Farkas, Donka 1995. Specificity and Scope. In: L. Nash & G. Tsoulas (eds.). *Actes du Premier Colloque Langues & Grammaire*. Paris, 119-137.
- Fernández Soriano Olga. 1999. El pronombre personal. Formas y distribuciones. Pronombres átonos y tónicos. In: I. Bosque & V. Demonte (dir.). *Gramática descriptiva de la lengua española. 1: Sintaxis de la clases de palabras*. Madrid: Espasa Calpe, 1209-1273.
- Fodor, Janet & Sag, Ivan 1982. Referential and Quantificational Indefinites. *Linguistics and Philosophy* 5, 355-398.
- Heim, Irene 1981. *The Semantics of Definite and Indefinite Noun Phrases*. Ph.D. Dissertation. University of Massachusetts, Amherst. Ann Arbor: University Microfilms.
- Heim, Irene 1991. Artikel und Definitheit. In: A. von Stechow & D. Wunderlich (eds.). *Semantik. Ein internationales Handbuch der zeitgenössischen Forschung*. Berlin; New York: de Gruyter, 487-535.
- von Heusinger, Klaus 1997. Saliency and Definiteness. *Prague Bulletin of Mathematical Linguistics* 67, 5-23.
- von Heusinger, Klaus 2002. Specificity and Definiteness in Sentence and Discourse Structure. *Journal of Semantics* 19, 245-274.
- Hintikka, Jaakko 1986. The Semantics of 'a certain'. *Linguistic Inquiry* 17, 331-336.
- Jaeggli, Osvaldo 1982. *Topics in Romance Syntax*. Dordrecht: Foris.
- Jaeggli, Osvaldo 1986. Three issues in the theory of clitics: case, doubled NPs, and extraction. In: H. Borer (ed.). *The Syntax of Pronominal Clitics*. Orlando: Academic Press, 15-42.
- Kaiser, Georg A. 1992. *Die klitischen Personalpronomina im Französischen und Portugiesischen. Eine synchronische und diachronische Analyse*. Frankfurt: Vervuert.
- Kamp, Hans 1981. A Theory of Truth and Semantic Interpretation. In: J. Groenendijk & T. Janssen & M. Stokhof (eds.). *Formal Methods in the Study of Language*. Amsterdam: Amsterdam Center, 277-322.
- Kamp, Hans & Reyle, Uwe 1993. *From Discourse to Logic. Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Dordrecht: Kluwer.
- Kany, Charles E. 1951. *American-Spanish Syntax*. Chicago; London: The University of Chicago Press, second edition.
- Karttunen, Lauri 1976. Discourse Referents. In: J. McCawley (ed.). *Syntax and Semantics 7: Notes from the Linguistic Underground*. New York: Academic Press, 363-385.
- Kayne, Richard S. 1975. *French Syntax. The Transformational Cycle*. Cambridge: The MIT Press.
- Laca, Brenda 1987. Sobre el uso del acusativo preposicional en español. *Romanistisches Jahrbuch* 38, 290-312.
- Lazard, G. 1984. Actance Variations and Categories of the Object. In: F. Plank (ed.) *Objects: Towards a Theory of Grammatical Relations*. London: Academic Press, 269-292.
- Leonetti, Manuel 1990. *El artículo y la referencia*. Madrid: Taurus.
- Leonetti, Manuel 1999. El artículo. In: I. Bosque & V. Demonte (eds.). *Gramática descriptiva de la lengua española. 1: Sintaxis básica de las clases de palabras*. Madrid: Espasa Calpe, 787-883.
- Leonetti, Manuel 2003. Specificity and Object Marking: the Case of Spanish *a*. This volume, S. 67-101.
- Lewis, David 1979. Scorekeeping in a Language Game. In: R. Bäuerle & U. Egli & A. von Stechow (eds.). *Semantics from Different Points of View*. Berlin; Heidelberg; New York: Springer, 172-187.
- Löbner, Sebastian 1985. Definites. *Journal of Semantics* 4, 279-326.
- Milsark, Gary 1974. *Existential Sentences in English*. Ph.D. Diss. MIT, Cambridge/Mass.

- Parodi, Teresa. 1998. Aspects of clitic doubling and clitic clusters in Spanish. In: R. Fabri, O. Albert & Teresa Parodi (eds.). *Models of Inflection*. Tübingen: Niemeyer, 85-102.
- Pérez Saldanya, Manuel 1999. El modo en las subordinadas relativas y adverbiales. In: I. Bosque & V. Demonte (eds.). *Gramática descriptiva de la lengua española. 2: Las construcciones sintácticas fundamentales. Relaciones temporales, aspectuales y modales*. Madrid: Espasa Calpe, 3253-3322.
- Real Academia Española 1973. *Esbozo de una gramática de la lengua española*. Madrid: Espasa Calpe.
- Rivero, Maria Luisa 1975. Referential Properties of Spanish Noun Phrases. *Language* 51, 32-48
- Russell, Bertrand 1905. On Denoting. *Mind* 14, 479-493.
- Silverstein, Michael 1976. Hierarchy of Features and Ergativity. In: R. Dixon (ed.) *Grammatical Categories in Australian Languages*. Canberra: Australian Institute of Aboriginal Studies, 112-171.
- Suñer, Margarita 1988. The role of agreement in clitic doubled constructions. *Natural Language and Linguistic Theory* 6, 391-434.

Specificity and Object Marking: the Case of Spanish *a*

Manuel Leonetti

manuel.leonetti@uah.es

1. Preliminaries

It is well known that several languages use certain grammatical devices as object markers that are in some sense associated to the *Specific / Non-Specific* distinction. Spanish is one of them, and Romanian, Turkish, Persian or Hindi are frequently mentioned as other representative cases of the general phenomenon which, following Georg Bossong's proposal, we call *Differential Object Marking* (DOM)¹. My aim in this paper will be that of determining what is the particular contribution of DOM to utterance interpretation in Spanish, and how specificity is related to it. I would like to begin by presenting some basic assumptions on the Semantics / Pragmatics interface and on the notion of specificity, to concentrate later on the analysis of the Spanish prepositional accusative. I intend to show that, although the correlation between the accusative marker and specificity is far from clear, basically because animacy –and not specificity- is the dominant trigger for DOM in Spanish, a number of facts still indicate that the prepositional accusative tends to be associated with specific readings, in a way which is not unrelated to what happens in scrambling and clitic doubling constructions. Far from deriving from some [+ specific] feature inherent in the meaning of *a*, such facts can be shown to be a consequence of a different basic semantic feature that should allow us to bring together most of the grammatical phenomena that are sensitive to specificity. Some closing remarks on specificity in grammatical structure will sum up the discussion.

I will try to present the general hypothesis that when natural languages encode specificity, they do it inside DP structure only (basically by means of determiners), and not in other positions; as a consequence, other alleged specificity markers outside DP structure are actually modality indicators or information structure markers. In these cases, the markers simply force a specific reading in certain DPs without encoding specificity themselves. Clarifying this issue should allow us to have a more precise view of the role of specificity in grammatical systems.

I am grateful to the participants at the workshop "Semantic and Syntactic Aspects of Specificity in Romance Languages" (Konstanz, October 2002) for useful comments, and to Vicky Escandell-Vidal and Olga Fernández-Soriano for helpful discussion on an earlier draft. Thanks also to Begoña Vicente for patiently correcting my English. All remaining errors are exclusively mine. This research is supported by the Spanish Ministerio de Ciencia y Tecnología through a DGICYT project (PB98/0707 "Gramática e interpretación en la teoría de la relevancia").

¹ See Bossong (1997) and Aissen (2000) for an overview.

The basic assumptions are as follows:

a) First, I assume Sperber and Wilson's distinction between *linguistic meaning* (*logical form*, in Relevance-theoretic terms) and *explicatures* as two different semantic levels (Sperber and Wilson 1986). While linguistic meaning yields a linguistically encoded semantic representation –the output of the grammatical system– and only provides an incomplete schema for the inferential construction of a full proposition, explicatures are the propositions explicitly communicated by an utterance and are obtained by fleshing out the logical forms encoded in it. Explicatures are derived from logical forms by linguistic decoding, on the one hand, and by pragmatic inference based on contextual information, on the other. In short, we have two levels of meaning: one linguistically encoded in lexical items or in syntactic positions, and the other inferentially obtained from the first, thus constituting a partially pragmatically derived level of representation.

It is important to clarify that it is explicatures, and NOT linguistic meaning, that formal semanticists work on when they try to give an account of the semantic content of utterances. But we should bear in mind that an important part of the explicature is pragmatically inferred, and thus is not to be considered as linguistically encoded meaning. A brief example will be enough to illustrate this theoretical distinction. In (1) there is a typical case of associative anaphora:

- (1) It is a very good restaurant, but the prizes are too high.

The definite DP *the prizes* is interpreted as anaphorically related to the DP *the restaurant*, but they are not coreferential. The anaphoric link between them is easily recovered by any speaker from the encyclopaedic knowledge corresponding to the frame activated by the lexical item *restaurant*. Such a link becomes a part of the propositional representation –the explicature– of the utterance (i.e. "... but the prizes of the aforementioned restaurant are too high"). But we would not want to say that the anaphoric dependence is encoded in the meaning of the definite article *the* (even if we were to accept the familiarity hypothesis for the linguistic meaning of *the*). This dependence is in fact inferred on the basis of some general pragmatic principle. In any case, recovering an associative anaphoric link in (1) is the easiest way to obey the semantic instruction encoded by *the*. The article simply encodes an instruction to find a referent that is uniquely identifiable (and this is linguistic meaning), and the remaining part of the interpretation, i.e. the recovery of the associative link between *restaurant* and *prizes* and the anaphoric dependence, is pragmatically inferred in the development of the logical form for (1) into an explicature. This division of labour between semantics and pragmatics is one of the assumptions I will rely on in the analysis of specificity. The crucial question is, thus, what is encoded by the grammatical system and what is inferred, when we talk about specific readings of DPs.

b) The second basic assumption concerns the status of indefinites, in particular specific indefinites. I will assume that they are not lexically ambiguous, despite their exceptional scope properties, and that, unlike definites, they do not guarantee that there is a uniquely identifiable referent, so that "the only condition... (they) impose is that the value assigned to their discourse referent be an element of the set denoted by the description" (Farkas

2001b:90). Indefinite DPs may or may not contribute conditions that specify such a value: specific readings arise when the choice of referent is presented as heavily constrained, and non-specific readings arise when the choice is presented as being relatively free. As Farkas (2001b:85) puts it, "...specificity turns out to be an epiphenomenon connected to a family of distinctions that are marked differently in different languages". I am especially interested in this characterization of specificity as an **epiphenomenon**, because it is often not encoded by grammar or by lexical items, but just inferred in the development of propositional explicatures. From this perspective specificity is described as a derived pragmatic effect, and not as an interpretable feature. Of course, I do not intend to deny the semantic nature of specificity in certain cases: there are languages that show clear manifestations of the *specific / non-specific* distinction in their determiner systems, and Haspelmath (1997) and Lyons (1999), for instance, present several examples of this. But in many languages specificity is not a feature of the linguistic semantics of any item or construction. Thus, languages can display so called "specificity effects", but that does not imply that specificity is a feature of their semantics. The point I want to make is that the situation we find with the notion of specificity is essentially the same that we find with the notion of definiteness: it is not encoded in the same way and to the same extent in all languages. In languages that lack definite articles, like Russian, definiteness has to be inferred on the basis of different grammatical markers (case markers, word order, agreement...), and the same happens with specificity in many of the cases we are analysing. The crucial question is: when is specificity grammatically or lexically encoded, and when is it inferred?

At least since Farkas (1995), it is usual to distinguish among different types of specificity: partitive specificity, epistemic specificity, scopal specificity, and we could add "relative specificity", following von Stechow (2001a&b). I would like to discuss this issue briefly in order to make my point clearer. First, with respect to partitive specificity, and against Eng (1991), I believe that partitivity cannot be the central property in a general definition of *specificity*, simply because it is not necessarily included in every specific interpretation (Abbott 1995); it is rather a way to obtain a specific interpretation when syntax forces it and contextual information makes available the recovery of a given domain of quantification. Thus, partitivity is only one of the factors favouring specificity, without precluding, nevertheless, the possibility that an indefinite partitive DP might have a non-specific reading (cf. *Usaré una de esas sartenes* "I am going to use one of those saucepans").

As for epistemic specificity, I agree with the observation that having something in mind or being able to identify it are not necessarily prerequisites for specificity (Geurts 2001, von Stechow 2001a&b). I follow Rouchota (1994) in assuming that, instead of the speaker's mental state, what is decisive for specificity is the speaker's intention to make manifest to the audience that the DP is employed to refer to a particular referent.

The definition of specificity in terms of scope faces well known problems too: the *specific / non-specific* distinction may arise in contexts where indefinites do not interact with any operator or quantifier. Of course, the speaker's knowledge of the referent and the indefinite having wide scope with respect to other elements are still very common properties of specific expressions, but they are not obligatory ingredients of specificity. Thus it seems quite difficult to determine what it means for an indefinite to be *specific* on the basis of one of these properties alone. I would rather use a very general notion of specificity that covers the

mentioned properties without taking one of them as the basic and central one, for two reasons: first, these specificity types are nothing more than contextual results of the inferential development of the linguistic semantics in indefinites; second, none of the grammatical devices I will discuss is exclusively linked to a particular type of specific interpretation, as far as I can tell.

c) The third basic assumption is that, as many authors have pointed out (Jäger 1995a, Erteschik-Shir 1997, Yeom 1998, Büring 2001, Portner and Yabushita 2001, Geurts 2002), specific readings of indefinites are in some way related to the topical status of the DP. Two very influential works paved the way for such a point of view: Enç (1991) and Diesing (1992). These two contributions popularised the idea that specificity is essentially connected with discourse linking or presuppositionality; I have already mentioned some reasons for not accepting such a restrictive perspective, but the core proposal was no doubt loaded with far-reaching, interesting consequences. In the nineties the idea spread that Diesing's findings were basically a matter of *information structure*, and this caused the notion of *topic* to become a crucial concept for the analysis of specificity in several contexts. So the central question should now be *Why is the concept of "topic" important for an understanding of specificity?* I assume that it is important, although not exactly for the same reasons that one finds in the literature: instead of taking for granted that all specific indefinites are topical, or presuppositional, or background information, I would simply say that the topical status forces the specific interpretation of an indefinite². This is a crucial assumption when looking at the interpretive effects of syntactic devices such as scrambling, clitic doubling, subject preposing and DOM. I will come back to this issue later (section 3).

Once these assumptions have been presented, I can proceed to the analysis of the prepositional accusative in Spanish and its interpretive role.

2. *a* + direct object

2.1. In favour of *a* as specificity marker

It is a well known fact that specificity is often reflected in the morphological marking of the direct object, and Spanish is usually presented as a language that distinguishes specific objects and non-specific objects by means of the preposition *a*. *A* is the same preposition used for indirect objects or datives; this is a widespread feature among the languages that resort to prepositions for DOM. It is true that the insertion of *a* has significant effects on specificity, but the whole issue deserves a detailed analysis before we can clarify the nature of the contribution of *a* to utterance interpretation. Let me first present some data that seem to indicate that the prepositional marker is indeed a specificity indicator. The examples are mainly taken from Brugè (2000), Brugè and Brugger (1996), Laca (1987), Leonetti (1999) and Torrego (1998) and (1999):

² In certain contexts the topical status forces a generic reading, instead of a specific one, but recall that both specific and generic interpretations of indefinites belong to the family of *strong interpretations*, while non specific interpretations are typically *weak*. So it is the strong / weak distinction that actually correlates with topic-focus structure. See Cohen and Erteschik-Shir (2002) for an extension to the interpretation of bare plurals in English.

a) *a* in opaque contexts

Examples like those in (2) are the most frequently mentioned argument for a connection between the prepositional accusative and specificity. They seem to indicate that the reading of the object DP is in fact specific with *a* and non-specific without *a*³:

- (2) a. Necesita a una enfermera que pasa la mañana con ella /
 Heneeds to a nurse that spends_{S_{IND}} the morning with her
 Necesita una enfermera que pase la mañana con ella
 Heneeds a nurse that spends_{S_{SUBJ}} the morning with her
 “He needs a nurse {that spends the morning with her / to spend the morning with her}”
- b. *Necesitan a camarero / Necesitan camarero
 They need to waiter / They need waiter
 “They need a waiter”
- c. Busca a un médico/ Busca un médico
 (S)he looks for to a doctor / S(he) looks for a doctor
 “(S)he is looking for a doctor”

Notice that in (2a) the mood of the subordinate verb in the relative clause reinforces the effect of DOM: the object is preceded by *a*, and specific, in the first case, where the verb is in the indicative mood, whereas the absence of *a* makes the object non-specific in the second case, where the verb is in the subjunctive mood. The contrast in (2b) is due to the incompatibility between *a* and a bare singular noun, which is an expected result, if we assume that bare nouns can never receive a specific reading and therefore should not be preceded by *a*. Finally, the contrast in (2c) is a classical example of the correlation of *a* with a specific reading, and of the absence of the preposition with a non-specific one. Nevertheless, it must be pointed out that *a* is not incompatible with a non-specific reading in *Busca a un médico*, and so the sentence remains ambiguous.

b) partitive constructions

It has been repeatedly pointed out that partitive object DPs take *a* obligatorily, as is shown in (3) (from Brugè and Brugger 1996):

- (3) He visto *(a) muchas de esas estudiantes
 I have seen *(to) many of those students
 “I have seen many of those students”

This is a natural consequence of the typically specific interpretation that partitives receive. Since partitives include a definite or familiar domain of quantification for the indefinite, they are expected to behave like definite DPs with respect to accusative marking (Recall that definites always require the insertion of *a*, except in certain generic interpretations).

³ From now on I will translate *a* as *to* in the English word-by-word version of the examples.

c) incompatibility with existential constructions

The construction with *haber* plus postverbal DP represents the canonical existential context in Spanish, with its associated definiteness effect. If *haber* takes only weak DPs and specific indefinites are a type of strong DP, then one should expect that *a*, being a specificity marker, should be ungrammatical in such contexts, and in fact it is (this is a very clear and strong restriction, in the sense that *a* is completely excluded).

- (4) Había (*a) una enfermera
 There was (*to) a nurse
 “There was a nurse”

Another existential context that gives rise to a definiteness effect is provided by the verb *tener*, “have” (cf. *?Este coche tiene el airbag* “This car has the airbag”), and again it excludes the presence of *a*, as shown in (5):

- (5) Ella tenía (*a) unhermano
 She had (*to) a brother
 “She had abrother”

In this case both the restriction against definites and the correlated restriction against the presence of *a* disappear when *tener* is followed by a secondary predication structure, an effect I will come back to later:

- (6) Este coche tiene el airbag estropeado/Ellatenía a un hermano en la cárcel
 This car has the airbag broken/ She had to a brother in the prison
 “This car’s airbag does not work.” / “She had a brother in prison.”

The data in (4), (5) and (6) indicate that there is a systematic correlation between the presence of *a* and the definiteness effect, and this is a relevant fact because definiteness restrictions in general have to do with specificity or some related notion, as indicated by Enç (1991) (Recall that partitive DPs are not fully acceptable in existential contexts, which allows me to include the data in (3) under the same generalization that covers (2) and (4)-(6)). Thus, the provisional conclusion should be something like this: *a* appears with specific DPs (or strong DPs) –in a very general sense of *specific*- . This reflects the central intuition that is usually mentioned in Spanish grammars, although it is obviously not the whole story about *a*.

2.2 Against *a* as specificity marker

The main problem for a characterization of *a* as a specificity marker is also well known: as many authors have concluded, the basic feature the prepositional accusative is correlated to is **animacy**, and not specificity or referentiality (see, for instance, Brugè and Brugger 1996). The strongest argument we can offer in support of this idea is the possibility of inserting *a* with non-specific indefinite DPs that still require the preposition because of their animacy

feature. The examples in (7) to (10) illustrate the combination of *a* with non-specific indefinites:

- (7) a. Está buscando a alguien / No está buscando a nadie
 (S)he is looking for to someone / (S)he is not looking for to anyone
 “(S)he is looking for someone” / “(S)he is not looking for anyone”
- b. Necesitan (a) un ayudante que sepa inglés
 They need (to) an assistant that speaks_{SUBJ} English
 “They need an assistant that speaks English”
- (8) a. Cada estudiante entrevistará a un personaje conocido
 Each student will interview to a celebrity
 “Each student will interview a celebrity”
- b. Todas las niñas admiraban a algún cantante
 Every child (fem) admired to some singer
 “Every child admired some singer”
- (9) a. Toda persona que contrata (a) un inmigrante...
 Every person that hires (to) an immigrant worker...
 “Every person that hires an immigrant worker...”
- b. Conoces (a) muchas personas para llevar aquí tan poco tiempo
 You know to a lot of people to be here not much time
 “You know a lot of people considering you haven’t been here for long”
- (10) a. La empresa ha contratado a trabajadores con experiencia
 The company has hired to workers with experience
 “The company has hired experienced workers”
- b. No conozco a candidatos con esas características
 I don’t know to candidates with those features
 “I don’t know candidates with those features”

In (7) *a* precedes indefinite pronouns like *alguien* and *nadie* that are interpreted non-specifically (but notice that they denote human beings only), or indefinite DPs with subjunctive relative clauses -one of the classical ways of signalling non-specific readings in Romance languages- (again with human *denotata*). In (8) the prepositional marker is obligatory and still the object DPs (with human *denotata*) can be naturally interpreted as narrow scope and as non-specific indefinites. In (9), following Bosque (2001) and Sánchez López (1995), I have reproduced two grammatical contexts that require non-specific indefinites, namely a “donkey sentence” fragment (9a) and an indefinite DP with a concessive

subordinate clause (9b)⁴; in both cases *a* is perfectly grammatical (with human *denotata*). And finally in (10) *a* combines with bare plurals, which are not supposed to get specific readings (at first sight, this is contradictory with what we observe in (2b), i.e. that bare nouns cannot be preceded by *a*, but the crucial factor is *singularity vs plurality*: bare singulars show a more constrained distribution than bare plurals).

We see then that there are good grounds for rejecting accounts of *a* as a specificity marker and supporting instead animacy as the semantic feature governing its insertion. Nevertheless, some link with specificity must still exist, if the data in the previous section are correct. So the question is how to integrate two sets of facts that seem to be contradictory in a coherent explanation.

For a correct understanding of the semantic contribution of *a* we should bear in mind three general points.

First, it seems natural to assume that the linguistic meaning encoded by the preposition is some abstract feature that underlies the interaction or combination of specificity and animacy (given that a mixed scale of definiteness / specificity and animacy controls the distribution of DOM in Spanish and many other languages); thus the marker cannot simply mean “animate direct object” or “specific direct object”. Something else must be responsible for the interpretive effects of accusative morphology, something that underlies other syntactic phenomena that systematically involve both animacy and specificity, like clitic doubling and object shift.

Second, the puzzling situation described above is partly due to the obligatory nature of accusative marking with several verbs in Spanish, and partly to the predominant role of animacy. In the literature on the topic it has been frequently observed that certain verbs require *a* obligatorily and others just allow its insertion without imposing it (see for instance Pensado 1995:33-35 and Torrego 1998:23). Thus, verbs like *saludar* (“greet”), *odiar* (“hate”), *insultar* (“insult”), *castigar* (“punish”), *sobornar* (“bribe”) or *atacar* (“attack”) impose *a* on their animate objects, but verbs like *encontrar* (“find”), *buscar* (“look for”), *esconder* (“hide”) or *ver* (“see”) admit non-overtly case-marked objects. The two classes of verbs differ in several respects, basically in their aspectual properties, the affectedness of the object and the agentivity of the subject –three of the usual parameters that control the presence of *a*. Bearing in mind this fact may help us to look again at the examples in (7) to (10).

As for the use of *a* with non-specific indefinites, it is important to notice that the verbs in the examples in (8), *entrevistar* and *admirar*, require the obligatory presence of *a*, which implies that the semantic contribution of the marker in these cases is no longer relevant for interpretive purposes. When *a* is automatically selected by the verb, it becomes an inert, non-distinctive morphological device: this is one of the reasons why *a* occurs with non-specific indefinites as well. Cases of partial grammaticalization of object markers are attested in other languages with identical effects⁵: Lidz (1999), for instance, shows that obligatorily case-

⁴ This kind of subordinate clause is licensed by the presence of indefinites like *mucho(s)* “many” or *demasiado(s)* “too much / too many” only if they are non-specific.

⁵ The process of grammaticalization of the object marker is slowly spreading in Spanish, and certain dialects are extending the use of *a* even to some inanimate and abstract nouns, according to Company (2002). This further complicates the analysis of its interpretive effects. See Aissen (2000: §5.4) for an overview of the historical expansion of DOM in Spanish.

marked objects in Kannada have the same range of interpretations as non-case-marked objects, a fact that seems to be a result of the historical expansion of DOM in the language, and Franco and Mejías-Bikandi (1999:115), focusing on clitic doubling in Basque Country Spanish, state that

“Overt object morphology, whether case-marking morphology on the noun, or interpretation of indefinite objects in those cases where object morphology is object agreement morphology on the verb, has a systematic effect on the apparently optional.”

From a wider perspective, it can be observed that the obligatory nature of a syntactic rule usually cancels the semantic effects the rule could have. As Adger (1996:117) puts it in his analysis of the interpretation of subjects in Italian,

“...optional movement seems to correlate with different interpretations for the resulting structures; when movement is obligatory, on the other hand, the single resulting structure seems to have both of the possible interpretations assigned to the two structures given by optional movement.”

Such considerations on movement rules apply, *mutatis mutandis*, to case-marking as well. As a consequence, the semantic contribution of *a* (and its association with specificity) is not systematic across all contexts. In order to throw some light on the specificity issue, then, the contexts that are worth concentrating on are those where there is a possibility to choose between using and non using the preposition: only there will the basic properties of *a* be revealed. In the next sections I will focus on such contexts.

In (7), (9) and (10) *a* is not obligatorily selected by the verb, but again we find no systematic semantic effects: with the intensional verbs in (7) it allows non-specific readings, and with extensional verbs such as *contratar* (“hire”) and *conocer* (“know”) in (9) and (10) the non-specific reading is in fact the only possibility. In the first case the presence of *a* is forced by the [+animate] feature of indefinite pronouns like *alguien* (“someone”) and *nadie* (“noone”), which blocks the choice between *a* and \emptyset , thus cancelling the semantic contribution of the preposition. In the second case the sentential context excludes any plausible specific interpretation, with similar consequences for the contribution of *a*: in (9a) it is genericity that precludes a specific interpretation of the indefinite object, in (9b) it is the insertion of the concessive clause, and in (10) it is the impossibility to assign a specific interpretation to a bare plural. Why does *a* occur, then? Because, as mentioned before, animacy overrides all other factors in the DOM system in Spanish, and all the indefinite objects in the example are [+animate].

Finally, it must be noticed that even in the limited contexts where the choice between *a* and \emptyset is allowed there is usually not an absolute contrast, but rather a sort of “privative opposition”: while *a*-case-marked objects may admit both strong and weak interpretations in most cases, non-case-marked objects take only weak interpretations. In (11), for instance, a specific reading arises only with *a*, as has often been pointed out, but the non-specific reading is available both with *a* and without it.

- (11) Necesitaban (a) un especialista.
 They needed(to) a specialist
 “They needed a specialist”

The privative opposition between *a* and \emptyset is hardly surprising, if one thinks of similar facts like those we find in the interpretation of preverbal and postverbal indefinite subjects in Romance: when both positions are available for the subject, i.e. typically with unaccusative predicates, the preverbal one tends to force strong readings for indefinites, while the postverbal one allows strong or weak readings, so that there is one kind of reading that is compatible with the two positions (Adger 1996).

Facts like these show how the semantic contribution of *a* is obscured and distorted both by the predominance of animacy and by the selection properties of verbs. Bearing in mind such issues, we should now address two crucial questions: a) Is there a notion that allows us to put together specificity and animacy in DOM systems in some principled way? And b) How can we deal with the (non-systematic) connection that still exists between *a*-marking and specificity, i.e. why are specific interpretations linked in some way to *a*-marked objects?

2.3 *a* as a Topic Marker

A path that is worth exploring to find a suitable answer to the aforementioned questions is the one that links specificity and topicality. Several authors have resorted to the notion of *Topicality* in their research on specific readings of indefinites. This allows me to exploit an obvious connection with the idea that *a* is actually a sort of topicality marker, an idea which is not new and has been already defended in Laca (1987). Before reviewing the arguments offered by Laca and some other additional data, it is important to recall that *topic* is being used here in the “aboutness” sense, i.e., as an anchor for new assertions, rather than in the sense of “familiar, given or old information”. As referentially independent expressions, topics introduce prominent participants in the discourse. A topical DP indicates that the individualization of the referent is relevant for utterance interpretation. Given this, if, on the one hand, it is possible to show that specificity is an effect of topicality and, on the other, there are reasons to consider *a* as a topic marker, a plausible answer could be given to our questions. In what follows I am going to review the reasons for taking *a* to be a topic marker.

2.3.1 Clitic Left-Dislocation

First of all, as Laca (1987), Pensado (1995) and Melis (1995) have pointed out, the preposition is overwhelmingly present in syntactic topicalization structures, i.e., in clitic left-dislocation constructions, even when the non-topicalized counterpart of the sentence allows both a marked or an unmarked object, as in (12) (cf. *Ya conocía (a) muchos estudiantes*; *Habían incluido (a) dos catedráticos en la lista*):

- (12) a. *(A) muchos estudiantes, ya los conocía
 *(To)many students, I already knew them
 “Many students I already knew”
- b. *(A) dos catedráticos, los habían incluido en la lista
 *(To)two professors, they them included in the list
 “Two professors they included in the list”

The obligatory presence of *a* in (12) is related to the fact that clitic left-dislocation tends to impose strong interpretations on DPs, and such interpretations are typically associated to *a*. The fact is that *a* is indeed correlated with clitic left-dislocation, which supports the idea that it behaves like a topic marker. According to Pensado (1995), clitic left-dislocation with personal pronouns is in fact the origin of *a* as a case-marker for objects in Modern Spanish; thus, *a* would originally be a topicalization mechanism, and the strong connection it shows now with clitic doubling is a related effect.

2.3.2 Bare plurals

Another significant piece of evidence has to do with bare plurals. Spanish does not admit them in preverbal subject position, unless they are modified by restrictive modifiers (adjectives, relative clauses, prepositional complements) or marked as contrastive focus; see the data in (13):

- (13) a. ??Guerrilleros atacaron ayer un puesto de policía en.../
 Guerrilla attacked yesterday a police station at.../
 Guerrilleros de las FARC fuertemente armados atacaron ayer un
 FARC guerrilla strongly armed attacked yesterday a
 puesto de policía en...
 police station at...
 “Guerrilla attacked a police station yesterday at... / Strongly armed FARC
 guerrilla attacked a police station yesterday at...”
- b. *Ratas han entrado por este agujero /
 Rats have entered through this hole /
 RATAS han entrado por este agujero
 RATS have entered through this hole

A striking parallelism between preverbal subject position and case-marked object position is that the same factors seem to license bare plurals in both of them: in fact, bare plurals are excluded in *a*-marked objects, unless they include some kind of restrictive modifier or are focused, as the examples in (14) show (bare plurals may occur as unmarked objects even when they do not obey such conditions).

- (14) a. ??Conocemos a profesores / Conocemos a profesores que se pasan el
 We know to teachers/ We know to teachers that spend the
 finde semana trabajando
 weekend working
 “We know teachers / We know teachers that spend their weekend working”
- b. ??Detuvieron a hinchas / Detuvieron a hinchas peligrosos del
 They arrested to supporters / They arrested to supporters dangerous of
 Atlético
 Atlético
 “They arrested supporters / They arrested dangerous Atlético supporters”
- c. ??En el poblado vi a pescadores / En el poblado vi a
 In the village I saw to fishermen/ In the village I saw to
 PESCADORES, no a turistas extranjeros
 FISHERMEN, not to tourists foreign
 “In the village I saw fishermen / In the village I saw FISHERMEN, not
 foreign tourists”

Following Laca (1996), I assume that the constraints on bare plurals in preverbal subject position are due to the topical nature of such a position and to the impossibility of interpreting bare plurals as topics –with strong readings– in sentence-internal positions in Romance languages (but see Moreno and Pérez 2001 for some exceptions). If this is basically correct⁶, then the obvious way to explain what happens in marked objects vs unmarked objects is to consider marked objects as topical positions in some sense: their topical nature is the reason why they exclude unmodified bare plurals. The immediate advantage of this account is that it captures the basic aspects of the distribution of Spanish bare plurals in a simple and unified way.

2.3.3 *Secondary predicates*

Another parallelism with subjects is the correlation of the presence of *a* and the occurrence of a secondary predicate, even when the object has no [+animate] feature; in such cases the object is taken as the predication topic, and the preposition seems to act like a topic marker. I reproduce some examples from Laca (1987) where *a* forces the hearer to assign a secondary predication structure to the sentence:

⁶ For a different approach based on configurational principles, see for instance Brugè (2000). On the occurrence of bare plurals with *a*, see Martín (1999), Bleam (1999a, 1999b) and Torrego (1999).

- (15) a. La tormenta dejó a treinta heridos y a muchos arruinados
 The tempest left to thirty wounded and to many ruined
 “The tempest left thirty persons wounded and many ruined”
- b. Juan tiene a un hermano enfermo
 Juan has to a brother ill
 “Juan has a brother that is ill now”

The generalization that emerges from the observation of all these facts is that marked objects behave in many respects like subjects, in particular topical subjects. Such an analogy with subjects is actually at the heart of the DOM phenomenon: as Aissen (2001) puts it, the higher in prominence a direct object is (in the scales of animacy and definiteness), the more likely it is to be overtly case-marked, which means that the closer it is to the typical properties of subjects (referential autonomy, agentivity), the more likely it is to be overtly case-marked (Laca 1987: 72-74). Topicality seems to be a plausible way to characterize what underlies the interaction of animacy and definiteness, as well as the similarities between subjects and marked objects.

2.3.4 *Individualization and genericity*

A look at the sentence pairs in (16)-19) confirms that the presence of *a* has an effect on the interpretation of the object DP, which is not always clearly reducible to specificity (notice that the preposition is not obligatory in these sentences, except in (19a))⁷.

- (16) a. Juan ha visto a muchas chicas
 Juan hasseen to many girls
- b. Juan ha visto muchas chicas
 Juan has seen many girls
- (17) a. Juan mató a un tigre
 Juan killed to a tiger
- b. Juan mató un tigre
 Juan killed a tiger

⁷ The contrasts in (16) and (19) are taken from Brugè (2000), and the one in (17) from Laca (1987).

- (18) a. Estaba dibujando a una niña
(S)he was drawing to a child
- b. Estaba dibujando una niña
(S)he was drawing a child
- (19) a. Pilar siempre contrata *(a) un chico cuando es guapo
Pilar always hires to a boy when he is handsome
- b. Pilar siempre contrata (a) una canguro cuando se va de viaje
Pilar always hires to a babysitter when she travels

According to Brugè (2000), the presence of *a* in (16a) forces a D-linked reading of the object, a reading that is impossible in (16b). I would rather say that the D-linked (or partitive) reading of *muchas chicas* seems to be just one possible interpretation for (16a), but not the only one - though probably the most natural one-. The girls mentioned in the example could be a group of specific girls, but not necessarily taken from an already mentioned set. In any case, *a* clearly favours a strong reading of the indefinite object, be it D-linked or partitive, or specific in any other sense, while its absence in (16b) is associated to a weak or existential reading. What is at stake here as well as in (17), as several authors had already noted, is the emphasis on the individualization of the referent triggered by *a*, compared to the emphasis on quantity or descriptive content that predominates in unmarked objects. So, when *a* appears, the relevant features are the independent, autonomous reference of the object DP, and its discourse prominence. As for (17a), Laca (1987) points out that the emphasis on individuation of the referent creates an expectation for additional information about it, thus presenting a specific tiger as a possible discourse topic; on the contrary, in (17b) there is simply a “tiger-killing” event. The theoretical distinction we need to capture the subtle semantic contrasts in (16) and (17) is *strong / weak*, and in both examples we are dealing with specificity in some sense.

In (18), on the other hand, *a* is used to distinguish one interpretation equivalent to *She was portraying a child* from another equivalent to *She was drawing a child*. A very similar contrast obtains in *Describió (a) una chica de ojos castaños* (“(S)he described a girl with brown eyes”). The problem had been already studied in Fauconnier (1984) as part of an analysis of referential ambiguities based on the notion of *mental spaces*. In a few words, the difference between (18a) and (18b) is related to the world where the child exists: the real world, in the *a* version, and the world of the drawing, in the non-*a* version. The natural way to interpret (18a) is to assume that the speaker is referring to a particular child; in (18b), the hearer/reader resorts to a different interpretation in which an event of “child-drawing” is being reported. It is a matter of some discussion whether this should be treated as an instance of the *specific / non-specific* distinction: on the one hand, it reminds us of the usual intensional contexts where the existence of an entity may be placed inside the world of beliefs, desires and expectations of some individual, or outside it, but on the other the opaque reading of (18b) does not seem to display all the usual features of non-specific indefinites (for instance,

it differs in its anaphoric properties). Nevertheless the contrast can be related to the individualization contrasts in the previous examples.

The facts I really want to draw attention to, to my knowledge first mentioned in Brugè (2000:272), have to do with the interaction between *a* and an adverbial quantifier like *siempre* (“always”), as in (19). The contrast involves the obligatory / optional presence of *a*, and the crucial factor to account for it is the status of the temporal subordinate clause introduced by *cuando* (“when”). In (19a), the temporal clause indicates that the interpretation of *un chico* (“a boy”) must be generic, given that a) *ser guapo* is an individual-level predicate and denotes a defining property, b) an indefinite subject with an individual-level predicate is typically generic, and c) the null subject of the temporal clause takes *un chico* as its antecedent. The essential condition for the generic interpretation of an indefinite DP is to be a topic, i.e. to be projected in the restrictive clause in the logical form of the sentence (Cohen & Erteschik-Shir 2002); if such a condition is not met, an indefinite in a generic context is likely to be interpreted as non-specific, and not as generic. The interesting question with respect to (19a) is why should *a* be obligatory: at first sight it may seem surprising, because *contratar* is one of those verbs that do not require the obligatory insertion of *a*, and the indefinite DP cannot receive a specific reading. The key to this puzzle is the generic interpretation of *un chico*: briefly, *a* is required for the generic interpretation to arise. Its semantic contribution lies in its role as a trigger for the mapping of the object into the restrictive clause in the logical structure. The preposition is thus indirectly constraining the specification of the value assigned to the discourse referent.

As for (19b), where the temporal clause introduces an episodic predicate (*irse de viaje*) and its null subject takes *Pilar* –and not the object– as its antecedent, no generic interpretation arises in the indefinite object *una canguro* (“a babysitter”). The absence of *a* correlates with a weak reading of the indefinite, while its presence is compatible with both strong (specific) and weak readings. Once the necessity to license a generic reading in the object disappears, the prepositional marker is again optional, as in (16) or (17). The relevance of the contrast in (19) lies in the fact that in (19a) *a* has to be inserted in order to get a generic reading of the object –not a specific one–. Moreover, this is a systematic fact that can be observed also in the following examples (built on the small group of verbs that allow the two options, *a/∅*, with animate objects):

- (20) a. La junta escoge *(a) un conferenciante extranjero si es
 The board chooses *(to) a foreign speaker if he is
 de reconocido prestigio
 really prestigious
 “The board chooses a foreign invited speaker if he is really prestigious”
- b. La junta escoge (a) un conferenciante si el congreso se celebra en
 The board chooses (to) a speaker if the conference is held in
 Madrid
 Madrid
 “The board chooses a speaker if the conference is held in Madrid”
- (21) a. Sólo admitimos*(a) un profesor nuevo cuando tiene el título superior
 We only accept *(to) a teacher new when he has a degree
 “We only accept a new teacher when he has a degree”
- b. Sólo admitimos (a) un profesor cuando hay una plaza vacante
 We only accept (to) a teacher when there is a vacant position
 “We only accept a teacher when there is a vacant position”

Again, if the sentential context forces the generic reading of the object, as in (20a) and (21a), *a* becomes obligatory, while it remains optional when the object is non-specific, as in the most natural reading in (20b) and (21b); a generic reading of the object with *a* is not excluded, in any case, in (20b) and (21b). The point is that, whatever our favourite interpretations for the examples may be, if the object is generic it requires the presence of the case-marker⁸. Thus, we have an additional problem that a characterization of *a* as a specificity marker cannot solve. An account is needed both of the particular facts in (19)-(21) and of the rest of interpretive mechanisms that the preposition triggers in indefinite objects, illustrated in (16)-(18). As mentioned before, a more abstract and general notion than specificity must underlie all these interpretive effects. On the one hand, *a* is associated with strong readings of indefinites, be they specific or generic, but not in a completely systematic way, because it admits weak readings as well. We can now make the following generalisation: *a* is a prerequisite for strong readings, although it does not exclude weak ones. On the other hand, the linguistic content attributed to *a* must be compatible with facts such as the licensing of bare plurals, the licensing of secondary predication, and the discourse prominence of the object, both in specific and in generic interpretations. The natural way to capture all these facts under a single account is to assume that *a* is a topic marker, and that so-called specificity effects are inferentially derived from the meaning of the preposition. In particular, this seems the only way to reconcile specificity effects with “genericity effects” such as those in (19)-

⁸ Moreno and Pérez (2001) demonstrate that information structure can force generic interpretations even in bare plurals in Spanish, in sentences like *Correos admite giros urgentes hasta las ocho* (“The post office admits urgent giros until eight o’clock”): the crucial factor is the processing of the bare plural as topic or part of the background. What *a* triggers in examples like (19)-(21) is exactly this: the mapping of the indefinite onto the restrictive clause in logical form. There is a close parallelism between the two sets of data.

(21), given that *specific* and *generic* are varieties of strong interpretations for indefinites, and both are favoured by topic positions. This account of the semantics of object marking could be easily extended to most other cases of Differential Object Marking in the languages of the world.

A look at the syntactic literature on object marking and object agreement shows that, when trying to express the properties of marked objects in configurational terms, a widely accepted hypothesis is locating them in a higher structural position than the basic one by means of Object Raising or some similar formal device⁹. This seems to me just another way to give expression to the same intuition I am dealing with here: object marking and object agreement are mechanisms that speakers use to emphasize the discourse prominence of certain arguments. In what follows I intend to remain neutral with respect to the necessity of resorting to syntactic solutions to account for the behaviour of marked objects.

Apart from the possibility of capturing all the previous data under a single generalization, there are some other arguments worth mentioning that favour an approach based on topicality. I will devote the next section to them.

3. Specificity and Topicality

3.1 Topic and referential autonomy

So far I have tried to show that the insertion of *a* does not consistently mark specificity in Spanish, in spite of being somehow connected to the occurrence of strong interpretations in indefinite DPs, and that its contribution to sentential meaning may reasonably be conceived of as a sort of topicality marking. In the remaining part of the paper I intend to support this perspective by collecting a variety of arguments, some empirical, some theoretical. The arguments are, briefly, the following ones.

First, it is crucial to bear in mind that topicality plays a decisive role not only in the interpretive properties of indefinite objects, but, even in a more evident way, in the interpretive properties of indefinite subjects; it seems therefore natural to rely on topicality for a unified account of indefinites.

Second, the whole picture emerging from the data is compatible with Van Geenhoven's theory of indefinite interpretation (Van Geenhoven 1998), which allows us to integrate the role of information structure into an elegant general framework that seems to work successfully in different languages.

Third, topicality seems to be the natural connection between object marking and other grammatical mechanisms that involve animacy and referentiality / definiteness, like clitic doubling and object scrambling. A comparison with Germanic scrambling raises interesting issues. In the following sections I will elaborate on these points and comment on some difficulties surrounding the notion of topic.

Before discussing such issues, it is convenient to return to the essential connection between topicality and specificity to give a more precise picture of it. Indefinite DPs in topical positions tend to receive strong readings –generic or specific–: there seems to be no disagreement on this generalization. The most frequently discussed case is the one that

⁹ See Torrego (1998) for Spanish *a*, Rapoport (1995) for Hebrew, Karimi (1990), (1996) and (1999) for Persian.

involves subjects of Individual-Level predicates. I assume that this is to be accounted for by means of the interplay of semantic and pragmatic principles¹⁰. No obligatory or conventional link exists between specific DPs and topics, given that specific readings may arise in non-topical positions as well. The basic idea is rather that topics orient the interpretive task towards strong readings. It can be sketched informally as follows. The variety of interpretations for indefinites stems from their sensitivity to sentential contexts, and from the role of different syntactic and semantic factors in constraining the assignment of a value to the discourse referent that indefinites introduce. Indefinites, in their weak reading, establish the cardinality of the intersection of the set denoted by their restrictor and the set denoted by the predicate. Their existential properties come from an external source (the verb, or the sentential context). This makes them non autonomous expressions from a referential point of view; therefore, weak indefinites are not good topics, in general. The notion of topic is relevant here because being a topic constrains the choice of value for indefinites. If topics require referentially autonomous DPs (i.e. expressions whose reference obtains independently of the cardinality of other denotations at the clausal level), then topics require indefinites to be assigned some kind of strong reading (i.e. an interpretation where a value for a discourse referent is established independently of the cardinality of other denotations). In order to build a strong reading, speakers resort to different procedures that turn the indefinite DP into a referentially autonomous expression (by establishing previously mentioned sets of entities to obtain a partitive reading, or accommodating some contextual assumption on the existence of a particular entity being referred to).

The contribution of *a*, then, is the encoding of an instruction to process the object DP as an internal topic, that is, as a prominent and referentially autonomous argument. As *a* is a functional head (Brugè 2000), and thus a procedural element in Sperber and Wilson's terms¹¹, its semantic content can be considered as a procedure that must be strictly obeyed in the interpretive process: *a* guides and constrains the inferential phase of comprehension that maps logical forms into explicatures (in the direct object domain). The occurrence of specific or generic interpretations is triggered by the necessity to obey the instruction encoded by *a*. Thus, when *a* is not obligatory, it favours a strong interpretation of indefinite objects as a result of the inferential processes activated by its procedural nature as topic marker.

3.2 Subjects and objects

Topicality is responsible for the availability of different kinds of interpretations in indefinite subjects. As Milsark (1977) already pointed out, specific readings of indefinite subjects arise both with individual-level (IL) and stage-level (SL) predicates, but non-specific readings are incompatible with IL predicates, while generic ones are usually favoured by them. Recent research on this issue has tried to reduce the distribution of readings to basic distinctions in information structure: IL predicates only give rise to categorical (bipartite) structures with topical subjects, while SL predicates may give rise to thetic structures as well. Thus, the notion of topic happens to be crucial to explain the occurrence of non-specific readings (typically unavailable in topics), generic readings (typically associated to topic constituents)

¹⁰ For a defense of this perspective, see Jäger (1995a), Leonetti (1998), Büring (2001).

¹¹ I am assuming here that functional categories always have procedural content, unlike lexical categories, that can be conceptual or procedural (see Escandell-Vidal and Leonetti 2000).

and specific readings (favoured in topics, but not excluded at all in other positions). It would be much surprising if its role in the interpretation of indefinite subjects and the role it may play in the interpretation of indefinite objects were not related. Given this, a unified approach that tries to derive the behaviour of indefinite subjects and the behaviour of indefinite objects from the same underlying notion is clearly preferable to another one that resorts to different criteria for the two cases.

In fact, the whole argument is reminiscent of the arguments put forth in Aissen (2000) for an analysis of split ergativity patterns as an instance of Differential Subject Marking (DSM), the mirror image of DOM in the subject domain. By means of the operation of *harmonic alignment*, Aissen establishes constraints that express the relative markedness of various DP types in the *object* function as well as in the *subject* function, and predicts

“that there should be case-marking systems in which some subjects are marked, but not all and (...) that the factors that favour differential subject marking (DSM) will be the mirror image of those that favour DOM.” (Aissen 2000:27)

Once the connection between DOM and DSM is established in this way, it seems just a natural extension to draw a further connection between the referential properties of marked objects and those of unmarked subjects. The class of expressions that count as marked objects (i.e. animate, referential, definite) is the class of expressions that typically occur as unmarked subjects. And the type of subject that is most likely to be marked in DSM systems (i.e. inanimate, non referential, indefinite) is the same type of subject that is most likely to be penalized by grammatical or interpretive constraints in other systems: some languages show a tendency to exclude indefinite subjects (Arabic, Chinese), some impose bans on certain readings in certain circumstances (Spanish, English, Chamorro). This last option is currently known under the label of *Property Predication Restriction* (PPR): properties (Individual-Level predicates) may only be predicated of strong DPs. Thus, DSM, the PPR and the constraints on indefinite subjects happen to be reactions to the same kind of pressure and are controlled by the same kind of scales or hierarchies. They represent a way to draw a borderline between unmarked and marked subjects. If this story is correct, it should be taken as an argument favouring the decisive role of the notion of *topic* in explaining DOM patterns with indefinites. Being poor candidates for topics, indefinites (and, more specifically, indefinites in their basic weak readings) should be the worst candidates for DOM, as well as the best candidates for DSM and for being subject to constraints like the PPR.

3.3 Semantic Incorporation

Van Geenhoven’s theory of *semantic incorporation* is one of the most successful theoretical proposals on indefinites in recent research. In the spirit of a previous analysis in McNally (1995), she develops the idea that bare nouns and certain indefinite DPs are property-denoting expressions characterized by their inherent narrow scope. Semantic incorporation by a verb licenses this kind of predicative indefinite expressions: they are absorbed or *semantically incorporated* by the verb as the restriction of the verb’s internal argument, and their existential properties are contributed by the governing verb. Semantically incorporated indefinites, then, do not have quantificational force of their own, as it is the verb that introduces the new discourse referent. The basic predicative meaning of indefinite expressions is their default interpretation, according to Van Geenhoven (1998).

The possibility of having semantic incorporation for an indefinite essentially depends on the lexical properties of the verb. Some verbs are semantically incorporating (the core cases are existential predicates), some are nonincorporating (the core cases are individual-level predicates), and finally some may or may not be incorporating. The three classes are illustrated in the following Spanish examples, where bare plurals seem to be licensed only as objects of incorporating verbs: in (22) the existential *tener* allows a bare plural as its object, but rejects a definite DP, giving rise to the well known definiteness effect (see 2.1); in (23) *adorar* “love”, as an individual-level transitive verb, prevents the occurrence of a bare plural, while accepting a definite DP; and finally in (24) *enterrar* “bury” admits both options.

- (22) Tiene grandes ventanas/ ?Tiene las grandes ventanas
It has large windows / ?It has the large windows
- (23) Adora las patatas fritas / *Adora patatas fritas
(S)he loves the fried potatoes / (S)he loves fried potatoes
- (24) Entierra huesos / Entierra los huesos
He buries bones/ He buries the bones

In fact semantic incorporation seems to be the crucial factor in the distribution of Spanish bare plurals. McNally’s (1995) and Van Geenhoven’s (1998) theories are able to explain their main properties (narrow scope, lack of anaphoric readings, lack of partitive readings, discourse transparency) by means of such a notion. Moreover, they extend their approach to non-specific indefinites and opaque contexts in Van Geenhoven and McNally (2002), treating all weak readings of indefinites as instances of semantic incorporation and accounting for strong readings in terms of accommodation. Although there are alternative ways to develop a theory of semantic incorporation, such a framework is undoubtedly a promising approach to the interpretation of indefinite expressions cross-linguistically. It should prove a useful tool also for the analysis of indefinite objects in languages that exhibit DOM systems. In fact the expectations begin to get confirmed as soon as the existential contexts exemplified in (4)-(5) are taken into account. According to McNally and Van Geenhoven, definiteness effects in these contexts are due to the incorporating nature of the predicate and the resulting weak interpretation of indefinites. The exclusion of *a* after *haber* and *tener* in the standard existential construction follows naturally from the ban that DOM imposes on semantic incorporation¹². The ungrammaticalities in (4)-(5) result from the clash between the incorporating verb and the anti-incorporating properties of the prepositional marker. In addition, there are other facts that can be covered under the same approach, as we will see now.

¹² It is worth pointing out that the behaviour of existential predicates and the group of incorporating verbs I am considering here is cross-linguistically homogeneous in its rejection of all the devices associated with specificity and object prominence: the same as Spanish *haber* rejects *a*, incorporating or “definiteness-effect verbs” in Ostyak exclude object agreement (Nikolaeva 2001: 21-22), the equivalent of *have* in Persian rejects the particle *-râ* (Karimi 1990: 174), and its equivalent in West Greenlandic is a morphologically incorporating affix (Van Geenhoven 1998: chapter 5).

Although the interaction between semantic incorporation and topic-focus structure is not investigated in detail in Van Geenhoven (1998), it is not difficult to find certain correlations between them that help to illuminate certain aspects of the semantics of Spanish *a*. Assuming a) with McNally (1995) and Laca (1996), that bare plurals in Spanish are predicative expressions, and b) that they are interpreted by semantic incorporation, it follows that their syntactic distribution must be regulated by the lexical properties of the verbs and predicates that take them as arguments (as shown in (22)-(24)) and by the resistance that certain argument positions offer to incorporation processes. These positions are basically preverbal subject and *a*-marked object (leaving indirect objects aside): in a few words, sentence-internal topical positions. It is quite natural to think that an argument inside a topic position should not make a good candidate for incorporation or absorption: the more prominent an argument is, the more reluctant it is to being interpreted as a predicate modifier or as “part of the predicate”. The constraints imposed by individual-level predicates on bare plurals (cf. (23)) are reducible to the topicality factor as well: as often stated in the literature, subjects of individual-level predicates are topical (cf. 3.2), and, furthermore, their objects are topical too, as Laca (1990) demonstrated in her study on Spanish bare plurals. This is one of the reasons why individual-level predicates do not give rise tothetic judgements, and is also responsible for the contrast in (23). So the general notion of topicality underlies both the diverging selection properties of verbs like *tener*, *adorar* and *enterrar* and the discourse properties that differentiate categorical andthetic judgements, and accounts both for the interpretations of subjects and direct objects.

The generalization that relates these facts concerning bare plurals and the previous data about DOM in Spanish is the following: internal topic positions tend to block incorporation processes. Notice that the generalization is about internal topic positions, because external topics may receive a weak reading quite naturally, as the following examples of left dislocation in Spanish (25) and split topicalization in German (26) demonstrate¹³:

(25) Bueno, un libro habréis leído, este verano ¿no?
Well, a book you will have read, this summer, won't you?

(26) Hausaufgaben haben die Studentennicht mal zwei gelesen
home tasks have the students not even two read
“As for homework sets it is not the case that the students read even two”

The indefinite expressions *un libro* “a book” in (25) and *hausaufgaben* “homework” in (26) are non-specific. This suggests that the interpretive constraints on external topics are weaker than those operating on internal topics (subjects and objects). If we assume that non-specific indefinites are interpreted via semantic incorporation à la Van Geenhoven, this operation must be available in some sense for external topics, but it seems forbidden, or at least disfavoured, in internal topics or prominent arguments. This is an issue that should be worth investigating

¹³ Example (26) is from Van Geenhoven (1998:5). Van Geenhoven takes German split topicalization as a clear instance of semantic incorporation. A sort of split topicalization is also possible in Spanish:

(i) Ejercicios, los estudiantes no han leído ni siquiera dos
Exercises, the students have not read even two

in more detail, but I cannot even sketch an explanation for it here¹⁴. I will simply limit my speculations to internal topic positions, trying to offer some additional support for the idea that “(non)specificity” effects in subject and object positions are related to topicality. A plausible answer to the problem of the asymmetry between the two types of topical positions could be based on the intuition that we are actually dealing with two different notions of *topicality*: I shall address the issue briefly in section 3.5.

Once semantic incorporation has been introduced as a tool for the study of indefinite descriptions, the obvious question that comes to mind is how it is related to DOM and its effects in Spanish. The answer is obvious as well: the insertion of *a* should have some consequence for the incorporation process, if *a* functions as a topic marker and topical positions constrain the possible readings of indefinite descriptions¹⁵. There are some proposals in this direction in the literature. Bleam (1999a:180; 1999b:33-34) considers *a* as an indicator of a type shifting operation on the indefinite: *a*-marked bare plurals are shifted to a kind reading or to an existential reading instead of being semantically incorporated to the verb, and indefinite DPs are shifted to a generalized quantifier status when *a* occurs (Bleam’s *A-Marking Hypothesis*). This seems a natural way to interpret the facts: semantic incorporation is the default interpretation procedure for unmarked indefinites, and the insertion of a special marker triggers a different interpretation procedure (accommodation, choice functions or any other mechanism for deriving specificity); the result is a variety of strong readings, while unmarked indefinite objects, as incorporated predicative expressions, get only weak readings.

However, two points should be stressed: first, accepting an analysis along these lines does not mean accepting that *a* encodes a specific procedure of the type “Shift to generalized quantifier”, which I think would be a mistaken interpretation of the facts; second, there is a residual problem with the availability of weak or non-specific readings in marked indefinites as well (as shown in (7)-(11)). Recall that *a* may appear with non-specific DPs, especially when it is obligatory and when there is no specific reading contextually available: this is an indicator of how far the expansion of DOM has gone in modern Spanish and how much it distorts the relationship between *a* and specificity. One must conclude that nowadays *a* cannot be said to block semantic incorporation systematically. The only safe generalization is that the absence of the preposition marks incorporation. In the central cases where *a* and \emptyset freely alternate, it is true that *a* represents a non-incorporated reading and \emptyset an incorporated one. At any rate, a link between DOM and the impossibility of semantic incorporation must still survive, with topicality, again, underlying such a link.

An additional argument in favour of such a link, and one that has gone unnoticed, as far as I know, is the fact that in Spanish the group of psychological verbs that require *a* in their objects (which are all individual-level predicates) correlates quite closely with the group of verbs that reject bare plurals as objects (an expected behaviour in individual-level predicates). As the correlation does not seem to be arbitrary, the most straightforward account is one that

¹⁴ A crucial factor for the acceptability of (25) and related structures is the absence of a clitic in a sentence-internal position. A parallelism between internal and external topic positions can be invoked only when there is such a clitic (i.e. in clitic left dislocation structures).

¹⁵ Recall that the expected consequences will only be visible in those contexts where the insertion of *a* is fully significant.

characterizes those verbs as non-incorporating; incorporating verbs are expected to accept bare plurals as well as unmarked objects. The relevant facts are represented in the grammaticality patterns in (27) and (28), where non-incorporating verbs are compared to a small group of typically incorporating verbs¹⁶:

- (27) a. odiar *(a)una persona / admirar *(a)una persona / despreciar*(a)una persona / amar *(a) una persona / aborrecer *(a)una persona / soportar *(a)una persona
 “to hate a person / ...admire... / ...despise... / ...love... /...detest... / ...put up with...”
- b. llevar (a) una persona / curar (a) una persona / contratar (a) una person
 describir (a) una persona / encontrar (a) una persona / ver (a) una persona
 “to take a person / ...cure... / ...hire... / ...describe... / ...find.../...see...”
- (28) a. *odiar personas / *admirar personas / *despreciar personas / *amar personas / *aborrecer personas / *soportar personas
 “to hate persons / ...admire... / ...despise... / ...love... /...detest... / ...put up with”
- b. llevar personas / curar personas / contratar personas / describir personas / encontrar personas / ver personas
 “to take persons / ...cure... / ...hire... / ...describe... / ...find... / ...see...”

Therefore, bringing semantic incorporation into an account of Spanish DOM, we obtain some benefits, both at the descriptive and the theoretical level: 1) the distribution of unmodified bare plurals is reduced to a simple mechanism, controlled by a variety of lexical and syntactic factors; 2) the same mechanism is responsible for the readings of subject and object indefinites, in languages with explicit DOM systems and in languages that do not mark semantic incorporation explicitly; 3) the insertion of *a* fits into a general theory of indefinite descriptions and its connection with topicality is preserved.

3.4 Clitic Doubling and Scrambling

It is well known that clitic doubling and object scrambling are among the grammatical phenomena that usually interact with definiteness and specificity¹⁷. One could wonder what is the origin of such an interaction. Indeed that is the right question to pose if one is looking for a principled account of the behaviour of *a* in Spanish, because a look at the similarities among all these grammatical devices yields a number of interesting results for the study of DOM. In

¹⁶ The two paradigms in (26) and (27) cannot be taken as representative of the remaining verbs without paying attention to the bundle of overlapping factors that control the insertion of *a*. In fact, things become much more complicated as soon as other groups of verbs are examined. In any case, the basic correlation presented here reinforces the idea that, in spite of the expansion of DOM in modern Spanish, there are still areas where the link between the *a* / \emptyset alternation and semantic incorporation can be clearly perceived.

¹⁷ See Lyons (1999: chapter 5) for a survey and for general considerations on the nature of the phenomena. This section owes much to Alexiadou and Anagnostopoulou (1997), Delfitto and Corver (1998), Bleam (1999a), Meinunger (2000) and Nikolaeva (2001).

this section I intend to examine the way in which the occurrence of *a* is related to clitic doubling and scrambling, in order to find out what lies behind the three constructions. And the unifying notion will be, unsurprisingly, the marking of internal topics.

At least since Richard Kayne proposed what has been known as “Kayne’s generalization” (*Clitic doubling requires the object to be case-marked*), clitic doubling and DOM have been considered as related syntactic mechanisms. In Spanish in fact direct object doubling seems to depend on the insertion of *a* before the object DP¹⁸. The two grammatical processes are triggered by the same kind of features: animate and referential / definite / specific. Nevertheless in standard modern Spanish the conditions for direct object doubling are more restrictive than the conditions for object marking: only personal pronouns trigger doubling, while any kind of animate and definite / specific DPs activate the insertion of *a* (a significant difference is that negative quantifiers cannot be clitic doubled, but can be *a*-marked). As Bleam (1999a: 199) correctly points out, “the semantic properties which give rise to clitic doubling form a subset of the semantic properties which give rise to the prepositional accusative...”. In some varieties of Spanish, like those spoken in Río de la Plata (Argentina) and the Basque Country¹⁹, clitic doubling is governed by less restrictive conditions and extends to definite and specific indefinite DPs, but it still obtains in a subset of the cases where *a*-marking obtains. This kind of dialect variation cannot disguise the fundamental similarity of the two phenomena. Both are expanding along the same scales and hierarchies. Moreover, it is precisely in some of the non-standard varieties that the parallelism emerges strikingly. Franco and Mejías-Bikandi (1999) show that in Basque Country Spanish the condition for an indefinite object to be clitic-doubled is to receive a strong interpretation: in (29a) there is only a strong (presuppositional, in Franco and Mejías-Bikandi’s terms) interpretation, imposed by clitic doubling, while in (29b), where the object is not doubled, the indefinite DP is ambiguous between a strong and a weak interpretation (notice that *a* is present in both cases).

- (29) a. Le he visto a un marinero
 To-him I have seen to a sailor
- b. He visto a unmarinero
 I have seen to a sailor
 “I have seen a sailor”

The subtle contrast in (29) confirms that explicit object morphology –clitic doubling is an instance of object agreement- usually forces strong interpretations in indefinites. Bleam (1999a: 44-55) obtains the same results in her analysis of Leísta Spanish (a variety that resorts to the dative clitic *le* for animate direct objects, and for doubling), Alexiadou and Anagnostopoulou (1997) report similar interpretive effects in Greek, and Nikolaeva (2001) describes how object agreement is associated with strong interpretations in Ostyak, an Uralic

¹⁸ Detailed analyses and discussions are provided by Torrego (1998, 1999) and Bleam (1999a, 1999b).

¹⁹ See Suñer (1988) and Franco and Mejías-Bikandi (1999) for a study on clitic doubling and specificity in these varieties of Spanish. A tighter correlation between clitic doubling and object marking can be observed in Roumanian.

language spoken in Western Siberia (and the correlation between object agreement and specificity is well attested in many other languages).

Given that clitic doubling (or object agreement) and *a*-marking display several common properties (they are triggered by the same features, they are both optional –in the cases we are interested in-, they have similar effects), it is reasonable to ask whether some abstract and general property might not be at the origin of such parallelisms. Some authors have tried to give an answer in terms of formal or configurational analogies, usually related to some kind of movement of the object to higher nodes in the syntax. I do not intend to go into a critical analysis of such proposals, but my impression is that the particular interpretive properties of the constructions are simply stipulated in them, rather than accounted for²⁰. A number of problems have still to be faced by these accounts: Where do specificity effects come from? Why would certain functional heads be associated with strong readings? Why does animacy correlate with specificity? Why is it direct objects that typically show the aforementioned common properties? In a few words, the way syntax is related to semantics (and pragmatics) needs a more principled explanation.

On the other hand, starting from the assumption –which I find quite natural- that human languages often resort to different types of object marking or object agreement to indicate a high degree of discourse prominence in direct objects (i.e. in arguments that typically tend to be less prominent than subjects or indirect objects), the facts can be accommodated into a coherent picture that highlights the role of topicality –again- and its associated interpretive properties. But, before trying to sketch such a picture, it is convenient to bring scrambling or object shift into the scene.

The extensive literature on scrambling²¹, particularly in Germanic languages, has convincingly shown that it is systematically associated with strong readings of indefinites (both generic and specific / referential), and, in addition, that it can have parallel consequences in the interpretation of definites (an issue I do not intend to discuss here²²). I reproduce in (30) a representative contrast in German (from Meinunger 2000: 66):

- (30) a. sie weil bestimmt schon mal eineSinfonie gehört hat
 she since surely already a symphony heard has
- b. weil sie eineSinfonie bestimmt schon mal gehört hat
 since she a symphony surely already heard has
 “since she surely has already heard a symphony”

While in (30a) the indefinite object *eine Sinfonie* is inside the VP, in its base position, and it only gets an existential or weak reading, in (30b) it is outside the VP boundary and gets a strong reading. Scrambling shares a number of properties with clitic doubling, as argued by Alexiadou and Anagnostopoulou (1997:144-153), and, what is more important here, with

²⁰ Meinunger’s (2000) theory of Agreement nodes as topic hosts is an exception, in that it combines movement to functional nodes and the derivation of all the interpretive effects from the topical status of such nodes.

²¹ I rely on the data presented in Diesing (1992), Delfitto and Corver (1998), Neeleman and Reinhart (1998), Choi (1999) and Meinunger (2000).

²² But see Delfitto and Corver (1998), Meinunger (2000: chapter 3) and Jäger (1995b) for a review of the data.

object case-marking in Spanish and other languages: apart from the well known definiteness / specificity constraints, there are interesting properties related to binding and scope, obviously not independent from the aforementioned constraints. Scrambling and clitic doubling may increase the discourse prominence of direct objects with respect to indirect objects, and so does case-marking. The following examples involve ditransitive verbs and illustrate binding / scope asymmetries stemming from the presence or absence of *a*: the contrast in (31) shows how case-marking of the direct object can turn it into a prominent binder for a pronoun inside the indirect object; those in (32) and (33) show how case-marking favours wide scope readings of the direct object with respect to the indirect object.

- (31) a. Devolvieron un prisionero a su tribu
They returned a prisoner to his tribe
- b. Devolvieron a un prisionero a su tribu
They returned to a prisoner to his tribe
- (32) a. Devolvieron un prisionero a cada tribu
They returned a prisoner to each tribe
- b. ??Devolvieron a un prisionero a cada tribu
They returned to a prisoner to each tribe
- (33) a. Enviamos un especialista a todos los departamentos afectados
We sent an expert to all the affected departments
- b. Enviamosa un especialista a todos los departamentos afectados
We sent to an expert to all the affected departments

All the differences between *a* and \emptyset arise from the prominence the indefinite object acquires when case-marked. As for the examples in (31), \emptyset favours the reading where the object *un prisionero* is not the antecedent of the possessive *su*, and yet *a* imposes the opposite reading, with *un prisionero* as the antecedent of the possessive. In a strictly configurational approach, this seems to be the result of the raising of the object to some higher node (i.e. the counterpart of scrambling in a language without scrambling like Spanish), but one does not need to postulate a raising rule triggered by *a*, if it is assumed that *a* marks an internal topic, thus turning the indefinite object into a prominent antecedent for an anaphoric pronoun (and, of course, if one chooses to depart from the classical binding principles). In any case, the parallelism between case-marking, scrambling and clitic doubling is descriptively clear in the binding facts. I believe that it holds for scope facts such as those shown in (32)-(33) too.

In (32) the contrast is particularly robust. The distributive quantifier *cada* in the indirect object forces a distributive reading in the indefinite object; such a reading is acceptable in the first case, without *a*, but unacceptable in the second one, with *a*. Again the prominence of the *a*-marked object is responsible for the difference: it follows from the raising of the object to a

position where it can no longer be bound by the quantifier -in case you like the configurational approach- , or, alternatively, from the fact that DPs in topic positions are rarely inside the scope of any quantifier –in case you choose an account in terms of topicality. Therefore, the oddity of (32b) is the result of a clash between the inherently distributive nature of *cada* and the effects of *a*-marking.

Finally, (33) involves the scope interaction between the indefinite object *un especialista* and the quantified indirect object *todos los departamentos*. As expected, the indefinite DP has wide scope when it is preceded by *a*, as in (33b) (where a unique expert is supposed to have been sent to every department), and narrow scope when it is not case-marked, as in (33a), where a different expert has been sent to each department.

The facts in (32)-(33) are strikingly similar to the well known “scope-freezing effects” that arise in English or Japanese when the indirect object precedes the direct object in examples like those in (34), from Nakanishi (2002: 141):

- (34) a. The teacher assigned one student every problem
- b. John-ga [sannin-no onna]- ni [futari-no otoko]-os yookaisita
 John-NOM[three-GEN woman]- DAT [two-GEN man]-ACC introduced
 “John introduced to three women two men”

In (34) the Indirect Object - Direct Object order blocks one of the readings available in the opposite order, Direct Object – Indirect Object, leaving as the only possibility the surface scope reading (IO > DO). The preposing of indirect objects in ditransitive constructions, thus, has a “scope freezing effect”, which, according to Nakanishi (2002), cannot be due to general properties of movement rules, but to the specificity of the indirect object in IO-DO order, semantically encoded as a choice function interpretation. Her analysis brilliantly covers several facts in the interpretation of ditransitive structures, and is compatible with the idea that the specificity responsible for “scope freezing” derives from the topical status of the indirect object in IO-DO order. This is relevant for an analysis of *a*-marking because it reinforces the parallelism between DOM and the variety of object shift rules operating in natural languages, as well as the central role of topicality in the distribution of the interpretations for indefinite descriptions. In fact the null hypothesis should be that we have “frozen scope” both in dative shift and in object case-marking because the two constructions encode the same kind of procedure: an indication to process a constituent as an internal topic (this is the basic meaning of scrambling in Germanic languages, according to the recent literature²³).

Building partially on Meinunger’s (2000) claims about the topic / agreement connection and the way it relates to scrambling, I would like to sum up by stating that an analysis of DOM in Spanish cannot ignore the analogies that relate it to all the other devices that trigger strong

²³ Neeleman and Reinhart (1998) rely on *destraining* and *discourse-linking* to account for the basic property of scrambling, Delfitto and Corver (1998) take *familiarity* as the key feature, Choi (1999) uses a general notion of *prominence* that covers both topic and contrastive focus, and Meinunger (2000) relates scrambling to topics, in the sense of constituents that bear familiar information and act as anchors for the new information to be linked to the old information. Although these approaches are not equivalent, I assume that all of them point to some aspect of what it means to be a topic.

interpretations in direct objects. *A*-marking in Spanish bears evident resemblances with grammatical mechanisms such as a) morphologically different cases for objects or special particles (accusative / partitive in Finnish, accusative / genitive in Russian, accusative / absolutive in Turkish, *-râ* in Persian), b) object agreement (Hindi, Hungarian, Bantu languages, Macedonian, Ostyak), and c) scrambling or object shift (German, Dutch, Korean, the *ba*-construction in Chinese). All three phenomena (case-marking, agreement, positional differences) have essentially the same interpretive consequences, and the natural way to integrate them all in a coherent explanation is assuming that they behave as topic markers, and that the higher a constituent is in the animacy and definiteness / specificity scales, the better it fits in a topic position. Moreover, this seems the only way to account for the role of animacy in the aforementioned mechanisms, bearing in mind that there is a natural correlation between being a topic and referring to an animate entity²⁴. Finally, this also seems the only simple way to integrate the representation of specificity constraints on subjects and objects into the same picture. What remains to be discussed is why different languages choose different “cut-off” points along the scales of animacy and definiteness / specificity, and why some languages extend the range of case-marking or agreement till the end of the scales, thus making no distinction between prominent and non-prominent arguments (see Meinunger 2000: 177-178). Answers to such questions are not easy to find.

3.5 Some problems

As we have seen in the previous section, there is enough supporting evidence linking the analysis of *a*-marking (and DOM in general) to topicality, and this deserves serious consideration. Nevertheless, there is still an obvious difficulty: the analysis is based on a notoriously vague and elusive notion. As already pointed out in section 3.3, the distinction *internal topic* / *external topic* needs to be investigated carefully if we want to have a more precise notion of *topic*. But even dealing exclusively with internal topics, a need emerges for an accurate specification of the way the term is used.

In the preceding sections I have been using *topic* basically as an information structure concept, in the sense of “anchor for new assertions” and “referentially autonomous expression”, typically –but not necessarily– conveying given information. If my approach is correct, case-marked objects in Spanish should share most of their properties with similar cases of internal or secondary topic marking in other languages. Nikolaeva (2001) provides us with a very interesting example of a language, Ostyak, where object agreement marks secondary topics. A comparison with Spanish can help to understand the nature of the problems surrounding the concept of *topicality*. Nikolaeva (2001: 26) defines a secondary topic as “an entity such that the utterance is construed to be about the relationship between it and the primary topic”. Secondary topics tend to be encoded cross-linguistically as direct objects. In Ostyak object agreement is optional: when the object does not agree with the verb, it is interpreted as focus, but when agreement appears the object acquires certain topical properties (existential presupposition²⁵, activation / definiteness, and the need for an explicit

²⁴ For the interaction between topicality and animacy, see Dahl and Fraurud (1996) and Yamamoto (1999).

²⁵ This is relevant for a comparison with Spanish *a*, as one of the basic constraints on object agreement in Ostyak has to do with specificity: non-specific expressions never trigger agreement (Nikolaeva 2001: 20-21).

primary topic in the utterance²⁶). These are all properties that secondary topics in Ostyak share with most cases of case-marked objects. But the fact I want to concentrate on is the incompatibility of object agreement in Ostyak with focus on the object: if the object gets narrow focus (contrastive or not), there is no agreement. Nikolaeva (2001: 29-31) shows that object agreement is thus extremely sensitive to focus structure, as one should expect if it marks secondary topic as a relation in information structure. Moreover, this is a strong argument for her analysis in terms of secondary topics. The problem for my approach to *a*-marking is that the occurrence of *a* is not sensitive at all to focus structure: it is compatible with all kinds of focus –contrastive or informative. In other words, the Spanish-Ostyak parallelism fails because DOM in Spanish is not motivated solely by information structure, whereas object agreement in Ostyak is, according to Nikolaeva (2001). The same problem arises in a comparison of Spanish *a* and Germanic scrambling, in spite of the similarities mentioned in section 3.4: scrambling is controlled by information structure factors (destressing, defocusing, familiarity...) that do not seem to play a prominent role in Spanish DOM. Maybe this is related to another difference I did not mention before –the fact that animacy is not relevant for scrambling in German or object agreement in Ostyak, but it certainly is for DOM, in Spanish as well as in many other languages.

At this point I can only offer some speculations about the possibility of integrating these problematic facts into a coherent picture. The topicality hypothesis for *a* seems to be worth maintaining, but there is a need to reconcile the two different notions of *prominence* that underlie the whole array of data I collected in the previous sections.

On the one hand, I have been talking about topics as prominent constituents in information structure, i.e. as anchors for new assertions, mostly discourse-linked, and opposed to focused constituents. This is what governs German or Dutch scrambling, Ostyak object agreement, or the readings of preverbal indefinite subjects and generic indefinite objects in Spanish. Animacy is not involved in this kind of informational prominence: scrambling, for instance, does not seem to be sensitive to animacy as case-marking is.

On the other hand, I have been talking about topics as prominent arguments in event structure, in the sense that they are referentially autonomous with respect to the verb and denote direct participants in the event denoted by the predicate²⁷. This is obviously independent of focus structure, but frequently intertwined with factors like animacy and affectedness. It is the basic feature underlying DOM, as well as many instances of clitic doubling and phenomena like semantic incorporation, case-marking and the Stage / Individual distinction. Furthermore, it seems to be involved in syntactic operations that are known to modify the relative prominence of certain expressions, such as Dative Shift and Possessor Raising²⁸. In any case, it is relevant for the occurrence of specific interpretations.

²⁶ Here we have another parallelism with Spanish *a*, since it also occurs predominantly in sentences with agentive and topical subjects (i.e. prototypical external arguments).

²⁷ See Laca (1987) and Martín (1999) for remarks on this kind of prominence, and the analogy with subjects and indirect objects. It is worth pointing out that Karimi (1999: 708), in her analysis of Persian *-râ*, describes the difference between non-specific and specific indefinite objects stating that in the first case the event is the focus of attention, whereas in the second one it is instead the participants in the event. Not by chance, she uses the same words that other authors have used to describe the effects of *a*-marking in Spanish.

²⁸ Bleam (1999a) studies how dative shift and clitic doubling share the property of presenting an entity as a direct participant in the event.

Given this, one could reasonably reach the conclusion that, due to the inherent vagueness of terms like *topic* or *topicality*, I have been mixing and confusing two different notions in the analysis of *a*-marking in Spanish: more precisely, two different kinds of topicality or prominence, one pertaining to information structure or focus structure, the other related to event structure and the distinction *autonomous* vs *incorporated*. In fact, I have been relying sometimes on one sense of *topical* and sometimes on the other. But, even accepting that better tools are needed for a complete account of the facts, I believe that there are at least two reasons to approach the interpretive effects of DOM along these lines: first, the two kinds of topicality actually interact in a number of contexts (for instance, in the generic reading of *a*-marked objects, and in the strong readings of individual-level predicates²⁹), so that it is not possible to keep them separated in every case; second, the price to be paid for maintaining “informative” topicality and “participant” topicality as strictly independent notions is considerable, as the specificity effects common to all the constructions considered here –as well as other common properties– can no longer be captured under a unified explanation. In other words, we would miss the main generalizations on the place of specificity in grammatical constructions.

The way in which the two kinds of topicality / prominence interact deserves careful investigation. It might well be that some languages are globally more sensitive to one of them, and some to the other. It could be that the two are different manifestations of a more abstract notion of salience / prominence³⁰. At any rate, a great amount of research is still needed on the topic, and at this point I will have nothing more to add on this.

4. Specificity in Grammatical Theory

4.1 More grammatical facts related to specificity

Once a treatment of specificity in DOM contexts has been sketched, it should be integrated into a wider perspective on the role of specificity in grammar. In particular, something has to be added on other alleged specificity markers in Spanish that do not seem to be amenable to an analysis based on topicality, and on the ways through which specificity appears to be an active feature in syntactic configurations. In this section I will concentrate on these issues, before offering some concluding remarks.

After a look at case-marking and clitic doubling, there are two well studied grammatical elements correlated with specificity that have to be mentioned: mood in relative clauses and adjective position inside the DP (Leonetti 1999: 865-868). To insert such factors in a general

²⁹ As for the generic reading, it seems to depend solely on information structure; but then, several questions arise: why is *a*-marking associated to it (if *a* is related to prominence as participant in the event)? In the case of the constraints imposed by individual-level predicates on their arguments, why is information structure relevant if the stage / individual distinction is essentially lexical and related to event structure?

³⁰ Very often linguists resort to quite abstract characterizations of the principles underlying the hierarchies of definiteness and specificity. When discussing the nature of such hierarchies, Lyons (1999: 215) suggests that “what we are dealing with is the subjective prominence or salience, in some sense, of entities in the domain of discourse. (...) Languages will then differ as regards what kinds of noun phrase conventionally count as prominent.” Lazard (1982) suggests that the function of the postposition *-râ* in Persian is polarizing the object in a sentence and distinguishing it from “depolarized” objects; this is another intuitive way of talking about prominence, in some sense. Unfortunately, in the present state of our knowledge, a more precise definition of *prominence* or *salience* is not easy to obtain.

framework for the study of specificity, it is necessary to bear in mind that grammar biases specific or non-specific readings of indefinites by means of three different mechanisms:

a) including a certain amount of descriptive information inside the DP structure, in noun complements and modifiers;

b) inserting modality markers either inside the DP (for instance, the subjunctive in relative clauses) or outside the DP (in verbal morphology);

c) by means of some syntactic mechanism usually external to the DP structure, such as word order, agreement or case-marking.

These three kinds of grammatical operations can be characterized as ways of constraining (or keeping unconstrained) the assignment of a value to the discourse referent associated to an indefinite DP. As is well known, accumulating expressions with a rich descriptive content inside the DP considerably reduces the possibility of having non-specific interpretations, whereas the absence of such expressions strongly favours it, by keeping open the choice among different values. Modality markers, particularly those that give rise to intensional contexts (i.e. future, imperative...), create linguistic environments where value assignments to discourse referents are only minimally constrained. On the other hand, the syntactic devices I have discussed in previous sections have a clearly restrictive effect on such value assignments, and lead to the assumption that the discourse referent must be salient in discourse, and thus probably specific.

There are good reasons to think that none of these grammatical operations encodes (non-)specificity in any sense. The subjunctive mood obviously does not encode it: characterizing subjunctive as expressing *non-specificity* would not allow us to account for its syntactic distribution and its contribution to utterance interpretation. The same could be held of any modality marker. Having some of them more or less systematically associated with certain readings of DPs does not mean that they encode such readings.

As for adjective position inside DPs, Picallo (1994) and Bosque (2001) demonstrated that the prenominal position of epithets and relative adjectives in indefinite DPs forces the specific reading, while the postnominal position is compatible with strong and weak readings. In (35)-(36), the DPs *una interesante novela* and *un famoso actor* must be specific; hence, a sentence like *Busco un famoso actor* in (36) is ungrammatical without *a*, but is grammatical when the order is N - Adj: *Busco un actor famoso*.

(35) Quiero leer {una novela interesante / una interesante novela}
I want to read {a novel interesting / an interesting novel}

(36) a. Busco a {un actor famoso / unfamoso actor}
I look for to {an actor famous / a famous actor}

b. Busco un actor famoso / *Busco un famoso actor
I look for an actor famous / *I look for a famous actor

Again there is no encoding of specificity. The contrasts have rather to do with the referential properties that epithets and relative prenominal adjectives require in the DP. The crucial fact is that they cannot be interpreted as restrictive modifiers. The only available interpretation is an

explicative or appositive one. This has important consequences for the referential status of the DP, simply because explicative modifiers operate on referentially autonomous expressions, i.e. on expressions whose reference is established independently of the property denoted by the prenominal adjective. To avoid a semantic clash between the modifier and the host phrase, the whole DP is assigned a strong interpretation (a specific one, usually). I want to stress that the notion that relates this kind of facts to the other facts discussed here is *referential autonomy*: just the basic property of topics. In a few words, the trigger for the inferential specification of a strong reading is essentially the same for *a*-marking, clitic doubling, subject and object raising, and finally prenominal relative adjectives, which allows for a unified account of specificity effects in all these constructions.

4.2 Does grammar encode specificity?

All the preceding discussion leads me to advance the general hypothesis that in natural languages the grammatical system does not encode features like specificity, familiarity or referentiality, but more abstract features related to information structure and processing instructions. More precisely, I do not want to say that certain lexical items (determiners and quantifiers) cannot encode specificity; in fact we know that several languages have determiners or quantifiers that are systematically associated with specific readings. What I want to suggest, contrary to Delfitto and Corver (1998) and Karimi (1996), among others, is that functional categories external to DPs (i.e., agreement, case, focus, and so on, leaving aside tense and other deictic categories) –the categories responsible for word order phenomena and basic constructions in natural languages- do not encode specificity. Semantic and pragmatic notions like specificity, familiarity, rigid designation or discourse-linking do not seem to play any role in the computational system. This idea is not new, at least for DOM and related phenomena. It takes up Neeleman and Reinhart's (1998: 346) conclusions about scrambling and its relationship to specificity and D-linking: "There is, then, no reason to assume that these discourse options are coded in any way in the computational system." The proposal is also in the spirit of Meinunger's (2000) work on scrambling, topicality and agreement nodes: in his analysis it is the topic status of an argument, and not specificity or definiteness, that activates agreement projections.

I cannot develop a complete discussion of the general issue here. Nevertheless, I would like to mention that one of the reasons why proposals like this must be studied and evaluated is that we need to constrain the possible range of meanings that functional nodes are able to code. We need to make progress in that direction to understand what kind of meanings syntax can express, and it is reasonable to think that it should be a highly restricted series of meanings. We need to put restrictions on the features associated to functional nodes in order to strengthen syntactic theory and limit the proliferation of new categories that sometimes may not be sufficiently justified. The price to be paid for carrying out this operation is, at least for the moment, the subsuming of specificity under a vague notion of *prominence / topicality*³¹, but some progress has been made in the specification of what is encoded by the grammatical system and what is pragmatically inferred.

³¹ See Lyons (1999: 226) for the same conclusion on certain aspects of definiteness marking.

To sum up, the central idea I have been arguing for is that Spanish has no grammatical device that encodes specificity; the devices that grammars usually describe (*a* + direct object, mood in the relative clause, adjective position, syntactic position of the DP) encode other meanings, mostly related to information structure, modality and the prominence of arguments. Specificity is pragmatically inferred on the basis of the procedural semantics of such devices and information taken from sentential context and communicative situation. The inferential process is one of the fundamental tasks in the determination of explicatures: reference assignment to DPs and other referential expressions. In this sense, accessing a specific reading is just a way to obey the instructions encoded by certain elements and developing an incomplete logical form into a complete explicature. Specificity appears to be an epiphenomenon, the indirect result of the interaction of several different factors. So too are features like affectedness or discourse-linking.

If this idea is extended to the analysis of other languages, it leads us to the general hypothesis that syntax does not encode features like [specificity] or [familiarity] in functional nodes; as a consequence, phenomena like scrambling, differential object marking, clitic doubling, agreement and so on are triggered by other kinds of features. The immediate advantage I seek to obtain is a better understanding of the distinction between those aspects of utterance interpretation that are semantically encoded in the logical form and those aspects that are pragmatically inferred. On the other hand, the main difficulty is posited by the necessity of 1) defining the abstract linguistic meaning of different syntactic positions and operations, and 2) mapping such a meaning into full fledged explicatures by means of pragmatic principles. Much research remains to be done on these issues.

5. References

- Abbott, B. (1995): "Some Remarks On Specificity". *Linguistic Inquiry* 26. 341-347.
- Adger, D. (1996): "Economy and Optionality: Interpretations of Subjects in Italian". *Probus* 8. 117-135.
- Aissen, J. (2000): "Differential Object Marking: Iconicity vs Economy", Manuscript, Ucsd.
- Alexiadou, A. & E. Anagnostopoulou (1997): "Toward a Uniform Account of Scrambling and Clitic Doubling". In: W. Abraham & E. Van Gelderen (Eds.): *German: Syntactic Problems, Problematic Syntax?*. Tübingen: Max Niemeyer. 143-161.
- Bleam, T. (1999a): *Leísta Spanish and the Syntax of Clitic Doubling*, Phd Dissertation, University of Delaware.
- Bleam, T. (1999b): "Object Bare Plurals in Spanish and the Semantics of Personal *a*". In: J.M. Authier, B. Bullock & L. Reed (Eds.): *Formal Perspectives on Romance Linguistics*. Amsterdam: John Benjamins. 21-37.
- Bosque, I. (2001): "Adjective Position and the Interpretation of Indefinites". In: J. Gutiérrez-Rexach & L. Silva-Villar (Eds.): *Current Issues in Spanish Syntax and Semantics*. Berlin: Mouton de Gruyter. 17-37.
- Bossong, G. (1997): "Le Marquage Différentiel de L'Objet dans les Langues d'Europe". In: J. Feuillet (Ed.): *Actance et Valence dans les Langues d'Europe*. Berlin: Mouton de Gruyter. 193-258.
- Brugè, L. (2000): "La Realizzazione Morfologica del Caso Accusativo". In *Categorie Funzionali del Nome Nelle Lingue Romanze*. Milano: Cisalpino. 193-304.
- Brugè, L. & G. Brugger (1996): "On the Accusative *a* in Spanish". *Probus* 8. 1-51.
- Büring, D. (2001): "A Weak Theory of Strong Readings", Unpublished Paper, Cologne University.
- Choi, H.-W. (1999): *Optimizing Structure in Context*. CSLI: Stanford.
- Cohen, A. & N. Erteschik-Shir (2002): "Topic, Focus, and the Interpretation of Bare Plurals". *Natural Language Semantics* 10. 125-165.
- Company, C. (2002): "El Avance Diacrónico de la Marcación Prepositiva en Objetos Directos Inanimados". In: A. Bernabé Et Al. (Eds.): *Presente Y Futuro de la Lingüística en España*, Vol. II. Madrid: Sel. 146-154.

- Dahl, Ö. & K. Fraurud (1996): "Animacy in Grammar and Discourse". In: T. Fretheim & J.K. Gundel (Eds.): *Reference and Referent Accessibility*. Amsterdam: John Benjamins. 47-64.
- Delfitto, D. & N. Corver (1998): "Feature Primitives and the Syntax of Specificity". *Rivista di Linguistica* 10.2. 281-334.
- Diesing, M. (1992): *Indefinites*. Cambridge, Mass.: MIT Press.
- Enç, M. (1991): "The Semantics of Specificity". *Linguistic Inquiry* 22. 1-25.
- Erteschik-Shir, N. (1997): *The Dynamics of Focus Structure*. Cambridge University Press.
- Escandell-Vidal, M.V. & M. Leonetti (2000): "Categorías Funcionales y Semántica Procedimental". In: M. Martínez Hernández et al. (Eds.): *Cien Años de Investigación Semántica: De Michel Bréal a la Actualidad*, Vol. I. Madrid: Ediciones Clásicas. 363-378.
- Farkas, D. (1995): "Specificity and Scope". In: L. Nash & G. Tsoulas (Eds.): *Actes du Premier Colloque Langues & Grammaire*. Paris. 119-137.
- Farkas, D. (2001a): "Vers une Typologie Sémantique des Syntagmes Nominaux". In: G. Kleiber, B. Laca & L. Tasmowski (Eds.): *Typologie des Groupes Nominaux*. Presses Universitaires de Rennes. 17-46.
- Farkas, D. (2001b): "Specificity Distinctions". *Zas Papers in Linguistics* 23. 85-101.
- Fauconnier, G. (1984): *Espaces Mentaux*. Paris: Minuit.
- Franco, J. & E. Mejías-Bikandi (1999): "The Presuppositionality Condition and Spanish Clitic-Doubled Objects". In: J.M. Authier, B. Bullock & L. Reed (Eds.): *Formal Perspectives on Romance Linguistics*. Amsterdam: John Benjamins. 107-119.
- Geurts, B. (2002): "Specific Indefinites, Presupposition and Scope". In: R. Bäuerle, U. Reyle & T. Zimmermann (Eds.): *Presuppositions in Discourse*. Oxford: Elsevier.
- Haspelmath, M. (1997): *Indefinite Pronouns*. Oxford: Clarendon Press.
- Heusinger, K. Von (2001a): "Cross-Linguistic Implementations of Specificity". In: K. Jaszczolt & K. Turner (Eds.): *Meanings in Contrast: the Cambridge Papers*. Amsterdam: John Benjamins.
- Heusinger, K. von (2001b): "Specificity and Definiteness in Sentence and Discourse Structure". *Zas Papers in Linguistics* 24. 167-190.
- Jäger, G. (1995a): "Weak Quantifiers and Information Structure". *Nels* 25. 303-318.
- Jäger, G. (1995b): "Topic, Scrambling and Aktionsart". In: I. Kohlhof, S. Winkler & H.B. Drubig (Eds.): *Proceedings of the Göttingen Focus Workshop*. Tübingen. 19-34.
- Karimi, S. (1990): "Obliqueness, Specificity, and Discourse Functions: *Râ* in Persian". *Linguistic Analysis* 20, 3-4. 139-191.
- Karimi, S. (1996): "Case and Specificity: Persian *Râ* Revisited". *Linguistic Analysis* 26. 174-194.
- Karimi, S. (1999): "A Note on Parasitic Gaps and Specificity". *Linguistic Inquiry* 30. 704-713.
- Laca, B. (1987): "Sobre el Uso del Acusativo Preposicional en Español". In: C. Pensado (Ed.): *El Complemento Directo Preposicional*. Madrid: Visor. 61-91.
- Laca, B. (1990): "Generic Objects: Some More Pieces of the Puzzle". *Lingua* 81. 25-46.
- Laca, B. (1996): "Acerca de la Semántica de los Plurales Escuetos del Español". In: I. Bosque (Ed.): *El Sustantivo sin Determinación. La Ausencia de Determinante en la Lengua Española*. Madrid: Visor. 241-268.
- Laca, B. (1999): "Presencia y Ausencia de Determinante". In: I. Bosque & V. Demonte (Eds.): *Gramática Descriptiva de la Lengua Española*, I. Madrid: Espasa-Calpe. 891-928.
- Lazard, G. (1982): "Le Morphème *Râ* en Perse et les Relations Actanciennes". *Bulletin de la Société de Linguistique de Paris* 73. 177-207.
- Leonetti, M. (1998): "A Relevance-Theoretic Account of the Property Predication Restriction". In: A. Jucker & V. Rouchota (Eds.): *Current Issues in Relevance Theory*. Amsterdam: John Benjamins. 143-169.
- Leonetti, M. (1999): "El Artículo". In: I. Bosque & V. Demonte (Eds.): *Gramática Descriptiva de la Lengua Española*, I. Madrid: Espasa-Calpe. 787-890.
- Lidz, J. (1999): "The Morphosemantics of Object Case in Kannada". In: *Proceedings of Wccfl 18*. Sommerville (Mass.): Cascadilla Press.
- Lyons, C. (1999): *Definiteness*. Cambridge: Cambridge University Press.
- Martín, J. (1999): "The Syntax and Semantics of Spanish Accusative *a*". In: J. Gutiérrez-Rexach & F. Martínez-Gil (Eds.): *Advances in Hispanic Linguistics*, Vol. 2. Sommerville: Cascadilla Press. 469-485.
- McNally, L. (1995): "Bare Plurals in Spanish are Interpreted as Properties". In: G. Morrill & R. Oehrle (Eds.): *Formal Grammar*. Barcelona: Universitat Politècnica de Catalunya. 197-222.
- Meinunger, A. (2000): *Syntactic Aspects of Topic and Comment*. Amsterdam: John Benjamins.

- Melis, C. (1995): "El Objeto Directo Personal en *El Cantar de Mio Cid*. Estudio Sintáctico-Pragmático". In: C. Pensado (Ed.): *El Complemento Directo Preposicional*. Madrid: Visor. 133-163.
- Milsark, G. (1977): "Towards an Explanation of Certain Peculiarities in the Existential Construction in English". *Linguistic Analysis* 3. 1-30.
- Moreno, N. & I. Pérez (2001): "Information Structure and the Referential Status of Bare Plurals". *Zas Papers In Linguistics* 23. 159-170.
- Nakanishi, K. (2002): "Scope Encoding of Indefinite Nps in Japanese". In: K. von Heusinger, R. Kempson & W. Meyer-Viol (Eds.): *Proceedings of The Workshop "Choice Functions and Natural Language Semantics"*. Arbeitspapier Nr. 110, Fachbereich Sprachwissenschaft, Universität Konstanz. 141-165.
- Neeleman, A. & T. Reinhart (1998): "Scrambling and the Pf Interface". In: M. Butt & W. Geuder (Eds.): *The Projection of Arguments*. Stanford: Csl. 309-351.
- Nikolaeva, I. (2001): "Secondary Topic as a Relation in Information Structure". *Linguistics* 39-1. 1-49.
- Pensado, C. (1995): "El Complemento Directo Preposicional: Estado de la Cuestión y Bibliografía Comentada". In: C. Pensado (Ed.): *El Complemento Directo Preposicional*. Madrid: Visor. 11-59.
- Picallo, C. (1994): "A Mark of Specificity in Indefinite Nominals". *Catalan Working Papers in Linguistics* 4.1. 143-167.
- Portner, P. & K. Yabushita (2001): "Specific Indefinites and the Information Structure Theory of Topics". *Journal of Semantics* 18. 271-297.
- Rapoport, T. (1995): "Specificity, Objects and Nominal Small Clauses". In: A. Cardinaletti & M.T. Guasti (Eds.): *Syntax and Semantics, 28: Small Clauses*. New York: Academic Press.
- Rouchota, V. (1994): "On Indefinite Descriptions". *Journal of Linguistics* 30. 441-475.
- Sánchez López, C. (1995): "Construcciones Concesivas con *Para*". *Revista Española de Lingüística* 25. 99-123.
- Sperber, D. & D. Wilson (1986): *Relevance*. Oxford: Blackwell.
- Suñer, M. (1988): "The Role of Agreement in Clitic-Doubled Constructions". *Natural Language and Linguistic Theory* 6. 391-434.
- Torrego, E. (1998): *The Dependencies of Objects*. Cambridge, Mass.: Mit Press.
- Torrego, E. (1999): "El Complemento Directo Preposicional". In: I. Bosque & V. Demonte (Eds.): *Gramática Descriptiva de la Lengua Española, II*. Madrid: Espasa-Calpe. 1779-1805.
- Van Geenhoven, V. (1998): *Semantic Incorporation and Indefinite Descriptions*. Stanford: Csl.
- Van Geenhoven, V. & L. McNally (2002): "On The Property Analysis of Opaque Complements", Unpublished Paper.
- Wilson, D. & D. Sperber (1993): "Linguistic Form and Relevance". *Lingua* 93. 1-25.
- Yamamoto, M. (1999): *Animacy and Reference*. Amsterdam: John Benjamins.
- Yeom, J.-I. (1998): *A Presuppositional Analysis of Specific Indefinites*. New York: Garland.

Clitic Doubling and Clitic-Left Dislocation in Spanish and Greek as Native and as L2 Grammars

Teresa Parodi

tp209@cam.ac.uk

1. Introduction¹

This contribution deals with the status of clitic doubling and clitic-left dislocation in Spanish and Greek both in native and in second language grammars.

Based on syntactic and morphological evidence object clitics in Spanish and Greek are analysed as agreement markers along the lines of Suñer (1988) and Franco (2000), among others. Features of definiteness, specificity and case are shown to rule the cooccurrence of clitics with a coreferent NP in clitic doubling structures. Specificity also conditions clitic-left dislocation.

The conditions on the cooccurrence of a clitic and a coreferent NP in clitic doubling and clitic-left dislocation have been tested on native speakers and on learners of Spanish and Greek. The results will be compared to those of a similar test carried out on Macedonian, a language close to Spanish and Greek with respect to clitic doubling. English, on the other hand, has no clitics and uses strong pronouns (or empty categories) instead. Due to the different nature of clitics and pronouns there is no one-to-one correspondence between them, which gives rise to acquisition problems.

Under the acquisitional perspective the paper discusses morphological development in the acquisition of a language morphologically richer than the L1.

2. On the phenomenon of clitic doubling

Among pronominal elements clitics present curious research problems, such as the cooccurrence of a clitic with a coreferent NP in clitic doubling (CD) and clitic-left dislocation (CLLD). This phenomenon raises questions both about the status of the clitic and of the NP.

One of the many questions clitics raise is whether they should be analysed as arguments or as agreement markers and, as such, half-way between pronouns and inflectional elements. Traditionally in studies such as Kayne 1975 and Strozer 1976 among others, they have been treated as arguments although their behaviour differs from that of „normal arguments“ in various aspects: they have a fixed place in the sentence (preverbal with finite verbs, postverbal with nonfinite ones), while the placement of arguments is freer. As opposed to

¹ The paper is a revised version of a talk given at the workshop “Syntactic and semantic aspects of specificity in Romance languages” in Konstanz in October 2002. I would like to thank the participants for helpful discussion. The study is partly based on joint work with Ianthi Tsimpli for Greek and with Daniela Karadzovska for Macedonian. Both of them deserve thanks for long and fruitful cooperation.

arguments, clitics cannot combine freely with one another. Of special interest is the cooccurrence in a sentence of a clitic and a coreferent NP, difficult to explain if the clitic is the pronominal realisation of the corresponding argument.

Two lines of analysis are offered for the cooccurrence of clitic and NP, a movement and a base-generation analysis. The movement analysis is basically the line taken by Kayne 1975 and Strozer 1976; it assumes that clitics move from an argument position to the one in which they surface. The next question is where the coreferent NP is. Placing it in the trace left by the clitic would give rise to ungrammaticality. The other option is that the NP is extraposed, in a A' -position. However, assigning the same argument to both the clitic and the NP would violate the theta-criterion (Chomsky 1981).²

An alternative are base-generation analyses which take clitics to represent a separate category, basically as markers of an agreement relation with the verb, and not to form canonical object positions (Kaiser 1992, Kaiser & Meisel 1992, Suñer 1988, Franco 1993, 2000, Cummins & Roberge 1994, Torrego 1995, Sportiche 1996, Parodi 1998, for Greek Tsimpli 1996, Miseška Tomić 2000 for Macedonian, among others). The type of category (a head, a feature, etc.) varies according to different authors. For Torrego (1995) and Uriagereka (1995) clitics are D-heads and the doubled element is in SpecDP. For Sportiche (1996) clitics head their own projections, in this case a Clitic-Voice projection.

2.1 Spanish

In the current study I will follow Parodi (1998) and Franco (2000) who see clitics as instances of object agreement in a continuum which includes inflectional affixes on one end and pronouns on the other. As opposed to subject agreement, we do not observe agreement with the object across the board: object agreement follows certain conditions related to case and specificity. If we take Spanish, agreement with the dative is unrestricted, while there are stricter conditions for the accusative. These conditions are shown to be related to specificity and definiteness.

In Spanish if the element coreferent with the clitic is a pronoun, the presence of the clitic is obligatory, both in the dative and the accusative and in both varieties under consideration.

- (1) a. *accusative* la veo a ella / *veo a ella
 (I) her see K her / *(I) see K her
 'I see her'
- b. *dative* le doy la carta a él / *doy la carta a él
 (I) him give the letter K him / (I)give the letter K him
 'I give the letter to him'

² Further counterarguments are the extraction properties shown by these constituents, as pointed out by Jaeggli (1986), Suñer (1988) and Franco (2000), such as subjacency effects (Jaeggli 1986) or the "unorthodox" starting point for wh-movement if the NP is in an A' -position (Suñer 1988).

The cooccurrence of a clitic with a dative NP is optional in both dialects but, at least in Río de la Plata Spanish, strongly preferred (see 2). The situation does not change when the NP is definite, as in (2a) and (2b) or indefinite as in (2c) and (2d). Its status as + or –specific (cf. 2a, 2c vs 2b, 2d) does not make any difference either.

- (2) a. le doy la carta al vecino
(I) him give the letter K the neighbour
- b. al primer vecino que aparezca
K the first neighbour who might turn up
- c. a un vecino que usa bastón
K a neighbour who uses a walking stick
- d. a un vecino
K a neighbour

Clitic doubling with an accusative NP is disallowed in European Spanish. It is, however, possible in Río de la Plata Spanish, provided the NP is definite and specific as in (3a), but not if the definite NP is non-specific (3b), or if the NP is indefinite, whether + or –specific (3c-d).³

- (3) a. la veo a la mujer (RPlata European *) +definite
3accfs see.1s K the woman +specific
'I see the woman'
- b. *la busco a la mujer que sepa turco +def, -spec
3accfs search.1s K the woman who might know Turkish
'I am looking for the woman who might know Turkish'

³ If the -definite NP has a generic reading, then the cooccurrence of clitic and NP is possible, both in doubling and in clitic left-dislocation. Cf. the example below.

te la roban enseguida una bicicleta cara
2dat.s 3acc.f.s steal.3pl at once a bicycle expensive (fem.)
'They steal it (fem.) at once an expensive bicycle (fem.)'

- c. *la veo a una mujer -def, +spec
 3accfs see.1sK a woman
 'I see a woman'
- d. *la busco a una mujer que sepa turco -def, -spec
 3accfs search.1s K a woman who might know Turkish
 'I am looking for a woman who might know Turkish'

The degree of acceptability does not change according to animacy in either dialect, i.e. animacy is not one of the conditions ruling doubling. Compare the examples under (4) with inanimate NPs in the accusative (4a-b) with the animate counterparts in (2).

- (4) a. la veo la carta (RPlata European *)
 b. *la veo una carta

The conditions which guide the cooccurrence of a clitic and an NP are seen in Parodi (1998) as the interaction of a case and a specificity hierarchy. The first one is based on the mapping of thematic roles on syntactic functions following Larson (1988) as in (5) below:

- (5) agent >recipient/goal >theme/patient
 nominative >dative >accusative (indirect object)(direct object)

With respect to specificity I follow Comrie (1981), who establishes an “animacy hierarchy” similar to that in (6) below. A literal reading of “animacy” is in this case inaccurate. In this scale animacy, definiteness and specificity all represent ways of identifying an entity. As Franco (2000:169) puts it, animacy is here a cover term that includes the notions of saliency and definiteness / referentiality.

- (6) 1+2 pronoun > 3 pronoun>
 +animate NP > -animate NP
 +definite NP > -definite NP
 +specific NP > -specific NP

One could ask at this stage whether the fact that semantics is playing a role speaks against the view of clitics as agreement markers. The answer is no, given that Comrie (1981) shows that many agreement relations are driven by an animacy hierarchy which holds cross-linguistically. Swahili and Macedonian, the latter very close to Río de la Plata Spanish in this respect, also exclude agreement with the lowest elements of the Animacy Hierarchy, i.e. indefinite inanimate nouns.

We end up with some sort of parametric account of agreement depending on which elements of the scale are selected by different languages.

2.2 Clitic doubling in Greek

The conditions described above also apply to Greek. Greek is a language with rich overt morphology. NPs are marked for gender, number, case and definiteness, the latter by means of articles. Greek has unrestricted use of clitic doubling in the genitive, but again restricted to definites in the accusative, as shown in the examples under (7).

- (7) a. accusative, +definite
 I Eléni **ta** diábase **ta biblía** pou tis ícha dósi
 the Helen 3accnpl read.past the books that 3genfs had given.1sg
 ‘Helen read the books I gave her’
- b. accusative, -definite
 *O Kóstas **to** agórase chthes **ena avtokíneto**
 the Costas 3MS-acc bought yesterday a car
 ‘Costas bought a car yesterday’

2.3 Clitic doubling in Macedonian

Macedonian is a South Slavic language which makes extensive use of clitic doubling. Macedonian has a definite article which is affixal, attached to the first element of an NP and not necessarily to the head noun (cf. 8). Articles in Macedonian carry deictic and agreement features (as opposed to English where it does not carry either). They are also marked for gender (masculine and feminine), for number (singular/plural) and for case on the pronouns (on the full and on the reduced form, i.e. the clitic). In Macedonian there is concord within the nominal. Furthermore Macedonian marks speaker/hearer orientation on the article (cf. 9).

- (8) ja vidov kniga-ta
 I saw book-the
 ‘I saw the book’
- (9) -ov, -va, -vo, -ve, -va (near the speaker)
 -ot, -ta, -to, -te, -ta (near the hearer)
 -on, -na, -no, -ne, -na (far from both speaker and hearer)

Macedonian has unrestricted use of clitic doubling with the dative but restricted to definites in the accusative. These conditions are fairly similar to those in Río de la Plata Spanish. One difference is that in Macedonian clitic doubling is practically obligatory if possible, while it is slightly more restricted in Río de la Plata Spanish.

- (10) Macedonian
- a. accusative, +definite
 Mira **ja** donese **tetrakta-ta**
 Mira it.acc brought notebook-the
 ‘Mira brought the notebook’
- b. accusative, demonstrative
gi procitav vera **ovie napisi**
 3accpl read.1s yesterday these articles
 ‘I read these articles yesterday’
- c. accusative, proper name
gi vidov **Tome i Vera**
 3accpl saw Tome and Vera
 ‘I saw Tome and Vera’
- d. dative, +definite
 toj **im** isprati pismo **na roditeli-te**
 he 3datpl sent a letter to his parents

3. Clitic left dislocation

A clitic and a coreferent NP also cooccur in clitic-left dislocation constructions. For clitic doubling I have assumed, following base-generation analyses, that the NP is in an argument position. Clitic-left dislocation represents a case of A'-movement involving the doubled NP. Different analyses have been offered, most of which point out differences between clitic doubling and clitic-left dislocation for a variety of reasons (Cinque 1990, 1997, Escobar 1997, Zubizarreta 1998). Agouraki (1992), on the other hand, proposed a unified analysis for both clitic doubling and clitic-left dislocation in Modern Greek: according to her the clitic and the NP are in a Spec-head agreement configuration, with the NP in the Spec of a CliticPhrase. The latter is higher than IP and lower than a FocusPhrase in the tree. In clitic doubling the verb (and the clitic) are assumed to raise to the head F of the FocusPhrase, accounting for the fact that the verb is focussed in CD.

Interesting for the current discussion is the role of specificity in clitic-left dislocation. Relevant for the analysis is the notion of sentence topic. In cases like (11) below a gap is related to the sentence topic (as opposed to the discourse topic).

- (11) Juan lo veo
 Juan acc.3.s.m see.1s
 ‘Juan I see’

This NP displacement is not unrestricted. According to Zubizarreta (1998), sentence topics in clitic-left dislocation are constrained by a *specificity condition* (where specificity is understood as in Enç 1991). Specifically Zubizarreta claims that definites (12a) and specific

indefinites (12b) can be topics in clitic-left dislocation (but not non-specific indefinites (12c)).⁴ Consider the examples below:

- (12) a. a la mujer que tiene pelo azul Juan la busca.
 the woman who has.indic. hair blue Juan acc.3.s.f look for.3s
 ‘the woman who has blue hair Juan is looking for’
- b. a una mujer que tiene pelo azul Juan la busca.
 a woman who has. indic. hair blue Juan acc.3.s.f look for.3s
 ‘a woman who has blue hair Juan is looking for’
- c. *a una mujer que tenga pelo azul Juan la busca.
 a woman who has.subj hair blue Juan acc.3.f.s look for.3s
 ‘a woman who would have blue hair Juan is looking for’

A specificity/definiteness holds also for Greek, where there is no difference according to case for clitic-left dislocation. See (13) and (14) below.

- (13) accusative, +definite
To Gianni ton ida sto párti to sábbato
 the Iánnis.acc 3accmssaw at the party on Saturday
 John I saw at the party on Saturday
- (14) genitive, +definite
Tis Eléniso Pétrou **tis** chárise éna portofóli
 the Helen-gen the Peter-nom 3FS-gen gave a purse
 to Helen Peter gave a purse as a present

To sum up, the present study works with the following assumptions:

- a. Clitics in Spanish and Greek are agreement markers, spell-out of agreement features, head of their own projections (a.o. Suñer 1988, Franco 1993, Cummins & Roberge 1994, Torrego 1995, Sportiche 1996).
 - b. Clitic left dislocation involves A-bar movement and is subject to a specificity condition (Cinque 1990, Escobar 1997, Zubizarreta 1998).
 - c. Clitic doubling is conditioned by case and specificity (Parodi 1998, Franco 2000).
- The level of acceptability of doubling varies between obligatory, optional and disallowed under different conditions summarised in table (I) below:

⁴ Zubizarreta (1998) claims that the specificity requirement cannot be attributed to the clitic; if this was the case the specificity requirement would also hold for clitic doubling with the dative, which is in fact unrestricted. The fronted phrase in clitic-left dislocation, on the other hand, must be specific even with a dative clitic.

Table I: Clitic doubling in the languages under consideration

	European Spanish	Río de la Plata Spanish	Greek	Macedonian
obligatory	accusative, pronoun dative, pronoun	accusative, pronoun dative, pronoun dative, NP	genitive acc, NP, +def	dative acc, NP, +def
optional	dative, NP	accusative, NP +def, +spec		
disallowed	accusative, NP	accusative, NP +def, -spec accusative, NP -def	acc, NP, -def	acc, NP, -def

Spanish, Greek and Macedonian, unlike English, have object clitic pronouns which are, arguably, functional heads consisting of Case and AGR-O features, on a V-related projection (Uriagereka 1995, Sportiche 1996, Parodi 1998). That clitics are verbal rather than nominal categories, distinguishes them from English pronouns in a variety of morphosyntactic and semantic ways (Cardinaletti & Starke 1994).

This analysis was tested on native speakers of Spanish and Greek and L2 learners of the same languages with English L1 for clitic doubling and clitic left dislocation. Clitic doubling was also tested for Macedonian, again with native speakers and L2 learners.

The results of the test as carried out with the native speakers should provide information on the validity of the analysis. If clitics are arguments and the NPs coreferent with them extrasentential elements, all combinations of clitic and cooccurrent NP should be equally acceptable or unacceptable; variations along the grammatical conditions proposed in the described analysis are at least not expected. These variations are, however, expected if clitics are instances of object agreement, as well as an effect of case, definiteness and/or animacy.

Differences and similarities between natives and L2 learners will further provide information on the mechanisms which operate on the acquisition of a second language.

4. The test

Table II below gives information about the subjects tested.

Table II: subjects included in the study

	native speakers	L2 learners
Spanish	9	18
Greek	14	15
Macedonian	6	6

The subjects were assigned to different levels (lower, upper intermediate and advanced) according to the results of a cloze test. The learners of Spanish and Greek were recruited from local language schools, local schools, evening classes and university students.

The methodology used for Spanish and Greek was an acceptability judgement as described in Bard, Robertson & Sorace (1996). This is basically a gradability judgement, i.e. subjects

are required to classify utterances on a scale they establish themselves. Bard et al. (1996) validate the estimations then in terms of self-consistency. In order to make the different scales comparable the individual values assigned by the subjects are expressed in logarithms. The stimuli were presented in a random order visually and aurally, having been recorded by native speakers. The test sentences contrast case (accusative/dative), definiteness (+/-definite) and animacy (+/-animate).

For the analysis of Macedonian I follow Karadzovska (1999). She used a grammaticality judgment with a 1-5 scale. The graphs are based on the information in her paper. The data was collected from L2 learners who are at the pre-intermediate level of proficiency in Macedonian. The L2 learners are 6 English and American students living in Macedonia with an average of 4 months intensive classes of Macedonian (6-8 classes per week). However, all of them have lived in Macedonia for about 1.5 to 2 years, that is they have been constantly exposed to the L2 input and they are considered to be at a pre-intermediate to intermediate level of knowledge of the language. The control group consists of 6 Macedonian students all native Macedonian speakers. Karadzovska tested only definite NPs with the accusative and both definites and indefinites with the dative. The test items are actually controlled for specificity.

5. The test results

5.1 The clitic doubling test

The results are presented in graphs, which display on the x-axis the categories tested. On the y-axis we can read the ranking assigned to the different items by the subjects in logarithms. What we take into account is the ranking within each group: e.g. how the native speakers of Spanish rank accusative pronouns with respect to accusative NPs which are –definite and –animate. We also compare this pattern with the pattern found with the different groups of learners.

Consider the graph corresponding to clitic doubling in Spanish. The graph reveals a significant preference for clitic doubling with the dative over the accusative for native Spanish speakers, together with an effect of definiteness with the accusative, i.e. interaction of case and definiteness on clitic doubling. Both groups of English learners of Spanish, on the other hand, accept clitic doubling to a lesser degree than the natives. No effect of case or animacy and at least not a clear effect of definiteness can be observed for either group of learners.

With respect to Greek (see graph) we observe a clear effect of definiteness both for the native Greek speakers and for the for the upper group of learners, but not for the two lower groups.

The graphs for Macedonian are to be read similarly to those for Spanish and Greek. Recall, however, that Karadzovska (1999) used a 1-5 scale for her test. There is one graph for accusative and another one for dative contexts.

The graphs indicate a clear requirement of a clitic in specific contexts for native Macedonian speakers, both in accusative and in dative contexts. The learners on the other hand show a preference for (ungrammatical) options without clitics in accusative contexts, which could be interpreted as transfer from their English L1. Furthermore in dative contexts

their judgments indicate either no preference (+/-definite, +/-clitic; -definite, quantifier, +/-clitic; -definite, negative quantifier, +/-clitic) or a preference for an ungrammatical option without a clitic (-definite, -clitic) or the opposite choice to what they had opted for in an accusative context (proper name, +clitic). This cannot be interpreted as transfer: if it was transfer, they would choose the options without a clitic, as required by English.

5.2 The cooccurrence of a clitic and an NP in left-dislocation

As seen on the relevant graph, in clitic-left dislocation contexts there is no significant difference according to case in the native Spanish speakers' judgments, as opposed to clitic doubling. There is, however, a slight effect (not significant) of definiteness and animacy: they disfavour clitic/left-dislocation with an accusative, non-definite NP if this is also inanimate. No such effect of definiteness can be observed in the upper group of learners of Spanish; there are also no clear differences between clitic doubling and clitic left dislocation for them. Together with the results on clitic doubling, these results are compatible with the categorization of clitics as pronouns by the English learners of Spanish.

The results of the Greek test, as seen on the graph, reveal no effect of case either for the native speakers or for the upper group of learners. At the same time there is a clear definiteness effect: clitics are required in definite contexts and with clauses, again both by native Greek speakers and by the upper group of learners. The lower group of learners, on the other hand, do not seem to make clear choices, similar to their judgments of clitic doubling.

6. Summary and discussion

The test results from the native speakers groups are compatible with the analysis of clitics as agreement markers. Their judgments in all language groups point to differences according to case and definiteness, but not according to animacy. The definiteness effect is also obvious in clitic-left dislocation, confirming the line of analysis represented by Cinque 1990, 1997, Escobar 1997, Zubizarreta 1998. The difference in clitic-left dislocation in the accusative versus the dative or genitive predicted in Zubizarreta's (1998) analysis does not obtain.

The results from the learner groups, however, reveal that in their system doubling and left-dislocation are subject to partially different conditions as those operative in the L2 grammar. In Clitic-Doubling structures the English learners of Spanish, Greek and Macedonian allow for clitics regardless of definiteness and case features of the doubled noun phrase, contrary to the target languages. On the other hand, in Clitic Left-dislocation structures, the definiteness of the doubled noun phrase for learners of Spanish and Greek does seem to play a role. A possible explanation for this is a transfer effect from English which allows for Left-dislocation structures and fills the object or subject position with a pronoun. No direct transfer effect is possible in Clitic Doubling structures since English does not have a similar structure. Thus, the definiteness constraints responsible for the availability of a Left-Dislocation structure in English L1 are transferred to the L2 Spanish and Greek grammars. Lower learner groups exhibit direct transfer effects in all structures, thus disallowing clitics across the board.

Another linguistic property which appears to affect the process of L2 acquisition of different pronominal systems is overt agreement morphology. Spanish, Greek and Macedonian, in contrast with English, are languages with rich overt morphology both for

subject and for object agreement, realized in the form of a verbal suffix for subject agreement and object clitics for object agreement. Learners of Spanish, Greek (and Macedonian) have to change from a phonetically null category into the use of overt agreement morphology in subject and object position. The data indicate that English learners of Spanish or Greek, especially the upper groups, show a tendency to prefer the clitic option. This can be interpreted as the use of a generalized strategy that dictates the use of overt agreement morphology in the form of clitics in these structures. We can assume that the saliency of rich agreement morphology in the learners' input guides the restructuring of their grammar in changing from null to overt elements. Syntactically, the choice of the clitic could be triggered by the requirement to identify the empty category in the object position. The fact that case does not appear to play a role supports this suggestion. Features such as definiteness and specificity which rule the choice of a clitic or an empty category in the target language do play a role in these L2 learners' grammars. However, this role differs from the one they have in native grammars. Their use appears to be based on the saliency of overt morphology in the L2 grammars.

7. References

- Agouraki, Y. (1992): Clitic-left-dislocation and clitic doubling: a unification. *UCL Working Papers in Linguistics 4*: 45-70.
- Bard, E.G., Robertson, D. and Sorace, A. 1996: Magnitude estimation of linguistic acceptability. *Language 72/1*, 32-68
- Cardinaletti, A. & M. Starke (1994): The typology of structural deficiency. On the three grammatical classes. *Working Papers in Linguistics. Centro Linguistico Interfacoltà, Università degli Studi di Venezia 4,2*.
- Cinque, G. (1990): *Types of A'-dependencies*. Cambridge, Mass.: MIT Press.
- Cinque, G. (1997): 'Topic' constructions in some European languages and 'connectedness'. In: Anagnostopoulou, E.; van Riemsdijk, H.; Zwarts, F. (eds.): *Materials on Left Dislocation*. Amsterdam: John Benjamins. 93-118.
- Comrie, B. (1981): *Language Universals and Linguistic Typology*. Oxford: Basil Blackwell.
- Cummins, S. & Y. Roberge (1994): Romance inflectional morphology in and out of syntax. *MIT Working Papers in Linguistics 22*, 53-70.
- Enç, M. (1991): The semantics of specificity. *Linguistic Inquiry 22*, 1-26.
- Escobar, L. (1997), Clitic left dislocation and other relatives. In: Anagnostopoulou, E.; van Riemsdijk, H.; Zwarts, F. (eds.): *Materials on Left Dislocation*. Amsterdam: John Benjamins. 233-273.
- Franco, J. (1993): Conditions on clitic doubling: The agreement hypothesis. *Anuarios del Seminario de Filología Vasca "Julio de Urquijo", XXVII/2*, 285-298.
- Franco, J. (2000): Agreement as a continuum. The case of Spanish pronominal clitics. In: Beukema, F.; den Dikken, M. (eds.): *Clitic Phenomena in European Languages*. Amsterdam: Benjamins. 147-189.
- Kaiser, G. (1992): *Die klitischen Personalpronomina im Französischen und Portugiesischen*. Frankfurt: Vervuert.
- Kaiser, G. & J.M. Meisel (1992): Subjekte und Null-Subjekte im Französischen. In: Olsen, S.; Fanselow, G. (eds.): 'DET, COMP und INFL': *Zur Syntax funktionaler Kategorien und grammatischer Funktionen*. Tübingen: Niemeyer. 110-136.
- Kayne, R. (1975): *French Syntax. The transformational cycle*. Cambridge, Mass.: MIT Press.
- Karadzovska, D. (1999): Clitic doubling in Macedonian. Difficulties for English L2 learners. Research Centre for English and Applied Linguistics. *Working Papers in English and Applied Linguistics 6*: 153-178.
- Larson, R. (1988): On the double object construction. *Linguistic Inquiry 19*, 335-391.
- Miseška Tomić, O. (2000): On clitic sites. In: Beukema, F.; Den Dikken, M. (eds.): *Clitic Phenomena in European Languages*. Amsterdam: John Benjamins. 293-316.
- Parodi, T. (1998): Aspects of clitic doubling and clitic clusters in Spanish. In: Fabri, R.; Ortmann, A.; Parodi, T. (eds.): *Models of Inflection*. Tübingen: Niemeyer, 85-118.

- Sportiche, D. (1996): Clitic constructions. In: Rooryck, J.; Zaring, L. (eds.): *Phrase Structure and the Lexicon*. Dordrecht: Kluwer. 213-276.
- Strozer, J. (1976): *Clitics in Spanish*. PhD dissertation, UCLA.
- Suñer, Margarita (1988): The role of agreement in clitic-doubled constructions. *Natural Language and Linguistic Theory* 6, 391-434.
- Torrego, E. (1995): On the nature of clitic doubling. In: Campos, H. (ed.): *Evolution and Revolution in Linguistic Theory*. Georgetown University Press. 399-418.
- Tsimpli, I.M. (1996): Resumptive pronouns, learnability and L2A: a minimalist account. *Working Papers in English and Applied Linguistics*. Vol. 3. 23-42.
- Uriagereka, J. (1995): Aspects of the syntax of clitic placement in Western Romance. *Linguistic Inquiry* 26, 79-123.
- Zubizarreta, M. L. (1998): *Word Order, Prosody and Focus*. Cambridge, Mass.: MIT Press.

Specifically Brazilian Pronouns

Uli Reich

uli.reich@uni-koeln.de

1. Specific readings, coding of specificity and the discursive ensemble

Specificity is a semantic property of NPs that lies across the level of discourse and the level of the sentence. Therefore, it is difficult to describe as a whole. While we can describe definiteness very well in purely discourse-pragmatic terms, specificity can be determined by arguments within the sentence or by the speaker in discourse. In a recent paper, Klaus von Heusinger (2002) joins these two levels by the integration of the speaker into the analysis of the semantics of a given sentence:

[...] specificity indicates that an expression is referentially anchored to another object in the discourse. 'Referentially anchored' means that the referent of the specific NP is functionally dependent on the referent of another expression. Furthermore, I assume that this relation is sentence bound, i.e. a specific NP can only be anchored to discourse items that are explicit in the same sentence (or to the speaker of the sentence). (Heusinger 2002, 268).

Von Heusinger's characterization thus combines the classical understanding of specificity as 'identifiable for the speaker' with the possibility of its determination, or 'anchoring', by arguments of the sentence. This characterization is very convincing, but it is hard to understand why the relation to the speaker should be 'sentence bound'. In my view, the speaker cannot be regarded as part of the sentence, but only as part of the discourse. Thus, the knowledge of the speaker has to be analyzed at the level of discourse, defining 'discourse' as the interactive rearrangement of the ensemble of actual and epistemic knowledge, situation and linguistic expression in time.¹ As Heusinger puts the possibilities of anchoring, they are sentence bound if the anchor is an argument of the sentence, but discourse bound if the anchor is the speaker. The difference between definiteness and specificity, then, would be that definiteness is a relation of referential identity of discourse items, while specificity is a relation of referential assertion either by the speaker or by another nominal expression within the same sentence.

The cross-over property of specificity is mirrored by the fact that it can be mapped into the semantic structure of linguistic expression by more or less explicit means. It is less explicit, for example, in the most prominent examples of the classical literature on the topic: indefinite NPs like the famous student who cheated in the syntax class (Fodor & Sag 1982, 355). In these cases we speak of ambiguities on the level of sentence semantics that we turn transparent by continuing the sentences in question with an additional sentence, construing thus a kind of a discourse around it:

¹ Cf. Reich 2002, 46-53.

- (1) A student in the syntax class cheated on the final exam.
 a. His name is Noam.
 b. The wicked fellow used recursive rules and I want to know who it was.

We say that sentence (1) gets a *specific reading* with contexts like (1a) and a *non-specific reading* with contexts like (1b). The semantic property ‘specific’, here, is clearly not coded by a special morphological form. In the absence of linguistic coding within the sentence, it cannot be analyzed on the level of the sentence, but only inferred from extra-linguistic parts of the discursive ensemble.

Specificity is most directly conveyed by linguistic expression in the form of lexical carriers like *a certain* or *a particular*, etc. Non-specificity is expressed overtly by quantifiers like *every*, *all*, *some*, etc., if they have scope with respect to the NP in question. I expect these quantifiers to be universal.

In a third group, like subjunctive verb forms in Spanish relative clauses, specificity appears to be coded by morphological and/or syntactic forms that “normally” serve other functions:

- (2) a. Una mujer que sabe 5 lenguas va a coger el puesto.
 INDEF woman REI know3P 5 languagePl FUT PREP take DEF job.
 ‘A woman who knows 5 languages will get the job.’
 b. Solo una mujer que sepa 5 lenguas puede coger el puesto.
 OnlyINDEF woman REI know3P**Subj** 5 languagePl can take DEF job.
 ‘Only a woman who knows 5 languages can get the job.’

The indefinite NP in (2a) is specific, whereas in (2b) it is not. Maybe the best way to explain these effects is to say that the speaker indicates explicitly by the subjunctive form of the verb that he does not presuppose the individual existence of the NPs referent.

Finally, a fourth group is set up by instances of special morphological or syntactic coding, like the treatment of direct objects in languages like Turkish² and Spanish. In (3) we are forced to a specific interpretation in (3b) because of the preposition *a*:³

- (3) a. Voy a pintar una mujer muy linda. (not specific)
 FUT PREP paint INDEF woman very pretty
 ‘I will paint a very pretty woman’.
 b. Voy a pintar **a** una mujer muy linda. (specific)
 FUT PREP paint PREP INDEF woman very pretty
 ‘I will paint a very pretty woman’.

Thus, as specificity can be determined in the sentence or in other parts of the discursive ensemble, its manifestation can also either be expressed linguistically or inferred from discourse.

The coding of specific, animated direct object in Spanish leads us directly to the discussion of specificity and animacy as conditions for the use of anaphoric pronouns.

² See Enç 1991, Heusinger 2002.

³ See Manuel Leonettis contribution to this volume for a view that challenges this classical interpretation of the *complemento directo preposicional*.

2. The pronoun game

2.1 Differential object marking and clitic doubling

The *complemento directo preposicional* that we can observe in (3b) is a prominent example of what Georg Bossong (1998 and prior literature cited there) calls *differential object marking*. According to this theory, the direct object is a syntactic function that is split in two semantic subcategories. In many languages, these subcategories are treated differently: direct objects whose referents have agentive semantic features are marked in a special way in order to distinguish them from subjects, which of course show agentive features prototypically. Bossong situates these features in two semantic dimensions, an inherent one and a referential one:

La dimension inhérentielle (la hiérarchie de l'humanité)
[déictique] > [propre] > [humain] > [animé] > [discret]

La dimension référentielle (la hiérarchie de la définitude)
[défini] > [identifiable] > [indéfini] > [non-référentiel] (Bossong 1998, 203; 204)

Bossong's feature 'identifiable' corresponds largely to what we call 'specific'. In Bossong's theory it is not treated as a category that is independent from definiteness, but as a point on a graded scale of referentiality. Von Heusinger (2002, 248-253), however, citing insights from Quine, points to the fact that not only indefinite, but also definite NPs can have non-specific readings and that specificity therefore has to be considered an independent category:

I assume that specificity is a 'referential property' of NPs. This property cuts across the distinction of definite v. indefinite, like genericity. (Heusinger 2002, 252)

Thus, differential object marking is triggered by a conspiracy of three independent dimensions of the semantics of NPs: definiteness, specificity and animacy.

The clearest evidence of differential object marking in Spanish, however, is not the use of prepositions for full NPs, but instances of *clitic doubling* like the following:

- (4) a. **Lo** has visto a mi hermano?
Him have2PERS seePART PREP POSS1PERS brother
'Have you seen my brother?'
b. ***Lo** has visto mi libro?
It have2PERS seePART POSS1PERS book?
'Have you seen my book?'
c. ***Lo** has visto un niño?
Him have2PERS seePART INDEF boy
'Have you seen a boy?'
d. **Lo** has visto a un niño con una guitarra en la
Him have2PERS seePART PREP INDEF boy with INDEF guitar in DEF
mano
hand
'Have you seen a boy with a guitar in his hand'?

Clitic doubling is only possible with animate, specific referents of NPs, as in (4a) and (4d). Differential object marking with clitic pronouns in spoken Spanish seems to be related to a combination of the dimension of animacy and the dimension of specificity: neither NPs with specific, inanimate referents (4b) nor non-specific, animate referents (4c) receive clitic doubling, definiteness doesn't seem to play any role.

Clearly, the use of anaphoric pronouns is sensitive to the specificity of the NP it refers to. I would like to continue with some general remarks about this relation before I reach the main purpose of this paper, the discussion of the specificity of Brazilian pronouns.

2.2 Pronouns, specificity – and fiction

Anaphoric pronouns depend on the reference of their antecedents. As non-specific NPs do not refer, Martin Haspelmath (1997) assumes that only specific NPs can be referred to by anaphoric pronouns. He uses this notion as one of three ‘specificity tests’:⁴

(i) Only a specific NP can have a ‘discourse referent’, i.e. can be referred to by an anaphoric pronoun in a present indicative clause.

- (5) (a) Cheolbai bought a bicycle [specific]. It is black.
 (b) Cheolbai wants to buy a bicycle. [non-specific]. *It is black.

(Haspelmath 1997, 38 [my numeration])

In my view, however, the oddness of (5b) is not caused by a restriction of anaphoric pronouns to refer only to specific NPs, but by the impossibility to make attributive predications about a non-specific referent. Thus, it is quite normal to use pronouns referring to non-specific antecedents in propositions which do not describe the referent in question:

- (6) a. Cheolbai wants to buy a bicycle but he doesn’t have enough money for **it**.
 b. A student in the syntax class cheated on the final exam but we couldn’t identify **him**.

These examples suggest that the specificity of the antecedent is not a universal condition on pronominalization. In my opinion, it is and it is not. The contradiction between Haspelmath’s intuition and (6) can be explained as an effect of an old witchcraft of human discourse that easily passes many restrictions: fiction.

The examples in (6) are instances of what Lauri Karttunen (1976, 377-379) called *short term referents* which allow for coreferential pronouns and NPs within a limited domain. This domain can be extended in *suppositions*. In such a mode of referring, a non-specific NP can be elaborated with attributes (as long as the supposition is marked as such by some device, e.g. a modal operator) and serve as an antecedent for coreferring pronouns and other anaphora. Karttunen illustrates this fact with the following example:

- (7) If Mary had a car, she would take me to work in **it**. I could drive **the car** too.

(Karttunen 1976, 378)

More than only being possible, definite pronouns in suppositions contribute to establish the intended fiction of reality. Consider the difference between the treatment of the short term referent *a rich man* in (8a) and (8b):

- (8) Fritz wants to marry a Japanese girl.
 a. He is always looking for **her**.
 b. He is always looking for **one**.

⁴ Haspelmath (1997, 38, note 10) notes that the NP in (4b) can also have a specific reading and that in this case the pronoun would be fine. The other tests are (ii) paraphraseability with an existential sentence and (iii) the possibility of the insertion of *a certain*.

By this rhetoric technique, the referent in question gains (the possibility of) concrete existence: The possibility of really finding a rich man to marry sounds more realistic in (8a) than in (8b). In (8a) *the speaker treats the non-specific antecedent NP as if it was a referring, specific expression*. The moment of the use of an anaphoric pronoun to corefer with an NP that is non-specific at the time of its occurrence precisely marks the passage from a realistic to a fictitious discourse. In the non-specific reading of the NP *a japanese girl*, no referent is fixed by the knowledge of the speaker or of the referent of another nominal expression within the sentence. Nevertheless, the speaker tacitly assumes its existence and therefore uses an anaphoric expression, in this case a pronoun, to continue its relevance in the text.

Let us now have a closer look on the effects of specificity on pronouns in Brazilian Portuguese (henceforth BP).

3. Specificity and animacy as triggers of Brazilian pronouns

3.1 Overview of the pronominal system in contemporary BP

Clitic doubling is not a game Brazilians play. Due to a prosodic change that is bound to optimize the salience of syllables and feet, this variety⁵ omits the (sub-)segmental object clitics *o/a* which themselves are the results of historical word-internal lenition processes (Reich 2002). As a consequence of this loss, BP allows so called null objects (9a) in structural positions in which other Romance languages, including European Portuguese (henceforth EP), would use clitic pronouns (9b):

- (9) a. João descascou a banana, mas Pedro não comeu \emptyset .
 João peel-3P-Perf. DEF banana, but Pedro not eat3P-Perf.
 ‘João peeled a banana, but Pedro didn’t eat **it**’.
- b. João descascou a banana, mas Pedro não **a** comeu.
 João peel-3P-Perf. DEF banana, but Pedro not **it** eat3P-Perfect.
 ‘João peeled a banana, but Pedro didn’t eat **it**’.

However, we will see in the following sections that there are semantic conditions which favour the use of the free pronoun *ele/ela* over the null object. Thus, the free pronoun is used preverbally as a subject and postverbally as an object:⁶

⁵ ‘These varieties’, to be precise, but it is not *varietätenlinguistik* that interests us here.

⁶ Note that the use of *ele/ela* as an object pronoun is an option of Portuguese that can be traced back until the early middle ages. Cf. Penna 1998 for rich documentation of this fact.

- (10) A Joana_i é uma mulher muito bonita. Eu conheci [ela]_{obj}
 DEF Joana be3PINDEF woman very pretty. I meet1PPERF she
 na praia. [Ela]_{subj} estava lá com o namorado dela.
 in+DEF beach. She be1P-IMPERF there with DEF boy-friend of+she
 'Joana is a very pretty woman. I met her at the beach. She was there with her
 boy-friend.

Together with the grammaticalization of the new pronouns *você(s)* and *a gente*, which also can appear in both syntactic functions, this contributes to the general picture of a change of the Brazilian system of pronouns from a morphological case marking system to a configurational system. Just as with full NPs, the syntactic function of the pronouns is attributed by the topological position and not by the morphological form, a unique feature within romance languages.

Pronouns of BP can be put into two different, idealized systems, one of which is older, more conservative and characteristic for formal styles, the other being newer, innovative and characteristic for relaxed styles.

Table 1: Pronouns in BP

Old system: morphological case marking		New system: configurational case marking	
Subject	Object	Subject	Object
eu	me	eu	eu ⁷
tu	te	você/cê	você
Ø/ele, ela	o, a	ele, ela	Ø/ele, ela
nós	nos	a gente	a gente
vós	vos	vocês/cês	vocês
eles, elas	os, as	eles, elas	eles, elas

Note that these systems are idealized: Brazilian written and spoken texts nearly always show a mixture of them.⁸

3.2 Specificity and Brazilian pronouns in the recent discussion

The first prominent contemporary work on the choice between free pronouns and zero anaphora in object position was Maria Eugênia Duarte's sociolinguistic approach (1986). She highlights the effect of animacy: in her corpus, 92,4 % of free pronouns in object position had [+animate] antecedents (Duarte 1989, 24). We will see in section 3 that in contemporary, relaxed, spoken Portuguese in São Paulo, this is not a condition on the use of *ele/ela*.

Sonia Cyrino (1997) was the first to discuss the effects of specificity on the diachronic process of the loss of the third person object clitics. However, she did not consider these effects on the synchronic choice between the free pronoun and the null object in

⁷ Note that *eu* as an object still is restricted to very marked occasions, like imperatives in emotional styles: *solta eu, deixa eu*, etc.

⁸ Cf. the corpora in Reich 2002. It should also be noted that the relation of the 'new pronouns' with some verbs is marked by a preposition: *ele te/nos liga* v. *ele liga pra você/prá gente*. Purely configurational case marking is restricted to the 'most directly transitive' verbs as *see, love, kiss, kill* etc. The 'new pronominal system' also is not meant as the predictable goal of change in BP, rather as a structural possibility of spoken contemporary BP. There is no reason to assume the loss of the clitic object pronouns for the deictic persons in the future.

contemporary spoken BP and rather explained the null object within a theory of syntactic reconstruction without going much into the details of its variation with the free pronoun.

An effort to correlate specificity with the use of the free pronoun or the zero anaphora in object position was made later in a coauthored paper from the three Brazilian investigators who most contributed to the description and explanation of these phenomena: Sonia Cyrino, Maria Eugênia Duarte and Mary Kato. In Cyrino/Duarte/Kato (2000), the authors introduce what they call the ‘Implicational Mapping Hypothesis’, that regulates the distribution of empty and full pronouns in both of the core arguments, direct object and subject. BP is special within the Iberoromance languages not only for its null objects and configurational pronoun system, but also for what has been called the loss of the pro-drop property: pronouns in subject positions are nearly obligatory:

(11) a. [+human, + specific]:

Quando **ela** acordou, **ela** estava em Hong Kong.
When sheawake_{3P-PERF}shebe_{3P-PERF} in Hong Kong

‘When she awoke, she was in Hong Kong’.

b. [+human, -specific]:

Se a criança não recebe uma alimentação eficaz, **ela**
If DEF child not receive_{3P-PRES} INDEF alimentation efficient, she

fica em desvantagem pro resto da vida.
stay_{3P-PRES} PREP disadvantage PREP+DEFrest PREP+DEF life

‘If the child isn’t properly fed, it is handicapped the rest of its life’.

c. [-human, +specific]:

Por que você não aprontou o almoço_i na hora? - **Ele_i** está pronto.
Why you not get ready DEF lunch in time It be_{3P-PRES} ready

‘Why didn’t you get the meal ready on time? – It is ready’.

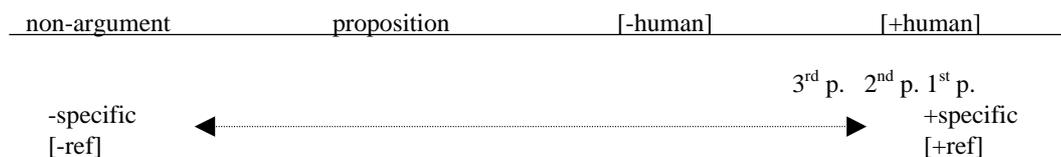
(Cyrino/Duarte/Kato 2000, 61-62)

In the examples (11), no dependency on any semantic property of the antecedent can be observed: all discourse topics are continued with pronouns if they are in subject position.⁹

In order to account for both phenomena, the preference for null objects *and* for subject pronouns, Cyrino/Duarte/Kato (2000) put forward their hypothesis based on a referential hierarchy that resembles similar ones from Bossong or Silverstein:

⁹ A diachronic study on the rise of this obligatoriness maybe could shed some light on the question if pronouns were first used in the case of antecedent NPs with specific referents.

I. Referential Hierarchy



II. The Implicational Mapping Hypothesis

- a. The more referential, the greater the possibility of a non-null pronoun.
- b. A null variant at a specific point on the scale implies null variants to its left in the referential hierarchy. (Cyrino/Duarte/Kato 2000, 59)

As we will see in the following sections, the predictions made by the hypothesis seem to be widely corroborated by diachronic and synchronic facts in BP. If we discuss subject and object realizations separately, Spanish also confirms the predictions made by the Implicational Mapping Hypothesis: only objects with specific referents can have a coreferential clitic within the same sentence (cf. section 1). On the other hand, subject pronouns are realized in this ‘pro-drop language’ only for purposes of contrast or focus and referents that are highlighted in discourse and sentence structure appear to be ‘specific by nature’:

(12) Un estudiante ha engañado en la clase de sintaxis.

‘A student cheated in the syntax class.’

a. **Él** no sabe resolver las tareas, mientras los otros lo saben. [+spec]
 He not know solve DEF exercises, while DEF others it know
 ‘He doesn’t know how to solve the exercises, while the others do know it.’

b. \emptyset se ha aprovechado de reglas recursivas. [-spec]

c. * **Él** se ha aprovechado de reglas recursivas. [-spec]

\emptyset /He himself AUX take advantage-PART of rules recursive

‘He took advantage of recursive rules.’

Problems appear in two misleading theoretical assumptions about the nature of specificity and referentiality that can be deduced from the form of representation of this hierarchy. First, referentiality and specificity appear as poles of a continuum. Both semantic concepts, however, can only be understood as discrete and binary: a linguistic expression either is referential or it is not, just as it is either specific or not. Second, specificity seems to be correlated to the semantic feature [+/- human]. In linguistic reality, this is not the case. Inanimate referents can be as specific as human referents, just as human referents can be as non-specific as inanimate referents. Note that Bossong put the two semantic properties into different dimensions, an ‘inherent’ one and a ‘referential’ one.¹⁰

In two recent papers, Scott Schwenter and Gláucia Silva (2002; forthcoming) explicitly investigate the effects of specificity and animacy on the two possibilities in BP and try to explain the choice in terms of Bossong’s differential object marking.

¹⁰ Cf. section 2.1.

They start with the observation that null objects are not possible for deictic persons, a claim that is not very surprising since the phonological process that omits the (sub)segmental clitics *o/a* does not affect the clitics for first and second person which prosodically are ‘optimal’ open syllables and, as deictic signs, have a completely different semantic status.

Schwenter and Silva then show very interesting effects of animacy, specificity and definiteness on the choice between free pronouns and null objects. Based on the judgments of five native speakers, they compare introspective examples that contrast in animacy, definiteness and specificity:¹¹

(13) [-anim, +spec, +def]:

Sabe a árvore grande que tinha na minha rua?
 Know DEF tree big REL be-IMPERF in+INDEF my street
 A prefeitura derrubou Ø/?ela.
 DEF City Hall knock down3P-PERF Ø/it.
 ‘You know the big tree that was on my street? City Hall knocked it down.’

(14) [-anim, +spec, -def]:

Na minha rua tem uma casa antiga linda,
 In+DEF my street be-PRES INDEF house old beautiful
 mas eles vão derrubar Ø/?ela.
 but they FUT knock down Ø/it.
 ‘On my street there is a beautiful old house, but they’re going to knock it down.’

(13) and (14) show that definiteness plays no role in the choice between free pronouns and null objects. In both cases, the informants of Schwenter and Silva apparently prefer the null object. In (15), they test an antecedent with an animated referent, choosing in this case a pet dog, an animal that is supposed to be something between an animal and a human being. The preferred choice in these cases is the free pronoun:

(15) [+anim, +spec, +def]:

O cachorro da Ana adora ir na rua.
 DEF dog of+DEF Ana love3P-PRESgo in+DEF
 Ela sempre leva ?*Ø/ ele para passear.
 She always take Ø/him PREP walk.
 ‘Ana’s dog loves to go out in the street. She always takes him for walks.’

(16) and (17) contrast specific with non-specific antecedents. For non-specific referents, the null object is preferred, while for specific referents the informants choose the free pronoun. Note that specificity is marked by the subjunctive mood of the relative clause in (16):

¹¹ Examples (13) to (17) are taken from Schwenter/Silva 2002. They are on pages 4-5 of the original manuscript that Scott Schwenter kindly sent me before its publication. The numeration follows that of the present paper.

(16) [+anim, -spec, -def]:

Eu estou procurando uma secretária que fale inglês,
 I be1P-PRES search-GERINDEF secretary REL speak3P-SUBJ English
 mas ainda não encontrei Ø/uma.
 but yet not find1P-PERF Ø/one.
 'I am looking for a secretary that speaks English, but I haven't found one yet.

(17) [+anim, +spec, +def]

Eu estou procurando uma secretária que fala inglês,
 I be1P-PRES search-GERINDEF secretary REL speak3P English
 mas ainda não encontrei ?*Ø/ela.
 but yet not find1P-PERF Ø/one.
 'I am looking for a secretary that speaks English, but I haven't found one yet.

Observing these examples, the authors conclude that

[...] null objects are permissible when the direct object referent is [-animate] or [-specific], or both. The acceptability of an overt pronoun varies depending on these same features, but an overt pronoun appears to be required, or at least comes closest to being required, only when the direct object is both [+animate] and [+specific]. Animacy and specificity, then, work in tandem in BP to determine the surface form of direct objects. (Schwenter/Silva 2002, 12)

Schwenter/Silva (forthcoming) basically corroborates this characterization quantitatively in the PEUL corpus of spoken Portuguese in Rio de Janeiro, conducted in the early 1980s. The recordings consist of interviews that were guided by researchers from the Universidade de Rio de Janeiro. The authors counted 1250 occurrences of anaphoric direct objects. The results confirm the characterization given above.

Table 2: Anaphoric direct objects in spoken Rio de Janeiro Portuguese

	+anim/+spec	+anim/-spec	-anim/+spec	-anim/-spec
Null	50 (24,4%)	102 (75,5%)	151 (87,3%)	604 (81,9%)
Pronoun	126 (61,5%)	12 (8,9%)	0 (0%)	13 (1,8%)
Full NP	29 (14,1%)	21 (15,6%)	22 (12,7%)	120 (13,3%)
Totals	205 (100%)	135 (100%)	173 (100%)	737 (100%)

(Schwenter/Silva forthcoming, 15)

We can see that animacy seems to be the decisive factor for the use of the free pronoun in this corpus. While the null object appears in every combination of the two semantic properties, the free pronoun is the preferred solution only for animated, specific referents. Thus, the authors basically come to the same results as Duarte (1986) concerning the importance of animacy.

3.3 Null objects and free object pronouns in contemporary spoken Portuguese of São Paulo

Interestingly, the corpora in Reich (2002) show a different picture of the importance of animacy. This corpus was not designed in the orthodox sociolinguistic style. The informants were asked to perform a set of linguistic tasks under a variety of different communicative conditions. These conditions were specially designed to check its influences on the choice

between the free pronoun and the null object. All 20 informants were born and raised in São Paulo, between 23 and 31 years old and had university degrees or were still studying. All of them were friends of the researcher who conducted the experiments. They also were familiar with the surroundings, so that a relaxed speaking style could emerge. The corpora were videotaped in December 1997 and January 1998. The analysis focused significant occurrences separately in its particular discursive embedding and deliberately refrained from quantitative methods.¹² The notation uses a vertical bar ‘ ’ to indicate the limits of intonation contours.

3.3.1 *The free pronoun with [-animated] referents of antecedent NPs*

Contrary to the results of Schwenter/Silva (forthcoming) and Duarte (1986), animacy is not a condition on the use of the free pronoun in object position. Examples of free pronouns whose antecedents have [-animated] referents abound in the corpora. (18) comes from a telephone corpus in which one of the informants asked about the locations of several things in the house, in this case a broom. Note that the speaker first uses the null object and then the pronoun for the same antecedent:

- (18) vassoura onde eu morava antes a gente deixava Ø
 broom where I live1P-IMPERF before we leave3P-IMPERF Ø
 atrás da porta né? então por isso eu coloquei **ela** atrás
 behind PREP+DEF door then therefore I put1P-PERF herbehind
 da porta
 DEF door.
 ‘Broom. Where I lived before, we left it behind the door, you know. So
 therefore I put it behind the door’

In (19), an example out of a corpus in which the informant was asked to explain how a steering-lock works, we can observe a similar realization. The speaker first uses two null objects, then the free pronoun and then turns to use the null object:

- (19) e depois tem uma outra parte que você passa em volta da
 and after be3PINDEFother part REL you put around PREP+DEF
 marcha engata Ø encaixa Ø na marcha deixa **ela** bem próxima
 gear arrange fix PREP+DEFgear leave herwell close
 do freio de mão e trava Ø
 PREP+DEF handbrake and lock
 ‘and then there is another part that you put around the gear, you arrange it, you
 fix it very tight to the handbrake and lock it’.

These examples could easily be continued. One might think that it is the referential distance that requires free pronouns, as in (18) and (19) the speaker starts with null constructions and then continues with the free pronoun, but (20), taken from a corpus in which the informants were asked to explain a Brazilian card-game, shows that this is not the case:

¹² Cf. Reich 2002, 80-82 for a full description of the method and arguments for the impossibility of quantitative studies on such material. Basically, I argue that occurrences out of corpora designed for sensitivity to a variety of pragmatic conditions cannot be compared quantitatively because the number of covariables increases in such an amount that it is not practicable to collect statistically representative samples.

- (20) A: se eu tivesse o ás de coração B: cê guarda ele
 if I have_{1P-PERF-SUBJDEF} ace_{PREP} heart you keep him
 A: 'If I had the ace of hearts? B: You keep it'.

We can resume the situation in contemporary spoken Portuguese from São Paulo in the following table:

Table 3: Null objects and free object pronouns in São Paulo

	+ animated	- animated
+ specific	ele	Ø / ele
- specific	Ø	Ø

(Reich 2002, 228)

Thus, the decisive factor for the use of free pronouns in object position in São Paulo is not animacy, but specificity.

How can the differences between the corpus of Reich (2002) and those of Duarte (1986) and the PEUL-group that used Scott Schwenter and Gláucia Silva be explained? I think there are three possible solutions to this question. Which one of them comes closer to the truth can only be determined by future research.

The first possibility is that realizations of free pronouns whose antecedents have [-animate] referents could have been omitted in the two corpora from Rio because of the more formal settings of the recordings. The use of the free pronoun in object position is still banned by conservative prescriptive grammars and many school teachers. In formal styles, Brazilians hardly ever use this option. In Reich (2002), all informants were friends of the researcher, knew the surroundings and had participated in other recordings of the same researcher, so that there was no stimulus for a switch to a different style. By contrast, the bigger corpora from Rio were recorded in rather anonymous situations with little familiarity between interviewers and informants. Against this hypothesis counts the fact that the free pronoun is banned for all referents of objects, though inanimated ones are considered to be 'worse'.

Second, there could be a difference between the varieties from Rio and São Paulo. The treatment of BP as a whole is a generalization that neglects the variation of a language with more than 170.000.000 speakers in an extremely complex society, to say the least.

The third possible explanation integrates the former two. The different results could be due to a change in the grammar of BP that happened between the early 1980s and the late 1990s. This is the most interesting possibility for the purpose of this paper and the scientific context that it tries to integrate. The hypothesis for this case would be that the loss of the third person clitic pronouns first led to null object constructions, an obvious claim that is shared by most researchers (Cyrino 1997, Kato 1994, Nunes 1993, Reich 2002). The next step would be that the free pronouns appear in object positions with animated antecedents. Animacy is a prototypical semantic feature of subjects, the traditional syntactic function of the free pronouns in Portuguese: the semantic similarity gives way to the change in the syntactic use. Finally, the free pronouns are generalized as object pronouns, dropping the semantic restriction of animacy: the syntactic similarity gives way to the semantic change. Specificity remains as the only restriction for the use of the free pronouns.

In this case, the different corpora simply show different stages of the change, independent of the question if these stages belong to different styles, different dialects or different times of recording the corpora. This question can only be answered by further research.

3.3.2 *Supposed specificity*

The corpora in Reich (2002) allow for a look on the use of pronouns with supposed referents in natural fictitious discourse, as has been sketched in section 2.3. In one of the corpora, a female informant (A) was asked to give a male second informant (B) advice on how to come in contact with girls. In the moment we enter the discourse, the informants already have spoken about *girls*, so that the NP *a(s) menina(s)* is definite, as marked by the definite article:

(21) B: o que seria bom
 what be-3P-CONDgood
 ‘What would be good?’

A: ah legal assim uma cantada diferente cê pode chegar na
 nice so INDEF singNOM different you can3P arrive in+DEF

menina convidar **ela** pra tomar uma coisa com você em outro
 girl invite shePREP take INDEF thing with you in other

lugar sei lá de repente fazer um programa legal um museu
 place know1P thereperhaps make INDEF programm nice INDEF museum

teatro depende da menina né assim se ela for meio
 theatredepend3P PREP+DEFgirl so if shebe3P-FUT-SUBJ half

intelectual chama **ela** pra ir num museu
 intellectual call-IMP she PREP go in+INDEF museum

‘A different way to come on to her. You can come close to her, invite her for a drink with you in another place, I don’t know, perhaps make a nice programme, a museum, theatre, it depends on the girl, if she is rather the intellectual type, call her to go to a museum’.

In (21), the referent of the NP is non-specific, but is referred to by the free pronoun, as the speaker sets up fictitious situations and possible descriptions of that referent. The pronoun thus corefers with a referent whose specificity is supposed in the ongoing fictitious discourse. In my view, this possibility shows that specificity is sensitive to the type of discourse and thus further contributes to the claim that specificity is a property of nominal expressions that can only be described across the levels of discourse and sentence semantics.

4. Conclusion

The distribution of null objects and free pronouns in object position in Brazilian Portuguese provides further evidence for the independence of the semantic dimensions animacy, specificity and definiteness. It is important to see that specificity is not bound to the speakers knowledge about some real world, but can also be supposed in fictitious discourse. This

explains why pronouns can corefer in some cases with non-specific NPs which normally do not establish discourse referents.

Contrary to the claim that the use of free pronouns in Brazilian Portuguese is restricted to antecedents with the inherent property animacy (Duarte 1986; Schwenter/Silva 2002; forthcoming), the study of corpora from contemporary spoken Portuguese from São Paulo clearly shows that the only restriction for object pronouns in this variety is the specificity of the referent of the antecedent NP. Definiteness doesn't play any role for the choice between null objects and object pronouns.

5. References:

- Bossong, G. (1998): Le marquage différentiel de l'objet dans les langues d'Europe. In: Feuillet, J. (ed.): *Actance et valence dans les langues de l'Europe*. Berlin, New York, 193-258.
- Cyrino, S. M. L. (1997): *O objeto nulo no Português do Brasil*. Londrina: UEL.
- Cyrino, S. M. L.; Duarte, M. E.; Kato, M. A. (2000): Visible subjects and invisible clitics in Brazilian Portuguese. In: Kato, M.; Negrão, E. V. (eds.): *Brazilian Portuguese and the Null Subject Parameter*. Frankfurt, 55-73.
- Cyrino, S. M. L.; Reich, U. (2002): Uma visão integrada ao objeto nulo no Português Brasileiro. In: *Romanistisches Jahrbuch 52*, 360-386.
- Duarte, M. E. L. (1986): *Variação e sintaxe: clítico acusativo, pronome lexical e categoria vazia no português do Brasil*. Unpublished masters thesis. Pontifícia Universidade Católica de São Paulo.
- Enç, M. (1991): The semantics of specificity. In: *Linguistic Inquiry 22/1*, 1-25.
- Fodor, J. D.; Sag, I. (1982): Referential and quantificational indefinites. In: *Linguistics and Philosophy 5*, 355-398.
- Haspelmath, M. (1997): *Indefinite Pronouns*. Oxford.
- von Heusinger, K. (2002): Specificity and definiteness in sentence and discourse structure. In: *Journal of Semantics 19*, 245-274.
- Karttunen, Lauri (1976): Discourse referents. In: McCawley, J. (ed.): *Syntax and Semantics 7: Notes from the linguistic underground*, 363-384.
- Kato, Mary (1994): A theory of null objects and the development of a Brazilian child grammar. In: Tracy, R.; Lattey, E. (eds.): *How Tolerant is Universal Grammar?* Tübingen, 125-153.
- Nunes, J. (1993): Direção de cliticização, objeto nulo e pronome tônico na posição de objeto em português brasileiro. In: Kato, Mary; Roberts, Ian (ed.): *Português Brasileiro: uma viagem diacrônica: homenagem a Fernando Tarallo*. Campinas, 207-222.
- Penna, H. M. (1998): *O Emprego do Pronome Tônico de Terceira Pessoa em Função Acusativa no Português: Mudança ou Retenção?* Unpublished masters thesis. Universidade de Minas Gerais.
- Reich, U. (2002): *Freie Pronomina, Verbalklitika und Nullobjekte im Spielraum diskursiver Variation des Portugiesischen in São Paulo*. Tübingen.
- Schwenter, S. A.; Silva, G. (2002): Overt vs. null objects in spoken Brazilian Portuguese: A semantic/pragmatic account. In: *Hispania 85*, 577-586.
- Schwenter, S. A.; Silva, G. (forthcoming): Anaphoric direct objects in spoken Brazilian Portuguese. In: *Revista Internacional de Lingüística Iberoamericana 3*.

Specificity and Countability in Old Italian

Elisabeth Stark

Elisabeth.Stark@romanistik.uni-muenchen.de

1. Introduction

The basic referential distinction in the semantic description of nominals in texts, namely the one between, informally speaking, ‘referring to one particular entity’ in contrast to ‘potentially referring to any arbitrary member of the class’ (cf. Lyons 1999, 165), is to be considered as fundamental when it comes to the emergence of systems of nominal determiners, e.g. in the Romance languages in the early Middle Ages. Noun phrases with the former interpretation are more likely to be explicitly determined than those with the latter (cf. Blazer 1979, Heinz 1982 for Old French, Elvira 1994 for Old Spanish or Selig 1992 for late Latin, Stark 2002 for Old Italian; see especially Givón 1981 who claims this to be universal in the grammaticalization from the numeral ‘one’ to an indefinite marker, spreading from those SPECIFIC uses to generics, predicative function, irrealis context and ultimately to direct negation, see below). This referential distinction is usually referred to as the distinction between SPECIFIC and NON-SPECIFIC interpretation of noun phrases. At least four different notions of SPECIFICITY can be found in the literature, i.e. „scopal specificity“, „epistemic specificity“, „partitive specificity“ and „relative specificity“ (for an useful overview and discussion cf. von Heusinger 2002).

The difference between ‘one particular entity’ (= SPECIFIC) and ‘any arbitrary member of the class’ (= NON-SPECIFIC) can semantically be conceived of as the non-varying (in the SPECIFIC case) or the varying (in the NON-SPECIFIC case) interpretation of the variable introduced by an indefinite NP in the ‘textworld’ (as in discourse representation approaches such as Heim 1988 or Kamp/Reyle 1993, cf. Farkas 2002). In other words: the interpretation of SPECIFIC NPs is the same throughout the whole text and cannot be altered by the linguistic context of the sentence containing the NP, while the interpretation of NON-SPECIFIC NPs depends directly on the (immediate) linguistic context. So, the NP *students of mine* in example (1) can refer to a certain group of students or to any of my students:

1) *I always meet students of mine in this bar.*

Slightly diverging, but still textual approaches to SPECIFICITY like the one proposed in von Heusinger 2002 explain this lacking variability or this interpretative independence of the new variable from the (immediate) linguistic context in the SPECIFIC case by “*referential anchoring*”, i.e. by the existence of some relation that the new text referent holds with an already introduced item in the ‘textworld’ (in example (1), this could also be the speaker of the sentence, remembering some funny facts about this special group or else, cf. also von Heusinger 1997, 13). This conception of SPECIFICITY (NON-SPECIFICITY being understood

accordingly as the lack of “referential anchoring”) highlights the ‘anaphoric aspect’ of SPECIFICITY.

There are important discourse-pragmatic consequences from this textual approach to SPECIFICITY: As speakers tend to choose SPECIFIC, i.e. ‘anchored’, referents as (central) topics of their discourse, signalling their SPECIFICITY can be crucial to a correct understanding or ‘parsing’ of the text by the hearer. More or less direct markers of SPECIFICITY are thus essential signals of textual thematic coherence, indicating not only ‘important anaphoric material’, but also those text elements that will be maintained (as topics) over longer portions of text. This issue relates to the well known ‘cataphoric aspect’ of SPECIFICITY and explains also quite directly the universally observable need of explicit determination of SPECIFIC rather than NON-SPECIFIC referents or their NPs, at least in early stages of determiner grammaticalization (see above). For modern article languages, however, SPECIFICITY or NON-SPECIFICITY of noun phrases is normally not directly marked by special determiners (with some exceptions, cf. *a certain, un certo* etc), but can be indicated by the context (some verb groups like “intensional predicates”, “world-creating predicates” such as *to look for, cercare* etc., also by the presence of some operators like negation etc, cf. Kamp 2001).

Coming back to the original motivation of explicitly determining an NP and of marking thereby the respective text referent, another argumentation, not directly based on SPECIFICITY, can be found, especially in the realm of nominal indefiniteness. Indefinite determiners, e.g. in the modern Romance standard languages French or Italian, can be seen as ‘classification devices’ in a broad sense. *Nominal classification* comprises the marking of the (NON-) COUNTABILITY of a noun phrase and belongs to the dimension of *apprehension* (which explains “how language grasps and represents concepts that correspond to objects or items”, Seiler 1986, 9). Apprehension is thus one of the means of reference which is activated before definitely determining the reference by nominal determination. In some languages, it can be the minimal condition for reference for indefinite noun phrases, whereas definite noun phrases can directly be identified by the speaker (via context or situation etc.) without being explicitly classified. In the case of nominal classification by indefinite articles, the potential referents of nominals become classified for example as ‘having a contour’ vs. ‘being diffuse’, like in the French example *un mouton* (‘one sheep’) vs. *du mouton* (‘mutton’, cf. Meisterfeld 1998 and 2000). This example shows that the term (NON-)COUNTABILITY used in this paper is not primarily concerned with extralinguistic physical features of the potential referents of nouns (cf. noun classes like “mass-denoting nouns” vs. “entity-denoting nouns”), but with a grammatically relevant feature of whole noun phrases, i.e. (in)compatibility with certain indefinite determiners and (im)possibility of having a morphological plural form (cf. Löbel 1993, Behrens 1995). In this sense, (NON-)COUNTABILITY is a typologically important grammatical feature of noun phrases in some languages, lacking in others, and correlated with other features of noun phrases like obligatory determination, morphological plural etc. (cf. Chierchia 1998, Longobardi 2001a and 2001b), possible in theory for any noun phrase with any kind of noun in its nucleus (cf. ital. *comprò un’acqua minerale speciale*, ‘he bought a special kind of mineral water’ vs. *comprò dell’acqua minerale*, ‘he bought mineral water’). (Classical) Latin did not have a classification and determination system via obligatory nominal determiners nor COUNTABILITY as a grammatical feature of its noun phrases, but was able, with its rich morphological declension system (declension classes and three genders), to

mark the difference between ‘contoured’ (*caseus*, ‘one (piece or sort of) cheese’) and ‘diffuse’ (*caseum*, ‘cheese as a substance’, cf. Lehmann 1991). The loss of this rich inflectional ‘classification system’ in the Romance languages might then have provoked the appearance of new ‘classification devices’ and the evolution of a new Romance classification system by indefinite determiners.

With these assumptions in mind, the aim of the present study is now to examine the precise function of indefinites (determiners and pronouns) in Old Italian texts and to discover why and where they appear and what they mark. It will be of central interest to see why there is a whole paradigm of indefinites at possibly any stage of (indefinite) article grammaticalization around the numeral ‘one’.

Old Italian or *Old Tuscan* (see below) is the term for some important Italo-Romance varieties from the 5th century until the first successful attempt of standardization (of the literary language) by Pietro Bembo in 1525. It comprises varieties with nominal determiners (in opposition to classical Latin) clearly less grammaticalized than in the modern Romance languages and represents thus an emerging system of nominal determination. The question will be if all indefinite determiners simply mark SPECIFIC referents and therefore if SPECIFICITY is thus the only feature necessary to describe adequately the position of ‘one’ and functionally related items on their grammaticalization path.

Our hypothesis is that there are in fact *two* key features organizing the indefinite paradigm: (NON-)SPECIFICITY for some elements (as a pragmatic indication for the hearer: ‘pay attention to this referent’, ‘this other one is not so important’) and COUNTABILITY as a basic category of *nominal classification* for the ‘core elements’ of the determiners’ paradigm: the indefinite articles *uno*, the partitive and ‘zero’.

These two features combine in the early uses of ‘one’ as an indefinite marker: ‘one of the X’ is SPECIFIC by referential anchoring to an already mentioned group in the text and COUNTABLE, ‘contoured’, by the original meaning of the numeral ‘one’. There seems to be an early ‘implicational relation’ between the two features in the beginning of determiner grammaticalization in that only SPECIFIC referents need to be explicitly classified. The question will be if this holds true for our texts, and what happens when *uno* and its satellites spread to ‘NON-SPECIFIC contexts’.

The paper is organized as follows: After shortly presenting the data base in section 2.1, sections 2.2 and 2.3 will then discuss the distribution of the 22 indefinite elements in the corpus texts relatively to some typical ‘SPECIFICITY’ and ‘NON-SPECIFICITY indicators’, in order to find out their eventual affinity to SPECIFIC or NON-SPECIFIC NPs. Section 2.4. will discuss the distribution of *uno*, the partitive and ‘zero’ and their potential as classification devices, before the synthesis of the results of part 2 will lead to a short conclusion (section 3) as to the actual stage of determiner grammaticalization and SPECIFICITY marking in the Old Italian corpus texts.

2. The Data

2.1 The Data Base

The data for the present study come from nine texts from the 14th to the 16th century, all written in Old Tuscan. The nine corpus texts can be grouped together in three ‘text groups’:

novellas, historiographical texts (in a rather broad sense) and ‘expositional-argumentative texts’ (cf. Stark (to appear)). According to their date of compilation, they belong to three subsequent periods, improperly but for simplicity’s sake called in the following *Trecento* (for the period from 1250 to 1350), *Quattrocento* (1350-1450) and *Cinquecento* (1450-1550; cf. table 1 in the appendix).

In these 9 texts, more than 8600 occurrences have been analyzed (up to 200 occurrences per lemma). They derive from 22 different lemmata, used as indefinite pronouns (marked as *_pro*) and/or determiners (marked as *_det*) in the position *_N* (not: *Det_N*), plus 100 bare singular (“*Zero Singular*”) and 100 bare plural noun phrases (“*Zero Plural*”) in each text. The 22 lemmata are listed in table 2 in the appendix.

2.2 Indefinites with Indicators of SPECIFICITY

In the abundant literature on SPECIFICITY, several indicators are listed which would all more or less induce or even directly provoke a SPECIFIC reading of the noun phrases cooccurring with these indicators. Since the ‘ambiguity approach’ of Fodor/Sag 1982 (cf. von Heusinger 1997 and 2002), the following characteristics of either the NP itself or of the surrounding linguistic material have been recurrently mentioned to be relevant ‘SPECIFICITY indicating factors’: rich lexical material inside the NP, in particular non-restrictive relative clauses (modified by Kleiber 1984: “*relatives spécifiantes*”: [attr.], [rel. clause]), position in topicalization constructions ([top.]) or subject position ([subject]). Furthermore, the only unequivocal marker of SPECIFICITY seems to be (according to Haspelmath 1997, 39) the perfectivity of the verb ([perfect.]; the elements in square brackets indicate the denomination of the different indicators used as analysis criteria for our indefinites in the corpus texts).

Table 3 in the appendix gives an overview over the percentage of occurrences of the 22 Old Italian indefinites with indicators of SPECIFICITY. The first two lines, reproduced here for reasons of clarity, contain the following information:

	A t t		R e l.		S u b j e c t		T o p .		P e r f e c	
<i>Alc_det</i>	383	31,84%	69	5,74%	216	17,96%	2	0,17%	196	16,29%
<i>Alc_pro</i>	96	30,19%	44	13,84%	193	60,69%	-	-	57	17,92%

Excerpt from Table 3

383 occurrences of *alcuno* in determiner position (*Alc_det*) are found in NPs with attributes (adjectives etc.), i.e. 31,84% of all *alcuno_det* occurrences in the corpus texts. 96 pronominal occurrences of *alcuno* are also attributed, i.e. 30,19% of all pronominal *alcuno_pro* occurrences in the corpus texts. 69 *alcuno_det* occurrences and 44 *alcuno_pro* occurrences have a non-restrictive relative clause as an attribute (the different categories are thus partly overlapping and cannot be summed up to 100%), which represents a 5,74% share with all *alcuno_det*-occurrences and of 13,84% with all *alcuno_pro* occurrences and so on.

The main results of the analysis for our discussion of SPECIFICITY are the following: 1) *Certo*, especially as a pronoun, *diverso*, *uno* and the **partitive** article seem to be rather clear markers of specificity. *Certo* and *uno* show both relatively high percentages for occurrences with attributes (49,40% for *certo_det*, 58,06% for *certo_pro*; 57,80% for *uno_det*, 47,87% for *uno_pro*), also with non-restrictive relative clauses (20,16% for *certo_det*, 29,03% for *certo_pro*; 18,64% for *uno_det*, 24,47% for *uno_pro*), in subject position (17,34% for

certo_det, 58,06 for *certo_pro*; 21,86% for *uno_det*, 42,55% for *uno_pro*) and with perfectly marked verb forms (*passato remoto* in Italian, see below; 38,31% for *certo_det* and 61,29% for *certo_pro*; 33,12% for *uno_det* and 26,60% for *uno_pro*). *Diverso* is a rather rare element, but seems to be specialized on perfective contexts (41,43% for *diverso_det*, 100% for *diverso_pro*). The partitive article is in part comparable with *uno* in its distribution (48,72% with attributes, 15,38% with non-restrictive relative clauses, 35,04% in subject position and 17,95% with perfectly marked verb forms). For each category, there may be higher rates for some other determiner or pronoun, but positive correlations between the five categories hold inconsistently only for these four elements.

2) Rather clear markers of NON-SPECIFICITY seem to be: *alcuno*, *chiunque*, *cosa*, *punto*, *persona*, *qualche*, *qualcosa*, *qualcuno*, *qualunque*, *roba*, *veruno* (the indefinites in **bold characters** will be subject to more intense discussion below). They have in common that they are rare or not attested at all in ‘SPECIFICITY contexts’ (*chiunque_det*, *punto*, *qualche_pro*, *qualcosa*, *qualcuno*, *roba*, *veruno_pro*) and in perfective contexts (from 16,29% for *alcuno_det* and 17,92% for *alcuno_pro*, 18,06% for *chiunque_pro*, 9,24% for *cosa* and 8,43% for *persona* down to 4,26% for *qualunque_det* and 4,55% for *qualunque_pro*, 3,85% for *veruno_det* and to 2,05% for *qualche_det*), while being possible or even being relatively frequent with attributed NPs, also with NPs with non-restrictive relative clauses as attributes.

3) These two groups are opposed to a rather large group of indefinites which seem to be reluctant to any classification according to their ‘SPECIFICITY affinity’: the negative indefinites (cf. *niente*, *nullo*: [+attr.], [- rel. clause], [+perfektiv], but [-subject]; *nessuno*, *niuno* [+attr.], [- rel. clause], [+top.], [+subject]; *niuno* [+perfektiv]) and some elements of „free-choice“ (cf. Haspelmath 1997): *qualsisia*, *qualsivoglia*: [+attr.], [- rel. clause], [+subject], [+perfektiv]. Besides these indefinites, bare singulars (“**zero singular**”) and bare plurals (“**zero plural**”) behave also not consistently: They are frequent in attributed NPs (45,11% and 43,89%), but rare with non-restrictive relative clauses (2,78% and 5,56%); they appear surprisingly often in perfective contexts, the most important indicator of SPECIFICITY (24,78% and 22,33%), but not as often in subject or topicalisation position (6,89% for bare singular subjects, 9,67% for bare plural subjects).

Pure statistical distributional evidence cannot, however, lead in every case to a satisfying description of the function of an indefinite, as can be shown in the following two examples:

2) *...per la qual cosa, essendo messer Torel di Stra per la sua nobiltà per lo essercito conosciuto, chiunque udì dire “Messer Torello è morto” credette di messer Torel di Stra e non di quel di Dignes...*

...by reason whereof, Messer Torello di Stra being renowned throughout the army of his magnificence, whosoever heard say ‘Messer Torello is dead’ believed it of Messer Torello di Stra, not of him of Dignes...

(Boccaccio, X, 9; all translations of Boccaccio taken from the revised John Payne translation from 1982, cf. references at the end of the paper)

The indefinite pronoun *chiunque* (‘whoever’), a ‘free-choice’ element, appears in subject position of a verb marked for perfective aspect (*udì*), but remains of course a ‘NON-SPECIFIC element’ by its special semantics. Lexical semantics play a major role also in the second example:

3) *I due fratelli, come che molta speranza non prendessono di questo, nondimeno se n' andarono a una religione di frati e domandarono alcuno santo e savio uomo che*

*udisse la confessione d' un lombardo che in casa loro era infermo; e fu lor dato **un frate antico di santa e di buona vita e gran maestro in Iscrittura e molto venerabile uomo, nel quale tutti i cittadini grandissima e speciale divozione aveano, e lui menarono.***

The two brothers, although they conceived no great hope of this, nevertheless betook themselves to a brotherhood of monks and asked for some holy and learned man to hear the confession of a Lombard who lay sick in their house. There was given them a venerable brother of holy and good life and a past master in Holy Writ, a very reverend man, for whom all the townsfolk had a very great and special regard, and they brought him to their house.

(Boccaccio, I, 1)

In this example, both *alcuno* and *uno* appear postverbally in internal argument NPs of a perfectly marked verb (*domandarono* and *fu dato*): The decisive factor determining the respective choice of *alcuno* in the first and *uno* in the second case is verbal semantics: *domandare* ('to ask for') is a "world-creating predicate" without existential presupposition of its arguments or referents and thus without any possibility of 'referential anchoring', whereas *fu lor dato* ('they were given') implies the existence of a SPECIFIC referent in argument position, which is also referentially linked to the before mentioned brotherhood.

To sum up: The 'indicators for SPECIFICITY' mentioned in the literature do not lead to clearcut classifications among the chosen paradigm of Old Italian indefinites. Perfective aspect and non-restrictive relative clauses seem to be rather important indicators of SPECIFICITY, but a close examination of the single occurrences of the indefinites is indispensable to achieve a full understanding of their (textual) functions.

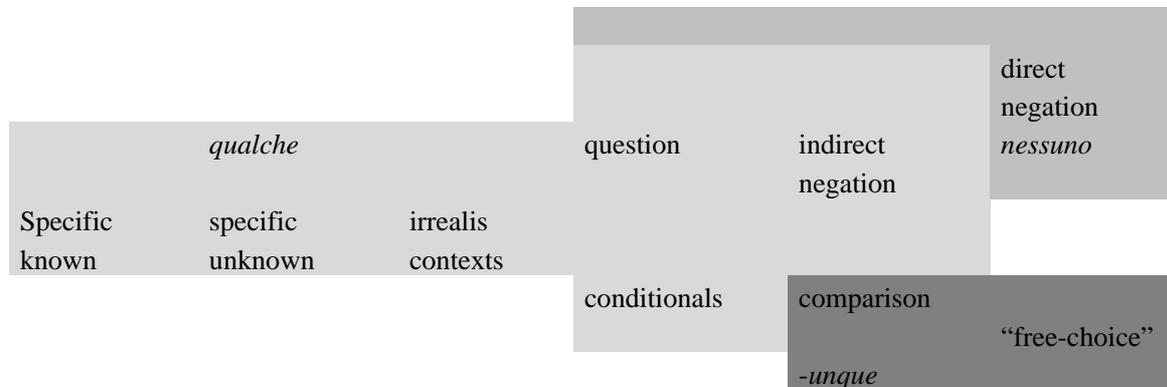
2.3 Indefinites with Indicators of NON-SPECIFICITY

Just like SPECIFICITY, NON-SPECIFICITY or at least ambiguity seems to be induced by certain linguistic factors in the immediate context of the NPs under discussion. The most prominent of these factors include "opaque contexts", i.e. "world-creating predicates" ([World-Cr.]), propositional predicates (*verba dicendi* etc.; [Propos.]); IRREALIS contexts (future, imperative; [Irr.Con.]), NON-FACT contexts (direct and indirect negation, [Dirneg] and [Indneg.], questions [Question], conditionals [Protasis], which are also contexts of negative polarity like the standard of comparison, [Comparison], cf. Givon 1978; again the elements in square brackets indicate the different indicators used as analysis criteria for our indefinites in the corpus texts). Most of these are, according to Haspelmath 1997, contexts of NON-SPECIFICITY and allow a description of the main functions of indefinite pronouns in the world:

SPECIFIC	Known to speaker	
	Unknown to speaker	
NON-SPECIFIC	Irrealis context	
	Negative polarity	Conditionals
		Questions
		Standard of comparison
		Indirect negation
		Direct negation
Free-choice		

Main functions of indefinites, Haspelmath 1997, 52

The following map for modern standard Italian represents an “implicational map” of “increasing hypotheticality” from left to right (cf. Haspelmath 1995), in which the indefinites mark continuous areas and spread diachronically from one function to another, always neighbouring function:



„Implicational map“ of the main functions of Italian indefinites, Haspelmath 1997, 70 (for an explanation of the main functions of indefinite pronouns cf. Haspelmath 1997, 52ff.)

The following English examples illustrate each of the IRREALIS and NON-FACT contexts sensitive to the distribution of indefinites in different languages:

- Irrealis contexts, for example imperatives: *Please try **somewhere** else!*
- Questions: *Did **anybody** tell you **anything** about it?*
- Conditionals: *If you see **anything**, tell me immediately.*
- Standard of comparison: *In Freiburg the wheather is nicer than **anywhere** in Germany.*
- Indirect negation: *I don't think that **anybody** knows the answer.*
- Direct negation: ***Nobody** knows the answer. (Haspelmath 1997, 2.)*

Table 4 in the appendix gives an overview over the distribution of the Old Italian indefinites in ‘NON-SPECIFICITY contexts’ in the corpus texts. It is to be interpreted in an analogous way to table 3:

	Worl	Prop	Irr.	Que	Prot a	Cop	Indn	Dirn								
Alc_d	60	4,99	86	7,15	179	14,8	33	2,74	65	5,40	60	4,99	48	3,99	303	25,1
Alc_p	13	4,09	17	5,35	51	16,0	8	2,52	25	7,86	30	9,43	17	5,35	76	23,9

Excerpt from Table 3

60 occurrences of *alcuno* as a determiner appear in the scope of “world-creating predicates”, i.e. 4,99% of the total amount of *alcuno_det* occurrences in the corpus texts, whereas only 13 pronominal occurrences of *alcuno* (4,09%) occur in the same contexts. 86 *alcuno_det* occurrences are found with propositional predicates (7,15%), and 17 pronominal *alcuno*-occurrences (5,35%). The different ‘NON-SPECIFICITY factors’ can be combined, as in the following example:

- 4) *E mai desse a tutti licenza in modo che in casa non fusse al continuo **qualcuno** a guardia delle cose, a ciò che, se caso avvenisse, sempre vi sia **qualcuno** apparecchiato.*
 And he would never let all of them go, so that there would not be anybody continuously in the house to guard the things, so that, in case of emergency, someone would always be prepared.
 (Alberti, III, 345)

The first occurrence of *qualcuno* in this example appears in the subordinate object clause of a ‘permission verb’ in the **conjunctive**, thus in an IRREALIS context, which is negated (indirect negation), just like the predicate verb of the subordinate clause (direct negation).

Table 4 shows the following distributional facts:

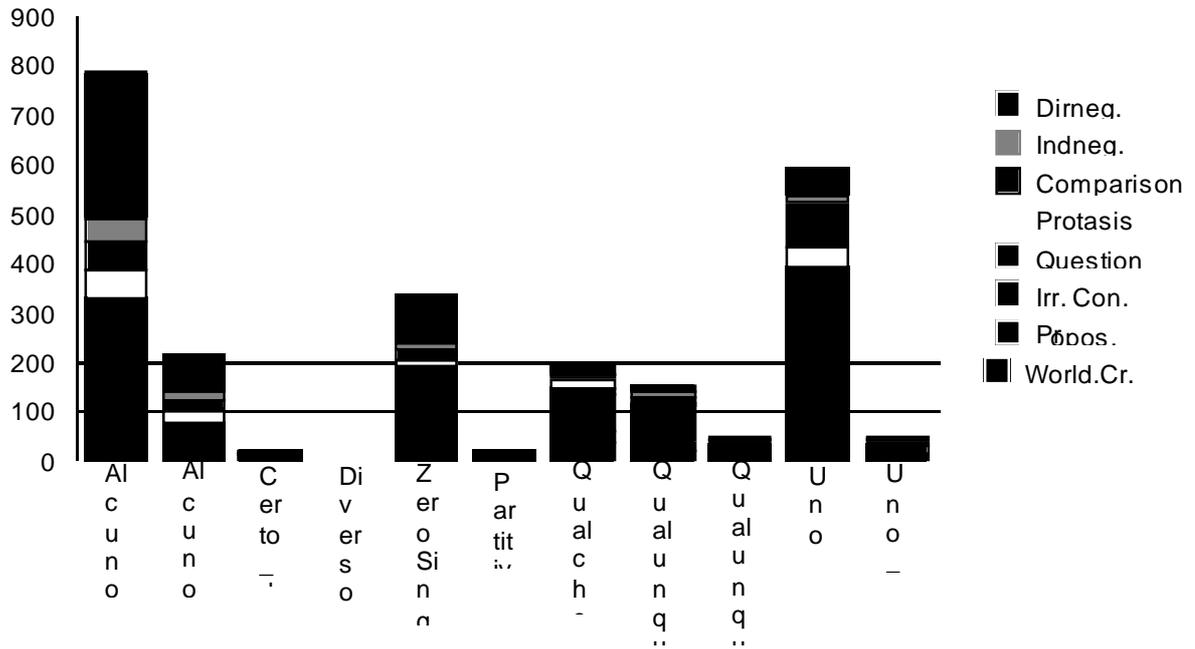
1) *Alcuno* and *veruno*, together with the pronominal *cosa*, *persona*, *punto* and *qualche*, appear particularly often in explicite negative contexts (more than 20% of occurrences with direct negation; and the aforementioned pronouns more than 50%), besides the negative indefinites, which indicates their clear status of negative polarity items (cf. Ladusaw 1996, Hoeksema 2000).

2) *Alcuno*, *cosa*, *niente*, *niuno*, bare singular and bare plurals, *nullo*, *persona*, *qualche*_det and *qualunque* occur in every single ‘NON-SPECIFICITY context’ analysed in table 4 and can therefore be considered as having a strong affinity to NON-SPECIFIC NPs. This corresponds perfectly to the findings of 2.2 for these indefinites.

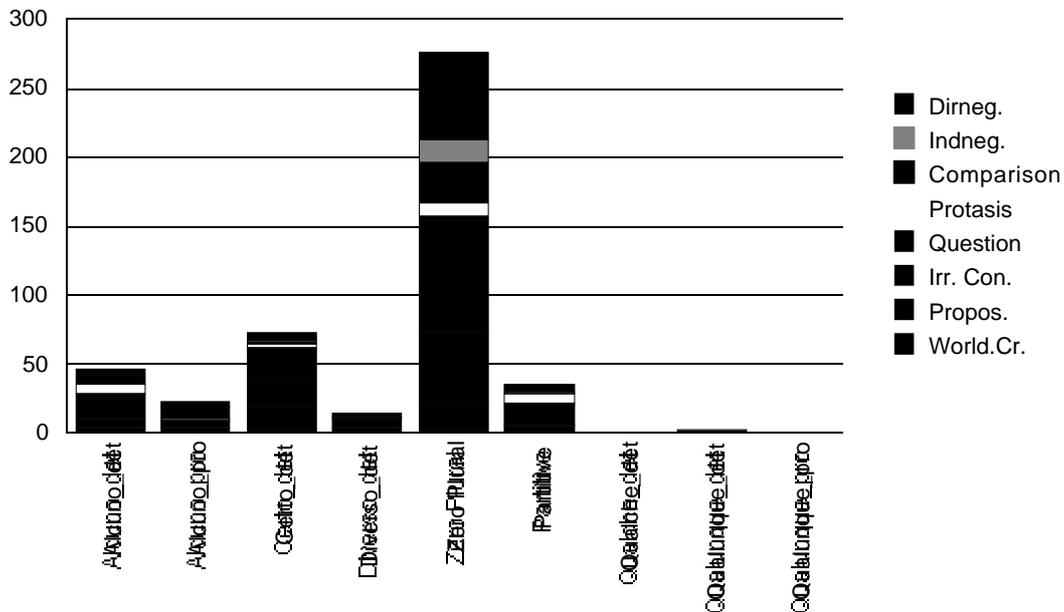
3) Occurrences in each single ‘NON-SPECIFICITY context’ hold equally for *certo*_det, the partitive and *uno*, but these latter indefinites show either very few occurrences in the more hypothetical contexts like the standard of comparison, indirect and direct negation (*certo*, partitive) or possess a mirror image distribution to *alcuno*, i.e. are relatively frequent in IRREALIS contexts (e.g. *uno* with 12,49% and 18,09% vs. *alcuno* with 14,88% and 16,04%), but have continuously decreasing frequency in more hypothetical contexts (e.g. *uno* in direct negation with 3,11% and 7,45% vs. *alcuno* with 25,19% and 23,90%). It is important to notice that *uno* and the partitive article are attested in all ‘NON-SPECIFICITY contexts’ and are thus already strongly grammaticalized. Out of the group of ‘SPECIFIC indefinites’ in 2.2., only *diverso* appears rarely in ‘NON-SPECIFICITY contexts’, both in absolute numbers and percentually, and can therefore remain classified as highly ‘SPECIFIC’.

As *alcuno* in its singular is a negative polarity item and means normally ‘none’ in modern standard Italian, in opposition to its plural (still meaning ‘some’), furthermore as *uno* is missing in the plural and as bare plurals are much more easily admitted in modern standard Italian than bare singulars (cf. Longobardi³1991, Renzi³1991), a separate analysis of singular and plural occurrences of the 8 most important indefinites in ‘NON-SPECIFICITY contexts’ seems to be indispensable. And indeed we find completely different distributions for singular or plural noun phrases, as can be seen from table 5 and 6 in the appendix and the two following tables (based on absolute number of occurrences):

SINGULAR



PLURAL



Singular and plural noun phrases with *alcuno*, *certo_det*, *diverso_det*, 'zero', partitive, *qualche_det*, *qualunque* and *uno* in 'NON-SPECIFICITY contexts'

Comparing the distribution of the indefinites in the singular and the plural, several findings deserve to be mentioned:

- 1) *Alcuno*, *uno* and bare singulars are the most frequent indefinites in the singular, whereas bare plurals lead by far in plural NPs in 'NON-SPECIFICITY contexts', especially after direct negation. This looks as if lacking determiners marks lacking SPECIFICITY (cf. Renzi ³1991).
- 2) *Alcuno* is the most important indefinite in the singular, often in parallel distribution with bare singulars, especially in IRREALIS contexts (15% and 11%) and with direct negation (27,49% and 11%).
- 3) Clearly 'specific' are *diverso* and *certo* with singular noun phrases. The plural of the latter, *certi*, is quite frequent with „world-creating predicates“ (19 occurrences, highest percentage with 4,97%) and “propositional predicates” (17 occurrences) and outranges *alcuni*, the morphological plural of a clearly NON-SPECIFIC indefinite. Yet, the more 'hypothetical' the contexts or the existence of the referent in the text world get, the more important *alcuni* gets, *alcuni* remaining in total much more SPECIFIC than its singular.

The following examples illustrate the distribution of bare singulars, *uno*_det and *alcuno*_det in IRREALIS contexts:

- 5) *...per ciò che ancora vivon di quegli che per questo si caricherebber di sdegno, dove di ciò sarebbe con risa da trapassare.*
...because there are still some of those alive who would feel disdain for this, where you should instead pass over it with laughter...
(Boccaccio, S.193)
- 6) *Figliuol mio, cotesta è buona ira, né io per me te ne saprei penitenza imporre...*
My son, this is a righteous anger, nor for my part might I enjoin you any penance therefor.
(Boccaccio, I, 1.)
- 7) *E appressandosi l'ora della cena, verso il palagio tornatesi con diletto cenarono; dopo la qual cena, fatti venir gli strumenti, comandò la reina che una danza fosse presa e, quella menando la Lauretta, Emilia cantasse una canzone da' leuto di Dioneo aiutata.*
...till suppertime drew near, when they returned to the palace and there supped merrily. Supper ended, the queen called for instruments of music and bade Lauretta lead up a dance, while Emilia sang a song, to the accompaniment of Dyoneo's lute.
(Boccaccio, I, Conclusione)
- 8) *La giovane, che prima la borsa d'Andreuccio e poi la contezza della sua vecchia con lui aveva veduta, per tentare se modo alcuno potesse a dovere aver quelli denar...*
The young woman, who had noted first Andreuccio's purse and then her old woman's acquaintance with him, began cautiously to inquire of the latter, by way of casting about for a means of coming at the whole or part of the money...
(Boccaccio, II, 5)

The relevant factors guiding the distribution of bare singulars, *uno* and *alcuno* in examples 5) to 8) in IRREALIS contexts seem to be:

- 1) the lexical class and the internal structure of the nouns and noun phrases: *danza* and *canzone* in example 7) with *uno* are 'more concrete' than the bare NPs *sdegno*, *penitenza* (examples 5) and 6)) and *modo* with *alcuno* (example 8);
- 2) the textual importance of the text referents introduced in the 'textworld' by the indefinite NPs: *danza* and *canzone* in example 7) with *uno* will be mentioned again in the text, will thus

gain ‘real existence’ and a certain relevance for the ongoing story, which is marked already at their first appearance in the text by *uno*. Together with example 3), examples 7) and 8) illustrate the opposition of *alcuno* and *uno*: the former appears with text referents of ‘doubtful existence’ and without ‘referential anchoring’, the latter with text referents (already) existing in the ‘textworld’ and being of some textual importance.

2.4 ‘Zero’, *uno* and the partitive: Indefinite determiners in ‘NON-SPECIFIC contexts’ and their potential of classification

The distribution of the partitive article, classified in 2.2. as having a certain affinity to SPECIFIC NPs, is different for singular and plural noun phrases. At first glance, the partitive seems to be more ‘NON-SPECIFIC’ in the plural than in the singular (where it lacks completely in the standard of comparison, with direct and indirect negation), occurring in every ‘NON-SPECIFICITY context’. But in relative numbers, singular occurrences of the partitive article are more frequent than plural occurrences with “world-creating predicates” (20% vs. 2,44%), “propositional predicates” (6,67% vs. 4,60%), in IRREALIS contexts (40% vs. 17,24%) and in questions (6,67% vs. 2,30%). The partitive is in general quite rare in these contexts, but relatively important in the singular, especially in IRREALIS contexts.

The following examples represent typical occurrences of the partitive article in IRREALIS contexts and in the protasis:

- 9) *Se se’ ricco, sia contento comperare **degli amici** co’ tuoi danari, se non ne puoi avere per altra via; ingegnati d’imparentarti **con buoni cittadini** e amati e potenti...*
If you are rich, be content with buying friends with your money, if you cannot have them in other ways; try to become friends with good, loved and mighty citizens...
(Morelli, III, 1517)
- 10) *...e se pure ne lasci **degli incarichi**, dà loro ispazio di tempo, se senti e’ rimanghino male agiati **a danari**, e non lasciare mai nulla in perpetuo...*
...and if you have left some things to do: give them time, and when you feel they do not work out right, try to move on with money, and do not leave anything undone...
(Morelli, III, 1263)
- 11) *Ma sopra tutto, se vuoi avere **degli amici e de’ parenti**, fa di non n’averne bisogno. Ingegnati d’averne **de’ contanti** e sappigli tenere e guardare cautamente, e que’ sono i migliori amici si truovino...*
But above all, if you want to have friends and relatives, be careful nor to depend on them. Try to have money and know how to keep it and to look carefully after it, because it is the best friend you can find...
(Morelli, III, 1806)
- 12) *...e danne a’ fanciulli: fa d’averne in casa, e fresca, e **del zucchero e dell’acquarosa e del giulebbo**.*
... and give it to the children: try to have some of it in the house, fresh, and sugar and *acquarosa* and sirop.
(Morelli, III, 2036)

The first three examples contain occurrences of the partitive plural with ‘animated entity-denoting nouns’ (*amici*, *parenti* in examples 9) and 11)), with ‘inanimated entity-denoting

nouns' (*contanti* in example 11)) and with 'abstract entity-denoting nouns' (*incarichi*, in example 10)). Example 12) typically shows occurrences of the partitive singular with "mass-denoting nouns". The presence or absence of the partitive article with plural noun phrases (always in COUNTABLE NPs) seems, among others, to depend on factors like attribution or position inside a PP (cf. *degli amici* vs. *con buoni cittadini* in example 9) or *de' contanti* vs. *a danari* in examples 10) and 11)). This holds also for modern Standard Italian. Yet, the distribution of the partitive singular is particularly sensitive to the lexical class of the respective noun:

SINGULAR	<i>Uno</i>		Partitive		'Zero'	
Trecento	573		9		109	
"Mass"	21	3,66%	7	77,78%	26	23,85%
<i>Abstract</i>	169	29,49%	-	-	78	71,56%
	<i>Uno</i>		Partitive		'Zero'	
Quattro	552		19		126	
"Mass"	24	4,35%	14	73,68%	28	22,22%
<i>Abstract</i>	201	36,41%	1	5,26%	92	73,02%
	<i>Uno</i>		Partitive		'Zero'	
Cinque	580		2		154	
"Mass"	4	0,69%	2	100,00%	11	7,14%
<i>Abstract</i>	223	38,45%	2	100,00%	108	70,13%

Table 7: *Uno*, Partitive and bare singulars with different lexical classes in singular noun phrases

PLURAL	Partitive		'Zero'	
Trecento	12		120	
"Mass"	-	-	-	-
<i>Abstract</i>	1	8,33%	42	35,00%
	Partitive		'Zero'	
Quattro	34		120	
"Mass"	-	-	-	-
<i>Abstract</i>	9	26,47%	58	48,33%
	Partitive		'Zero'	
Cinque	42		112	
"Mass"	-	-	-	-
<i>Abstract</i>	-	-	36	32,14%

Table 8: *Uno*, Partitive and bare plurals with different lexical classes in plural noun phrases

Table 7 permits to consider the distribution of *uno*, the partitive article and 'zero' as a classification system, *uno* appearing most often with concrete "entity-denoting nouns" and thus indicating COUNTABILITY, the partitive being specialized on "mass-denoting nouns" and therefore indicating NON-COUNTABILITY, and 'zero' being the typical 'non-determination' of

abstract concepts (which would be difficult to classify either as ‘contoured’ or ‘diffuse’). The situation with plural noun phrases in table 8 shows furthermore that the partitive plural has to be considered as the functional plural of *uno*, a hypothesis sustained by its affinity to SPECIFIC noun phrases in modern standard Italian (cf. Renzi ³1991) together with the fact that it highlights important text referents in examples 9) to 11) just like *uno* in example 7).

4. Conclusions

The discussion of the distribution of 22 indefinite elements in Old Italian texts in different ‘SPECIFICITY’ and ‘NON-SPECIFICITY contexts’ (cf. section 2.1. and 2.2) showed for some of the elements clear affinities to SPECIFIC or to NON-SPECIFIC NPs. On the one hand, *certo* and *diverso* seem to indicate SPECIFICITY, like *certo* still does in modern standard Italian. On the other hand, *alcuno* singular is clearly ‘NON-SPECIFIC’, just like *qualche* and *qualunque* (appearing also with direct negation; *qualche* will only later become an indicator of SPECIFICITY, cf. Haspelmath’s claims above). The morphological plural of *alcuno*, *alcuni*, however, is much more SPECIFIC than its singular, which is only one result of our analysis which illustrated the relevance of number in the investigation of distributional and functional properties of (Italian) indefinites.

A large number of indefinites, especially the negative indefinites and “free-choice elements” like *qualunque* cannot be fully described with the ‘SPECIFICITY’ and ‘NON-SPECIFICITY indicators’ listed in the literature; they seem to have strong lexical features determining their distribution and function in the texts in different ways.

Finally, three elements, the partitive article, *uno* and ‘zero’, show a rather widespread distribution in the corpus texts, appearing both with SPECIFIC and with NON-SPECIFIC NPs (the partitive still less frequently, being the ‘youngest’ indefinite determiner in Italian). ‘Zero’ or the absence of explicit determination is no longer automatically to be considered a mere indicator of ‘NON-SPECIFICITY’ (as in late Latin texts, cf. Selig 1992), but acquires a proper value in nominal classification (marking abstract nouns or concepts), just like the partitive (marking NON-COUNTABILITY). This fact does not simply hint at a stage of rather advanced grammaticalization, but documents above all an ‘inverted implicational relation’ between SPECIFICITY and COUNTABILITY or *nominal classification* as NP features governing nominal determination in emerging article systems (cf. introduction). At the very beginning of determiner grammaticalization, only or preferably SPECIFIC indefinite NPs are explicitly determined and thus classified according to their (NON-)COUNTABILITY. But later articles, i.e. highly grammaticalized determiners, appear also with NON-SPECIFIC NPs, losing their function as ‘SPECIFICITY indicators’, but keeping their (original) function as markers of (NON-)COUNTABILITY in languages or language families that lost their former nominal classification devices such as declension classes or gender systems.

5. Selected references

Texts analyzed in OVI und BIT via Internet; Addresses:
Opera del Vocabolario Italiano; <http://www.csovi.fi.cnr.it>.
Biblioteca Telematica Italiana; <http://www.cibit.unipi.it>.
 und LIZ (*Letteratura Italiana Zanichelli* 2.0).

Translation of Boccaccio's Decameron:

- Singleton, Charles S. (1982): *Giovanni Boccaccio: Decameron. The John Payne Translation Revised and Annotated by Charles S. Singleton*, 3 Bde., Berkeley/Los Angeles.- London: University of California Press.
- Behrens, L. (1995): Categorizing Between Lexicon and Grammar. The MASS/COUNT Distinction in a Cross-Linguistic Perspective. In: *Lexicology* 1-1, 1-112.
- Blazer, E. D. (1979): *The Historical Development of Articles in Old French*. Diss. University of Texas at Austin.
- Chafe, W. L. (1976): Givenness, Contrastiveness, Definiteness, Subjects, Topics and Point of View. In: Li, Ch. N. (ed.): *Subject and Topic*. New York, 25-55.
- Chierchia, G. (1998): Reference to Kinds across Languages. In: *Natural Language Semantics* 6-4, 339-405.
- Christopherson, P. (1939): *The Articles: A Study of Their Theory and Use in English*. Copenhagen.
- Declerck, R. (1986): Two Notes on the Theory of Definiteness. In: *Journal of Linguistics* 22, 25-39.
- Delfitto, D. / Corver, N. (1998): Feature Primitives and the Syntax of Specificity. In: *Rivista di Linguistica* 10-2, 281-334.
- Elvira, J. (1994): Un(o) en español antiguo. In: *Verba* 21, 167-182.
- Enç, M. (1991): The Semantics of Specificity. In: *Linguistic Inquiry* 22-1, 1-25.
- Farkas, D. F. (2002): Specificity Distinctions. In: *Journal of Semantics* 19-3, 213-243.
- Fodor, J. D. / Sag, I. A. (1982): Referential and Quantificational Indefinites. In: *Linguistics and Philosophy* 5, 355-398.
- van Geenhoven, V. (1998): *Semantic Incorporation and Indefinite Descriptions. Semantic and Syntactic Aspects of Noun Incorporation in West Greenlandic*. Stanford.
- Givón, T. (1978): Definiteness and Referentiality. In: Greenberg, J. (ed.): *Universals of Human Language*, vol. 4. Stanford, 291-330.
- Givón, T. (1981): On the development of *one* as an indefinite marker. In: *Folia Linguistica Historica* II, 35-53.
- Givón, T. (1983): Topic Continuity in Discourse: An Introduction. In: Givón, T. (ed.): *Topic Continuity in Discourse: A Quantitative Cross Language Study*. Amsterdam/Philadelphia, 1-41.
- Haspelmath, M. (1995): Referential Distinctions in Indefinite Pronouns: a Typological Perspective. In: Geiger, R. A. (ed.): *Reference in Multidisciplinary Perspective. Philosophical Object, Cognitive Subject, Intersubjective Process*. Hildesheim, 626-647.
- Haspelmath, M. (1997): *Indefinite Pronouns*. Oxford.
- Hawkins, J. A. (1978): *Definiteness and Indefiniteness. A Study in Reference and Grammaticality Prediction*. London.
- Heim, I. (1988): *The Semantics of Definite and Indefinite Noun Phrases*. New York.
- Heinz, S. (1982): *Determination und Re-präsentation im Altfranzösischen*. München.
- von Heusinger, K. (1997): *Saliens und Referenz: der Epsilonoperator in der Semantik der Nominalphrase und anaphorischer Pronomen*. Berlin.
- von Heusinger, K. (2002): Specificity and Definiteness in Sentence and Discourse Structure. In: *Journal of Semantics* 19-3, 245-274.
- Hoeksema, J. (2000): Negative Polarity Items: Triggering, Scope, and C-Command. In: Horn, L.; Kato, Y. (eds.): *Negation and Polarity. Syntactic and Semantic Perspectives*. Oxford – New York, 115-146.
- Jacobs, J. (1988): Fokus-Hintergrund-Gliederung und Grammatik. In: Altmann, H. (ed.): *Intonationsforschungen*. Tübingen, 89-134.
- Kamp, H. (2001): Anchored Representations and Indefinite NPs (Paper presented at the Workshop “Information Structure and the Referential Status of Linguistic Expressions”, 23th Annual Meeting of the Deutsche Gesellschaft für Sprachwissenschaft in Leipzig, Leipzig, 28. Februar – 2. März 2001).
- Kamp, H. / Reyle, U. (1993): *From Discourse to Logic. Introduction to Model-theoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Dordrecht.
- Karttunen, L. (1969, Reprint 1976): Discourse Referents. In: McCawley, J. D. (ed.): *Syntax and Semantics* 7. New York, 364-384.
- Kiss, K. É. (1995): On Generic and Existential Bare Plurals and the Classification of Predicates. In: Rothstein, S. (ed.): *Events and Grammar*. London, 145-162.
- Kleiber, G. (1984): Remarques sur l'opposition relative restrictive / relative appositive et l'article indéfini *un* spécifique. In: *Travaux de linguistique et de littérature* 22-1, 179-191.
- Ladusaw, W. A. (1996): Negation and Polarity Items. In: Lappin, Sh. (ed.): *The Handbook of Contemporary Semantic Theory*. Oxford, 321-341.

- Lambrecht, K. (1994): *Information Structure and Sentence form. Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge.
- Lavric, E. (1990): *Mißverstehen verstehen: Opake Kontexte und Ambiguitäten bei indefiniten und definiten Nominalphrasen*. Graz.
- Lehmann, Chr. (1991): The Latin Nominal Group in a Typological Perspective. In: Coleman, R. (ed.): *New Studies in Latin Linguistics. Selected Papers from the 4th International Colloquium on Latin Linguistics, Cambridge, April 1987*. Amsterdam/Philadelphia, 203-232.
- Löbel, E. (1993): On the Parametrization of Lexical Properties. In: Fanselow, G. (ed.): *The Parametrization of Universal Grammar*. Amsterdam/Philadelphia, 183-199.
- Longobardi, G. (1991): I quantificatori. In: Renzi, L.; Salvi, G.; Cardinaletti, A. (eds.) (1988-1993): *Grande grammatica italiana di consultazione, vol. 1*. Bologna, 645-696.
- Longobardi, G. (2001a): The Structure of DPs: Some Principles, Parameters and Problems. In: Baltin, M.; Collins, Chr. (eds.): *The Handbook of Contemporary Syntactic Theory*. Oxford – Cambridge/Massachusetts, 562-603.
- Longobardi, G. (2001b): How Comparative is Semantics? A Unified Theory of Bare Nouns and Proper Names. In: *Natural Language Semantics* 9-4, 335-369.
- Lyons, Chr. (1999): *Definiteness*. Cambridge.
- Meisterfeld, R. (1998): *Numerus und Nominalaspekt. Eine Studie zur romanischen Apprehension*. Tübingen.
- Meisterfeld, R. (2000): Die unbestimmte Bestimmung: Zur Entstehung des unbestimmten Artikels in den romanischen Sprachen. In: Staib, B. (ed.): *Linguistica romanica et indiana. Festschrift für Wolf Dietrich zum 60. Geburtstag*. Tübingen, 303-332.
- Primus, B. (1993): Word Order and Information Structure: A Performance-Based Account of Topic Positions and Focus Positions. In: Jacobs, J.; von Stechow, A.; Sternefeld, W.; Vennemann, Th. (eds.): *Syntax: Ein internationales Handbuch zeitgenössischer Forschung, vol 1*. Berlin; New York, 880-896.
- Prince, E. F. (1981): Toward a Taxonomy of Given-New Information. In: Cole, P. (ed.): *Radical Pragmatics*. New York, 223-255.
- Renzi, L. (1991): L'articolo. In: Renzi, L.; Salvi, G.; Cardinaletti, A. (eds.) (1988-1993): *Grande grammatica italiana di consultazione, vol. 1*. Bologna, 357-423.
- Russell, B. (1905): On denoting. In: *Mind* 56, 479-493.
- Seiler, H. (1978): Determination: A Functional Dimension for Interlanguage Comparison. In: Seiler (ed.), 301-328.
- Seiler, H. (1982): Das sprachliche Erfassen von Gegenständen (Apprehension). In: Seiler; Lehmann (eds.), 3-11.
- Seiler, H. (ed.) (1978): *Language Universals. Papers from the Conference held at Gummersbach/Cologne, Germany, October 3-8, 1976*. Tübingen.
- Seiler, H.; Lehmann, Chr. (eds.) (1982): *Apprehension. Das sprachliche Erfassen von Gegenständen. Teil I: Bereich und Ordnung der Phänomene*. Tübingen.
- Seiler, H.; Stachowiak, F. (eds.) (1982): *Apprehension. Das sprachliche Erfassen von Gegenständen. Teil II: Die Techniken und ihr Zusammenhang in Einzelsprachen*. Tübingen: Narr.
- Seiler, H. (1986): *Apprehension. Language, Object, and Order. Part III: The Universal Dimension of Apprehension*. Tübingen.
- Selig, M. (1992): *Die Entwicklung der Nominaldeterminanten im Spätlatein. Romanischer Sprachwandel und lateinische Schriftlichkeit*. Tübingen.
- Stark, E. (1997): *Voranstellungsstrukturen und topic-Markierung im Französischen. Mit einem Ausblick auf das Italienische*. Tübingen.
- Stark, E. (2002): Indefiniteness and Specificity in Old Italian Texts. In: *Journal of Semantics* 19-3, 315-332.
- Stark, E. (to appear): Mutamento linguistico nelle tradizioni discorsive: indefiniti e tipi di testo nel Trecento e nel Quattrocento. In: Rainer, F.; Stein, A. (eds.): *I nuovi media come strumenti per la ricerca linguistica*. Frankfurt a.M.
- Tschida, A. (1995): *Kontinuität und Progression. Entwurf einer Typologie sprachlicher Information am Beispiel des Französischen*. Wilhelmsfeld.

Appendix

	1250-1350 (= <i>Trecento</i>)	1350-1450 (= <i>Quattrocento</i>)	1450-1550 (= <i>Cinquecento</i>)
Novellas	<i>Il Novellino</i> (anonymous), 1280-1300 (27029 words)	<i>Il Decameron</i> (Giovanni Boccaccio), written in its major parts after 1348 (269588 words)	<i>Il Novellino</i> (Masuccio Salernitano), 1440-1475/76 (135102 words)
Historiography	<i>Nuova Cronica</i> (Giovanni Villani) 1308-1348 (481607 words)	<i>Ricordi di Giovanni di Pagolo Morelli</i> (Giovanni Morelli) 1393-1411 (372628 words)	<i>Istorie fiorentine</i> (Niccolò Machiavelli) 1519 - 1525 (147424 words)
Expositional- argumentative texts (philosophy, linguistic and sociological discussion)	<i>Il Convivio</i> (Dante Alighieri) 1304-1307 (73236 words)	<i>I Libri della famiglia</i> (Leon Battista Alberti) 1433/34 - 1436/37 or 1440, final compilation 1475/1476 (118638 words)	<i>Prose della volgar lingua</i> (Pietro Bembo) written from 1512 on, published in 1525 (67590 words)

Table 1: The corpus texts

<i>Alcuno_det</i>
<i>Alcuno_pro</i>
<i>Certo_det</i>
<i>Certo_pro</i>
<i>Chiunque_det</i>
<i>Chiunque_pro</i>
<i>Cosa_pro</i>
<i>Diverso_det</i>
<i>Diverso_pro</i>
<i>Nessuno_det</i>
<i>Nessuno_pro</i>
<i>Niente_det</i>
<i>Niente_pro</i>
<i>Niuno_det</i>
<i>Niuno_pro</i>
,Zero Singular'
,Zero Plural'
<i>Nulla_det</i>
<i>Nulla_pro</i>
Partitive
<i>Persona_pro</i>
<i>Punto_det</i>
<i>Punto_pro</i>
<i>Qualche_det</i>
<i>Qualche_pro</i>
<i>Qualcosa_pro</i>
<i>Qualcuno_pro</i>
<i>Qualsiasi_det</i>
<i>Qualsivoglia_det</i>
<i>Qualunque_det</i>
<i>Qualunque_pro</i>
<i>Roba_pro</i>
<i>Uno_det</i>
<i>Uno_pro</i>
<i>Veruno_det</i>
<i>Veruno_pro</i>

Table 2: The analyzed indefinites

	A t t		Rel.		Subject		T o p .		Perfec	
<i>Alc_det</i>	383	31,84%	69	5,74%	216	17,96%	2	0,17%	196	16,29%
<i>Alc_pro</i>	96	30,19%	44	13,84%	193	60,69%	-	-	57	17,92%
<i>Certo_d</i>	245	49,40%	100	20,16%	86	17,34%	1	0,20%	190	38,31%
<i>Certo_p</i>	18	58,06%	9	29,03%	18	58,06%	-	-	19	61,29%
<i>Chiun_d</i>	-	-	-	-	-	-	-	-	-	-
<i>Chiun_p</i>	2	2,78%	62	86,11%	27	37,50%	1	1,39%	13	18,06%
<i>Cosa_pr</i>	30	25,21%	82	68,91%	17	14,29%	-	-	11	9,24%
<i>Div_det</i>	32	22,86%	13	9,29%	14	10,00%	-	-	58	41,43%
<i>Div_pro</i>	-	-	-	-	-	-	-	-	1	100,00%
<i>Ness_de</i>	5	35,71%	1	7,14%	7	5 -	1	7,14%	1	7,14%
<i>Ness_pr</i>	1	6,25%	1	6,25%	12	75,00%	-	-	-	-
<i>Niente_</i>	1	50,00%	-	-	-	-	-	-	1	50,00%
<i>Niente_</i>	16	7,92%	2	0,99%	17	8,42%	1	0,50%	42	20,79%
<i>Niuno_d</i>	276	36,95%	63	8,43%	318	42,57%	5	0,67%	131	17,54%
<i>Niuno_p</i>	38	16,59%	34	14,85%	179	78,17%	-	-	22	9,61%
<i>Zero</i>	406	45,11%	25	2,78%	62	6,89%	-	-	223	24,78%
<i>Zero</i>	395	43,89%	50	5,56%	87	9,67%	-	-	201	22,33%
<i>Nullo_d</i>	61	29,47%	12	5,80%	66	31,88%	1	0,48%	48	23,19%
<i>Nullo_p</i>	36	12,08%	8	2,68%	125	41,95%	1	0,34%	31	10,40%
<i>Partitive</i>	57	48,72%	18	15,38%	41	35,04%	1	0,85%	21	17,95%
<i>Persona</i>	7	8,43%	9	10,84%	31	37,35%	-	-	7	8,43%
<i>Punto_d</i>	-	-	-	-	2	100,00%	-	-	-	-
<i>Punto_p</i>	3	60,00%	-	-	1	20,00%	-	-	-	-
<i>Qualche</i>	106	36,30%	10	3,42%	42	14,38%	-	-	6	2,05%
<i>Qualche</i>	2	100,00%	-	-	-	-	-	-	-	-
<i>Qualcos</i>	-	-	-	-	-	-	-	-	-	-
<i>Qualcun</i>	4	57,14%	-	-	4	57,14%	-	-	-	-
<i>Qualsi_</i>	1	50,00%	-	-	1	50,00%	-	-	-	-
<i>Qualsv_</i>	5	100,00%	-	-	1	20,00%	-	-	1	20,00%
<i>Qualun_</i>	87	33,72%	39	15,12%	43	16,67%	3	1,16%	11	4,26%
<i>Qualun_</i>	24	27,27%	34	38,64%	35	39,77%	-	-	4	4,55%
<i>Roba_pr</i>	-	-	-	-	-	-	-	-	-	-
<i>Uno_det</i>	986	57,80%	318	18,64%	373	21,86%	2	0,12%	565	33,12%
<i>Uno_pro</i>	45	47,87%	23	24,47%	40	42,55%	2	2,13%	25	26,60%
<i>Veruno_</i>	7	26,92%	1	3,85%	5	19,23%	-	-	1	3,85%
<i>Veruno_</i>	-	-	-	-	2	66,67%	-	-	-	-

Table 3: Indefinites with 'SPECIFICITY indicators' in the texts

	Worl		Prop		Irr.		Que		Pr o t a		Cop		Indn		Dirn	
<i>Alc_d</i>	60	4,99	86	7,15	179	14,8	33	2,74	65	5,40	60	4,99	48	3,99	303	25,1
<i>Alc_p</i>	13	4,09	17	5,35	51	16,0	8	2,52	25	7,86	30	9,43	17	5,35	76	23,9
<i>Certo</i>	26	5,24	23	4,64	29	5,85	1	0,20	5	1,01	2	0,40	1	0,20	9	1,81
<i>Certo</i>	-	-	1	3,23	-	-	-	-	-	-	-	-	1	3,23	-	-
<i>Chiu</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>Chiu</i>	10	13,8	4	5,56	8	11,1	-	-	-	-	1	1,39	1	1,39	1	1,39
<i>Cosa</i>	14	11,7	17	14,2	31	26,0	4	3,36	6	5,04	7	5,88	9	7,56	64	53,7
<i>Div_</i>	1	0,71	3	2,14	7	5,00	1	0,71	-	-	2	1,43	-	-	1	0,71
<i>Div_</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>Ness_</i>	1	7,14	2	14,2	5	35,7	1	7,14	-	-	-	-	2	14,2	3	21,4
<i>Ness_</i>	2	12,5	1	6,25	3	18,7	-	-	1	6,25	-	-	1	6,25	5	31,2
<i>Nient</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>Nient</i>	2	0,99	9	4,46	26	12,8	3	1,49	5	2,48	1	0,50	3	1,49	40	19,8
<i>Niun</i>	56	7,50	83	11,1	117	15,6	10	1,34	14	1,87	18	2,41	10	1,34	188	25,1
<i>Niun</i>	22	9,61	25	10,9	58	25,3	5	2,18	3	1,31	10	4,37	7	3,06	41	17,9
<i>Zero</i>	33	3,67	47	5,22	99	11,0	13	1,44	15	1,67	20	2,22	14	1,56	99	11,0
<i>Zero</i>	22	2,44	51	5,67	81	9,00	3	0,33	9	1,00	30	3,33	16	1,78	64	7,11
<i>Nullo</i>	17	8,21	9	4,35	14	6,76	1	0,48	1	0,48	11	5,31	2	0,97	58	28,0
<i>Nullo</i>	15	5,03	19	6,38	53	17,7	5	1,68	11	3,69	10	3,36	5	1,68	66	22,1
<i>Partit</i>	7	5,98	6	5,13	27	23,0	4	3,42	7	5,98	1	0,85	1	0,85	5	4,27
<i>Perso</i>	7	8,43	11	13,2	22	26,5	5	6,02	6	7,23	1	1,20	4	4,82	63	75,9
<i>Punto</i>	-	-	-	-	1	50,0	-	-	-	-	-	-	-	-	2	100,
<i>Punto</i>	-	-	-	-	-	-	-	-	2	40,0	-	-	-	-	3	60,0
<i>Qual</i>	38	13,0	24	8,22	73	25,0	12	4,11	23	7,88	2	0,68	9	3,08	23	7,88
<i>Qual</i>	-	-	-	-	-	-	-	-	1	50,0	-	-	-	-	1	50,0
<i>Qual</i>	-	-	-	-	1	100,	-	-	-	-	-	-	-	-	-	-
<i>Qual</i>	-	-	-	-	3	42,8	-	-	1	14,2	-	-	1	14,2	1	14,2
<i>Quals</i>	-	-	-	-	-	-	-	-	-	-	1	50,0	-	-	-	-
<i>Quals</i>	-	-	1	20,0	1	20,0	-	-	-	-	1	20,0	-	-	1	20,0
<i>Qual</i>	18	6,98	23	8,91	62	24,0	10	3,88	6	2,33	14	5,43	10	3,88	14	5,43
<i>Qual</i>	9	10,2	5	5,68	22	25,0	1	1,14	2	2,27	4	4,55	2	2,27	4	4,55
<i>Roba</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>Uno_</i>	58	3,40	100	5,86	213	12,4	25	1,47	35	2,05	93	5,45	16	0,94	53	3,11
<i>Uno_</i>	5	5,32	8	8,51	17	18,0	1	1,06	6	6,38	5	5,32	2	2,13	7	7,45
<i>Veru</i>	5	19,2	3	11,5	1	3,85	-	-	-	-	-	-	3	11,5	19	73,0
<i>Veru</i>	-	-	2	66,6	2	66,6	-	-	-	-	-	-	2	66,6	-	-

Table 4: Indefinites with 'NON-SPECIFICITY indicators' in the texts

SIN	Worl		Prop		Irr.		Que		Prot a		Cop		Ind n		Dirn	
<i>Alc_d</i>	56	5,22	80	7,46	161	15,00	32	2,98	59	5,50	56	5,22	48	4,47	295	27,49
<i>Alc_p</i>	12	5,31	15	6,64	47	20,80	6	2,65	23	10,18	22	9,73	17	7,52	72	31,86
<i>Certo</i>	7	6,14	6	5,26	4	3,51	-	-	3	2,63	-	-	-	-	3	2,63
<i>Div_</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>Zero</i>	33	3,67	47	5,22	99	11,00	13	1,44	15	1,67	20	2,22	14	1,56	99	11,00
<i>Partit</i>	6	20,00	2	6,67	12	40,00	2	6,67	1	3,33	-	-	-	-	-	-
<i>Qual</i>	38	13,15	24	8,30	72	24,91	12	4,15	23	7,96	-	-	9	3,11	23	7,96
<i>Qual</i>	18	7,20	22	8,80	60	24,00	10	4,00	6	2,40	14	5,60	9	3,60	13	5,20
<i>Qual</i>	9	10,59	5	5,88	22	25,88	1	1,18	2	2,35	4	4,71	2	2,35	4	4,71
<i>Uno_</i>	58	3,40	100	5,86	213	12,49	25	1,47	35	2,05	93	5,45	16	0,94	53	3,11
<i>Uno_</i>	5	5,32	8	8,51	17	18,09	1	1,06	6	6,38	5	5,32	2	2,13	7	7,45

Table 5: Singular noun phrases with *alcuno*, *certo_det*, *diverso_det*, 'zero', partitive, *qualche_det*, *qualunque* and *uno* in 'NON-SPECIFICITY contexts'

PLU	Worl		Prop		Irr.		Que		Prot a		Cop		Ind n		Dirn	
<i>Alc_d</i>	4	3,08	6	4,62	18	13,85	1	0,77	6	4,62	4	3,08	-	-	8	6,15
<i>Alc_p</i>	1	1,09	2	2,17	4	4,35	2	2,17	2	2,17	8	8,70	-	-	4	4,35
<i>Certo</i>	19	4,97	17	4,45	25	6,54	1	0,26	2	0,52	2	0,52	1	0,26	6	1,57
<i>Div_</i>	1	0,74	3	2,22	7	5,19	1	0,74	-	-	2	1,48	-	-	1	0,74
<i>Zero</i>	22	2,44	51	5,67	81	9,00	3	0,33	9	1,00	30	3,33	16	1,78	64	7,11
<i>Partit</i>	1	1,15	4	4,60	15	17,24	2	2,30	6	6,90	1	1,15	1	1,15	5	5,75
<i>Qual</i>	-	-	-	-	1	33,33	-	-	-	-	-	-	-	-	-	-
<i>Qual</i>	-	-	1	12,50	-	-	-	-	-	-	-	-	1	12,50	-	-
<i>Qual</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Table 6: Plural noun phrases with *alcuno*, *certo_det*, *diverso_det*, 'zero', partitive, *qualche_det*, *qualunque* and *uno* in 'NON-SPECIFICITY contexts'

Addresses of the Contributors

Susann Fischer
Institut für Linguistik und Anglistik
Universität Stuttgart
Keplerstr. 17
70174 Stuttgart
Germany
sfisher@ifla.uni-stuttgart.de

Rodrigo Gutiérrez Bravo
CIESAS
Juárez 87
Col. Tlalpan
México D.F., C.P. 14000
Mexico
gutierrz@cieras.edu.mx

Klaus von Heusinger
Institut für Linguistik/Germanistik (ILG)
Universität Stuttgart
Postfach 10 60 37
70049 Stuttgart
Germany
vonHeusinger@ilg.uni-stuttgart.de

Manuel Leonetti
Departamento de Filología
Facultad de Filosofía y Letras
Universidad de Alcalá
c/ Trinidad, 5
28801 Alcalá de Henares (Madrid)
Spain
m_leonetti@hotmail.com

Georg A. Kaiser
Fachbereich Sprachwissenschaft
Universität Konstanz
D-78457 Konstanz
Germany
Georg.Kaiser@uni-konstanz.de

Teresa Parodi
University of Cambridge
Research Centre for English and Applied
Linguistics
Keynes House
Trumpington St
Cambridge CB2 1QA
UK
tp209@hermes.cam.ac.uk

Uli Reich
Universität zu Köln
Romanisches Seminar
Albertus-Magnus-Platz
D-50923 Köln
Germany
uli.reich@uni-koeln.de

Elisabeth Stark
Institut für Romanische Philologie
Ludwig-Maximilians-Universität München
Ludwigstraße 25
80539 München
Germany
Elisabeth.Stark@romanistik.uni-muenchen.de

Sekretariat des Fachbereichs Sprachwissenschaft
Frau Ursula Haase,
Fach D 185, D-78457 Konstanz, Tel. 07531/88-2465,
Email: Ursula.Haase@uni-konstanz.de.

Hefte mit einem Stern * sind vergriffen. Einige Arbeitspapiere sind elektronisch verfügbar. Sie liegen als Dateien im PDF-Format vor und können per ftp bezogen werden. Zum Betrachten und Ausdrucken wird ein Acrobat Reader benötigt. Eine Liste der bis 1999 erschienenen Arbeitspapiere ist ebenfalls elektronisch verfügbar:

<http://ling.uni-konstanz.de/pages/publ/arbeitspapiere.html>
<http://www.ub.uni-konstanz.de/serials/fb-sprach.htm>

100. Veronika Knüppel: *Two lexical-functional analyses of French*. Mai 1999, DM 4,00.
101. Milena Vegnaduzzo: *I verbi italiani in -sc: morfologia e semantica in prospettiva diacronica*. Juni 1999, DM 3,50.
102. Vieri Samek-Lodovici: *The Internal Structure of Arguments: Evidence from Italian Nominalization-based Complex Predicates*. August 1999, DM 3,50
103. Matthias Weisgerber: *“Öffnen” und “aufmachen”. Eine empirische Studie zum Gebrauch zweier Verben*. Oktober 1999. DM 4,50.
104. Regine Eckardt: *On the underlying mechanics of certain types of meaning change*. November 1999, DM 4,50.
105. Zvi Penner, Petra Schulz and Karin Wymann: *Normal and Impaired Language Acquisition II. Studies in Lexical, Syntactic, and Phonological Development*. November 1999, DM 10,-.
106. Regine Eckardt & Klaus von Heusinger (eds.): *Meaning Change – Meaning Variation. Workshop held at Konstanz*. Feb. 1999, Vol. I, Feb. 2000, DM 7,-.
107. Christoph Schwarze: *Aspetti semantici della formazione delle parole*. April 2001, DM 2,-.
108. Nikolaus Schpak-Dolt: *Bibliographische Materialien zur romanischen Morphologie 1875–1950*. Oktober 2001, DM 5,-.
109. Klaus von Heusinger: *Italian Nominalization of -ata: Derivation and the Structure of the Lexicon*. Januar 2002, EUR 1,-.
110. Klaus von Heusinger & Ruth Kempson & Wilfried Meyer-Viol (eds.): *Proceedings of the Workshop “Choice Functions and Natural Language Semantics”*. Januar 2002, EUR 3,50.
111. Klaus von Heusinger & Christoph Schwarze: *Underspecification in the semantics of word-formation. The case of denominal verbs of removal in Italian*. Oktober 2002, EUR 1,-.
112. Götz Wienold & Christoph Schwarze: *The Lexicalization of Movement Concepts in French, Italian, Japanese and Korean: Towards a Realistic Typology*. Oktober 2002, EUR 1,-.
113. Klaus von Heusinger & Georg A. Kaiser (eds.): *Proceedings of the Workshop Semantic and Syntactic Aspects of Specificity in Romance Languages*. April 2003, EUR 3,50.