Spanish Clitic Doubling:

A Study of the Syntax-Pragmatics Interface

by

Valeria A. Belloro
October 26, 2007

Acknowledgements

I would like to thank first and foremost my thesis director, Dr. Robert Van Valin, and the members of my dissertation committee, Dr. Jean-Pierre Koenig and Dr. David Zubin. They all were extremely generous with their time and insights, and I am most grateful to them for their constant guidance and support. It would be difficult to find an academic environment where intellectual rigor and collegiality are combined as well as they are in the Linguistics Department at Buffalo. It has been a privilege and a pleasure to study there.

While working on this dissertation, I presented my ongoing work at several Role and Reference Grammar Conferences whose participants provided me with many significant observations. I would like to especially thank Delia Bentley, Lilián Guerero, Mitsuaki Shimojo, Rolf Kailuweit and Carlos González Vergara.

I also benefited from the comments and camaraderie of my fellow graduate students at the Linguistics Department, in particular Rodrigo Romero and Carolyn O'Meara.

Most of the writing process took place at the Instituto de Investigaciones Filológicas of the Universidad Nacional Autónoma de México. I am very thankful for the generous welcome I received there.

I benefited greatly from discussions with the linguists I had to opportunity to meet in México. I am particularly grateful to Ricardo Maldonado, Cecilia Rojas, Enrique Palancar, Sergio Ibáñez, Armando Mora, and Sergio Bogard, who generously shared with me their careful observations and directed me to relevant literature.

Contents

List of Figures

List of Tables

Abbreviations

1	First person
3	Third person
A	Actor
ABS	Absolutive
ACC	Accusative
ACS	Accessible
ACT	Active
AFD	Actual focus domain
AGX	Agreement index
ANIM	Animate
ARG	Argument
ASP	Aspect
DAT	Dative
DEF	Definite
DO	Direct object
ERG	Ergative
F	Feminine
FOC	Focus
INA	Inactive
INGR	Ingressive
IO	Indirect object
LDP	Left-detached position
LS	Logical structure
LSC	Layered structure of the clause
M	Masculine
N (P)	Noun (phrase)
NMR	Non macrorole

NOM	Nominative
NUC	Nucleus
PFD	Potential focus domain
PL	Plural
PoCS	Post-core slot
PrCS	Pre-core slot
PRED	Predicate
PRO	Pronoun
PSA	Privileged syntactic argument
PURP	Purposive
RDP	Right-detached position
S	Singular
SEML	Semelfactive
SPEC	Specific
U	Undergoer
V (P)	Verb (phrase)

Abstract

This dissertation is a study of the interpretive effects of so-called "clitic doubling" constructions in Spanish. Clitic doubling constructions are those in which a weak pronoun, cliticized to the verb, co-occurs in the clause with a nominal phrase coreferential with it, as illustrated in (0).

(0)   a.   **La**   invité   **a**   **Mabel**.
             ACC.3FS invited.1S ANIM Mabel
             'I invited Mabel.'

        b.   **Le**   di   un regalo **a**   **Mabel**.
             DAT.3S gave.1S a   gift   DAT Mabel.
             'I gave a gift to Mabel.'

In recent decades there have been many studies which propose grammatical analyses of these structures. These studies are for the most part based on examples constructed *ad hoc* and, as it will be shown, cannot adequately account for the empirical data presented here. On the other hand, corpus-based accounts of the "clitic doubling" phenomenon are scarce and advance seemingly contradictory conclusions, some linking doubling with the participants' topicality, and others with the participants' newness.

In this dissertation it is argued that Spanish clitic doubling constructions are dependent on the particular *cognitive accessibility* of the target referent. Specifically, that clitic doubling structures mark the deviation of the target referent from its expected accessibility level given the case-role chosen to encode it.

This claim is supported by the analysis of a corpus of spontaneous interactions among native speakers of a prototypical "doubling dialect" (Buenos Aires Spanish), and formalized using the tools of Role and Reference Grammar.

1 INTRODUCTION

Situated within the domain of research on the interface between syntax and pragmatics, this dissertation is an investigation of so-called "clitic doubling constructions" in Spanish, and of the conditions that govern the occurrence of these constructions in naturally occurring texts. Likewise, this dissertation proposes a grammatical account of these constructions informed by the corpus data.

Clitic doubling constructions are those in which a given entity is grammatically encoded within the clause by both a pronoun cliticized to the verb and an independent phrase coreferential with it. In Spanish there are no "subject" clitics, and doubling constructions involve exclusively direct or indirect objects.

These constructions raise a series of issues which have to be addressed. First, there is the question of which of the two elements, the pronominal clitic or the independent phrase, should be considered as the verb's argument and, depending on the solution of choice, what is the grammatical status assigned to the other element.

Second, it is not *a priori* clear whether the doubling of accusative and dative arguments obeys the same principles. There are dialects of Spanish in which accusative doubling of lexical phrases is claimed to be very infrequent or even nonexistent, whereas dative doubling occurs productively in all dialects. On the other hand, even in those dialects where accusative doubling is more widespread, the contexts in which it may occur are more restricted than those allowing dative doubling.

Third, there is the question of whether those restrictions are syntactic, semantic or pragmatic in nature.

Finally, given that doubling constructions are always grammatically optional, one might want to determine in which contexts they do effectively occur and, to the extent that it is possible to provide plausible answers for such questions, why.

Because of all the challenges they present, clitic doubling constructions have been a pet topic of grammatical analyses for the past two decades. The studies devoted to these constructions, generally produced within the framework of generative grammar, have provided valuable insights into the complexities involved. At the same time, however, it is often unclear whether alternative hypotheses about the interpretive import of doubling constructions are due to progressive refinements of the linguists' intuitions or to grammatical differences in the dialects they speak.

It is also the case that many of these grammatical studies have started to incorporate linguistic notions into the analysis of doubling constructions whose scope go beyond clause-internal syntax (e.g. "presuppositionality") and point to the particular discourse context in which these constructions occur.

All of these reasons have motivated this attempt to provide an analysis of clitic doubling constructions in Spanish based not only on naturally-occurring examples, but further on examples which can be related to the broader communicative context in which they occur. Given the prominence of the Buenos Aires dialect for the phenomenon under study, the claims presented in this dissertation are based on the analysis of spontaneous discourse produced by speakers of this variety of Spanish.

With the support from the empirical evidence provided by the corpus, I will argue that the grammatical, semantic and information-structure variables which have been

proposed to correlate with clitic doubling are epiphenomena of a more general mechanism by which interlocutors cue their communicative partners about the *cognitive accessibility* of the referents involved in the exchange. The possibility of examining this as a plausible hypothesis follows from a view of language organization in which pragmatics mediates between sentence meaning and form, creating pragmatically structured propositions which reflect the speaker's assumptions about the addressee's state of knowledge at the time of the utterance (Lambrecht 1994).

In this dissertation I will show that clitic doubling constructions of both accusative and dative arguments serve to mark the deviation of the relevant referent from its expected activation state given its grammatical role: DO-doubling marks participants as more accessible than expected for canonical DOs; and IO-doubling marks participants as less accessible than expected for canonical IOs. This amounts to saying that in both cases doubling constructions are predominantly used to talk about entities which have *intermediate* levels of cognitive accessibility. This generalization, as it will be shown, is made possible by the particular perspective presented in this work, and it allows to capture the single communicative function that accusative and dative doubling serve.

Clitic doubling constructions resemble, in many respects, agreement relations between a verb and its arguments. It is has been shown that these systems of grammatical agreement are a diachronic evolution of anaphoric relations between incorporated pronouns and their antecedents (Givón 1976). In many languages one can show the transition from one stage to the next, represented by forms whose behavior share characteristics of grammatical agreement and pronominal inflections (Bresnan 2001). In

terms of the chronology proposed by Creissels (2001), for instance, one can recognize three stages in the evolution of pronominal markers. In Stage I pronominal markers are in complementary distribution with the corresponding free phrases, and whether a given argument is encoded with the bound or free forms depends on its degree of topicality or recoverability from context. In Stage II pronominal markers have become obligatory. Now the corresponding free phrases are optional, and occur in order to clarify reference or express emphasis. In Stage III pronominal markers are also obligatory, but have lost their referential properties. In this last stage the free forms that encode the verb's arguments must always co-occur. Spanish object clitics illustrate some intermediate stage between I and II, with dative clitics closer to stage II than accusative clitics.

Assigning these forms their proper place in a grammatical model can be challenging, precisely because of their ambiguous grammatical nature. As it is the role of speakers to eventually solve the potential ambiguities present at any diachonic stage in the evolution of a language, the proposed grammatical analysis should be able to capture these categorical ambiguities, while at the same time providing an analysis consistent with the basic nature of the language and with similar facts about other languages with which it could be compared. I have found the tools for advancing such an analysis in the framework of Role and Reference Grammar, as presented in Van Valin and LaPolla (1997) and Van Valin (2005). Defined from its outset as a structural-functionalist framework (Van Valin 1993:2), it makes possible to provide a formal analysis of Spanish clitics sensitive to the discourse-pragmatic factors which motivate their use.

With these objectives in mind, the organization of this dissertation is as follows. The next section of this Introduction presents the general characteristics of direct object- and indirect object-doubling constructions in Spanish. Chapter 2 is a summary of the most relevant studies on Spanish clitic doubling: those realized from a more theoretically oriented standpoint, and those more concerned with descriptions based on corpus data. Chapter 3 presents evidence of "doubling" constructions in other languages, so as to examine a broader spectrum of factors that may correlate with analogous constructions cross-linguistically. Chapters 4 and 5 present the analysis of the corpus, and are focused on DO-doubling and IO-doubling, respectively. Chapter 6 provides a Role and Reference Grammar analysis of these constructions, incorporating the pragmatic categories that arose from the corpus analysis. The Conclusions in 7 summarize the findings and proposals presented in this work, and advance the lines of future research.

1.1. OVERVIEW: DOUBLING VS. DISLOCATION

The traditional grammatical roles of direct object and indirect object can be realized, in Spanish, by means of pronominal morphemes cliticized to the verb, as in the following examples.

(1) a. Carolina **le**   dio un libro.
          Carolina DAT3S gave a book
          'Carolina gave him a book.'[1]

---

[1] Morpheme-by-morpheme glosses are provided when necessary for the point at hand. When examples are taken from another author, they are adapted for consistency. The terms "direct object" and "indirect object" are provisionally used interchangeably with "accusative argument" and "dative argument" to ease the presentation. To facilitate the identification of the relevant coreferential structure, boldface is preferred instead of the more traditional sub-

b. Juan **lo** leyó.
   Juan ACC3MS read
   'Juan read it.'

These functions may, also, be realized by nominal phrases, as below.

(2) a. Carolina dio un libro **a Juan**.
        Carolina gave a book DAT Juan
        'Carolina gave a book to Juan.'

     b. Juan leyó **el libro**.
        Juan read the book
        'Juan read the book.'

Finally, the language provides the option of encoding these participants by means of both the pronominal clitic and the nominal phrase.

(3) a. Carolina **le** dio un libro **a Juan**.
        Carolina DAT3S gave a book DAT Juan
        'Carolina gave a book to Juan.'

     b. Juan **lo** leyó **el libro**.
        Juan ACC3MS read the book
        'Juan read the book.'

The referential chains formed by clitics and nominal phrases result in what are commonly termed "clitic doubling constructions". It is the analysis of this kind of structures that is the object of this work.

As mentioned above, one challenge that these constructions immediately pose is what the grammatical role of the elements involved is. The fact that the clitics seem to be

_____

indexation. The source of each example is provided right beside it, or indicated in the text preceding it. No source indication means the example is constructed.

6

able to satisfy the valence of the verb, as in (1), has been at the origin of their treatment as pronominal arguments; an approach that, within the generative tradition, was pioneered by Kayne (1975). Under this view, clitics are treated as pronominal heads that originate in argument position. It is posited that, due to their weak phonological nature, they "move" in order to attach to the verb, leaving behind a "trace" which prevents a coreferential phrase from further occupying this syntactic slot.

This analysis fares well for most Romance languages, in which clitics and lexical phrases are considered to be in strict complementary distribution within the clause. In French, for instance, it is generally agreed that an object noun phrase may only co-occur with a coreferential clitic if the phrase is in a dislocated position. The clearest mark of dislocation is an intonation break, as represented in (4)c.

(4) a. Jean-Paul **l'** a aimée.
         'Jean-Paul has loved her.'

     b.* Jean-Paul **l'**a aimée **Simone**.
         'Jean-Paul has loved Simone.'

     c. Jean-Paul **l'**a aimée, **Simone**.
         'Jean-Paul has loved her, Simone.'

The hypothesis that the same situation applies to Spanish, with coreferential phrases occurring in dislocated, non-argumental positions, was among the first to be advanced (Aoun 1981, Hurtado 1984). Further studies, however, challenged this analysis on a number of grounds. In terms of the principles of the framework where the hypothesis was proposed (i.e. Government and Binding), to consider doubled nominals as dislocated

constituents raised problems with respect to the principles governing wh-extraction (Suñer 1988) and anaphor binding (Franco 1993).

On the other hand, Jaeggli (1986) was among the first to point out that in Spanish there is no intonation break evidencing the extra-clausal position of the doubled noun phrase, which is in fact may be followed by additional clausal material. An example of this is presented in (5) (from Jaeggli (1986: 33).

(5)   Parece que tuvieron que llevar**la a la hija del Cnel. Martínez** de urgencia a los Estados Unidos.
        'It seems they had to take the daugther of Cnel. Martínez urgently to the US.'

Similar claims are made by Suñer (1988:400), who also argues that doubled noun phrases are uttered with the same unbroken intonation curve they would have without the clitic, and may occur in embedded clauses or be otherwise non sentence-final. She presents the following examples ((6)a) originally from Barrenechea and Orecchia 1977).

(6)   a.   Lo último que escuché, claro que **la** encontré pesada **la audición**, fue el reportaje.
             'The last thing I listened to, of course I found boring the radio-program, was the interview'.

       b.   Yo **lo** voy a comprar **el diario** justo antes de subir.
             'I am going to buy the newspaper just before coming up'.

Since these pioneering studies, there is general agreement in the literature that Spanish allows the clause-internal co-occurrence of a nominal and a coreferencial clitic, i.e. "clitic doubling", and that this structure is different from cases of right-dislocation[2].

1.2. THE GRAMMATICAL STATUS OF PRONOMINAL CLITICS

The discussion about the status of the doubled lexical phrase has been accompanied by varying positions about the status of the clitics. The grammatically ambiguous nature of these forms, which share features with words and affixes, has consequently allowed views which highlighted either their nominal or their affixal characteristics. Jaeggli (1981), for instance, argued for clitics to be considered as nominal elements. He based this position on three features of Spanish clitics: (i) they can "climb" in the syntactic structure, (ii) they do not affect the stress pattern of the verb, and (iii) they show nominal inflectional morphology. Still, these reasons do not always hold. Whereas the phenomenon of "clitic climbing" does not generally have any correlate among affixes[3], it is less clear that clitics never affect the stress pattern of the verb to which they attach. On the contrary, it has been frequently noted that in Argentinean Spanish some enclitic cluster constructions cause the stress to shift to the ultimate syllable (e.g. *poné* 'put' vs.

---

[2]   We will see below that this assumption is consistent with the typical information structure of doubled sentences attested in the corpus. In terms of information structure, dislocated phrases are expected to be topical, whereas non-dislocated are usually focal. Given this, the occurrence of doubled noun phrases coinciding with the focus of the sentence suggests that Spanish has "clitic doubling" structures that should be kept distinct from topicalized dislocated ones.

[3]   But see Franco (2000:182) who refers to Laka (1993) as presenting "evidence from languages which full-fledged verbal agreement [in which] agreement markers can be displaced within the inflectional amalgam".

poneteló 'put it on you'). Likewise, there are other phonological processes generated by the presence of the clitic. Fernández Soriano (1993) notes that Standard Peninsular Spanish eliminates, in the imperative, the second person plural –*d* and the first person plural –*s* from the verb (e.g. *poned* 'put' – *poneos* 'put on you (all)'; *vamos* 'we go' - *vámonos* 'we go (inceptive)'). Finally, it is also acknowledged the tendency, in colloquial speech, to attach the plural subject agreement morpheme to the cluster formed by the verb and the clitic (e.g. *denle* → *delen* 'give (you all) to him/her').

With respect to the issue of whether the clitics retain inflectional similarities with nominals (such as the –*s* marker for plurals and the contrast –*a/o* for marking feminine and masculine gender) one must note that number inflection in –*s* only applies to third person clitics (*le-les*; *lo-los*), and gender inflection only to third person accusative ones (*lo(s)/la(s)*). The rest of the paradigm presents neither number nor gender inflection.

Alternatively, there seem to be more straightforward characteristics likening Spanish clitics to inflectional affixes. Following the criteria laid out by Zwicky and Pullum (1983) and Monachesi (2003), we can mention the following:

1) *Degree of selection with respect to the host:* Spanish clitics do not alter the lexical category of the host and are constrained in terms of their combinatory possibilities. On the other hand, it is the morpho-syntactic properties of the verb which determines whether the clitic will precede or follow the nucleus. In modern Spanish, clitics precede finite verbs but follow infinitives and imperative forms, an alternation that is not typical of affixes.

2) *Rigid order*: Like in the case of affixes, clitic clusters are arranged into an idiosyncratic rigid order, with the dative clitic preceding the accusative one.

(7) a. Carolina se   **lo**   dio.
        Carolina DAT3 ACC3MS gave
        'Carolina gave it to him/her/them.'

     b. *Carolina **lo**   se   dio.
        Carolina ACC3MS DAT3 gave
        'Carolina gave it to him/her/them.'

3) *Coordination*: It is generally impossible for a clitic to have wide scope over coordinated verbs (8). One marginal exception concerns verbs whose concatenation is often perceived as expressing a conventionally associated sequence. In this case a clitic may have scope over both verbs (9)a, but only if it precedes them (9)b.

(8)   **Lo** pinté y *Ø/lo** colgué en la pared.
        'I painted it and hung it on the wall.'

(9) a. **Las** pierdo y encuentro a cada rato.
        'I loose them and find them all the time.'

     b. *Me gustaría poder perder y encontrar**las** a cada rato.
        'I'd like to be able to loose them and find them all the time.'

4) *Arbitrary gaps*: As in inflectional paradigms, there are arbitrary gaps affecting the clitics' combinatory possibilities. Spanish does not allow a sequence formed by first or second person accusative clitics followed by a third person dative (10)a. In order to express the intended meaning, the strong pronoun must be used instead (10)b[4].

---

[4] For a functional analysis of this phenomenon, see Haspelmath (2004)

(10) a. *Me   **le**   entregaron.
        ACC1S DAT3S gave.3PL
        'They handed me to him.'

    b. Me   entregaron **a él**.
        ACC1S gave.3PL   DAT him
        'They handed me to him.'

5) *Verb left-detachment*: As it has been noted for Italian (Beninca 1988) and Catalán (Vallduví 2001), Spanish bare infinitives and cliticized infinitives can be left detached (11)a, (11)b, but the same construction is ungrammatical if the infinitive is followed by a complement (11)c.

(11) a. Trabajar, trabajo todos los días.
         'To work, I work every day.'

    b. Comprar**lo**, **lo** compro siempre.
         'To buy it, I buy it always.'

    c. *Comprar el periódico, **lo** compro siempre.
         'To buy the newspaper, I buy it always.'

6) *Morpho-phonological idiosyncrasies*: As it is commonly the case with inflectional formations, the phonological shape of Spanish clitics may be affected by other clitics with which they combine. An example from Spanish is the allomorph of the dative clitics *le, les* which occurs before accusative clitics; i.e. the so-called "spurious *se*" (Perlmutter 1971) as illustrated in (7) above.

The characteristics that liken Spanish clitics to inflectional affixes have underscored their view as (object) agreement markers[5]. In effect, it has been noted that the kind of feature erosion displayed by the clitics is typical of agreement systems (Franco 2000). Consider, for instance, the neutralization of number marking on the "spurious *se*" form (**ses*), and the widespreading use of the singular dative form *le* with plural reference (12) (For additional arguments about the "de-pronominalization" of dative clitics see also Bogard 1992 and Company 2006).

(12)   **Le**   compré un libro **a   los chicos**.
         DAT3S bought a   book DAT the children.
         'I bought the kids a book.'

The analysis of Spanish clitics as agreement markers is, of course, crucially supported by their possible co-occurrence with lexical arguments, which is interpreted as analogous to the co-occurrence of lexical subjects and the obligatory subject-agreement suffix on finite verbs. In this respect, object clitics are considered to be governed by the same kind of mechanisms that account for agreement phenomena in general (Suñer 1988, García-Miguel 1991, Franco 1993).

The assimilation of these pronominal clitics to agreement markers has the advantage of allowing a rather straightforward account of some of their distributional properties. For instance, agreement phenomena are driven by so-called "animacy hierarchies", with cross-linguistic evidence showing that arguments which rank higher on this hierarchy display stronger agreement relations than arguments which rank lower

---

[5] Even though, as argued in Auger (1994), the morphophonological and morphosyntactic dimensions are, in principle, independent of one another.

13

(Comrie 1989). This predicts that indirect objects, which typically refer to animate referents, should enter in "agreement" relations more frequently than direct objects, which typically refer to inanimates. As we will see, there are asymmetries in the distribution of DO-doubling and IO-doubling that seem to follow exactly this pattern.

1.3. ASYMMETRIES ON INDIRECT VS. DIRECT OBJECT DOUBLING

It is generally acknowledged that IO-doubling is much less constrained that DO-doubling. This asymmetry is reflected both on its relative degree of dialectal spreading as well as on the degree of grammatical restrictions that it is subject to.

IO-doubling yields grammatical sentences across all varieties of Spanish; and there is no Spanish dialect in which IO-doubling does not occur to some extent. Moreover, the clitic is in fact considered as obligatory when in a chain with a strong pronoun (13)a, and in constructions introducing a benefactive (13)b or an experiencer (13)c, or involving inalienable possession (13)d, among other contexts (cf. Chapter 5).

(13) a. ***Ø/Le** hablaron a ella.
          'They spoke to her.'

      b. ***Ø/Le** preparé la cena **a Carolina**.
          'I prepared dinner for Carolina.'

      c. ***Ø/Le** gusta el cine **a Juan**.
          'Juan likes the movies.'

      d. ***Ø/Le** duele la muela **a Ernesto**.
          'Ernesto has a toothache.'

On the other hand, dialects seemingly differ in the obligatoriness of the dative clitic when the coreferential phrase encodes a "goal". In Peninsular varieties of Spanish the clitic is assumed to be optional (Demonte 1994, Mayer 2003). Even though this is often generalized for all Spanish dialects, corpus studies have shown that at least in some varieties of American Spanish the co-occurrence of the clitic has become the norm in these contexts as well. This is the case for the Spanish of Santiago, Chile (Silva-Corvalán 1981), Caracas, Venezuela (Bentivoglio 1978) and México (Company 2006). Jaeggli (1981:12) makes a similar claim for Argentinean Spanish, pointing out that the co-occurrence of the dative clitic with thematic goals is "optional, but highly preferred". He illustrates this with the example in (14).

(14)   A las doce en punto, la dirección (**les**) entregó las notas **a los estudiantes**.
         'At 12 o'clock sharp, the administration gave the grades to the students.'

With respect to DOs, on the other hand, doubling is considered obligatory across all dialects if the accusative phrase is realized by a strong pronoun.

(15)   ***Ø/Lo** vi **a él**.
         'I saw him.'

There are, in addition, some contexts where it is marginally tolerated. The *Diccionario Panhispánico de dudas* (RAE) affirms that DO-doubling is "normal" if the tonic complement is the pronoun *todo* 'everything' (16)a); when it is a numeral preceded by an article and refers to an animate entity (16)b); when it is realized as the indefinite

uno 'one' and its referent is the current speaker (16)c) and if has an emphatic value (16)d).

(16)   a.   (**Lo**) sabe **todo**.
        'She knows everything.'

    b.   (**Los**) invité **a los cuatro**.
        'I invited the four of them.'

    c.   Si (**la**) ven **a una** vacilar, enseguida se aprovechan
        'If they see one doubt, they immediately take advantage.'

    d.   Ya (**lo**) creo **que vendrá**.
        'I believe that s/he will come indeed.'

In other contexts, dialects once more are believed to differ. According to the *Diccionario*, DO-doubling is "alien to the cultured norm of most of the Hispanic world"; although "normal" in some regions of the Americas, especially those around the *Río de la Plata*.

1.4. A NOTE ON DIALECTAL VARIATION

At this point a word must be said concerning dialectal differences. As we saw above, it is generally acknowledged that there is a considerable degree of variation with regards to clitic doubling (which is not surprising, considering Spanish is a language spoken in over twenty countries). Despite this, the literature is often ambiguous in terms of whether hypotheses or descriptions are advanced for one dialect in particular, or assumed to be valid cross-dialectally. These descriptions and hypotheses are often based on constructed examples, and thus one is left to wonder whether any incompatibility

among them may be due to differences in the accuracy of the intuitions or on the dialects being described. To make matters worse, where claims are made with respect to some particular dialects, the lack of agreed upon labels for them is source of additional misunderstanding.

Consider the following example as a case in point. Jaeggli (1981) proposes that in "River Plate Spanish" DO-doubling is only possible if the coreferential phrase is preceded by *a*. As we will see below, the occurrence of this form is crucial for the theoretical constructs proposed to explain DO-doubling in this dialect. Thus, a sentence like (17) is considered as ungrammatical.

(17)   Yo **lo** voy a comprar **el diario** justo antes de subir.
         'I am going to buy the newspaper just before coming up.'

This particular example, however, is presented as grammatical by Suñer (1988), who discusses what she refers to as "Porteño" dialect. Is it possible then that non *a*-marked phrases can be doubled in Porteño but not in River Plate Spanish? It may be that these are different dialects governed by different principles and that, say, "Porteño dialect (…) takes certain River Plate forms to an extreme"[6].

This hypothesis reduces its plausibility when one notes that there are actually no bases for considering that River Plate and Porteño are two different dialects. After all "Porteño" is the term generally used to refer to the inhabitants of Buenos Aires, a city that sits on the shores of the "Río de la Plata", often referred to as "River Plate" in the

---

6   Mayer (2003:12). For analogous conclusions, see also Everett (1996:71) or Guttierrez-Rexach (2000:315, 331).

English literature. The different labels used by Suñer and Jaeggli are just a matter of taste, and it is unfortunate that they have lead to erroneous assumptions.

Claims about DO-doubling have also been made with respect to "South American Spanish" (Lyons 1999: 209) or "Southern Cone Spanish" (Franco 2000). The danger here is to err on the other side, assuming generalizations that are not necessarily warranted. As we will see, there is no evidence that DO-doubling occurs in all South American dialects; and in fact there is data that suggests that it may occur in some North American varieties as well (e.g. Mexican Spanish, cf. § 3.1.1).

What this makes apparent, I think, is the importance of unambiguously defining the dialect or dialects that serve as the object of the analysis. Again, unless this condition is met, it is impossible to establish whether contradictory grammaticality judgments (as for instance those regarding the obligatoriness of *a*-marking exemplified above) should be interpreted as a matter of diverging descriptive adequacy on the part of competing proposals or, as may in principle also be the case, as a result of different rules at play in different Spanish dialects.

One way to solve this problem is to work with naturally occurring texts. Since they provide empirical data and can be traced back to particular dialects, they automatically solve the issue of diverging grammaticality judgments and potential dialectal differences. The study of Spanish clitic doubling constructions seems an ideal candidate to take under this approach; and following this premise the analysis presented in Chapters 4 and 5 will be based on corpus data from a paradigmatic "doubling" dialect.

2 PREVIOUS STUDIES

This Chapter examines the most relevant proposals that have been put forth in the literature, in relation with the phenomenon of clitic doubling in Spanish. The first section of this Chapter reviews grammatical proposals; in particular those advanced by Margarita Suñer (1988, 1999a, 1999b, 2000) and Jon Franco (1983, 2000; Franco and Mejías Bikandi 1997, 1999). These proposals explain the more restricted nature of DO-doubling over IO-doubling in terms of an extra condition that should be met for DO-doubling to be posible; namely the specificity or presuppositionality of the intended referent[7].

The second part of this Chapter summarizes corpora-based studies of these constructions, which correlate doubling with the relative topicality or newness of the target referents.

2.1. CLITIC DOUBLING IN GENERATIVE GRAMMAR

One of the phenomena that received attention from early on, within generative approaches to clitic doubling, is the asymmetry between dative and accusative arguments manifested in interrogatives. It was noted that wh-words in IO function may be doubled by a dative clitic (18)a; whereas doubling is apparently banned for wh-words in DO function (18)b.

---

[7] The grammatical literature on clitic doubling in Romance languages is vast enough that even a mere list of references risks omitting relevant works. For reasons of space and ease of exposition I will review here those works which have a direct bearing on the interpretive effects that Spanish doubling constructions may have. For a more comprehensive review of the theoretical issues involved, see also Torrego (1995), Uriagereka (1995), Sportiche (1995), Belletti (1999), Kayne (1994) and the relevant references cited therein.

(18) a. ¿A quién$_i$ **le** regalaron un auto e$_i$?
         'Whom did they give a car to?'

     b. *¿A quién$_i$ **lo** condecoraron e$_i$?
         'Whom did they decorate?'

Within the generative framework this asymmetry was originally explained in terms of "theta-roles" (Aoun 1981). It was argued that accusative clitics absorb the theta-role assigned by the verb, thus leaving none left to the DO constituent. A consequence of this proposal was assigning doubled DOs to non-argumental positions. As mentioned in §1.1., however, this proved to be an undesirable result. Alternatively, Borer (1984) and Jaeggli (1981) advanced analyses in terms of "Case". Borer suggested that the constraints on doubled DO extraction resulted from the properties of the preposition *a*, which was argued to assign only dative case, thus yielding a clash of features in the case of doubled accusatives. Jaeggli, in turn, suggested that the asymmetry resulted from a difference in the nature of the clitics: accusative clitics were treated as "obligatory case absorbers"; and dative clitics were treated as "optional case absorbers". This proposal amounted to the claim that dative clitics would behave differently depending on the occurrence of a coreferential phrase in the clause. On the other hand, it explained the more restricted nature of DO-doubling positing that, since accusative clitics always absorb the case features assigned by the verb, a coreferential accusative phrase was only possible when case features were assigned by some other means; specifically, the preposition *a*. From this it follows that DO-doubling was predicted to be possible only if the lexical phrase was marked by *a*, and that "accusative clitics in Spanish are in strict complementary

distribution with direct objects which are not preceded by the marker *a*" (Jaeggli 1986:19).

Jaeggli's proposal rested on an assumption regarding the prepositional status of the *a* marking doubled phrases; an assumption challenged by Suñer (1988). In her argument, Suñer starts by summarizing the different uses that this morpheme has in Spanish. She notes that it may pattern with prepositions (19)a), but also introduce idiomatic expressions (19)b), or be selected by the verb (19)c). It may also mark accusative arguments, which are canonically considered as NPs. In this case, the occurrence of *a* is linked to the animacy or individuation of the referred entity (20).

(19) a. Lía se sentó **a** (delante de, sobre, bajo, etc.) la mesa.   (Suñer 1988)
          'Lía sat at (in front of, on, under) the table.'

       b. Lo mataron **a** sangre fría.
           'They killed him in cold blood.'

       c. Asistió **al** (=a+el) simposio.
           'He/she attended the symposium.'

(20)   Oían a Paca / a la niña / a una niña / *a niñas/ a la gata / *a la radio.   (op.cit.)
           'They heard Paca / the girl / a girl / *girls / the cat / *the radio.'

Suñer argues in particular against the prepositional status of *a*-marked IOs (21), claiming that its behavior is closer to the markers in (20) than the more canonical prepositions illustrated in (19).

(21)   **Les** compraron la casa **a los Ramírez**.   (op.cit.)
           'They bought the house from/to the Ramírez.'

Suñer points out that *a*-marked IOs behave differently from prepositional phrases in a series of respects. Pronominal forms inside PPs can refer back to a clausal subject (22)a, suggesting that Spanish PPs behave as governing categories. Doubled IOs, however, must be disjoint in reference with respect to the subject (22)b.

(22) a. Ese señor$_i$ nunca se cansa de PRO$_i$ hablar [PP de él$_{i/j}$].   (op.cit.)
         'That man never tires himself of speaking about him/himself.'

      b. Mara$_i$ **le**$_j$ aceptó la invitación [**a ella** $_{*i/j}$].
         'Mara accepted the invitation from her/*herself.'

Further, PPs cannot function as binders for anaphors (23)a, a restriction that does not hold for *a*-marked IOs (23)b.

(23) a. Paco$_i$ habló con el profesor$_j$ de sí mismo $_{i/*j}$.   (op.cit.)
         'Paco talked with the professor about himself/*him.'

      b. Paco$_i$ (**le**) habló **al profesor**$_j$ de sí mismo $_{i/j}$.
         'Paco talked with the professor about himself/him.'

Finally, within the government and binding framework, prepositions assign specific theta roles to their complements. On the contrary *a*-marked IOs are semantically unrestricted, and can be interpreted with the same theta roles as each of its prepositional counterparts.

(24) a. Los niños compraron chucherías [de la vendedora]. $_{SOURCE.}$ (op.cit.)
         Los niños **le** compraron chucherías [**a la vendedora**].
         'The children bought trifles from the vendor.'

b. Conseguí un empleo [para María]. ₍GOAL₎
   Le conseguí un empleo [**a María**].
   'I obtained a job for María.'

c. Arreglaron la televisión [de mi madre]. ₍POSSESSOR₎
   Le arreglaron la televisión [**a mi madre**].
   'They fixed the TV set of my mother.'

d. Lavó el auto [por ella]. ₍BENEFICIARY₎
   Le lavó el auto [**a ella**]
   'He/she washed the car for her (=for her sake/instead of her).'

From these tests Suñer concludes that clitic doubling only occurs in nominal chains, and not prepositional ones. The clitics, in turn, are assigned the role of (object) agreement markers, a view that since has become widely accepted.

2.1.1. Specificity Restrictions

In an influential paper, Suñer (1988) argued that the grammaticality of doubling constructions is governed by a "Matching Principle", determining the features that the clitics have to agree with. It was proposed that the lexical entry for clitics contain a series of features ([animacy], [gender], [number], and [person]), whose values have to agree with those of the lexical phrase co-occurring with them. A crucial difference between accusative and dative clitics is that the former are considered as additionally marked as [+specific], which predicts that IO-doubling should occur with specific or non-specific referents, whereas DO-doubling should be possible only if the relevant referent is intended as specific. Suñer (op.cit. 394ff.) presented the following examples.

(25) a. **Le** ofrecí ayuda **a la niña** / **a una estudiante**.
'I offered help to the girl / to a student'. [+specific,+/- definite]

b. **Les** ofrecieron queso y leche **a familias de pocos medios**.
'They offered cheese and milk to low-income families.' [-specific, -definite]

(26) a. (**La**) oían **a Paca** / **a la niña** / **a la gata**.
'They listened to Paca / the girl / the cat.' [+anim, +spec, (+def)]

b. Diariamente, (**la**) escuchaba **a una mujer que cantaba tangos**.
'Daily, he/she listened to a woman who sang tangos.' [+anim, +spec, (-def)]

c. *No **lo** oyeron **a ningún ladrón**.
'They didn't hear any thieves.' [+anim, -specific]

d. ***La** compramos (**a**) **esa novela**.
'We bought that novel.' [-anim, +spec, (+def)]

The sentences in (25) show that IO-doubling is allowed regardless of the specificity / definiteness of the noun phrase. Doubling of DOs, in turn, occurs with definite nouns (26)a) or indefinite-specific ones (26)b), but is ungrammatical if the noun is non-specific (26)c), or specific but inanimate (26)d).

In this work, the notion of specificity implies that "the referent(s) of a [+spec] NP can be identified with a particular x in the linguistic context" (op.cit. 397). Suñer acknowledges that, in this sense, her notion of specificity is close to some definitions of definiteness, as that in Chafe (1976). Evidence for the relevance of specificity over definiteness is, however, presented by her in (26)b).

The relevance of animacy is less clear. Suñer generally considers it a necessary value (26)d), although it is recognized that doubling of inanimate DOs may marginally occur, at least in the "Porteño" dialect, where "of the [+spec, +anim] features needed for

24

DO-CL animacy is disregarded -thus *a* omitted- and specificity by itself provides a sufficient condition for CL-D" (op.cit. 400, footnote 12).

Suñer applies this proposal to the wh-extraction asymmetries mentioned at the beginning of this section. She notes that, contrary to the traditional intuitions, wh-extraction of doubled accusatives is in fact possible, provided that the DO is interpreted as specific. The specificity of the doubled DO is in this case deduced from its partitive interpretation, which identifies a set of entities within which the referent is chosen. Thus, it is proposed that a non-partitive interrogative cannot be doubled (27)a), but a partitive interrogative can (27)b).

(27) a. *¿**A quién lo** condecoraron?   =(18)
          'Whom did they decorate?'

      b. ¿**A quién/cual de ellos lo** condecoraron?
          'Whom/which one of them did they decorate?'

On the other hand, Fernández Soriano (1993: 322, 377), argued that doubling of non-partitive interrogatives is in fact possible in "leísta" dialects[8] (28). She also argued that doubling of interrogative phrases is also possible if the interrogative phrase makes explicit a gender feature which matches that of the clitic (29).

(28)   ¿**A quién le** vieron en televisión?
          'Who did they see on TV?'

(29)   ¿**A cuántos chicos los** pillaste copiando?
          'How many guys did you catch cheating?'

---

[8] "Leísta" dialects are those in which the accusative clitic is replaced by the form *le, les*, generally when in reference to human participants.

Whereas these examples seem to belie the relevance of specificity as presented by Suñer, there are other tests she provides in order to validate her hypothesis. For instance, she cites Hurtado (1984) as the first to recognize that quantified doubled IOs are ambiguous between a wide and a narrow scope reading (30); but quantified doubled DOs can only receive a wide scope interpretation (31). If this intuition is correct, the only interpretation of (31) is that in which every voter chose *the same set of candidates* (examples from Suñer 1988: 423). (Note that in this case "specificity" is associated with wide scope readings. The different senses in which the notion of specificity can be understood will be addressed in Chapter 4).

(30)   Todos los candidatos **les** han dicho la verdad **a algunos electores**.
        'Every candidate has told the truth to some voters.'
        a. ∃y, y a voter, ∀x, x a candidate (x has told the truth to y)
        b. ∀x, x a candidate, ∃y, y a voter, (x has told the truth to y)

(31)   Todos los electores **los** eligieron **a algunos de los candidatos**.
        'Every voter elected some of the candidates.'
        ∃y, y a candidate, ∀x, x a voter (x elected y)

Since Suñer's original proposal, the relevance of specificity as a necessary condition for DO-doubling has been endorsed, for instance, in Sportiche (1995), Bleam (1999) and Gutiérrez Rexach (2000), among others, and it has been adopted in the analysis of doubling constructions in other languages, as we will see in Chapter 3.

2.1.2. The Presuppositionality Condition

Under the specificity hypothesis, DO-doubling was possible only if the doubled phrase was interpreted as specific. The less constrained nature of IO-doubling, in turn, depended on the lack of a particular feature value for specificity on the lexical entry for the dative clitic. This hypothesis was challenged by the finding of some non-specific IOs that seemed to ban doubling constructions. Fernández Soriano (1989:318) presented the following examples[9].

(32) a. Creo que **Ø/*les** daré todo mi dinero **a personas necesitadas**.
'I think I will give all my money to people in need.'

b. **Ø/*le** hablaré del asunto **a gente interesada**.
'I will talk to interested people about the matter.'

On the other hand, Franco (1993) argued that DO-doubling was in fact possible with certain non-specific referents:

(33) a. Juan **lo** invitaba **a uno** y luego se olvidaba.   (Franco op.cit. 86)
'Juan used to invite people and then forget all about it.'

b. En ese departamento **lo** admiten **a cualquiera**.
'In that department, they admit anyone.'

It is claimed that *uno* in (33)a functions as a prototypical pronoun of arbitrary reference. Because of its pronominal nature it requires, however, the co-presence of the

---

[9] In fact, whether the clitic should be considered ungrammatical in sentences such as (32) is not completely agreed upon. See Cuervo (2003) and Chapter 5 below.

clitic. Likewise, in (33)b), *cualquiera* encodes non-specific referents, and the clitic is nevertheless claimed to be allowed[10].

Further, Franco also argues that the inherently non-specific complements of existential sentences with *haber* (34)a) may in fact be referred to by accusative clitics in subsequent discourse (34)b), something which would not be possible if the accusative clitic were inherently specific.

(34) a. *Había los hombres.   vs.   Había unos hombres.   (op.cit. 88)
      'There were the men.'   'There were some men.'

    b. A: -Parece que hoy no había **mejillones** en la pescadería.
        'It seems that today there were no mussels in the fish market.'

        B: -Sí que **los** había, pero estaban ya vendidos
        'Yes, there were indeed, but they were already sold.'

Finally, in certain contexts, DO-doubling is claimed to be ungrammatical despite the specificity of the lexical phrase. It has been noted that there is a general correlation between the mood in an embedded relative clause and the referential interpretation of the antecedent head: If the relative clause occurs in the indicative, the antecedent is interpreted as specific, whereas if it occurs in the subjunctive it is understood as non-specific (Rivero 1977). Thus, *un candidato* is understood as specific in (35)a) and as non-specific in (35)b):

---

[10] The claim that *uno* in (33) has arbitrary reference is, in fact, arguable. When doubled by the clitic, *uno* is normally used as a polite form of self-reference. In this context, its reference is not arbitrary (see also Gutiérrez Rexach 2000:352ff).

(35) a. Carolina votará por un candidato que **impulsa** [indicative] reformas políticas.

  b. Carolina votará por un candidato que **impulse** [subjunctive] reformas políticas.
      'Carolina will vote for a candidate that promotes political reform.'

With this as general background, Franco points out that Suñer's hypothesis would incorrectly predict the grammaticality of doubling in the sentences in (36) which, given the indicative mood of the embedded verb, trigger the specific interpretation of their antecedent.

(36) a. ***Lo** busca **a un médico que vino de N.Y. ayer**.   (op.cit. 89)
        'S/he is looking for a doctor that came from N.Y. yesterday.'

  b. ***Lo** busca **a un hombre que lleva camisa azul**.
        'S/he is looking for a man wearing a blue shirt.'

Based on these facts, Franco concludes that neither the notions of specificity or, as it is often also claimed, animacy, can capture the conditions underpinning DO-doubling. He proposes instead that DO-doubling responds to a "Presuppositionality Condition" requiring doubled accusative indefinites to be interpreted as presupposed.

In arguing for his proposal, Franco draws on Diessing's (1992) "Mapping Hypothesis", which suggests that presuppositional interpretations can be related to particular syntactic positions mapped onto structural configurations. The Mapping Hypothesis depends on a distinction between presuppositional indefinites (those which presuppose the existence of the entity they are applied to), and non-presuppositional ones (those which assert it). Along the lines of Milsark (1977), Franco and Mejías-Bikandi (1997, 1999) allege that non-doubled weakly quantified phrases, i.e. those introduced by

29

un, unos, *alguno*, etc. are ambiguous between presuppositional and non-presuppositional interpretations. On the other hand, when the clitic appears, it is assumed to force the presuppositional interpretation of the relevant referent. Consider the following examples from Franco and Mejías-Bikandi (1999:108)[11].

(37) a. **Lo** he visto **a un marinero**.
'I have seen one of the sailors.'

   b. He visto a un marinero.
'I have seen a sailor.'

As hinted by the English gloss the authors provide, their hypothesis is that whereas referents of non-doubled DOs may be discourse-new, referents of doubled-DOs are members of a previously introduced set. It is in this sense, then, that notion of "presuppositional" should be interpreted in this context.

According to these authors, another effect of DO-doubling is to disfavor a restrictive reading of certain adjectival complements. They provide the following examples[12].

(38)   Vi a los estudiantes borrachos.
'I saw the students drunk.' [small clause reading]
'I saw the drunk students.' [restrictive reading]

---

[11] Most of their data come form Basque Spanish in which, as in other *leísta* varieties, animate accusative clitics appear as *le(s)* instead of *lo/a(s)*. In the following examples the canonical accusative clitics are substituted for the *leísta* variants occurring in Basque Spanish. Nothing in Franco and Mejías-Bikandi's analysis hinges on this distinction, nor they state that their hypothesis is assumed to be applicable only to some particular Spanish dialects.

[12] It is unclear how the author's argument about the small reading triggered by doubling is related to the presuppositionality hypothesis.

(39)   **Los** vi **a los estudiantes** borrachos.
        'I saw the students drunk.' [small clause reading]

It is claimed that the adjective in (38) is usually understood as a secondary predicate (small clause reading), although its potential interpretation as a restrictive modifier is somehow present. The restrictive reading is even more difficult to obtain in (39), due to the occurrence of the accusative clitic.

Whereas it has been argued that agreement phenomena should not affect the semantics of the sentence (Laka 1993, Sportiche 1995), it is also the case that, as Franco acknowledges, the presence of agreement morphology has been correlated with semantic features in many languages, of which Swahili (Wald 1979) and Turkish (Diesing 1992) are two examples. In this regard, Franco (2000) suggests that DO "agreement" does not, in fact, affect the semantics of the sentence when it occurs with a definite phrase. He proposes, on the contrary, that doubling has a semantic effect only in relation to human indefinites, and that it is through the analysis of these kinds of referents that the restrictions over DO-doubling should be explained. Consider the following pair.

(40) a. (**Lo**) he llamado **a Pedro**.   (op.cit. 183)
         'I have called Pedro.'

      b. **Lo** he visto **a un estudiante**.
         'I have seen a student.'

Franco argues that the presence or absence of the clitic in (40)a does not have any semantic import, and therefore definite NPs are indeed insensitive to clitic doubling. On the contrary, in (40)b "the direct object *a un estudiante* can only be [interpreted as

referring to] any student whose face I have previously seen but cannot associate with a name; however, never anyone who happens to look like a student" (op.cit. 183).

Thus, examples such as (40)b serve as a crucial piece of evidence for the proposed contention that "Spanish direct object clitic doubling –in addition to being an instantiation of object agreement- is mainly a strategy to scope objects out the VP and allow them to escape the nuclear scope. In this way, the clitic does not only guarantee that the object will not be interpreted existentially, but also functions as a disambiguator between the existential and presuppositional readings that non clitic-doubled indefinite objects would otherwise have" (op.cit. 184).

Note that these statements are made with respect to DO-doubling. In the case of IOs, doubling does not change the interpretation of the target referent in neither the case of definites or indefinites. Franco and Mejías-Bikandi (1999: 107) claim that doubled indefinite IOs can be interpreted as presuppositional or non-presuppositional, and they attribute the absence of interpretive differences "to the greater freedom and higher frequency of doubling with indirect objects".

In terms of the formal implementation of this proposal, it is argued that DOs move to the appropriate "Agreement-DO node" in "overt syntax", whereas IOs move "covertly" to the "Agreement-IO node" in "Logical Form". The principles of the Principles and Parameters framework predict that overt movements generate more costly derivations than overt ones and, from this, the greater productivity of IO-doubling over DO-doubling is deduced.

The insights behind the "presuppositionality" hypothesis have been adopted, with some modifications and to different extents by, for instance, Suñer (1999, 2000) and Gutiérrez Rexach (2000). Suñer proposes that doubling with strong pronouns is obligatory in all cases for all Spanish dialects precisely because strong pronouns refer to presupposed entities. Gutiérrez Rexach acknowledges that one effect of doubling DOs is the marking of the relevant referents as "presupposed". He further suggests that presupposed elements must be non-focal, concluding that "focused noun phrases cannot be doubled" (op.cit. 330, but see §6.1.3). He illustrates his claim with the following examples.

(41) a. *Yo **lo**   ví   a   [_FOC_ **Juan**]
            I   ACC3MS saw.1S ANIM [_FOC_ Juan]
            'I saw [_FOC_ Juan].'

      b. *Yo **la**   encontré a   [_FOC_ **María**]
            I   ACC3MS found.1s ANIM [_FOC_ María]
            'I found [_FOC_ María].'

      c. ***La**   saludé   a   [_FOC_ **su madre**]
            ACC3FS greeted.1S ANIM [_FOC_ his mother]
            'I greeted [_FOC_ su madre].'

The prediction that doubled DOs cannot be focal contrasts with Suñer's claims about doubling and focal structure. Contrary to Gutiérrez Rexach, Suñer (1988:420) maintains that "since the focused elements […] are specific, the grammaticality [of (42)] comes as no surprise". (We will address the relationship between doubling and focus structure in Chapter 6).

33

(42)   **La**   nombraron   **a**   **Mara**.   (op.cit. 420)
        ACC3FS nominated.3PL ANIM Mara
        'They nominated Mara.'

2.1.3. Summary and Discussion

In this section I discussed a number of predictions made with respect to the conditions under which clitic doubling is possible. In particular, that clitic doubling depends on the *a*-marking of the doubled phrase, its specificity and its presuppositionality. These proposals focus on DO-doubling, since none of those factors are presumed to affect the doubling of IOs. One remaining question, then, is whether the doubling of IOs generates any interpretive difference or correlates with any semantic or pragmatic factor. The corpora studies reviewed in the next section will provide some suggestions in this regard, and we will further analyze IO-doubling constructions in Chapter 5.

The specificity / presuppositionality hypotheses, as they are formulated, are not only silent with respect to the possible effect of doubling on IOs, but also with its possible effect on definite DOs, which are assumed to be inherently specific / presupposed. A second issue that needs to be examined, then, is whether there is any correlation between definiteness and doubling. This will be one of the goals of Chapter 4.

Finally, both specificity and presuppositionality are notions which can be (and have been) interpreted in different ways, ways which do not necessarily address the same conceptual distinctions. In the works presented here, the labels "specific" or "presupposed" are often times used as cover terms for different phenomena: the

implication of the existence of the referent, its familiarity to the speaker, the fact that it was previously mentioned, that it belongs to an identifiable set, among others. Note that this ambiguity makes these hypotheses particularly difficult to evaluate. Thus, before confronting their predictions with the empirical data, they will need to be more closely examined. This will be a second goal of Chapter 4.

Before tackling these tasks, however, we will review the kind of phenomena uncovered by discourse-based analyses of clitic doubling constructions in Spanish, which are the focus of the next section.

2.2. DISCOURSE-BASED APPROACHES

There is only a handful of works that analyze clitic-doubling constructions in naturally occurring texts. The first of these studies (Barrenechea and Orecchia 1977) is a report of quantitative data based on some of the interviews that were eventually published as *El Habla Culta de la Ciudad de Buenos Aires* (Barrenechea 1987) as well as some other recordings that were not included in the published version. Silva-Corvalán (1981) bases her study on transcriptions of oral interactions among speakers of Chilean Spanish. She interprets clitic-doubling as an instance of verb agreement with topical elements and, consequenly, sensitive to the relative topicality of each case-role. Weissenrieder (1995) analyzes instances of dative-doubling in a novel by the Argentinean author Manuel Puig, and also concludes that doubling tends to occur with highly identifiable, topical nouns. The opposite perspective is advanced by Colantoni (2002), who correlates dative and accusative doubling with referents that are new to the discourse context.

2.2.1. Barrenechea and Orecchia (1977)

This article takes a quantitative approach to direct and indirect object doubling in Buenos Aires Spanish. Barrenechea and Orecchia cross-analyze in different ways several variables, such as case (datives vs. accusatives), the internal structure of the objects (pronominals vs. full NPs), position (pre-verbal vs. post-verbal), and the features [human] and [definite]. Their data comes from recordings made as part of the project later published as *El Habla Culta de la Ciudad de Buenos Aires* (Barrenechea 1987), but it overlaps only partially with the 1987 corpus. Unfortunately, there is no information

regarding how the 1977 sample was constituted in terms of number of informants, characteristics of the recorded interactions (genre, level of formality, etc.), duration or number of clauses. The scarcity of examples provided by the authors to illustrate the interaction of variables in the charts, and the general succinctness of their descriptive interpretation, makes it ambiguous how some of their data should be interpreted. Despite these shortcomings, this article allows for a preliminary approach to the kind of factors that correlate with clitic-doubling in naturally occurring texts.

The data shows that the percentage of dative doubling is greater than the percentage of accusative doubling. The numbers lump together pre- and post-verbal phrases, as well as lexical and pronominal phrases. From a total of 1406 DOs in their sample, 61 (4.33%) appear in a doubling construction. For IOs, the percentage of doubling is 14.09, from a total of 518 instances (p.362). Their data suggests that most DOs occur as (non-doubled) NPs whereas most IOs occur as (non-doubled) personal pronouns (p.373).

Pre-verbal objects co-occur with a coreferential clitic more often than post-verbal objects. The presence of a dative clitic with pre-verbal full NPs in IO function is categorical (100% of the cases). There is also a high incidence of resumptive accusative pronouns with pre-verbal full NPs (86.04%). Based on the examples provided in the article (p. 367, presented in (43)), the set of pre-posed accusative NPs that do not co-occur with a resumptive pronoun seem to be limited to non-referential bare nouns, whose status as "direct objects" is in fact questionable (see for instance Bogard 2007b), and which cannot normally be replaced by a pronoun.

(43)   a.   Sí, **televisión**, vemos.
'Yes, television, we see.'

b.   Entonces dijeron que **esperanzas** no tienen…
'Then they said that hopes they don't have.'

c.   … **tiempo** para mí misma no tengo…
'…time for myself I don't have…'

There is also a chart illustrating the effect of definiteness. The authors consider as definite: phrases marked by definite articles, possessive adjectives, demonstratives, numerals, and the indefinite *todo* followed by an article, possessive adjective or demonstrative. Possessive pronouns, demonstratives and nominal numerals are also definite. On the other hand, they consider nouns preceded by indefinite articles or by qualifying, indefinite or quantificational adjectives, as well as indefinite and quantificational pronouns and nouns not preceded by a modifying adjective to be indefinite. Definite post-verbal DOs in their sample co-occur with the clitic more often than indefinite ones (2.70% vs. 1.09% respectively). Post-verbal IOs, on the contrary, are doubled more frequently when they are indefinite (70% vs. 45%).

Post-verbal full NPs (lumping together datives and accusatives) are doubled in only 3.91% of the cases. This contrasts with an 88.77% of doubled constructions when the objects occur in pre-verbal position. Analyzing the case of each object, they find a 51.11% of doubled post-verbal IOs, versus a mere 1.84% of doubled post-verbal DOs. From the 19 cases of post-verbal DOs in their sample, 13 (68.42%) are preceded by *a*, and 12 (63.15%) have human reference (op.cit. 368).

Their data generally agree with the expected trends, especially with regards to the greater frequency of doubling with datives than with accusatives and, for DOs, the greater frequency of doubling for those with human reference or when the noun phrase is preceded by *a*. It is also interesting that they found that DO-doubling involves definite phrases more often than indefinite ones and that the opposite results obtain for IO-doubling, a matter to which we will come back below.

2.2.2. Silva-Corvalán (1981)

Silva-Corvalán studies DO-doubling of full NPs from a corpus of recorded semi-directed conversations with 29 speakers from Santiago, Chile. She finds that there is a correlation between doubling and definiteness, which is consistent with the data presented by Barrenechea and Orecchia for Buenos Aires Spanish. Definite pre-verbal DOs are "doubled" in 91% of the cases (219/241). However, 21% of indefinite pre-verbal DOs are "doubled" as well (13/62)[13]. Since definiteness marks a tendency, but it is not a criterial feature, the presence of determiners is further examined. Silva-Corvalán argues that there is an increased continuum of specificity reflected in formal correlates, going from bare nouns, at one end, to to nouns modified by a definite determiner, at the other; with nouns modified by an indefinite determiner as an intermediate category. The result is a 4-point scale aimed at representing degrees of "specificity", which is predicted to correlate with the likelihood of "doubling" of preverbal accusatives (op.cit. 165).

---

[13] It is important to note that these are not instances of proper "clitic doubling" as defined in this dissertation, since they entail the left-dislocation of lexical object.

$$\overset{\varnothing}{\text{[-det., -def.]}} \quad \overset{\varnothing \:/\: ?lo}{< \text{[+det., - def.]}} \quad \overset{lo \:/\: ?\varnothing}{< \text{[-det., +def.]}} \quad \overset{lo}{< \text{[+det., +def.]}}$$

Table 1. *Definiteness, determiners and specificity*

The following examples illustrate the interaction of these features (op.cit. 165-166).

(44) a. **Padrastro** tengo   [-det.,-def.]
          'Stepfather I have.'

      b. **Ochenta mil pesos** dice que le dieron.   [-det., -def.]
          'Eighty thousand pesos she says they gave her.'

(45) a. **Un hogar mal constituido** no **lo** necesita nadie. [+det., -def.]
          'A poorly constituted home nobody needs it.'

      b. **Un mechoncito chico** pues le sacan nada más.   [+det, -def.]
          'A small lock of hair they cut him only.'

(46) a. **Eso** tomo de desayuno.   [-det, +def.]
          'That (I) have for breakfast.'

      b. **Eso lo** puse en la novela yo.   [-det, +def.]
          'That I put in the novel.'

(47) a. **La culpa la** tiene ella.   [+det, +def.]
          'The blame she has it.'

      b. Y **a los pascuenses los** amontonaron en la parte de pedregullo. [+det, +def.]
          'And the Easter Islanders they packed them in the area with rocky ground.'

The sentences in (44) show that pre-verbal indefinite NPs disfavor "doubling". In

(45) we find two pre-verbal indefinites modified by a determiner. This case represents an

increase in specificity, according to Silva-Corvalán's criteria, and is therefore predicted to favor "doubling" to a certain extent. This is illustrated with a "doubled" instance in (45)a), and a "non-doubled" one in (45)b). Yet more specific are bare definites, as illustrated in the examples in (46), also showing a "doubled" and a "non-doubled" instance. Finally, the examples in (47) illustrate sentences with pre-verbal DOs, the most "specific" group, and we see that both examples contain a resumptive clitic pronoun.

Silva-Corvalán's data shows that the percentage of clitics coreferential with pre-verbal DOs increases with the "specificity" of the object, again defined in terms of the interaction of definiteness and presence of determiner. She provides the following table (op.cit. 168).

1. [-determiner] [-definite]   6/47   =13%
2. [+determiner] [-definite]   6/14   =43%
3. [-determiner] [+definite]   68/81   =84%
4. [+determiner] [+definite]   153/161   =95%

Table 2. *Incidence of definiteness in the encoding of preverbal DOs*

Silva-Corvalán argues that these factors are actually expressions of the relative topicality of the participant involved in the "doubling" construction. In this context, "[*t*]*opicality* refers to the likelihood of a given constituent to qualify as the topic of a sentence, i.e. a constituent that has a series of features that are characteristic of the kinds of referents that people tend to talk about" (op.cit. 166).

An account of clitic "doubling" in terms of topicality also allows for explaining the greater frequency of IO-doubling with respect to DO-doubling, since over 90% of the IOs in Silva-Corvalán's corpus are definite and human, whereas DOs encode human referents in 15% of the instances, and are definite in 36%. Further, 43% of the IOs occur preverbally (i.e. in topic position), whereas this is the case for only 7% of the DOs, an asymmetry that reflects the status of DOs as least likely topics.

Since post-verbal DOs are lower in topicality than pre-verbal ones, it is expected that there will be a lesser number of doubling constructions with post-verbal objects. Of a total of 5,086 post-verbal DOs in the sample, only 65 occur in a doubling construction (i.e. about 1.3%). The same factors that affected the doubling of pre-verbal DOs, however, can be invoked for explaining the doubling of post-verbal DOs as well: all doubled post-verbal DOs are definite (although the relative percentage of bare definites is not stated) and 41.5% of them (27/65) are human.

The author also notes that out of the 65 doubling cases, 36 (55.4%) involve entities which had been referred to in one of the preceding two clauses, and provides a table showing that, within the subset of NPs with a close antecedent, the percentage of doubling increases for those involving a human referent (op.cit. 169).

1.	[+human] [+definite]	17/59	=29%
2.	[-human] [+definite]	19/193	=10%
Total:		36/252	=14%

Table 3. *Percentage of DO-doubling for human referents*

The fact that human objects are doubled more often than non-human ones is considered a reflection of their higher ranking in topicality. This study is therefore advanced as supporting Givón's (1976) hypothesis that the spreading of agreement, of which DO-doubling is interpreted as an instance, is governed by topicality hierarchies.

2.2.3. Weissenrieder (1995)

Weissenrieder's study is focused on IO-doubling in Argentinean Spanish. As in other approaches summarized above, here again clitics are assimilated to agreement markers: "[i]n traditional accounts the pronoun *le* in the sentence *Le dije a Abelardo la verdad* is viewed as the product of pronominalization. The pronoun *le* is oddly redundant because it 'substitutes' for a noun phrase that remains present. Reinterpreted as a token of IO-V agreement, however, the clitic *le* merely matches the information given in the IO noun phrase in much the way that verb endings match features of their subject noun phrases" (op.cit. 171). Thus, object agreement is viewed as analogous to subject agreement, with the only difference being that subject agreement is obligatory (in finite clauses), and object agreement is not. As Silva-Corvalán, Weissenrieder predicts that the occurrence of the object agreement marker depends on the relative topicality of the referent involved. She cites Croft (1988:175) in saying that "both pronouns and agreement markers are used to identify and maintain the identity of their referents across the discourse… Therefore, the speaker must make a choice as to which entities will continue to be cross-referenced and which ones will not. Naturally, the most important or

43

salient entities will continue to be cross-referenced, and those tend to be the most animate ones, the most definite ones, and the ones most central to the events being reported".

In accordance with this hypothesis, Weissenrieder analyzes the realization of third-person IOs in *El beso de la mujer araña*, a novel by the Argentinean author Manuel Puig. The results from this analysis are presented in the following table (op.cit. 173; "Vag" =dative clitic).

FORM	TOTAL/%	EXAMPLE
Vag+Ø	632/75%	**Le** pide disculpas.
Vag+NP	130/16%	**Le** grita de todo **a la chica**.
Ø+NP	38/5%	Pide **al ordenanza** un café doble.
Vag+Pro	32/4%	…**le** está haciendo bien **a ella**.

Table 4. *Relative frequency of encoding-types for IOs*

The high frequency of what is called the "null representation" (Vag+Ø=75%) is interpreted as resulting from the correlation between IOs and highly topical entities, since in these cases only a minimal token (the clitic) is generally sufficient for keeping the correct referent in mind. The least frequent form (Vag+Pro=4%) is used mostly in cases of ambiguity of the referent, or as a device to warrant a shift in topical emphasis. With regards to the occurrence of the clitic and a full NP (Vag+NP=16%), in the majority of these cases the noun is necessary to establish referent identity, something that wouldn't be possible if only the clitic were present. Finally, Weissenrieder found 5% of non-agreeing IOs, i.e. cases where only the lexical phrase occurres in the clause (Ø+NP=5%). These instances involve NPs that are interpreted as topically unimportant, and correspond to entities which are the least often mentioned in the text. The conclusions from this

analysis point to the correlation between IOs and topical participants. The frequent presence of the dative clitic in the clausal structure is viewed precisely as an effect of this correlation, given the sensitivity of agreement phenomena to discourse factors, such as the relative topicality of the referents involved.

2.2.4. Colantoni (2002)

Colantoni (2002) presents an analysis of dative and accusative doubling based on a corpus consisting of two-hour long interviews with six residents of different towns in the province of Corrientes, Argentina. The author provides a count of non-reflexive first singular, second singular and third person object pronouns, which total 575 instances in her sample. From them, 89 occur in a doubling construction, which is defined here as "the duplication of the clitic by a lexical expression" (op.cit. 322). Doubling is most common with third-person dative clitics (69/89=78%) than with any other form. In this dialect, however, dative clitics do not only refer to dative arguments. The particular variant analyzed is *leísta*, and thus some of the doubling constructions involving the *le* form actually refer to accusative arguments. An example is provided below (op.cit. 323).

(48)   Hay muchas que no **le** quieren amamantar **a sus hijitos**.
         'There are many (mothers) who don't want to nurse their children.'

From the 69 instances of doubling involving *le*, the clitic should be interpreted as accusative in 23 of them (33.3%). Note further that in the example in (48) the clitic does not show number agreement with the NP. Colantoni reports that in her sample the form *le*

is used indistinctly for singular and plural referents in most of the cases. It is argued that this extra *bleaching* of the form conspires for making doubling more frequent, since a coreferential NP is often needed to facilitate identification.

The second most frequent doubling construction involves accusative pronouns (10 instances=11%). From this percentage it is possible to deduce that doubling of DOs (with both the *lo* and the *le* clitic forms) amounts to 37% of all doubled clauses (23+10/89). In Colantoni's corpus the accusative clitic doubles animate NPs in only 3 cases, and none of them is human. This result is expected since human objects are referred to by the *le* form.

Since dative pronouns are more frequent than accusative ones, and dative pronouns mostly refer to animate entities, Colantoni concludes that if the referent is high in the animacy scale, it is more likely for the clitic to be doubled (op.cit. 325). Further, it is argued that one of the functions of clitic-doubling (for both datives and accusatives) is to establish reference when new entities are introduced in the discourse. Colantoni reports that a doubled lexical phrase is used to introduce a new referent in 59.6% of the instances of doubling in her corpus. In the remaining 40.4% of the cases, which include all the doubling instances with first or second person clitics, the doubled phrase is used with a contrastive or emphatic purpose. The author observes that doubling is more frequent with third person than with first or second person clitics, and also more frequent with lexical NP encoding new referents than with other NPs. Based on these facts, the author reaches a conclusion that seems to oppose that of the other researchers; namely that doubling constructions arise whenever the referent is unidentifiable or receives contrastive focus.

2.2.5. Summary and Discussion

Corpora-based studies, as reviewed in this section, generally agree in correlating clitic doubling correlates with topicality hierarchies (even if sometimes from opposite ends). In particular, Silva-Corvalán posits that DO-doubling typically occurs with entities which are topical, and Weissenrieder makes a similar claim with respect to IO-doubling. Apparently discordant are Colantoni's conclusions, which assimilate doubling with referents new to discourse. It is just not plausible to assume that this divergence is due to dialectal variations, since it wouldn't be simply the case of one dialect pushing further than others the conditions of use of certain construction, which is common, but of related dialects using a particular construction with the opposite function. As an alternative, I will propose that even if the conclusions the authors derive from these studies are not compatible, their data are. A consistent interpretation of their empirical facts depends, I believe, on analyzing DO-doubling and IO-doubling separately, and against the canonical form of encoding of each case-role (i.e. datives vs. accusatives), an idea that will be further developed in Chapter 5.

On the other hand, the diverging interpretations of the data seem to arise from an ambiguous use of the concept of "topic" to refer to both "inherent topicworthiness" and "context-imparted topicworthiness" (Payne 1992:51) or the related "intrinsic salience" vs. "plot salience" distinctions (Du Bois 1980:248). In Chapter 5 I will argue that the greater frequency of IO-doubling when compared to DO-doubling may depend on the greater "intrinsic salience" of dative arguments compared to accusative arguments; but what determines that a particular argument is doubled in a particular text is, on the contrary,

the "plot salience" of that argument or, more precisely, its assumed cognitive accessibility at a particular point in the exchange. I will try to show that conceptualized in this way, we may gain greater insight into the function of doubling constructions of both dative and accusative arguments.

3　'DOUBLING' IN OTHER LANGUAGES

In the previous chapter we reviewed the main hypotheses that have been advanced for explaining clitic doubling in Spanish. The three main theoretically oriented proposals we summarized addressed the factors constraining the occurrence of DO-doubling (i.e. the specificity, presuppositionality, or *a*-markingo the lexical phrase). It was proposed, on the other hand, that IO-doubling was not subject to any of those constraints.

We also reviewed works based on corpus data, more concerned with patterns based on frequencies of occurrence. In these studies, clitic doubling is related with topicality, although this notion was invoked to make opposite predictions: that that doubling marks topical entities (Silva-Corvalán 1981, Weissenrieder 1995) and that it marks referents newly introduced into the linguistic context or is used for emphatic purposes (Colantoni 2002).

In the present chapter we will see that the insights behind these hypotheses were adopted, to various extents and in different versions, in subsequent works devoted to clitic doubling in other Spanish dialects as well as in other languages.

It is well known that in many languages there may be object-verb agreement where objects are positively marked for certain features: specificity in Turkish (Enç 1991), definiteness in certain Uralic languages (Comrie 1981), identifiability in Persian (Roberts 2005), etc. These agreement markers do not carry the pronominal specifications of the arguments they agree with, and cannot "replace" these arguments in the clausal structure. Thus they are excluded from the survey of "doubling" languages presented here. (It is

relevant, however, to note that the kind of features governing these agreement processes correlate with the features proposed to govern clitic doubling as well).

Factors such as the characterization of a language as head-marking vs. dependent-marking may alter the labels chosen for certain categories, and the way in which they are analyzed. For instance, bound forms are more easily viewed as incorporated pronouns in head-marking languages and as agreement markers in dependent marking languages, even if in some cases there is no particular evidence of their distinct grammatical behavior. The distinction between head-marking and dependent-marking is also the basis for referring to these bound morphemes as instances of cross-reference vs. agreement, as we will see in Chapter 6. The summary presented here is, therefore, tentative; and limited to cases where a particular construction has been interpreted as an instance of "doubling", or where the available description suggest that an analogous phenomenon is at play, independently of the labels used to characterize it.

I will consider here that a language contains "doubling" constructions if (i) there is a bound form on the verb indexing pronominal features of some argument; (ii) this bound form is optional, at least in some constructions (e.g. there is obligatory verb indexing of strong pronouns, but optional indexing of lexical phrases); (iii) there are at least some contexts where the verb-marker can occur alone; the lexical phrase can occur alone, or both forms can co-occur.

Finally, there is the important issue of distinguishing between doubling and dislocation. In well studied languages, evidence has been presented for opting for one or the other analysis. For instance, Spanish and Greek, for instance, are traditionally

analyzed as allowing doubling, whereas superficially analogous strings in, for instance, Italian and French are traditionally analyzed as involving dislocation. (I will refer briefly to the issues related to the Italian and French examples at the end of § 3.1.). In less studied languages, the question is more difficult to solve. The characterization of some constructions as potential cases of doubling summarized at the end of this chapter should thus be taken with a grain of salt until more evidence becomes available.

With this caveat in mind, in this chapter we will first review clitic doubling in varieties of Spanish beyond the dialect of Buenos Aires and the surrounding areas. Then, we will examine analogous constructions in other Romance languages, Balkan-Slavic languages, and languages of the Bantu family. Some scattered evidence related to other languages will be briefly mentioned at the end of the Chapter.

3.1. ROMANCE LANGUAGES

3.1.1. Clitic doubling in other Spanish dialects

Since IO-doubling is assumed to be grammatically unrestricted, much of the discussion about clitic doubling in Spanish is focused on doubling of direct objects. And since DO-doubling seems particularly productive in the Spanish of Buenos Aires [BA][14], this dialect is the most often referred to in the literature on the subject. However, as Franco's examples from Basque Spanish and Silva-Corvalán's discussion of the Spanish of Santiago de Chile illustrated, DO-doubling occurs in other varieties as well.

---

[14] That is, the same dialect also called "Porteño", "Río de la Plata", or "River Plate" Spanish, and included in what is sometimes called "Southern Cone" Spanish (cf. §2.4).

With regard to Chilean Spanish, the only grammatical difference with Buenos Aires Spanish that we can note at this point is that in the former all doubled post-verbal DOs are definite. Recall that, on the contrary, much of the discussion of DO-doubling in Buenos Aires Spanish was based on doubling of indefinite NPs. With regard to Basque Spanish, it differs from Buenos Aires Spanish in that Basque, being an instance of so-called *leísta* dialects, substitutes *le, les* for *lo, los, la, las* when the doubled participant is human. The other difference is that in Basque Spanish the *a*-marking of the doubled DO is obligatory (Franco 2000:171), whereas opinions differ with respect to this requirement for Buenos Aires Spanish (cf. §1.4 and §4.2 below).

For other dialects the evidence is for the most part anecdotal. It is generally assumed that Peninsular Spanish bans DO-doubling. For instance Mayer (2003:16) notes that "Standard Peninsular Spanish prohibits clitic doubling of direct objects, even if they are animate". But Gutiérrez Rexach (2000:31) challenges this assumed view, and observes that DO-doubling "in spoken and colloquial Castillan Spanish, as well as in Southern Cone Spanish, (…) is quite regular". He cites examples from Llorente and Mondejar (1972).

(49) a. **Le** vimos **a Carrancido**.
          'We saw Carrancido.'

     b. Me ha costado despertar**lo a don Lorenzo**.
          'I had trouble waking up don Lorenzo.'

     c. Usted a de ver**lo a Lucas**.
          'You should see Lucas.'

Likewise, Suñer (1989b) provides examples of DO-doubling in the Spanish of Madrid (taken from Quilis *et al.* 1985:101), which suggest that in this variety DO-doubling is possible even with non *a*-marked and inanimate referents.

(50) a. No **le** he oído **a ese señor**.
          'I haven't heard that man.'

      b. **La** encuentro **a la tía**.
          'I meet this woman.'

      c. …sí, pero no **la** voy a pisar **la Universidad**.
          '…yes, but I'm not going to set foot on the University.'

Consistent with her hypothesis about the specificity of doubled DOs, Suñer suggests that the Spanish of Madrid behaves like Buenos Aires Spanish in requiring doubled DOs to be interpreted as "specific". Madrid Spanish allows so-called "extraction" of doubled DOs whenever the DO phrase is specific, regardless of whether the indication of specificity is overt, as in (51), or covert, as in (52).

(51) a. ¿**A cuántos de ellos les** entrevistaron?   (Suñer op.cit.)
          'How many of them they interview?'

      b. ¿**A cuál de los actores le** invitarán?
          'Which one of the actors will they invite?'

(52) a. ¿**A cuál le** entrevistaron?   (op.cit.)
          'Which one did they interview?'

      b. ¿**A cuántas candidatas las** invitaron?
          'How many candidates did they invite?'

When specificity (in the sense of partitiveness) is covert, as in (52), Suñer suggests that speakers visualize the quantifier as ranging over an implicit group, a conceptualization that is potentially available for those speakers who also accept (53).

(53)   **¿A quién le** viste?   (op.cit.)
         'Who did you see?'

The facts reported about Madrid Spanish, therefore, seem to be consistent with those posited for Buenos Aires Spanish (in both cases accusative doubling depends on the specificity of the doubled NP), although it is possible that DO-doubling is less productive in Madrid Spanish than in the Buenos Aires variety.

Suñer also provides examples of DO-doubling in the Spanish spoken in Quito, Ecuador. This dialect is also *leísta*, to a greater extent than varieties spoken in Spain (i.e. Basque Spanish and Madrid Spanish). Accusative clitics are replaced by *le(s)* in all contexts (i.e. for animates and inanimates, masculine and feminine objects). Like the varieties of Buenos Aires, the Basque country, Chile and Madrid, this dialect also permits doubling of inanimate DOs, as illustrated in (54).

(54) a. **Le** contrataré **al taxi**.   (op.cit. 388)
          'I will hire the taxi.'

      b. Ya **le** veo **a la camioneta**.
          'I already see the pick up.'

Again, consistent with the specificity hypothesis, it is claimed that wh-extracted doubled DOs are accepted only if they are interpreted as referring to some identifiable

set. Whenever indefinites are unambiguously marked as non-specific (e.g. by the subjunctive mood on the relative clause), clitic doubling is claimed to be disallowed (55)

(55)   ***Le** busco **a una estudiante que sepa japonés**.   (op.cit. 391)
          'I'm looking for a student who speaks [subjunctive] Japanese.'

Sanchez (ms.) reports evidence form the Spanish of Lima, Perú. In it, the specificity condition seems to be applicable only to DO interrogatives. In affirmative sentences, on the contraty, doubling of non-specific DOs is claimed to be allowed as well. This is illustrated with the example in (56).

(56)   **Lo** alabarán **al niño que termine su tarea primero**.   (op.cit. 14)
          'They will praise the boy who finishes [subjunctive] his homework first.'

Sanchez claims that whereas DO-doubling of non-specific referents is possible, a sentence like the one presented in (56) can be uttered only if the "potential antecedents" for the doubled NP have being previously mentioned in the discouse context. The author explains the conditions governing doubling constructions in this dialect in terms of Kiss' (1998) distinction between "information" vs. "identificational" focus. Kiss defines identificational focus as marking "a subset of the set of contextually or situationally given elements for which the predicate phrase can potentially hold; it is identified as the exhaustive subset for which the predicate phrase actually holds" (Kiss 1998:245).

Sanchez suggests that whereas non-doubled DOs carry information focus, doubled DOs carry identificational focus. Kiss proposes that whereas it is possible to negate exhaustive identificational focus, the negation of information focus results in

ungrammaticality. Accordingly, Sanchez presents the following examples as evidence that the distinction between non-doubled structures (i.e. information focus) and doubled structures (i.e. identificational focus) is operative in Limeño Spanish.

(57) A: -María la vio a Teresa.   (op.cit)
          '-María saw Teresa.'

      B: -No, **la** vio **a Tatiana** también.
          '-No, she saw Tatiana too.'

(58) A: -María vio a Teresa.   (op.cit)
          '-María saw Teresa.'

      B: **?**-No, vio a Tatiana también.
          '-No, she saw Tatiana too.'

The rational is that the doubled NP in (57) can be negated because it carries exhaustive identificational focus, whereas the one in (58) cannot be negated because it carries information focus. In this respect Limeño Spanish seems to differ from the Spanish of Buenos Aires, where (58) is fully acceptable[15]. Another difference between Limeño and Buenos Aires Spanish is that Limeño is argued to ban doubling of inanimate NPs, a possibility which was assumed to be marginally possible in the other variety (cf. Suñer 1988).

(59)   * **Lo** compraron **al libro**.   (Sanchez, op.cit. 14)
          'They bought the book.'

---

[15] Note that the author marks this variant with a question mark, possibly suggesting that the sentence is not fully ungrammatical in Limeño.

The examples presented by Sanchez contrasts with those provided by Mayer (2003), whose study is based on data collected from native monolingual Limeño speakers. For Mayer "Limeño is still very conservative […] barring almost all non-pronominal NPs from clitic doubling. There is one exception though; clitic doubling of proper names has become optional" (op.cit. 20).

There is also some evidence of DO-doubling in Mexican Spanish. Alarcón and Orozco (2004) report cases of what they characterize as DO-doubling among Mexican children. Their subjects produced doubling constructions in 13 of 588 analyzed clauses (2.2%). Some of their examples are provided below, with the available context.

(60) a. Después el niño fue a subirse a una piedra muy grande y le volvió a gritar y
después salió un venado y lo agarró y se lo llevó y el perro lo iba siguiendo. Iba
siguiendo al venado con todo y niño. Y después **lo** tiró **al niño** y tiró también al
perro.
'Then the child went to climb a very big stone and shouted to him again and
then a deer came out and grabbed him and took him away and the dog was
following him. It was following the deer with child and all. And then **threw the
child** and threw the dog also.'

b. Después la tortuga le avisó al niño y se fijaron y entonces la tortuga se enojó
con la rana y **la** buscaron **a la rana chiquita**.
'Then the turtle warned the child and they watched and then the turtle went mad
with the frog and they **looked for the little frog**.'

c. E iban nadando en esa tabla y el niño iba jugando como a los piratas. Después
el niño… luego la rana llegó y **la** aventó **a la ranita** al agua.
'And they were swiming in that table and the boy was playing pirates. Then the
boy… then the from came and **threw the little frog** into the water.'

d. Entonces la tiró de la tortuga, pero el niño se dio cuenta. Entonces le dijo… **la**
regañó **a la rana**… entonces **la** dejaron **a la rana** y se fueron todos.
'Then threw it from the turtle but the boy realizad it. Then he say… he scolded
the frog… then they left **the frog** and they all went away.'

57

e. **Lo** acompañaba **al niño**.
   'H/she went with the boy.'

f. **Los** tira **al niño y al perro**.
   'It **throws the boy and the dog**.'

In turn, Parodi and Santa Ana (2002) report the occurrence of DO-doubling in the rural region of Michoacán, México, based on collected data from adult native speakers.

(61) a. **La** tienen **la herramienta**.
        'The have/hold the tool.'

b. **Los** registran **a los muchachos** al otro lau.
   'They register the guys on the other side.'

c. Se **los** vendo **el animal**.
   'I sell the animal to you.'

d. Nunca **lo** he visto **que no es reconocido como guandari** [poeta].
   'I have never seen that he is not recognized as poet.'

Most of the examples categorized as instances of DO-doubling in Mexican Spanish available in the literature involve animates (and predominantly humans). The example in (61)a, however, suggests that doubling of inanimates may be possible, and that *a*-marking is not a criterial feature for DO-doubling in this dialect. Note, also, that the example in (61)d involves a clausal complement[16].

Groppi (2002) attests the occurrence of DO-doubling in the Spanish of Montevideo, Uruguay. She claims that Montevideo and Buenos Aires Spanish are alike in allowing DO-doubling regardless of the animacy of the referent, and in the absence of *a*-

---

[16] The doubling of clausal complements was also attested in the Buenos Aires corpus and will be discussed in Chapter 4.

marking. Montevideo Spanish seems to be like Chilean Spanish, however, in allowing doubling only if the object is definite. She provides the following examples.

(62) a. **Lo** preparé **el mate** con la yerba que me trajiste.
           'I prepared the *mate* with the *yerba* you brought me.'

      b. ***La** vi **una mujer** en el parque.
           'I saw a woman in the park.'

3.1.2. Romanian

There is general agreement in that Romanian is the only other Romance language, aside from Spanish, where DO-doubling is relatively productive. In fact, the obligatory co-occurrence of an accusative clitic, limited in Spanish to contexts in which the DO appears as a strong pronoun, extends in Romanian to names, definite descriptions and strongly quantified NPs, all of which *must* be clitic-doubled. On the other hand, doubling is limited to human referents (Gierling 1997). Note, incidentally, that doubled phrases must also be marked by the preposition *pe*.

(63) a. ***Ø/L-**am   vazut **pe Ion**.   (op.cit. 73)
               him-have(I) seen   PE John.
           'I've seen John.'

      b. Am   vazut-***Ø/o**   **pe ea**.
           Have-(I) seen-   her PE she.
           'I've seen her.'

      c. ***Ø/Il**   cunosc **pe postas**.
               him know(I) PE postman.
           'I know the postman.'

    d. ***Ø/O** pupa **pe fiecare fata**.
        her kisses PE every girl.
      'S/he kisses every girl.'

An interesting aspect pointed out by Gierling is that the addition of "extra material" (e.g. a modifier), by favoring the specific interpretations of the referent, diminishes the acceptability of non-doubled constructions in contexts where it would be otherwise optional.

(64)   Ion **?Ø/o** asteapta **pe servitoarea** *grasa*.   (op.cit.)
       John   her waits   PE servant-the fat
       'John is waiting for the *fat* servant.'

In Romanian, clitic doubling is claimed to be also compatible with weakly quantified NPs, but only if there is no material which would force a non-specific interpretation.

(65)   **Il** astept   **pe medic (*de orice specialitate)**.   (op.cit.)
       him wait-for(I) PE doctor (of any specialization)
       'I am waiting for a doctor (*of any specialization).'

It is argued that in Romanian any DO can be clitic-doubled provided that it is referential and denote a human participant. However, the fact that accusative clitics may refer to non-human entities when they are not in a doubling construction, suggests that the clitics are not inherently marked as [+human]. As a consequence, the ungrammaticality of doubling non-human referents cannot follow from a mismatch in feature values and probably results from discourse-pragmatic considerations

(Anagnostopolou 1995:4, Gutiérrez Rexach 2000:39), among them the type of focus structure they instantiate (Gierling 1997).

3.1.3. French and Italian

As mentioned in the introduction to this Chapter, in French and Italian the sequence of an object clitic plus a lexical phrase in canonical position is analyzed as involving dislocation. In effect, the literature on the topic has put forth some compelling arguments to show that these languages behave differently than Spanish or Romanian with respect to putative "doubling" constructions (cf. Auger 1994, Cecchetto 1999, Cardinaletti 2002, among others).

There have been, however, certain claims in the sense that, although object right-dislocation may be the most frequent case, object doubling is not completely non-existent. For instance Wanner (1987: 418) comments that "[i]n many Italian dialects and spoken variants of regional Italian the phenomenon of clitic doubling can be observed, i.e., the expression e.g. of an indirect object phrase by a nominal constituent (NP or strong pronoun) and a concomitant clitic pronoun with the same reference and attached to the verb". In turn, Capone (2001) suggests that Italian may further allow DO-doubling, provided that the DO expresses a "speaker / hearer presupposition".

Similar claims have been made for French. For instance Auger (1994: 39, 108) cites studies by Kaiser (1994) and Carroll (1982) reporting the (marginal) occurrence of object doubling in European and Quebec Colloquial French, respectively. Auger, who in turn argues for the dislocation hypothesis, recognizes that "sentences containing right-

61

dislocated objects and sentences involving doubled objects are also superficially very similar. Indeed (…) one cannot rely on the presence of an intonational break between the dislocated phrase and the rest of the sentence in order to distinguish the two types of constructions, since the presence of such breaks is not always necessary (p. 95).

Analogous observations were made by Duranti and Ochs (1979: 389) for Italian. The authors also consider "doubled" DOs as right dislocations in this language, but they point out that coreferential clitic pronouns might have started to act as "agreement markers", something that "is consistent with the fact that we did not find, in the majority of cases, any remarkable intonational break between the left-dislocated (or right dislocated) constituent and the rest of the utterance".

One should note, in this regard, that if there is no intonational break produced by the speaker, there is nothing in principle preventing the hearer to reinterpret the "dislocated" construction as a case of "doubling" and the clitic pronoun as an agreement marker; save probably the pragmatic value associated with the target referent. Since there is still certain disagreement regarding the pragmatic values of doubled vs. dislocated constituents (and of dislocation vs. "afterthoughts")[17] I will have to leave the question open as to whether at least some of the instances considered dislocation could be analyzed as in fact involving doubling constructions in these languages.

---

[17]   See for instance Lambrecht (1994) and Grosz and Ziv (1997).

3.2. BALKAN AND SLAVIC LANGUAGES

3.2.1. Albanian and Greek

According to Kallulli (2000), in Albanian and Greek DO-doubling is allowed not only with animate and definite NPs, but also with inanimate and indefinite ones. Doubled DOs need not be preceded by a preposition[18] (her examples).

(66) a. Do **ta** pija me kënaqësi **një uiski**.   [Albanian]
           FUT-it drink with pleasure a   whisky
           'I would gladly drink a whisky.'

      b. **To** pino eukharistos   **ena ouiskáki**.   [Greek]
            it drink with pleasure a   whisky
           'I would gladly drink a whisky.'

However, it is argued that both languages strongly disfavor DO-doubling when the NPs are marked as focal or are part of the focus domain, even in the NPs are definite.

(67)   -*What happened?*

      a. Jan-i   (***i**)   hëngri **fasule-t**.   [Albanian]
           Yannis them ate   beans-the
           'Yannis ate the beans.'

      b. O Yánnis (***ta**)   éfaye **ta   fasólia**.   [Greek]
           the Yannis them ate   the beans
           'Yannis ate the beans.'

---

[18] It has been argued that, in Greek, doubled indefinite NPs occur in dislocated positions (Anagnostopoulou 1994). A different perspective is argued for by Kallulli (2000), who also presents evidence of the non-dislocated position of doubled indefinites in Albanian.

This hypothesis is further strengthened by the fact that so-called "extraction" of wh-words results in ungrammaticality in both languages.

(68) a. Kë (*e) pe?   [Albanian]
          who it saw.you
          'Who did you see?'

      b. Pjon (*ton) idhes?   [Greek]
          who   it   saw.you
          'Who did you see?'

Based on this and related evidence, Kallulli argues that, in Albanian and Greek, clitic doubling can only be used to encode referents which were mentioned in the preceding segment[19]. This is illustrated with the following examples:

(69) -*Who read the book? / What did Ann do with the book?*

      a. Ana e lexoi **libr-in**.   [Albanian]
          Ann it read book-the
          'Ann read the book.'

      b. I   Ana **to** diavase **to vivlio**.   [Greek]
          the Ana it read   the book
          'Ann read the book.'

It is further claimed that Albanian is stricter that Greek in that, in Albanian, referents introduced in preceding segments *must* occur doubled in subsequent mentions. Thus, a non-doubled version of (69)a) would be infelicitous. The non doubled alternative of (69)b) is possible in Greek, although it is claimed to be dispreferred.

---

[19] Similar arguments are presented in Anagnostopoulou (1995:2). In this work it is argued that, regarding DO-doubling, "for Romanian the relevant property is specificity, for Spanish partitivity and for Greek, what Heim (1982) calls prominence", prominence understood as resulting from explicit mention in the previous linguistic context.

The greater productivity of clitic doubling in Albanian is also manifested in the obligatoriness of doubling for dative arguments. Obligatory doubling of datives occurs with definites (70)a, indefinites (70)b and wh-words (70)c.

(70) a. Ev-a ***Ø/i** dërgoi **An-ës** lule.
     Ev-the her sent Ann-the.DAT flowers
     'Ev sent Ann flowers.'

   b. Ben-i ***Ø/i** dërgoi një vajze lule.
     Ben-the her sent a girl.DAT flowers
     'Ben sent a girl flowers.'

   c. Kujt ***Ø/i** foli mësues-i?
     Who.DAT him/her talked teacher-the
     'Who did the teacher talk to?'

3.2.2. Macedonian and Bulgarian

In Macedonian, clitic-doubling is obligatory for IOs and DOs involving definite NPs (Schik 2000).

(71) a. Rada ***Ø/go** bara **Ivo**. (op.cit.)
     Rada ACC3MS is.looking.for Ivo
     'Rada is looking for Ivo.'

   b. Rada ***Ø/mu** pomaga **na deteto**.
     Rada DAT3MS is.helping PREP child-the
     'Rada is helping the child.'

This is the case independently of whether the doubled NPs are topical or not. Thus, it is argued that in Macedonian clitic doubling constructions do *not* interact with information structure but "are primarily and consistently used as a grammaticalised

means of emphasizing the referential definiteness of the doubled object phrase" (op.cit. 271).

On the contrary, in Bulgarian, referential humans in DO function can be doubled, even if indefinite (Franco 2000: 169), but only if they refer to "sufficiently individuated entities" (Schik op.cit. 265, her examples).

(72) a. Rada (**go**) tărsi **pismoto** / **edno pismo**.
        'Rada is looking for the letter / a certain letter.'

    b. Rada (**mu**) pomaga **na deteto** / **edno dete**.
        'Rada is helping the child / a certain child.'

This expresses, as Shick argues, that clitics "signal in a consistent way that a new interpretational perspective is introduced in which the referentially specific doubled object phrase is presupposed in the discourse as a given entity" (op.cit. 268). Thus, contrary to Macedonian, clitic doubling in Bulgarian is claimed to systematically interact with information structure. For instance (73) can only be felicitously uttered in response to the (a) question, but not the (b) question. If the DO is in focus, doubling is ungrammatical (74)a, but if the focus falls on the subject or the verb, doubling is obligatory (74)b.

(73)   Ivo (**go**) napisa **pismoto**.
        'Ivo wrote the letter.'
        a. What happened to the letter?
        b. What happened?

(74) a. Ivo **Ø/*go** napisa PISMOTO.
          'Ivo wrote THE LETTER.'

      b. IVO ***Ø/go** napisa pismoto / Ivo ***Ø/go** NAPISA pismoto.
          'IVO wrote the letter / Ivo WROTE the letter.'

3.3. BANTU LANGUAGES

In a number of Bantu languages there are series of so-called "object markers" which occur attached to the verb and encode the pronominal features of the object argument. Similar to Spanish DO clitics, these forms are often optional, and their occurrence depends on some semantic or pragmatic feature of the relevant referent.

For instance in KiRimi (Bantu, Tanzania) object agreement can only occur with definite animate objects (Hualde 1989); but in Swahili (Bantu, Tanzania) it is less constrained: Object marking is obligatory for first and second person objects and common with objects referring to humans. It may also occur with objects referring to inanimate participants, although these cases are rarer (Wald 1979). In terms of factors linked to the context of use, it has been claimed that referential persistence favors object marking of both human and inanimate referents, suggesting a correlation between object marking and topicality or discourse saliency (Seidl and Dimitriadis 1997).

In Ruwund (Bantu, Zaire and Angola; Woolford 2001) animate and specific direct objects trigger obligatory object markers on the verb. The absence of the object marker unambiguously indicates that the object is non-specific (examples from Nash 1992:565ff).

(75) a. ku-kimb   muntu.
INF-look.for person
'to look for a (any) person.'

b. ku-**mu**-kimb   **muntu**.
INF-OM-look.for person
'to look for a/the person (speaker has a particular person in mind).'

For inanimates, the presence of the agreement morpheme is optional. In these cases it is argued that the absence of the morpheme does not provide any indication with respect to the specificity of the object. However, if the agreement morpheme occurs, the object is interpreted as specific and, further, as focal (76).

(76) a. ku-land malong.
INF-buy plates
'to buy some/the plates.'

b. ku-**ma**-land   malong.
INF-OM-buy plates
'to buy THE PLATES.'

Chicheŵa (Bantu, Malaŵi; Bresnan and Mchombo 1987) requires a more detailed comment. This language has a series of "subject and object markers". The two markers differ in that the first is obligatory and the second is optional. When the object agreement marker occurs, the object NP can be omitted. On the other hand, it may be the case that what appears as a direct object NP co-occurs with the object marker. Up to here the facts from Chicheŵa seem analogous to those from Spanish. In both languages there is an obligatory marker on the verb which expresses the pronominal features of the subject and an optional marker which codes the pronominal features of the object NP. Under Bresnan and Mchombo's analysis, however, whereas the subject marker can perform double-duty,

serving as a marker of grammatical agreement or as an incorporated pronoun, the object marker is always an incorporated pronoun, co-occurring not with a lexical object but with a "topic" NP. Similar facts with regards to the object marker are reported for Setawana (Bantu, Botswana; Demuth and Johnson 1989). In this sense, thus, there is no "doubling" in Chicheŵa or Setawana as there is in Spanish, where the lexical NP does not function as topic. As we saw in Chapter 2, Spanish doubled objects may be questioned. Further, in declarative sentences, they normally occur within the focus domain and may even be the locus of narrow focus, as it will be argued in Chapter 6.

3.4. OTHER LANGUAGES

The distribution of "pronominal affixes" is sometimes sensitive to the type of phrase they "double". In Welsh pronominal affixes are in complementary distribution with full nominals, but may co-occur with personal pronouns (Lyons 1999a). In Palauan (Austronesian, Palau Islands), on the contrary, the complementary distribution is with pronouns and not with lexical phrases (Corbett 2003:23). In turn, in Kichaga (Bantu, Tanzania), a pronominal marker is obligatory with free pronouns and optional with lexical NPs (Bresnan and Moshi 1990:151-152).

In other cases, the occurrence of "pronominal affixes" depends on the grammatical function of the elements involved. In Tukang Besi (Austronesian, Indonesia) subject pronominal markers are obligatory, but clitics cross-referencing the object are optional (Donohue 1999).

Finally, in some languages the occurrence of these affixes is directly linked to sematic / pragmatic factors. In Siddhama (Cushitic, Ethiopia), pronominal markers occur only with animate affected entities (Kawachi 2007). In Kambera (Austronesian, Indonesia), markers may optionally co-occur with definite lexical objects or "replace" them (Klamer 2004). Tamazight (Berber, Morocco; Ouali 2006) has optional IO-doubling; the dative clitic may occur with animate and inanimate referents, but only if they are specific. In Larike (Austronesian, Indonesia) the co-occurrence of the lexical object plus the object marker is claimed to grant "maximum prominence" to the target entity; whereas subsequent lower levels of prominence are indicated by the use of the phrase alone, the marker alone, or neither (Laidig and Laidig 1990:106; as cited in Corbett 2003).

3.5. SUMMARY AND DISCUSSION

This chapter had two main goals. The first was to present evidence of doubling constructions (and in particular of DO-doubling constructions) in an assortment of Spanish dialects where these constructions are generally considered ungrammatical.

The second goal was to compile data about potentially equivalent constructions in other languages of diverse filiations, abstracting away from terminological differences, so as to open the way to future cross-linguistic research.

The resulting picture suggests a common range of factors correlating with "doubling" constructions, and some common patterns: "doubling" seems to be more common with IOs than with DOs; with animates than inanimates; with definites than

indefinites; and with specific than non-specific referents. In the next two chapters we will examine the relative role of each of these factors on clitic doubling constructions in Buenos Aires Spanish.

4   ACCUSATIVE DOUBLING IN BUENOS AIRES SPANISH

In previous chapters we reviewed the main existing hypotheses which have been proposed to account for DO-doubling constructions: *a*-marking, animacy, specificity, presuppositionality, topicality, and newness. In this Chapter I present a critical evaluation of these hypotheses and confront them with the empirical data provided by a corpus of interactional speech by native speakers of Buenos Aires Spanish. The corpus analysis shows that none of the existing hypotheses can account for the empirical data, even though they are consistent with an interpretation of DO-doubling as coding the relative degree of cognitive accessibility of the target referents, which is the hypothesis defended here[20].

4.1. METHODOLOGY

Most of the data used in this and the next chapter come from *El Habla Culta de la Ciudad de Buenos Aires* (Barrenechea 1987), which is a transcription of a series of interactions among native speakers of Buenos Aires Spanish[21]. The interactions presented in the Buenos Aires corpus vary in both the number of interlocutors and the level of formality, ranging from a university lecture to the informal dialogue among four speakers who interact while playing cards. In most cases the transcriptions capture spontaneous dialogues among two or three speakers. According to the parameters set for the project of

---

[20]   Sections of this Chapter were originally presented in Belloro (2006). The term "referents" should be understood as denoting assumed cognitive representations.

[21]   I want to thank Carlos Luis, from the Instituto de Lingüística, Universidad de Buenos Aires, for facilitaring my access to the digitalized version of the corpus.

which this corpus was part (Lope Blanch 1977), all participants had to be born in the target city, or live there since age 5 and for at least 75 per cent of their lives, and they must also have received their primary and higher education in the target city. The resulting Buenos Aires volume consists of thirty three interactions, amounting to about 24 hours of recordings and about 250,000 words.

This corpus presented itself as the best option available for the kind of qualitative analysis I whished to purse. First, it controls for the native proficiency of the speakers in Buenos Aires Spanish, unanimously considered as the paradigmatic variety of DO-doubling Spanish. I discussed the importance of unambiguously defining the dialect under analysis, since there may be, in theory, different principles at play in doubling varieties. Thus it seems particularly relevant, at this point in our knowledge about the contexts in which doubling constructions occur, to start with a corpus which is consistent in terms of the variety displayed. Second, the texts contained in this corpus make possible to capture the speech of multiple informants (about 40), thus providing more validity to the data obtained.

This corpus consists of spontaneous interactional speech. Clitic doubling has long been stigmatized in prescriptive grammars, and is therefore more likely to be avoided to a greater extent in writing. Theferefore, oral interactions seemed the most productive terrain where to look for instances of these constructions.

Finally, by providing transcriptions of complete interactions, this corpus allowed me to trace the referential continuity of the participants involved from the beginning to

the end of each text, and to place the tokens of "doubling" in their widest context of use, an option not available in other existing corpora.

In counting DO-doubling constructions I selected only the cases in which the DO is a post-verbal lexical phrase: in other words, those cases where there is general agreement that the clitic is grammatically "optional"[22]. In some cases the researchers who were eliciting the data participate in the interaction. I excluded instances of DO-doubling uttered by them, since, even though it is most likely that they are also native speakers of the dialect, there is no explicit information in this regard. Also excluded were the cases accredited to any participant for whom their status as native speaker of the dialect is not made explicit by the descriptions provided in the published work. This resulted in a set of 125 instances of DO-doubling consistent with these criteria, which were used as the basis for the hypothesis presented in this chapter. Additional examples attested in the media or through personal interactions were included to illustrate particular points.

4.2. A-MARKING AND ANIMACY

As mentioned in §2.1, one of the seminal studies of clitic doubling in Buenos Aires Spanish (Jaeggli 1981, 1986) proposed that accusative doubling is only possible if the doubled phrase is introduced by the particle *a*, and predicted that "accusative clitics in Spanish are in strict complementary distribution with direct objects which are not preceded by the marker *a*" (1986:19). This prediction posits the acceptability of sentences such as that in (77), and the ungrammaticality of those in (78).

---

[22] For an analysis of the discourse function of pre-verbal DOs, see Silva-Corvalán (1984) and Bogard (2007a)

(77) Allí se **lo** ve **a Jaime** sentado en el suelo…   (Página 12, Radar, 04/08/07)
      'There you can see Jaime seating on the floor…'

(78) a. **La** compré (***la casa**).   (Jaeggli 1986:19)
          'I bought it (*the house).'

      b. **Lo** vendí (***el libro**).
          'I sold it (*the book).'

There is an extensive literature on the conditions and functions of *a*-marked (or so-called "prepositional") direct objects (cf. Aissen 2000, Pensado 1995, Laca 2006, among others). At the present stage on the evolution of the language, *a*-marking typically targets those DOs which denote animate entities. However, *a*-marking of some inanimates may occur as well, subject to certain conditions. For instance, Jaeggli (1986:36) offers the following examples (from Lois 1982) as an illustration of inanimate DOs marked by *a*, and thus potentially subject to clitic doubling.

(79) a. Las aves saludan a la aurora.
          'The birds salute the dawn.'

      b. Los ácidos atacan a los metales.
          'Acids attack metals.'

      c. El invierno sigue al otoño.
          'Winter follows autumn.'

      d. El entusiasmo vence a la dificultad.
          'Enthusiasm wins over difficulty.'

In accordance with Jaeggli's predictions, the Buenos Aires corpus shows that inanimate *a*-marked DOs may be clitic doubled. We find, however, only four instances of doubled *a*-marked inanimates, which are presented below.

(80) a.   …**lo** estoy levantando **al curso**.   (HC:XI)[23]
       '…I'm bettering the course.'

   b.   …en realidad **la** subestimo en el fondo **a la profesión**.   (HC:XXIII)
       '…actually deep inside I underestimate the profession.'

   c.   …**lo** dejé así **al asunto**, ¿no?   (HC:XXIII)
       '…I left the matter at that, no?'

   d.   …**lo** llevan así **al diario**.   (HC:XXVII)
       '…they take the newspaper like this.'

The scarcity of instances of doubled *a*-marked inanimates can be interpreted as resulting from the extra conditions that must be met for inanimate DOs to be *a*-marked, among them being specific and topical (cf. Torrego 1999, Leonetti 2003). Note that specificity and topicality are features which have also been proposed to correlate with DO-doubling (cf. §2.1.1, 2.2.2), thus strengthening the link between doubled DOs and *a*-marked ones. In fact, the diachronic expansion of *a*-marking on accusative arguments started with strong pronouns (Laca 2006:426); a category that, as we mentioned, triggers obligatory accusative doubling in all Spanish dialects[24].

---

[23]   The letters HC indicate that the example is from the corpus *El Habla Culta de la Ciudad de Buenos Aires.* The roman numeral indicates the interaction from which the example is taken.

[24]   This suggests that it is in chains with strong pronouns that the accusative clitic may be showing the first signs of grammaticalization; a process that has extended, in the case of the dative clitic to lexical NPs (cf. Chapter 5).

76

(81) a. *Saludamos a ella.
     b. **La** saludamos **a ella**.
       'We greeted her.'

Given the similar conditions applying to *a*-marking and DO-doubling, it comes as no surprise that the two phenomena were associated in the first studies devoted to the subject. The question is, however, if only *a*-marked DOs can be doubled, as Jaeggli's hypothesis implies.

Since *a*-marking and DO-doubling seem to correlate with similar features, one would expect the corpus to show a correlation between those DOs which appear doubled and those marked by *a*. On the other hand, if DO-doubling and *a*-marking are independent of each other (e.g. both epiphenomena of an analogous set of semantic or pragmatic features) one would also expect this correlation not to be categorical, and find at least some doubled DOs which are not *a*-marked. It is this last scenario that we find in the Buenos Aires corpus. Even though most doubled accusatives are *a*-marked (77/119=65%), there is a significant percentage which is not (42/119=35%)[25]. Consider, for instance, the following examples.

(82) a. Ahora tiene que seguir usando**ló el apellido**.   (HC:XVI)
         'Now she has to keep using the last name.'

     b. …**lo** han aprendido **el predicativo**.   (HC:XI)
         '…they have learnt the predicative.'

---

[25] For these percentages I am considering only those DOs realized by nominal phrases. There are 6 DOs realized as clausal complements, which will be addressed below. This accounts for the difference between the 125 DO-doubling instances in the corpus and the 119 used as base here.

c. ¿Pero quién **la** hizo **esa película**? (HC:XXV)
   'But who made that movie?'

d. ¿**Lo** leíste **La Balada del Café**...*?* (HC:XXX)
   'Have you read The Ballad of the (Sad) Cafe?'

All the sentences in (82) involve doubled accusatives which denote inanimate entities. Inanimates are typically *not a*-marked, and therefore the absence of this particle in these examples is not surprising. We mentioned, however, that, in the case of inanimates, *a*-marking depends on the topicality or specificity of the referent. Thus these examples suggest, though weakly, that there is no strong correlation between DO-doubling and specificity / topicality. These examples also clearly show that neither *a*-marking nor animacy are criterial features for DO-doubling in the Buenos Aires dialect, contrary to the claims commonly presented in the literature.

In effect, the hypothesis that, in Spanish, doubled DOs must have some sort of prepositional marking, whose validity for the Buenos Aires dialect was originally questioned by Suñer (1988), is still commonly assumed as a fact in the literature (cf. Lyons (1999a:210 and Belletti 2005, among others). In this regard, the frequency of non *a*-marked doubled DOs present in the Buenos Aires corpus suggests not only that *a*-marking cannot be considered a criterial feature for DO-doubling in this dialect, but also that non *a*-marked doubled DOs cannot be set aside as representing only a handful of marginal cases.

On the other hand, as expected, those doubled DOs which are marked by *a* overwhelmingly involve human referents (73/77=95%). Some examples are presented in (83).

(83) a. El año pasado yo **la** invité **a Mabel**.   (HC:XIV)
          'Last year I invited Mabel.'

      b. Y **la** veíamos **a Sonia** bajar…   (HC:IV)
          'And we saw Sonia coming down…'

      c. Pero si por ejemplo... eh... **la** invitás **a la mucama**…   (HC:XXIII)
          'But if for instance…eh… you invite the maid…'

It is also interesting to note that about 40% of doubled DOs (74/119) involve inanimate entities. To the instances in (82) we add here some additional examples.

(84) a. Me **la** fabriqué yo **la prueba objetiva**.   (HC:XI)
          'I created the objective test myself.'

      b. Un médico amigo que estaba siguiéndo**la su especialidad**.   (HC:XXXII)
          'A doctor friend who was following his specialization.'

      c. Si a mí me **lo** hubieran dicho **eso** hace quince o veinte años…   (HC:XXIV)
          'If I had been told that fifteen or twenty years ago…'

Thus, the evidence from the corpus not only challenges the hypothesis about *a*-marking, but also the proposed correlation between DO-doubling and animacy (Suñer 1988) and human reference (Silva-Corvalán 1981).

4.3. SPECIFICITY

4.3.1. Specificity contrasts

As discussed in Chapter 2, it is often argued that DO-doubling is only possible if the doubled NP can be conceived of as specific. The literature is unclear, however, as to what it means, in this context, for an NP to be specific, or to denote a specific referent.

The notion of specificity addresses what appear to be straightforward interpretive contrasts; but the conceptual distinctions that are behind them have proven difficult to define. Thus, in order to be able to evaluate the specificity hypothesis we need first to distinguish among the criteria that form the basis of most approaches to specificity.

Traditionally, the division between specific and non-specific NPs has been used to distinguish different uses of indefinite NPs. It is often assumed that definite NPs are inherently specific, but indefinite NPs may be specific or non-specific. The various conceptual distinctions at play have received different names: "specific" vs. "non-specific", (Karttunen 1976); "referential" vs. "quantificational" (Fodor and Sag 1982); "strong" vs. "weak" (Milsark 1977); or "partitive" vs. "non-specific" (Enç 1991).

One long-established approach to the various interpretations of indefinites is in terms of scope. Under this view, specific expressions are those that have wide scope over other quantified NPs, operators such as negation, modals, or verbs of propositional attitude. The works that address the contrast in terms of scope interactions present examples like the following (from Karttunen 1976).

(85)   I intend to visit a museum.

    a. There is a certain museum that I intend to visit.
        $(\exists x)$ [museum $(x) \wedge$ intend $((I),$ visit $(I, (x))]$

    b. I intend to do a museum visit.
        intend $(I, (\exists x)$ [museum $(x) \wedge$ visit $(I, x)])$

In (a) *museum* is represented as having wide scope over *intend*, and this represents that the speaker plans to visit a specific museum, say the MoMA. In (b) *museum* has

narrow scope, and thus this representation corresponds to the reading in which there is no specific museum the speaker intends to visit.

Under an approach in terms of scope interactions, specific-indefinites (i.e. those that correspond to readings such as (85)a) are characterized as presenting the referent as existentially presupposed, and as being able to serve as antecedents for a coreferential pronoun or a definite noun phrase later in the text[26].

Whereas scope interactions allow for the formalization of these "specific" uses of indefinites, ambiguity between "specific" and "non-specific" readings may also arise in the absence of operators. Fodor and Sag (1982) present examples such as the one below. As the possible continuations in (a) and (b) reveal, the indefinite in this sentence also permits two interpretations, which the authors term "referential" vs. "quantificational", respectively:

(86) A student in Syntax 1 cheated on the exam.
      a. His name is John.
      b. We are still trying to figure out who it was.

In the reading that corresponds to the continuation in (a), the speaker is asserting something about one particular student, who she does not explicitly identify. Under the second reading, she simply asserts that *students in Syntax 1 who cheated on the exam* is not an empty set. Thus, under this view the two readings diverge in terms of whether the speaker intends to refer to one particular individual or not, but both assert the existence of a denoted entity.

---

[26] For arguments in support of distinguishing between specificity contrasts due to multiple quantifiers vs. contrasts due to intensional contexts, cf. Ioup (1977), Leonetti (1990:103ff).

Yet a third approach to specificity, due to Enç (1991), equates it with partitiveness. Enç presented evidence from Turkish, where the presence or absence of case-marking morphology in the object NP determines aspects of its interpretation. Consider the following examples (from op.cit. 6).

(87)   Odam-a   birkaç çocuk girdi.
         my-room-DAT several child   entered
         'Several children entered my room.'

(88) a. Iki   kiz-**i**   taniyordum.
         two girl-ACC I-knew
         'I knew two girls.'

     b. Iki   kiz taniyordum.
         two girl I-knew
         'I knew two girls.'

(89)   Kiz-lar-dan iki-sin-i   taniyordum.
         girl-P-ABL   two-AGR-ACC I-knew
         'I knew two of the girls.'

Imagine that (87) has been uttered. The speaker may continue with (88)a or (88)b. If she chooses (88)a, the version with the accusative case marker on the noun, the girls have to be interpreted as belonging to the set of the children established in the preceding context. If the speaker utters (88)b, on the contrary, the girls are unambiguously interpreted as external to the set introduced in (87) (which makes this a rather incoherent continuation of the first sentence). Since there is a relation of *inclusion* between the referent in (88)a and the set established in (87), this construction is presented as semantically equivalent to (89), with an overt partitive. It is posited that the only semantic difference between the two structures is that partitiveness is overt in one case and covert

in the other; but both overt and covert partitives serve to encode specific referents. In this context specificity amounts to "linked to previously established discourse referents" (op.cit. 7) and is therefore related to the prototypical contexts of uses of definite NPs. In fact, in formalizing her proposal Enç follows Heim's (1988) approach to definiteness: Heim's "Novelty Condition" posits that all indefinites must introduce referents that were not previously in the domain of discourse. In order to incorporate specific-indefinites into this account, Enç proposes that whereas non-specific indefinites obey the Novelty Condition in a strong sense, specific indefinites obey it in a weaker one: their referents cannot be *identical* to previously established ones, but they must be related to them, typically by inclusion. In turn, specific-indefinites are distinguished from definites in that definiteness is proposed to involve identity of reference, whereas specificity implies that the entity under analysis has to refer to a subset of a previously mentioned one, or to stand in some recoverable relation to it.

The idea that partitives are necessarily specific (Enç 1991:10) has, however, been challenged. Von Heusinger (2002:261ff) notes that the fact that partitives select their referent from a previously introduced set does not necessarily imply that they are always specific. He points out that all this implies is that "partitives are complex expressions that are formed by an indefinite choice from a definite set". Under this view, partitives have two referential aspects: there is a definite aspect, related with the fact that they chose their referent from a discourse-given set. But, within this set, the referent can be specific, non-specific, negative, etc.

A similar perspective is defended in Farkas (1995) and Leonetti (2003). Also similar is Abbott's (1995) proposal, which claims that some of Enç's own examples have (non-partitive) NPs which may receive specific interpretation even if uttered discourse-initially. On the other hand, Abbot also provides examples of NPs which meet Enç's conditions for partitivity and are, nevertheless, interpreted as non-specific.

The three kinds of "specificity" associated with indefinite NPs reviewed so far (i.e. "specificity" as reflexting scope interactions, referential intent, or covert partitiveness) are not equivalent. Certain uses of indefinites can be considered "specific" under one criterion and "non-specific" under another. Following Farkas we will refer to each kind of "specificity" as "scopal" (for that related with the interaction of quantifiers), "epistemic" (for that related with referential intent), and "partitive" (for that related with anaphoricity).

The tradition of limiting specificity contrasts to indefinites arose from the fact that, in many languages (including English), most non-specific interpretations are encoded with indefinites. The consequence is that definiteness and specificity are often assimilated. The assumption that definites are inherently specific, however, has been challenged on a number of grounds (Alcina 1999, von Heusinger 2002). Contrasts similar to those noted for indefinite NPs have been also noted to arise with definite NPs. Donnellan (1966) proposed to distinguish possible interpretations of definites in terms of "referential" and "attributive" interpretations. Consider for instance the following sentences (from Cole 1978).

(90) a. *The best doctor* spares no effort to save a patient.

  b. *The restaurant on Broadway between Grant and Stockton* serves a great dim sum.

Whereas (90)a) is typically understood as referring to any doctor describable as best (i.e. it triggers an "attributive" reading), (90)b) is typically interpreted as referring to one particular restaurant, and not to any restaurant which happens to be at that location (i.e. it triggers a "referential" reading).

These kinds of contrasts lead to the assimilation between attributive uses of definite descriptions and non-specific uses of indefinite NPs, on the one hand, and between referential uses of definite descriptions and specific uses of indefinites, on the other (Partee 1972). The two classes are distinguished in terms of whether the referent is known to the speaker or not. In referential uses of definite descriptions and (epistemic-) specific uses of indefinite NPs, the referent is assumed to be known to the speaker. In attributive uses of definite descriptions and (epistemic-) non-specific indefinites, it is assumed to be unknown.

In fact, Partee proposes that the "referential" / "attributive" contrast applies to indefinite phrases as well. She discusses the example in (91).

(91)   John succeeded in marrying a girl his parents didn't approve of.

The indefinite NP *a girl his parents didn't approve of* may be merely a way for the speaker to refer to John's wife, which would amount to a "referential" use. But the expression may also highlight the relationship between this person and the fact that

John's parents didn't approve of her. In the later case, which is the one that arises more naturally, it is claimed that the interpretation is analogous to that which takes place with attributive uses of definite descriptions.

It has been suggested that the general tendency to interpret indefinite NPs as attributive does not depend on some inherent semantic property of indefinites, but rather on general communicative concerns (Alcina 1999). In using an indefinite NP the speaker encodes her belief that the hearer does not know the object which is referred to. Thus, the speaker cannot freely choose any description which would fit the referent, but only one which is somehow related with the communicative context so as to increase the chances of referential success. This has also been noted by Abbott (1993), who posits that when the referent is not known, the speaker will tend to introduce it not only with an indefinite NP, but furthermore with one which is used attributively. As the example above shows, however, this isn't a necessity but a tendency, and the referential / attributive contrast may be argued to affect indefinite NPs as well as definite ones.

There is a series of formal features which favor specific or non-specific readings, of both definite and indefinite phrases. It has long been noted, for instance, that there is a correlation between descriptive richness and specific understandings. Fodor and Sag (1982) present examples such as the following:

(92) a. Sandy didn't see *a squirrel.*

 b. Sandy didn't see *a squirrel that was chasing its tail around the oak tree.*

(93) a. Everyone hates *a student in the syntax class.*

   b. Everyone hates *a particularly obnoxious student in the syntax class who shouts at the instructor and hogs the discussion.*

The most natural reading of the indefinite NPs in the (b) examples is specific, whereas the non-specific interpretation is more readily available in the descriptively poorer (a) examples. Thus, the richer the description of an entity, the more likely the speaker is referring to a specific one.

In Spanish, some examples involving relative clauses provide an extra mark for determining the specificity of the head noun. In particular, subjunctive verbs in restrictive relative clauses dominated by a non-factive verb trigger the non-specific interpretation of the relevant arguments, even if definite (Rivero 1977)[27], as illustrated in the example below, where *novel* is interpreted as non-specific.

(94)   Carolina leerá la novela que gane el concurso.
        'Carolina will read the novel that wins [subjunctive] the contest.'

In effect, it has been noted that the effect of generating non-specific readings can be generalized to all contexts in which there is a grammatical mark indicating non-factitivity. Among these marks are modals, interrogatives, imperatives, and future and conditional tenses, as illustrated in the following examples[28].

---

[27] For the limits of the correlation between mood and specificity see Leonetti (1990:60ff).
[28] The factors presented in what follows (as well as most of the examples) are taken from Leonetti (1999).

(95) a. ¿Quién me presta un bolígrafo?
          'Who will lend me a pen?'

      b. Me compraré / compraría una gabardina nueva.
          'I will / would buy myself a new jacket.'

      c. Cómprate una gabardina nueva.
          'Buy yourself a new jacket.'

      d. Quizás me compre una gabardina.
          'I may buy myself a new jacket.'

      e. Deben llegar unos agentes de refuerzo.
          'There must arrive some backup agents.'

Non-specific readings are also favored in the context of the negative morpheme, negative quantifiers, verbs of rejection or doubt, the preposition *sin* ('without') and adverbs like *raramente* ('rarely'):

(96) a. El departamento no dispone de un especialista que supervise el experimento.
          'The department does not have a specialist who supervises the experiment.'

      b. Se oponen a una reforma del reglamento.
          'They oppose a reform of the reglament.'

      c. Aprobó sin haber abierto un libro.
          'S/he approved without opening a book.'

      d. Raramente hacemos un viaje en Navidades.
          'We rarely make a trip at Christmas.'

The lack of *a*-marking on animate DOs, is also claimed to unambiguously signal non-specificity (in the sense of non-identifiability):

(97) a. Vimos unas mujeres en la plaza.
'We saw some [non-specific] women on the square.'

b. Vimos *a* unas mujeres en la plaza.
'We saw some [specific] women on the square'

Further, the absence of *a* is categorical with bare nouns in object position (e.g. *Vi mujeres* 'I saw women'), and this is correlated with the fact that bare nouns in object position cannot receive a specific interpretation (Leonetti 2003:71). In turn, when there is more than one quantifier (i.e. when scope differences may emerge), *a*-marked DOs are ambiguous between a specific and a non-specific interpretation; but this ambiguity disappears in the absence of *a*, in which case the DO is normally interpreted as having narrow scope with respect to the subject (i.e. as triggering a non-specific reading):

(98) a. Todos los encuestados vieron *a* una persona sospechosa.   [+/- specific]

b. Todos los encuestados vieron Ø una persona sospechosa.   [- specific][29]
'All the interviewees saw a suspicious person.'

In terms of the contexts which favor specific readings, Leonetti includes the presence of explicative relative clauses and appositions, which are normally incompatible with non-specific antecedents, as illustrated below.

(99)   a.*No encontró ni un taxi, que estaba libre.
'S/he didn't find a taxi, which was free.'

---

[29]   Leonetti argues that, in this case, more than non-specificity in its typical sense what arises is a distributive reading, in which there have been several seeing events, and several suspicious persons who were seen.

b. #No encontró ni un taxi, un medio de transporte insólito para él.
    'He didn't find a taxi, an uncommon transportation for him.'

Specific readings also arise when the modifier comes before the noun or when the noun is modified by a superlative. Thus, it is claimed that whereas (100)a is ambiguous, in (100)b and (100)c the indefinite NP can only be interpreted as specific.

(100) a. A las siete siempre ponían *un programa interesante* en la tele.
          'At seven they always put an interesting program on TV.'

      b. A las siete siempre ponían *un interesante programa* en la tele.
          'At seven they always put an interesting program on TV.'

      c. A las siete siempre ponían *un programa interesantísimo* en la tele.
          'At seven they always put a very interesting program on TV.'

Interrogative words also present some interesting contrasts. The interrogative *cuál* 'which one' is traditionally considered as "presuppositional", in the sense that its referent has to belong to a previously established set. The interrogatives *qué* 'what' and *cuántos* 'how many' do not carry this entailment. Accordingly, *cuál* cannot occur in existential contexts with *haber* (101)a), while the other interrogatives can (101)b). Thus, *cuál* is often associated to specific interpretations, in the *partitive* sense. (Recall Suñer examples of wh-extraction of doubled DOs).

(101) a. *¿Cuáles de los documentos había en la caja fuerte?
          'Which of the documents there was on the safe?'

      b. ¿Qué / Cuántos documentos había en la caja fuerte?
          'What / How many documents there was on the safe?'

The contexts summarized above either favor or determine specific or non-specific interpretations of the relevant NPs. It is disputed, however, whether Spanish has particular grammatical devices for encoding specificity. Leonetti (2002:99) argues that it does not, and that, on the contrary, "the devices that grammars usually describe (*a* + direct object, mood in the relative clause, adjective position, syntactic position in 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".

For the purposes of this study, we do not need to decide on one particular understanding of the notion of specificity, nor even settle on whether it should be interpreted as a semantic or pragmatic phenomenon. I would like to draw, however, two conclusions from the previous review: first, that we need to distinguish between specificity and definiteness, given that, as we saw, at least certain aspects of the specificity contrasts apply to definite as well as indefinite NPs. Second, that we need to keep in mind that, even within the restricted domain of indefinite NPs, there are at least three understandings of the notion of specificity (scopal, epistemic, partitive) which have been argued to target related but ultimately independent phenomena. We showed above that partitive specificity does not entail epistemic specificity. Ioup (1977) also argued for the distinction betweeen "espistemic" and "scopal" specificity: NPs may be interpreted as having wide scope and yet not entail that the speaker has one specific entity in mind. In Russian, for instance, whereas wide vs. narrow scope can be lexically encoded, the

epistemic contrast cannot. Nouns marked by *–to* strongly tend to take wide scope over other quantifiers, but *-to* cannot be used in contexts in which epistemic specificity is unambiguous. In other words, *-to* marks entities with wide scope which are *not* specific in the epistemic sense. The examples are from Kagan (2006).

(102) a. Dima ne   zametil kakogo-to studenta.
         Dima NEG noticed some   student
         'There is a student that Dima didn't notice (but I don't know which one)'

     b. #Ja xorošo znaju kakogo-to šveda.
         I   well   know some   Swede
         'I know some Swede well.'

At this point we can go back to our main concern and evaluate whether it is indeed the case that doubling of direct objects is only possible if they are specific and, if this is the case, what aspect of specificity is relevant to this phenomenon.

4.3.2. Evaluating the Specificity hypothesis

In light of the previous discussion, in order to evaluate hypotheses about the necessary specificity of doubled DOs we need first to examine what notion of specificity is meant by different authors. The original and most explicit presentation of this hypothesis appears in Suñer (1988), who also works with the Buenos Aires dialect. Because of this, I will evaluate the relevance of specificity as this notion is presented in her work.

In the first definition Suñer offers (1988:397), specific is meant to imply that "the referent(s) of a [+spec] NP can be identified with a particular x in the linguistic context".

In this regard, then, the prediction seems to be that doubled DOs must have a coreferential antecedent within the discourse segment which serves as this *x* that they identify with. We will call this the *anaphoric* hypothesis. This hypothesis predicts that accusative doubling may occur if the doubled referent has been previously introduced in the discourse segment, but should not occur for entities which are newly introduced.

We should note, however, that in some of the examples Suñer presents in the same article, "specificity" is meant epistemically. Consider for instance (26)b, repeated below as (103):

(103)   Diariamente, **la** escuchaba **a una mujer que cantaba tangos**. (Suñer 1988:396)
         'Daily, s/he listened to a woman who sang tangos.'

For this and other examples of the same kind the notion of "specific" seems to be meant to imply that the speaker intends to refer to one particular entity which she has in mind. In this case, to one particular woman who sang tangos, and not any female tango singer. There several reasons why this interpretation seems more plausible than one which coincides with the definition above (the *anaphoric* hypothesis). Among them, the fact that the NP is marked as indefinite (a mark that is normally used to introduce new referents) and that it has rich semantic content (as manifested by the embedded relative clause modifying the head noun). As mentioned above, descriptive richness correlates with specificity, but only in its "epistemic" interpretation. Finding complex descriptions of entities which have been already introduced in the discourse context is quite unexpected; but finding complex descriptions of discourse-new entities is not, given that the speaker may assume that this entity is unknown to the hearer.

93

Under this (epistemic) interpretation, the specificity hypothesis predicts that doubling may occur when the speaker intends to refer to a particular entity. Accordingly, DO-doubling should not occur with, for instance, generic reference.

Likewise, under this hypothesis, doubling should not occur in non-factive contexts, which generally signal that the identity of the referent is not known by the speaker. In fact, it is under this assumption that the ungrammaticality of sentences like (104) is explained.

(104)   (***Lo**) alabarán **al niño que termine primero**.   (Suñer 1988:396)
          'They will praise the boy who finishes [subjunctive] first.'

Besides the "anaphoric" and "epistemic" interpretations of specificity, this notion is also, as we saw in the previous section, sometimes associated with the wide scope reading of certain quantified phrases. Consider the following:

(105)   Todos los electores eligieron a algunos de los candidatos.
          'Every voter elected some of the candidates.'

          a. $\forall x$, x a voter, $\exists y$, y a candidate (x elected y)
          b. $\exists y$, y a candidate, $\forall x$, x a voter (x elected y)

(106)   Todos los electores **los** eligieron **a algunos de los candidatos**.   =(31)
          'Every voter elected some of the candidates.'
          a. $\exists y$, y a candidate, $\forall x$, x a voter (x elected y)

A sentence like (105) has two possible readings; one where *some of the candidates* has narrow scope with regards to the universal quantifier (a), and one in which it has wide scope (b). When the clitic appears (106), on the contrary, only the wide scope

94

interpretation of the object is claimed to be possible, something that is presented as further evidence in support of the necessary specificity of doubled DOs.

Finally, specificity is also associated with partitiveness, for instance when explaining the possibility of doubling constructions in contexts such as the one in (107).

(107)   El médico **los** examinó **a cada uno/ varios/ muchos de ellos/ de los pacientes**.   (Suñer 1988:411)
'The doctor examined each one/ several /many of them/ of the patients.'

In this case it is argued that "lexical partitives carry the feature [+specific], and since the DOs [in (107)] are [+animate], the Matching Principle is observed", and therefore doubling is allowed (1988:411).

With regard to definite NPs, there is the issue of whether all definite phrases (save those occurring in non-factive contexts) are considered as inherently specific (and therefore able to be doubled) or not. In this regard Suñer (1999a:246) appeals to the commonality of doubling in the Buenos Aires variety and notes that

> "In this dialect CL-D of DOs is common, which is not to say that all definite DOs are doubled. The doubled ones are those whose referents have previously been introduced in the linguistic or extra-linguistic discourse; that is, doubling occurs when the referent is presupposed, or [+specific] understood as identifiable in Suñer's (1988) terminology. In other words, a 'new' referent is not doubled even if definite; but if it is mentioned again, it is".

To illustrate the cases in point, Suñer provides the following examples (originally from a historic-fictional novel by de Miguel 1997).

(108) a. –Llámame a Doña María Josefa_j -le ordena. Hace días que quiere ver**la a la cuñada del Restaurador**_j.
'-Get Doña María Josefa for me -she orders to her. She has wanted to see the Restorer's sister in law for days.'

b. Ahora, según escuché, ha hecho ir a un tal Martínez_j [. . .]; yo presente **lo** tengo **al tal Martínez**_j [...]
'Now, according to what I've heard, he has called somebody named Martínez [...]; I have that Martínez very present in my mind [...]'

c. […] ¿qué puede extrañar que una noche hasta Sarmiento_j bailara una contradanza? 'Véan**lo al viejo**_j, bailando', dijo Urquiza muerto de risa.
'[…] why should it be strange that one night even Sarmiento danced a contradance? 'Look at the old guy, dancing', said Urquiza dying of laughter.'

It is unclear whether Suñer implies that, in the Buenos Aires dialect, all definite phrases encoding a previously introduced referent must be doubled. It is clearer, however, that within the set of definite NPs doubling is predicted to be allowed only for those which are coreferential with another phrase previously introduced (even if encoding the referent under a different descriptive guise, as in (a) and (c)).

Considering the different ways in which "specificity" is understood in the literature on clitic doubling, we can, for the purposes of evaluating this hypothesis against the corpus, subsume the notion into two basic understandings: the *epistemic* sense, and the *anaphoric* sense. In the *epistemic* sense, a phrase is considered specific if the speaker intends to refer to one particular referent. Of course, texts generally offer only indirect evidence about speakers' intentions, and there are often clearer formal marks if the speaker does *not* intend to refer to one particular entity. I therefore consider a phrase as (epistemically) non-specific if, because of the sentential or discourse context, it is unambiguously clear that the speaker is *not* using it to refer to one particular entity.

96

Applying this criterion, 18% (22/119)[30] of the cases of doubling in the Buenos Aires corpus involve non-specific entities. Some examples involve what can be considered "generic" reference, as the ones below.

(109) a. [nuestro voseo y eso] **los** divierte mucho **a los peruanos**.   (HC:XXIV)
          '[our 'voseo' and all that] amuses Peruvians a lot.'

    b. Hasta en el exterior uno inmediatamente, casi a veces esté... viéndolo de
          lejos, **lo** ve **al porteño**.   (HC:I)
          'Even abroad you immediately, almost sometimes eh… watching him from far
          away, you spot the *porteño.*'

    c. la novela produce una comodidad, **lo** deja vagar un poco **al individuo**.
          'novels produce a certain comfort, they let the person wander a little.'(HC:VII)

    d. Inf: -Ahora la mujer casada no puede sacarse el apellido, aun cuando esté
          divorciada y aun cuando esté unida por--- Méjico o--- por el Uruguay--- en otro
          matrimonio.
          Enc. -¿Tiene que seguir usándolo?
          Inf. -Ahora tiene que seguir usando**ló el apellido**.   (HC:XVI)

          'Inf: Now the married woman cannot take off [her husband's] last name, even if
          she is divorced and even if she is united in--- Mexico or--- or Uruguay in
          another marriage.
          Enc: -She has to keep using it?
          Inf: -Now she has to keep using **the last name**.'

The speakers refer to the class of Peruvians (a), the class of porteños (b), the class of people who read novels (c), and the class of what we may term "husband's last names" (d). In all of these examples the reference is non-specific, in the sense that the speaker does not intend to refer to one particular entity from the class she denotes. The last example, on the other hand, deserves a little more attention. It is clear from the context

---

[30] Again, this number excludes doubled DOs realized as clausal complements, which will be considered below.

that the use of the doubled definite NP *el apellido* is non-specific inasmuch as the speaker is using it to refer a situation applicable to any married woman, who has to keep using whatever last name their current (or former) husbands have. We should note, though, that there is a prior mention to the *class* in the first clause, and further that it is with respect to this first mention that the pronoun used by the *Encuestador* in the second turn should be linked. The doubling construction in the last turn, then, is preceded by other two mentions of the same "discourse referent" (Karttunen 1976). Can cases like this be accommodated by the specificity hypothesis or not? The answer depends on the sense in which the notion of specificity is used. Interpreted as the result of anaphoric linking in the domain of discourse, then the object in (109)d) may be considered "specific" in this particular sense. It is more difficult to argue that one can deem it specific under the "epistemic" interpretation, and therefore I am considering examples as this one as belying this particular interpretation of the specificity hypothesis.

There are also instances of doubling where the (epistemic) specificity hypothesis is challenged based on necessary non-factive interpretation of the sentence. Consider the following examples.

(110) a. **lo** podría hacer muy bohemio **el viaje**.   (HC:I)
      'I could make the trip in a bohemian style.'

   b. Yo **lo** mato **al que me dice eso**.   (XXIV)
      'I would kill the one who tells me that.'

   c. No **lo** he llegado a tener nunca **un diálogo inteligente**.   (HC:VII)
      'I could never have an intelligent dialogue.'

In these cases the referent is interpreted as non-specific by virtue of the opacity of the context in which it occurs. This context has the effect of cancelling the presupposition of existence that would otherwise be attached to the definite NP in the (a) and (b) examples, and blocking the potential specific interpretation of the indefinite NP in the (c) example.

Thus, there are many examples in the Buenos Aires corpus which show that specificity, in its epistemic interpretation, is not *necessary* for DO-doubling and that, in fact, DO-doubling constructions *often* involve non-(epistemically) specific referents.

But we have yet to examine whether the relevant criterion for DO-doubling is "specificity" in its *anaphoric* interpretation. As we mentioned, this criterion predicts that DO-doubling constructions should not occur unless the doubled referent has already been introduced in the discourse context. We should note that this hypothesis does not make any predictions regarding *where* in the preceding context the referent should have been mentioned. It shouldn't matter whether this happened in the preceding clause or fifty clauses prior. Thus, to evaluate this hypothesis, potential antecedents were looked for as far back as the beginning of each interaction.

Still, in the Buenos Aires corpus there is a considerable percentage of doubling constructions involving entities that had *not* been mentioned in the preceding context. In fact, this is the case for most of the DO-doubling constructions in the corpus: 57% (68/119). Some examples are presented below.

(111) a. ¿**La** vieron **a Niní Marshall**?   (HC:XXXII)
          'Did you see Niní Marshall?'

b. Tomémos**lo a Troilo**. (HC:II)
'Let us take Troilo.'

c. Hacía cuarenta años que estaba persiguiéndo**lo al último policía**.   (HC:XXX)
'It was forty years that he was chasing the last policeman.'

Thus, the evidence from the corpus indicates that specificity, in any of its most relevant interpretations, cannot be used as an accurate criterion for explaining the necessary conditions for accusative doubling in Buenos Aires Spanish[31].

4.4. PRESUPPOSITIONALITY

As reviewed in Chapter 2, another prominent account of DO-doubling in Spanish argues that the accusative clitic has the function of forcing a "presuppositional" reading of the doubled NP (Franco 1993, 2000; Franco and Mejías-Bikandi 1997, 1999). As with specificity, the notion of presuppositionality can have (and has had) various interpretations (cf. Karttunen and Peters 1977, Oh and Dinneen 1979, Abott 2000, and references therein). The clearest formulation of the sense in which the notion of presupposition is used in the context of accounting for clitic doubling appears in Franco and Mejías-Bikandi (1999), and is based on the following examples.

(112) a. **Lo** he visto **a un marinero**. =(37)
'I have seen one of the sailors.'

b. He visto a un marinero.
'I have seen a sailor.'

---

[31] For a recent critic of the notion of specificity as it has been applied to the analysis of clitic doubling see Leonetti (2007).

(113) a.   Todos los técnicos **lo** eligieron **a un jugador**.
           'Every coach chose one of the players.'

      b.   Todos los técnicos eligieron a un jugador.
           'All the coaches chose a player.'

(114) a.   ¿**A quién lo** han seleccionado?
           'Who among them did they select?'

      b.   ¿A quién han seleccionado?
           'Who have they selected?'

Franco and Mejías-Bikandi argue that the noun *un marinero* in (112)a) "makes reference to a previously introduced set of sailors. That is, it is interpreted as presuppositional. The objects in [the rest of the (a) examples] receive a similar interpretation: They presuppose a previously introduced set of entities. On the other hand, the object in [(112)b), as in the rest of the (b) examples] may be interpreted either as presuppositional or as existential non-presuppositional. That is, it may make reference to a previously introduced set of entities or it may introduce a new entity in the discourse" (op.cit. 108). Note that they hint at this interpretation in the glosses they provide, where the DO-doubled versions are translated into English with a partitive construction. In effect, under this interpretation, the notion of "presuppositional" in fact coincides with "partitive specificity" as defined above. The common contention is that doubled DOs must be referentially linked to a discourse participant previously introduced, whereas non doubled DOs may introduce new ones. In this sense, then, the presuppositionality hypothesis fails as much as the "anaphoric" interpretation of the specificity hypothesis in

accounting for all the cases in the Buenos Aires corpus where doubling occurs in reference to discourse-new entities[32].

In all fairness, in evaluating this hypothesis we should widen our criteria for potential discourse antecedents of doubled indefinite phrases, since this proposal predicts that the antecedent of the doubled indefinite does not need to be referentially co-extensive with it, but it may denote *the set* of which the doubled referent is part. However, this alternative also fails to account for the empirical data. There is only 1 case (0.8%) in which an indefinite doubling construction targets a member of a previously mentioned set.

(115)   **lo** trompea **a uno de los capos de la empresa donde él está**.   (HC:XXXI)
             'he beats one of the big shots of the company where he works.'

Note, moreover, that in this case the clitic doubles an *overt* partitive construction, and therefore under the present hypothesis the interpretation of the structure does not change as a consequence of the presence of the clitic. (Recall that overt partitives are considered inherently "presuppositional").

This proposal's reliance on examples involving indefinite noun phrases seems to arise from the traditional association between definite descriptions and existential presupposition, dating back to Strawson (1950). Thus, if the function of the clitic is, as claimed, to disambiguate between "presuppositional" and "non-presuppositional"

---

[32] There is another interpretation of "presupposed" which may be relevant for clitic doubling; namely, the idea that an entity may be part of the *pragmatic presupposition* evoked by an utterance and, hence, topical (Lambrecht 1994). This is *not* the hypothesis defended in the proposals reviewed here. The notion of pragmatic presupposition will be reviewed in §6.1.3.

readings, and definite noun phrases are inherently "presuppositional", there is no function that the clitic can serve while in a chain with a definite noun phrase.

In fact, we mentioned in §2.1.2. that Franco (2000) makes this statement explicit when arguing that the interpretation of sentences such as (116) do not change depending on the presence or absence of the clitic, and that therefore the key to understanding the clitic doubling phenomenon lies on the analysis of the doubling of indefinite phrases.

(116)   (**Lo**) he llamado **a Pedro**.   =(40)
          'I have called Pedro.'

The use of indefinite noun phrases in reference to previously introduced participants is, however, quite unexpected in terms of the organization of natural discourse. Instead, the use a definite noun phrase or, if partitiveness is involved, an overt partitive construction (as, in fact, is the case in the example from the corpus) appears as a simpler alternative.

Be that as it may, the reliance on indefinite noun phrases faces the more troublesome question of the low percentage of indefinites in doubling constructions. The Buenos Aires corpus shows that doubling of indefinite noun phrases occurs in less than 5% of the cases (5/119), whereas doubling of definite noun phrases is actually the norm[33]. This result suggests that if the presence of the clitic does have any effect, disambiguating

---

[33] I considered as (grammatically) definite proper nouns and nominals marked by definite articles (*el, la, los, las* and the neuter *lo*), demonstratives (*esta, aquella*, etc.) and possessives in determiner position (e.g. *nuestro libro*). Following Laca (2006:438) I included as definites the universal quantifiers *todo/a(s), cualquiera, ninguno*. I also considered definite the pronoun *uno* when used in reference to the speaker (cf. §1.4).

among potential interpretations of indefinite noun phrases cannot be considered the primary one.

On the other hand, what this correlation uncovers is that, in fact, grammatical definiteness is more descriptively relevant than any of the features reviewed so far. The next section, therefore, analyzes the kind of conceptual distinctions that underlie the use of definite phrases.

4.5. DEFINITENESS

The literature on definiteness traditionally recognizes two basic sources which motivate the use of definite noun phrases. One the one hand, there is the speaker's assumption about the familiarity of the referent for the addressee, or the assumed capacity of the addressee to identify it. On the other hand, there is the speaker's intention to refer to an entity which is conceived of as unique, or as including all the individual members which satisfy the description. These two sources need not coincide.

The intuition behind the appeal to the notion of familiarity is that definite noun phrases are used when their referent is known to the addressee, if nothing else because that particular entity was referred to in prior discourse, as when someone mentions *a movie* at the beginning of a narrative and anaphorically refers to it by saying *the movie* latter on. Following Lyons (1999b) we can contrast these "anaphoric" uses with "situational" ones. Situational uses arise when the situational context narrows down the search domain for the referent in question. Thus, if I were to refer to *the last sentence*, the

reader will most likely interpret that I mean the last sentence in the context of what she is reading and not, say, the last sentence to be uttered in the history of mankind.

Understood broadly, the notion of familiarity may also be invoked to account for some uses of definites in reference to entities that are not necessarily present in either the discourse or the situational context. The typical case involves concepts *stereotypically associated* with some other mentioned entity. Thus, if a speaker mentions *a job*, then she may go on to refer to *the boss, the colleagues* or *the salary*, marking them as definite. In this case it is the association between these entities and jobs that has rendered them "familiar" in this particular context, thus justifying the use of the definite article even for antities that are, strictly speaking, new to the discourse[34].

But whereas the notion of familiarity can be broadened to make it account for anaphoric, situational and associative uses of definite phrases, there are cases in which familiarity is clearly not a good criterion. If someone says *I'd like to speak with <u>the person who gave that order</u>, whoever she is*, the speaker may use a definite description without assuming that the addressee is familiar with the referent (or able to identify it). On the contrary, the relevant assumption seems to be that there is a *unique* individual who meets the description. The same criterion explains the use of definite phrases when referring to entities as *the North Pole*, or *the global economy*, assumed to be unique. It can also account for the incompatibility of indefinites with superlatives, which likewise evoke unique denotata (**a cleverest person*), and for those cases in which definite phrases, usually with a rich descriptive content, are used for introducing entities unlinked

---

[34] These uses, also called "inferrables" (Prince 1981), "bridging inferences" (Clark and Haviland 1977), or "associated anaphors" (Hawkins 1978), will be central to our discussion below.

to anything mentioned in the previous discourse (Farkas 2002:220). In fact, the richer the descriptive content of the definite phrase, the more likely it is used to introduce information which is new to discourse (Leonetti 1999:792), since the richer the descriptive content of a phrase, the less likely this phrase denotes a referent assumed to be familiar to the addressee. Uniqueness seems to be a good candidate for explaining the use of definite phrases in these cases.

But what happens with the uniqueness approach when definite articles are used with plural and mass nouns, as in *the linguists*, or *the water*, where there is no unique entity referred to? Here the alternative has been to re-conceptualize the notion of uniqueness in terms of "inclusiveness" (Hawkins 1978). Under this view, a definite phrase implies that reference is intended to encompass all the entities or all the mass satisfying the description. When there is one single individual meeting the description, *uniqueness* naturally follows from this inclusive approach, but it is not a necessary condition. Now, whereas in the clearest instances, the implication of unique (or "inclusive") referents is independent of the context of the interaction, in other cases it has to be relativized to particular communicative contexts. If someone is leaving a house and asks the host to open *the door*, the assumption will be that she refers to the door that leads to the street, despite the fact that there could potentially be other closed doors in the perceptual environment.

Other instances where the use of definite noun phrases resists a straightforward interpretation in terms of uniqueness generally involve locations of the type present in sentences like *We came to the bank of the river*, or *Billy scribbled on the living room*

wall, where there is clearly no implication of a unique bank or a unique wall (Christophersen 1939, Du Bois 1980, Birner and Ward 1994). Yet another case, especially relevant for Spanish, is the use of definite noun phrases instead of possessive constructions when referring to someone's body parts. Thus, in Spanish one can say *Carolina metió la mano en la cartera* (lit. *Carolina put <u>the hand</u> in the purse*), without carrying the implication that there is a unique referent. In fact, in these cases the use of an indefinite may trigger a reading in the object is dissociated from its possessor.

Thus, both familiarity and uniqueness may leave remnant instances they cannot explain. For these cases, one solution implies resorting to the notion of "accommodation" (Lewis 1979) and assuming that if there is some presupposition that is required for the correct interpretation of the definite phrase, and this presupposition didn't exist before that phrase was uttered, then it will be automatically created by the hearer[35]. A problem with the notion of accommodation is that, given its potency, it may render a theory of definiteness virtually vacuous (Abbott 2005, 2006). And the question remains as to what kinds of presuppositions need to come into existence when a definite phrase is used. Imagine a scenario where two students meet at some university campus. One is a senior, the other is a freshman and has just arrived. The senior wants to schedule an appointment with the freshman, and says *I'll meet you at <u>the Student Union</u>*. The use of the definite phrase can be considered felicitous even though the freshman does not know the Student Union or its location. One could claim that this is so because the freshman accommodates the presupposition that the definite phrase carries. But it is less clear whether the

---

[35] For instance, this is how "associative" uses are accounted for in Heim (1988): "associative uses" of definites satisfy her "Familiarity Condition" only indirectly through accommodation.

presupposition is that the *Student Union* exists, that there is only one, that she will be able to identify it or, most probably, a combination thereof.

The variety of factors that may be at play in the use of definite descriptions makes a notional theory of definiteness hard to characterize, but it may explain why, in fact, the use of definite phrases to refer to discourse-new entities is less exceptional than expected. Based on corpora of Swedish narrative texts, Fraurud (1996:68) shows that 85% of definite noun phrases in natural discourse introduce new entities. Gundel *et al.* (2001) show that "close to half of all definite article phrases in naturally occurring English texts, both spoken and written, have non-familiar [i.e. not previously mentioned] referents". Likewise, arguing for an elaborated version of the uniqueness approach, Leonetti (2000:3) notes that "first mention uses are very common in definite DPs, given that the article does not require a familiar or activated referent, and that the descriptive content is enough to satisfy uniqueness". Birner and Ward (1994) make similar claims, arriving at the conclusion that no single factor can account for all the felicitous cases of definite reference.

The existence of diverging semantic / pragmatic criteria at play has lead some authors to argue that definiteness should be considered as a *grammatical* category, serving as the formal (but imperfect) correlate of *identifiability* (cf. Bolinger 1977, Lyons 1999a, Lambrecht 1994:87). The argument is that the easiness with which an entity can be identified may be grammaticalized in some languages via its association with definite markers. As in other domains, nonetheless, the match between the grammatical category and the concept it represents does not need to be exact.

The idea that the definite nouns phrases cue the hearer to the identifiability of the intended referent have been also proposed by Gundel *et al.* (1993), Ariel (1988) and others. It is plausible, then, that accusative doubling in Spanish correlates with definite phrases because doubling constructions encode identifiable referents. This is the hypothesis examined next.

4.6. IDENTIFIABILITY

In the last section we established that DO-doubling overwhelmly involves definite nouns phrases (114/119=95.8%), and that definite noun phrases typically, although not necessarily, encode identifiable referents. In the core section of the analysis, presented here, I will argue that an approach to clitic doubling in terms of identifiability can open a new way for explaining the function of clitic doubling constructions and, in particular, that it fares better than the specificity / presuppositionality hypotheses reviewed above when accounting for the empirical data.

The focus on identifiability over (epistemic) specificity or presuppositionality is not a mere terminological shift. There are several respects in which identifiability crucially differs from specificity and presuppositionality. Firstly, specificity is basically a speaker-dependent notion, whereas identifiability is basically a hearer-dependent one. In terms of the specific / non-specific distinction what is at stake is whether the speaker intends to refer to one particular entity, or has one particular entity in mind. On the contrary, the notion of identifiability is focused not on the speaker's knowledge but on the speaker's assumptions about the knowledge of her interlocutor.

A basic hypothesis behind this approach is that speakers structure their messages in such a way as to minimize the cognitive effort the hearer has to make in order to interpret it. In other words, that "information packaging in natural language reflects the sender's hypotheses about the receiver's assumptions and beliefs and strategies" (Prince 1981:224), and that the speaker "tries, to the best of his ability, to make the structure of his utterance congruent with his knowledge of the listener's mental world" (Clark and Haviland 1977:5).

It is in accordance with this premise that identifiability is characterized, basically in terms of the speaker's assessment of whether the hearer has already stored in her mind a certain representation of the target concept (Lambrecht 1994:76). From the perspective of the addressee, in turn, identifiability concerns his ability "to establish a link between the referring expression used by a speaker and the concept it refers to" (Küntay 2002:78). Elaborating on these definitions, I will consider "identifiability" as concerning the assumption, on the part of the speaker, that the hearer can activate, with a minimum of processing effort, an appropriate mental representation of a certain entity at the time the phrase encoding that entity is uttered.

The distinction between speaker-dependent and hearer-dependent factors is important because hypotheses based on each criterion make different predictions about potential contexts for DO-doubling. For instance, if these constructions depended exclusively on the speaker-dependent notion of specificity, we would expect to have at least some cases of doubling involving entities which are known to the speaker but not assumed to be known by the addressee (i.e. which are *specific* but not *identifiable*).

Recall from our previous discussion that, in these contexts, we would expect to find relatively long, descriptively rich, formally complex phrases encoding the specific knowledge the speaker has about the referent in question, used for helping the addressee build an adequate representation of it. Consider the pair in (117).

(117) a.   I have just read *the most recent play by Clare Quilty.*
      b.   I have just read *the play.*

In both cases we can assume that the referent encoded as direct object is specific, in the sense that it is *identifiable for the speaker*. It has been noted in many discourse studies that the form of the referential expression tends further to signal whether the referent is assumed to be identifiable for the hearer as well, since "[i]nformation that has been previously evoked can frequently be identified on the basis of a relatively short phrase, with the limiting case being a pronoun or null argument for highly salient information; brand-new information, correspondingly, requires a sufficiently long or complex linguistic realization to enable the hearer to construct an appropriate discourse referent". (Ward and Birner 2004:158). Thus, the form of the referential expressions used in the previous example will typically signal that in (a) the referent is considered specific but not identifiable, whereas it is specific *and* identifiable in (b).

Going back to the analysis of the Buenos Aires corpus, we can track the different formal structures of the noun phrases that occur in a doubling construction and use this as a heuristics for its assumed identifiability. In accordance with our predictions, the analysis of the doubled examples in the corpus show that, in terms of the formal complexity of the phrases serving as direct objects, DO-doubling occurs with phrases

closer to (117)b) than to (117)a). The following table organizes the doubled nominals in terms of their formal structure.

Doubled-DO Structure	#	%
Definite Determ. + Common Noun	53	42.4%
Proper Noun	36	28.8%
Definite Pronominal	16	12.8%
Clause	6	4.8%
Definite Determiner + Noun + Adjective	4	3.2%
Indefinite Determiner + Noun	3	2.4%
Indefinite Determiner + Noun + Adjective	1	0.8%
Partitive + Prepositional Phrase + Clause	1	0.8%
Definite Determiner + Prepositional Phrase	1	0.8%
Definite Determiner + Noun + Prepositional Phrase	1	0.8%
Definite Determiner + Noun + Adjective + Clause	1	0.8%
Definite Determiner + Pronoun + Clause	1	0.8%
Definite Pronoun + Clause	1	0.8%
Total	**125**	**100%**

Table 5. *Correlation of DO-doubling and NP structure in the BA corpus*

From the table we see that 84% of the doubled DOs are realized by the simplest nominal structures of the language, consisting either of a bare nominal (for proper nouns), a nominal marked by a definite determiner (for common nouns), or some other form functioning as a definite pronominal. In this last class I am including two instances of *uno* in reference to the speaker (118)a); seven instances of the quantifier *todo/a(s)* 'all'

used pronominally (118)b), and another seven instances of the demonstrative pronouns *esto, eso* 'this, that' (118)c)[36].

(118) a. Cada uno ha puesto su parte en… en hacer**lo a uno** como uno es.   (HC:VI)
     'Each one has played their part in… in making one the way one is'.

    b. Pero… pero tienen que tener**las todas**.   (HC:XXVI)
     'But… but they have to have all (of them)'.

    c. **Lo** digo **esto** con el más profundo dolor de mi corazón…   (HC:I)
     'I say this with the deepest sadness in my heart…'

Note that not only doubled DOs are most frequently realized by the simplest forms but also, as mentioned above, the analysis shows that doubling of indefinites is extremely rare. We will come back to these data below. For now, I want to call attention to the few cases in which the direct object is realized by a complement clause, the fourth more common structure (6/125=4.8%), as in the examples below.

(119) a. A: -Bueno, tenés que saber--- qué--- ocurre con... con el juego, y que juegues
     con tu compañero.
     B: -Ya **lo** sé yo **que juego con mi compañero**.   (HC:XXV)

     'A: -Well, you have to know--- what--- is going on with… with the game, and
     that you play with your partner.
     B: -I already know **that I play with my partner**.'

    b. …mi hermano mayor, que después se recibió de médico esté...[…] desde
     chico, siempre que se quería disfrazar, siempre agarraba una valijita chica que
     nosotros... sacaba una cartera así, de esas valijitas redondas que usaban los
     médicos antiguamente, habrás visto algún dibujo vos, y él se ponía una galera

---

[36] Note, incidentally, that Suñer (1988: 408) specifically claims that *eso, esto* are incompatible with DO-doubling, whereas in the corpus these cases are relatively frequent, as compared with other forms (7/119=5.9%).

de papá y decía: "Soy el doctor." Desde chico me acuerdo haber**lo** visto siempre **que quería ser médico él**.   (HC:XXVII)

'…mi older brother, that then became a doctor eh… […] since he was little, every time he wanted to disguise himself, he would take a little case that we… he would take a purse like this, one of those round little cases that doctors used to carry, you must have seen a picture, and he would wear daddy's top hat and he would say: 'I am the doctor'. Since he was a boy I remember having always noticed **that he wanted to be a doctor**.'

The fact that these instances, although rare, do occur (actually with more frequency than the doubling of indefinite nominals) is another factor we can use for preferring identifiability over specificity. Specificity is, at least in the literature on clitic doubling, a property of nominal referring expressions, and thus the propositions encoded in the clausal complements in (119) cannot be evaluated in terms of it. On the other hand, it has been argued that the notion of identifiability can be applied to propositions as well as to referring expressions (Dryer 1996, Lambrecht 1994). Note that in both cases the clausal complements are linked to something that has just been said. In the (a) example the doubled clause is a virtual repetition of the previous turn (changing the pronominal features on the verb and the possessive to reflect the shift in point of view). In the (b) example the doubled complement clause synthesizes the anecdote the speaker has just told without adding any new information to it. The fact that the speaker knew that her brother wanted to be a doctor (the information contained in the doubled complement), is easily inferred from the information that she had repeatedly seen him playing a doctor when they were little. By being told this childhood anecdote, the hearers have also come to know the content of the doubled clause. Thus, what is identifiable in these examples is

not just a referent, but a proposition (*I play with my partner, he wanted to be a doctor)*, which reformulates something in the immediate linguistic context[37].

An approach in terms of identifiability also provides a natural explanation for the scarcity of doubling with indefinite noun phrases, something that would remain mysterious under the specificity / presuppositionality hypotheses. Recall that these hypotheses suggest that as long as an indefinite noun phrase is specific or presupposed, there are no restraints for its occurrence in a doubling construction. On the contrary, I believe that indefinite noun phrases are extremely rare in doubling constructions because indefinites typically encode non-identifiable referents, even if these referents are known for the speaker (i.e. specific). Below we will address what the source of the assumed identifiability of certain indefinites is.

The distinction between specificity and identifiability also allows us to account for doubling examples which involve generic reference, as the ones presented in (109) above. Although there is some room for debate in the literature on whether expressions with generic interpretation should be considered referential or non-referential, there is nevertheless agreement that generics cannot be interpreted as specific in any relevant sense of the term (Jackendoff 1972, Givón 1978, Leonetti 1990, among others). On the other hand, there is nothing that prevents generics from being considered *identifiable*;

---

[37] Identifiable propositions are *pragmatically presupposed*, if we distinguish between the logical/semantic and the pragmatic understanding of this notion (cf. §6.1.3). Here I am extending the reach of *identifiability* in order to avoid introducing a yet new sense in which the notion of presupposition is used in this study. Recall that for studies of clitic doubling, "presuppositional" is typically used in the sense of partitive specificity.

since all that is required from the hearer is to identify the category denoted by the phrase (Lambrecht 1994:82, 88).

There are also differences between identifiability and the notion of presupposition as it is commonly evoked in logical tradition (i.e. as existential presupposition). In order to consider an item identifiable it is not sufficient (and depending on how "existence" is defined, maybe not even necessary) that the referent is assumed to exist. The relevant criterion for identifiability is whether the interlocutor is believed to have a certain representation of the referent. As noted by Lambrecht (1994:78), "the relevant property of an identifiable referent is not that it is presupposed to exist, but that the speaker assumes that it has a certain representation in the mind of the addressee which can be evoked in a given discourse".

Finally, a crucial difference between identifiability and both presuppositionality and specificity is that identifiability is inherently a gradual category, whereas specificity and presuppositionality are not. It does not make much sense to ask how specific an entity is, or how much it is presupposed to exist. But it does make sense to ask about different degrees of identifiability and, in particular, to wonder what degree of identifiability may be needed for a given grammatical phenomenon to occur. We will examine this issue in detail in §4.6.2.

4.6.1. Frame-linked Doubling

Before looking into degrees of identifiability, we have to address the possible *sources* from which identifiability can be assumed. We have seen that much of the

116

previous proposals on DO-doubling depended on the condition that the doubled participant be coreferential with some referential entity previously mentioned (or that it be a member of a previously mentioned set). These cases are referred to as "anaphoric doubling".

In anaphoric doubling the source of identifiability is straightforward, as we can establish the discourse context as the basis of the hearer's ability to identify the referent that is being talked about. Notice that, as mentioned above, claims based on anaphoric linking do not place any constraints on where in the prior context the antecedent may be located. But this lack of constraint should not be, in principle, a problem for asserting the identifiability of the referent. Once a referent has been introduced in the discourse, its status as identifiable remains for longer than it may be imagined. Chafe (1976:40) makes this point clear with an example of a novel, in which a referent is introduced as *a letter* on page 13 and then not mentioned again until it is reintroduced with the definite phrase *the note* 105 pages later.

A more interesting challenge is posed by all those tokens for which there is no referential antecedent of the doubled element *anywhere* in the interaction. I will refer to them as instances of "non-anaphoric doubling". What is the source of identifiability in these cases? Consider a typical example.

(120)   Y... y cuando se toma el taxi **lo** mira **al taximetrista** y le dice: "¿Usted no pensará (…) que yo salgo de aquí por loca, ¿eh?"   (HC:XXXII)

'And when she takes the taxi she looks at **the taxi driver** and says: "You wouldn't think (…) that I am leaving from here because I'm crazy, eh?"'

When (120) is uttered, the "doubled referent" (*the taxi driver*) has not been mentioned before. If the identifiability hypothesis is correct, where does the identifiability of this participant come from? The answer is intuitively simple: there is a recent mention of a taxi, and we expect taxis to have drivers. We mentioned these kinds of stereotypical associations when discussing definiteness, and we saw that they license the use of definite articles for entities which are, strictly speaking, new to discourse. The *taxi driver* in (120) is a typical example of what Prince (1981: 236) defines as "inferrable", i.e. an entity such that "the speaker assumes the hearer can infer it, via logical –or, more commonly, plausible- reasoning, from discourse entities already Evoked or from other Inferrables". It is by virtue of its stereotypical association with another mentioned entity that *the taxi driver* in this example is treated as identifiable, something that licenses both its definite marking and the co-occurrence of the accusative clitic[38].

Consider now the example in (121).

(121)   ¿Y **la** habrá hecho con él **la tesis**?   (HC:XXI)
         'And would she have done **the thesis** with him?'

In this interaction the three interlocutors are talking about common acquaintances, among them one who has completed her Ph.D. They then refer to a different woman, who they say is about to return to Argentina from Edinburgh. Given the context, the inference is clear that this woman is returning from a *university* in Edinburg, and that she has been

---

[38] In effect, as will become clearer in what follows, what we find is, in terms of Prince's (1994) categories, that DO-doubling may involve "discourse-old" or "discourse-new" entities, as long as they are "hearer-old".

there completing her doctoral studies as well. Then there is a series of turns in which the speakers talk about an article by one particular scholar from that university. It is in this immediate context that one of the interlocutors asks the question in (120), where the third person pronoun refers to the scholar they were talking about, the subject agreement suffix on the verb signals the doctoral student, and the noun phrase mentioning the thesis that she presumably completed appears doubled by the accusative clitic. This case differs form the previous one in that there is no particular referential expression to which *the thesis* can be associated, as in the *taxi → taxi driver* example. In this case the identifiability of the doubled referent is warranted by the topic of the conversation as a whole. In particular, by our knowledge of the different stages involved in the process of completing doctoral studies. It is in this sense that we can say that "completing doctoral studies" functions as the conceptual "script" (Schank and Abelson 1977) that determines our expectation about the involvement of, among other elements, doctoral theses.

Another instance of non-anaphoric doubling is presented in (122).

(122)   …yo la invito a hacer este experimento: tomar todo lo que es sacando Piazzolla, tomar... tomar todo lo que es eh... yo no diría--- nueva ola, diría--- vanguardia en tango; en todos los movimientos musicales siempre nos interesa la vanguardia; es lo que va quedando, por supuesto. Tomémos**lo a Troilo**, tomémos**lo a Berlingieri**, tomémos**lo a Salgán**…   (HC:II)

          …I invite you to do this experiment: take everything except Piazzolla take all that is eh… I wouldn't say--- new wave, I would say--- avant garde in tango; in all the musical movements there is always an interest in the avant garde; is what remains, of course. Let's take **Troilo**, let's take **Berlingeri**, let's take **Salgán**…'

In this case the "doubled referents" are also linked with the topic of the interaction. They are not, however, part of a sequence of events, as in the example above. Instead, the

119

doubled participants are, at the time of the interview, culturally salient members of a category (tango musicians) which is the topic of discussion. This category serves as a conceptual "schema" or "frame" (to use Chafe's and Fillmore's terms, respectively) with which new referents can be linked, and it is this linkage what allows the speaker to treat them as identifiable.

As noted by Lambrecht (1994:90),

"the frame within which a referent becomes identificable can be so broad as to coincide with the speaker/ hearer's natural or social universe, accounting for the identifiability of *the sun* or *the President of the United States*. It can be narrower, as the personal frame within which the referent of *the cleaning lady* or *the car* becomes identifiable. Or it can be the physical environment in which a speech act takes place, making it possible to identify the referents of such noun phrases as *the women over there* or *those ugly pictures*. Finally, the text-internal discourse world itself can be such a cognitive frame [so that a referent] is identifiable to the hearer by virtue of the frame of reference established by the ongoing discourse alone."

Following Lambrecht, we will use here the terms "frame" and "schema" as synonyms, and as general terms encompassing inferentially-triggered identifiability of the sort illustrated in the previous examples[39].

Since these frames or schemata arise from our personal experience as members of a particular community, they are culturally dependent. Thus, for instance, the examples in (120) and (121) differ from the one in (122) in that the last one evokes a frame that is only available for members of a particular culture, linked to Buenos Aires and the world

---

[39] For critical reviews of the relationship between concepts like "frame" and "schema", see Casson (1983) and Tannen and Wallat (1987). For a historical account of the development of these concepts, see Nerlich and Clarck (2000). For potential processing differences between "schema"-linked inferences (of the sort of (120) and (121)) vs. "taxonomy"-linked inferences of the sort of (122) see, e.g., Mandler (1984).

of tango music. The first two examples, on the other hand, are less geographically constrained (although they still depend on certain socio-cultural groups for which doctoral studies, and taxis, are familiar objects). In all of these cases, it is this assumed familiarity for the addressee that licenses doubling.

The examples above are cases of non-anaphoric doubling involving definite noun phrases. But what happens with the few instances in which doubling occurs with an indefinite NP? We saw that the scarcity of these examples can be explained if we abandon the notions of specificity / presuppositionality and replace them with identifiability: Doubled indefinites are infrequent because indefinites are commonly used to introduce non-identifiable referents. However, I will suggest that if we are willing to broaden the concept of identifiability to encompass frame-linked referents, certain indefinites may need to be considered identifiable as well. The following example from the Buenos Aires corpus illustrates this point.

(123)　　Entonces estaban esperando que llegara la novia y m... el novio estaba ya. Y había gente allí afuera, y había un señor alto--- canoso, un señor como de sesenta años. Y de pronto el señor se fue al suelo, se cayó. Y bueno, ahí, estaba solo. Corrieron a... ella fue con el marido que estaba... y otra gente se juntó. Empezaron a ayudar, **lo** llamaron enseguida **a un médico** y todo, pero el señor--- se murió del corazón.　　　　　　　　　　　(HC:XXVII)

　　'So they were waiting for the bride to arrive and … the groom was there already. And there were people out there, and there was a tall man--- gray haired, about sixty years old. And suddenly the man fell down, dropped. And then, there, he was alone. They ran to… she went with her husband who was…and other people gathered. They started to help, they immediately called **a doctor** and all, but the man--- died from a heart attack.'

It is clear that, in this context, identity of the doctor is not known (and it is not even relevant). The mention of a man suffering a heart attack, however, plausibly triggers an expectation about the involvement of a doctor. Note that, contrary to what the presuppositionality hypothesis would predict, the referent of the doubled noun phrase had not been previously mentioned, the doubled phrase does not have a partitive meaning, and its referent cannot be considered to be presupposed in any meaningful interpretation of the term. However, it is frame-linked with the topic of the interaction, rendering doubling possible (although, as we mentioned, highly marked) even in relation with an indefinite noun phrase.

There is another common type of non-anaphoric doubling to which I want to refer in this section. In this type the doubled referent is used to "reify" some propositional content, in some cases with the intention of categorizing it in a particular way. Consider the following example.

(124)   Inf. A. -Pero si dice que cuando se… los ladrones ésos se... ¿cómo se dice? Se
         entrenan para robar, dice que lo hacen con muñecos con cascabeles.
         Inf. B. -Pero eso es de novela.
         Inf. A. -No, yo lo he visto así--- no sé si habrá sido en la televisión, no
         recuerdo.
         Inf. B. -[*irrecuperable*]
         Inf. A. -La policía creo que **lo** contaba **eso**.   (HC:XXVII)

         'Inf. A. -But if he says that when they… those thieves they… how do you say?
         they train to steal, he says that they do it with little puppets with sleighbells.
         Inf B. -But that sounds like from a novel.
         Inf. A. -No, I have seen it--- I don't know if it was on TV, I don't
         remember.
         Inf. B. [*irrecoverable*]
         Inf. A -I think it is the police who would say **that**.'

In this example the doubled element is simply a demonstrative pronoun, *eso* 'that', whose antecedent is the proposition expressed in the first turn, i.e. that *thieves would train using puppets with slighbells*. What the speaker does in the last turn is to create a referential expression out of a proposition. The fact that the speaker is creating a referent out of a proposition makes this process crucially different from the cases where the speaker refers back to a referential expression (the canonical anaphoric doubling cases). It is only by virtue of the speaker actual reification of a proposition or set of propositions that we can conceptualize them as a "thing", a "referential object" that can be accordingly manipulated in subsequent discourse. The next two examples illustrate the same point.

(125)   Inf. A. -Y en este momento---tengo que admitir que hay cosas que me parecen--
que, bueno, habrá sido un instrumento un poco prostituido de la... de... de...de
una--- realidad política que se estaba m... moviendo, pero... pero que hay que
reconocer que hizo cosas que debían hacerse.
Inf. B. -Sí señor.
Inf. A. -Y que fue instrumento de cosas que son inevitables y que tienen...
Inf. C. -Claro, claro.
Inf. A. -¿Te das cuenta? Pero si a mí me **lo** hubieran dicho **eso** hace quince o
veinte años…   (HC:XXIV)

'Inf. A. -And right now--- I have to admit that there are things that I think---
that, well, he could have been a prostituted instrument of the… of… of… of
a--- political reality that was ch… changing, but… but one has to
recognize that he did things that had to be done.
Inf. B. -Absolutely.
Inf. A. -And that he was an instrument for things that are inevitable and have…
Inf. C. -Of course, of course.
Inf. A. -Do you see? But if I had been told **this** fifteen or twenty years ago…'

(126)   Me dice la empleada--- así tratándome un poco de ingenuo, ¿no? "Pero doctor -
me dice- pero--- usted me dice que la... las llame a las patronas, pero usted no
conoce--- y hace veinte años que yo vengo hablando con patronas--- a las
patronas no les interesa nada si la... si las chicas... eh... eh... tienen un peso, si

dejan de tenerlo, dónde van a dormir a la noche, si están enfermas o no, eso es una cosa que para ellas es totalmente indiferente. Es decir, hablar con una patrona para explicarle una situación particular de una empleada es... es como predicar en el desierto." Y yo me tuve que... es decir, reconocí que había estado un poco ingenuo y que--- bueno --a ver que... **lo** dejé así **al asunto**, ¿no? Pero-- realmente --me he llevado una sorpresa en muchos aspectos.   (HC:XXIII)

'The clerk (fem.) tells me--- treating me like I was a little naïve, no? "But doctor she tells me- but--- you say that I should call the lady patrons, but you do not know--- and it's been twenty years that I keep talking to the lady patrons-- the lady patrons do not care at all if the… if the girls… eh… eh… have a nickel, if they don't, where they go to sleep at night, if they are sick or not, that is something that for them is completely irrelevant. I mean, to talk to a lady patron to explain her a particular situation of one employee is… is like preaching in the desert." And I had to… I mean, I recognized that I had been a little naïve and that--- well--- to see that… I left it at that **the matter**, no? But- really-- I had received a surprise in many respects.'

The reification of propositional content implies the creation of a category which generally involves a certain descriptive characterization. In the examples above the availability of a new referential entity in the context of the interaction was hinted at in purely "deictic" terms, by a form devoid of descriptive content. But speakers typically use this mechanism, at the same time a new conceptual entity is created, to characterize it in a certain way; and they do so by assigning it to a certain class. For instance, to the class of "problems", as in the example below.

(127)   Enc. -Entonces eso lo... ¿quién lo compraba? No lo compraban los
           intelectuales, lo compraba la gente... el hombre medio, en fin.
           Inf. -No, claro. Bueno, y...
           Enc. -No el proletariado, pero...
           Inf. -Por eso. No, yo **lo** estudié bastante **ese problema**.   (HC:XXXI)

'Enc. -Then that… who would buy it? It wasn't bought by the
Intellectuals, it was bought by the people… the average man, I mean.
Inf. -No, of course. Well, and…
Enc. -Not the proletarian, but…
Inf. -Exactly. No, I studied **that problem** quite a bit.'

The examples in this section involve cases where there is no referential antecedent for the doubled element. The doubled entity, however, is identifiable because of its relationship with some previously mentioned entity, with which it is conventionally linked, or by virtue its association with the topic of the interaction as a whole.

Identifiability may also be due to a more context-independent anchoring of an entity on assumed shared knowledge. A typical example of this is the introduction into the discourse context of a participant known to all interlocutors, which may be referred to simply by its proper name. There are also some instances like this in the corpus, where the new participant cannot be linked either to a linguistic antecedent or more generally to a schema triggered by some other mentioned entities. All involve a (bare, human) proper name (or the relational name *papi* 'daddy'). Given that there is no sign in the following turns that the referential intent was unsuccessful, and that unmodified proper names tend to encode identifiable referents (Ariel 1990), there is no evidence to doubt that these names refer to identifiable people, and that these instances still conform to the analysis proposed here.

Thus, the analysis of the corpus shows that DO-doubling is sensitive to the identifiable status of the target referent. This identifiability may be due to the prior mention of the same referent in the discourse context (i.e. the anaphoric doubling cases);

but it also may be due to the doubled referent's association with a frame triggered by another mentioned entitity, the topic under discussion or, more generally, the cultural or interpersonal knowledge shared by the interlocutors.

Considering those cases where there is a referential expression in the previous discourse context that targets the same entity denoted by doubled DO (or that refers to a set to which the doubled DO is part), there are 51 cases of anaphoric doubling in the corpus (51/119=43%). The remainder 57% are frame-linked cases of the sorts discussed above. The two groups, clearly, are not mutually exclusive, as the presence of an antecendent for the "doubled referent" does not imply that this referent is not also schematically linked to other elements in the text.

4.6.2. Cognitive Accessibility

I argued above that identifiability is inherently a gradual category. We said that a referent is identifiable if there is a certain mental representation of it that the hearer can access, but different referents may be more or less costly to access, depending on their inherent characteristics or the characteristics of the context where they are evoked. Accessing a referent that was mentioned in the preceding turn should be easier than accessing one that was mentioned many turns back, and accessing a referent that is stereotypically linked to what is being discussed should be easier than accessing a referent that it is not so linked. The degree of "easiness" with which a referent can be accessed represents its "activation state" (Chafe 1987), and the particular activation state a referent may have depends on attentional and short-term memory limitations. In other

words, we can draw a distinction between identifiable and non identifiable referents based on the (assumed) pre-existing representation of a referent in the mind of the interlocutor. Once a referent is deemed identifiable, however, speakers may further compute what their particular activation status is, depending on the easiness with which they believe the hearer will be able to access a cognitive representation of it and link it to the discourse context.

Although the relative activation of any referent forms a non-discrete cognitive continuum, there are divisions of that continuum that have particular grammatical correlates (cf. Prince 1981, Chafe 1987, Lambrecht 1994, Ward and Birner 2004, among others). In the following I will adopt Chafe's (1987, 1994) three way distinction between "active" concepts (i.e. those which are in the interlocutor's focus of consciousness), "semi-active" or "accessible" concepts (i.e. those for which the interlocutors have a peripheral awareness), and "inactive" concepts (i.e. those that reside in the interlocutors' long-term memory, and are neither focally nor peripherally active).

I suggested above that it is actually the distinction between different levels of activation (and not simply the difference between identifiable and non-identifiable referents) that may explain the most typical contexts occurrence of clitic doubling. In this connection, it is generally assumed that a concept may become semi-active (and be treated as accessible information) via two sources: either through deactivation from an earlier active state (due to attentional and short-term memory limitations) or by being activated because of its association with an idea that is or was active in the discourse (i.e. a conceptual schema). For instance Chafe (1987: 29) notes that "[when] a schema has

been evoked in a narrative, some if not all of the expectations of which it is constituted presumable enter the semi-active state. From that point on, they are more readily available to recall than they would have been as inactive concepts"[40].

In fact, for those cases in which the doubled DOs do have a discourse antecedent (anaphoric doubling) it is generally the case that there are several intervening clauses between the antecedent and the doubled NP, or several candidates competing for the reference, both factors that diminish the relative accessibility of the target referent. Consider the following example.

(128)   Mire el otro día sube al colectivo un un porteño, bien bien calibradamente
         porteño. Y el que manejaba el colectivo--- era otro bien calibradamente
         porteño, ésos que manejan de costado--- en ángulo de cuarenta y cinco grados
         con respecto al volante. Éste no sé si le pagó con cien pesos o con quinientos
         y el otro le dio un vuelto... esté... marcadamente en monedas--- cualquier
         cantidad. Entonces éste cuando recibió ese impacto de todas las monedas
         que no se lo esperaba, quedó ahí con la mano todavía en forma de balanza
         como pesándola y mirando ese paquete brutal de monedas. Y **lo** miró fijo y
         provocativamente **al colectivero**, y todo lo que le dijo es: "Mucho, ¿no?". Y el
         colectivero cancheramente lo miró de costado y perdonándole la vida le dijo:
         'Comprate caramelos'. Yo creo que de esto solamente nos podemos reír los
         porteños.   (HC:II)

         'Look, the other day it gets in the bus a 'porteño', a very typical porteño. And
         the guy who drove the bus--- was another very typical porteño, those that drive
         on the side--- in a forty five degree angle with respect to the wheel. And this
         guy   I don't know if he paid him with a hundred pesos or five hundred and the
         other guy gave him a change… eh… notoriously in coins--- any amount of
         them. Then this guy when he received the impact of all those coins, that he
         didn't expected, stayed there with his hand still in the shape of a scale like
         weighing them and staring at this brutal pack of coins. And he looked **the bus
         driver** straight and provocatively, and all he said is: 'A lot, no?' And the bus
         driver teasingly looked at him out of the corner of his eye and 'sparing his life'

---

[40] An analogous claim is made by Fillmore (1982: 111) when he notes, in relation to his notion
      of "frames", that "when one of the things in such a structure is introduced into a text, or into a
      conversation, all of the others are automatically made available".

said: 'Buy some candies.' I think this is something only porteños can laugh about.'

In this segment the speaker is telling an anecdote about a humorous exchange between a bus driver and the person who was buying a ticket from him. The anecdote involves two participants, plus the objects exchanged (the paper money, the coins); and it focuses on different aspects of the situation. Note that the bus driver is introduced in the second clause (as *the guy who drove the bus*), a mention that is followed by a brief description of his physical appearance connoting his general character. Then the attention shifts to the traveler, and the coins he received. By the time the speaker mentions the bus driver again, this participant is no longer the focus of attention. It is in this context that this referent appears encoded in a doubling construction.

Something similar occurs in the following example. Here the speaker is talking about something that happened to a friend of hers. The referent of the doubled NP had been introduced in the second clause, but between the two mentions several other participants intervene: another client, the clerk, Bahía Blanca, Betty and, crucially *another coat*:

(129)   ''…tenemos un problema porque <u>ese abrigo suyo</u> vino una clienta y dijo que le quedaba muy bien y se lo quería para ella. Es una clienta hace mucho nuestra, así que se <u>lo</u> vamos a tener que dar.' 'Ah, no', dice. 'Si quiere le hacemos otro, y después se lo mandamos a Bahía Blanca.' Y Betty le dijo: 'No –dice- yo lo elegí primero. Si ustedes no me **lo** dan **ese abrigo**, yo no compro nada…''
(HC:XXVII)

''…we have a problem because, <u>that coat of yours</u>, a client came and said that it fit her very nicely and that she wanted it for herself. She has been our costumer for a long time, so we'll have to give it to her.' 'Ah, no', she says.

129

'If you want we'll make you another one, and then we'll send it to Bahía Blanca' And Betty said: 'No–she says- I chose it first. If you don't give me **that coat** I won't buy anything…'.'

What these examples suggest is that DO-doubling is not just a matter of *knowledge* (whether the referent is identifiable or not), but also a matter of *attention* (whether the referent is accessible at a particular point in the interaction). What the anaphoric doubling examples illustrate is that, when the doubled participant is reintroduced, this participant is no longer the center of attention, even if it has not become completely deactivated. According to the three way typology proposed by Chafe, in anaphoric doubling cases the participant can also be assumed to be *accessible*, or in this "intermediate" activation status also proposed for referents which are exclusively frame-linked. What this suggests, then, is that DO-doubling in Spanish is used precisely to encode this intermediate level of activation, i.e. to signal *accessible* referents.

In this regard, the three-way categorization of attentional states, which translates into three activation statuses, correlates with three possible encodings of direct objects, either by means of a nominal phrase alone, by means of a clitic alone, or with a doubling construction. Studies on topic continuity (Givón 1983, Bentivoglio 1983, 1993) have shown that new referents tend to be realized by lexical NPs[41], whereas referents which continue as attentional foci are typically encoded by pronominal forms.

---

[41] The particular structure of the NP also tends to reflect the relative effort that establishing reference is assumed to take (cf. e.g. Ariel 1990, 2001, Gundel *et al.* 1993). In connection with this, recall that DO-doubling overwhelmly involves either a bare noun or a bare noun marked with a definite determiner (cf. Table 5).

In light of these patterns, clitic doubling can be interpreted as the formal correlate of an intermediate level of referent accessibility, along a continuum which has weak pronouns (i.e. clitics) and lexical NPs at either end. This can be schematically represented in the table below (based on example (83)a).

EXAMPLE	**La** invité *'I invited her'*	**La** invité **a Mabel** *'I her-invited Mabel'*	Invité **a Mabel** *'I invited Mabel'*
ENCODING FORM	clitic	doubling const.	NP
COGNITIVE STATUS	active	accessible	inactive

Table 6. *DO-doubling and activation levels*

4.7. SUMMARY AND DISCUSSION

In this chapter we reviewed several hypotheses that address the interpretive correlates of Spanish DO-doubling constructions. We saw that criteria based on animacy, specificity, or presuppositionality cannot account for the empirical data, since there are considerable examples in the corpus which involve inanimates, or cannot be considered specific or presupposed in any of the various senses in which these notions are used in the literature.

On the other hand, I proposed that DO-doubling can be accounted for in terms of the identifiability of the referent. More specifically, resorting to Chafe's (1994) three-way distinction of cognitive accessibility levels, and his claim that *accessible* referents acquire this state via de-activation from an earlier active state or association with a semantic

schema, I suggested that accusative doubling in Spanish is used to mark accessible referents and, in this sense, can be analyzed in light of its non-doubled alternatives (i.e. NP-only and clitic-only structures) as an extra point along the accessibility continuum.

The proposal presented here builds on (and is compatible to) the insights of previous works. And even though it does, I believe, account for empirical data more accurately than they do, it can still be explored further and accordingly refined. For anaphoric doubling cases, it would be interesting to determine the relative weight of each of the different factors that affect referents' accessibility (e.g. the number of intervening clauses that mediate between the doubled referent and its antecedent, the number of competing participants and the relevance of thematic boundaries, among others). For non-anaphoric doubling, much more work can be done on the analysis of the different types of inferential processes at play, which I have subsumed here under the notion of conceptual "frame".

For the time being, notwithstanding, I would like to note that the approach advanced here not only is compatible with the general insights behind previous proposals but it also opens the way for potentially reconciling some seemingly contradictory results, as those obtained by Silva-Corvalán (1981) and Colantoni (2002), discussed in Chapter 2 (§2.2.2, 2.2.4). Recall that the former associated doubling with topical chains, and the latter associated doubling with the introduction of new referents. Given that their conclusions were based on some relative predominance of one or the other case, it seems at least possible that a closer look at their data would reveal that the remaining cases conform to the opposite strategy. In other words that, just as we saw for the Buenos Aires

corpus, doubling constructions do occur with topical chains (but specifically when there are intervening clauses which introduce competing referents) *and also* when new participants are referred to for the first time (but only if they are identifiable via some schematic linking with the context of the interaction).

At this point, there is one more hypothesis I would like to advance, which is related to the "reification" of propositional antecedents presented in (124) to (127). I suggested then that one of the consequences of this process was that it allowed the speaker to "categorize" the newly created referent as a member of a particular class, and thus to assign it a particular descriptive content. I would like to suggest here that speakers not only can "categorize" newly created discourse-entities, but can also exploit the meaning associated with doubling constructions to *re-categorize* old ones.

The example I would like to present in support of this "recategorization" hypothesis requires some contextualization. At the relevant point in this interaction, the participants are talking about a possible article one of them would like to write for a political magazine. They discuss the idea of interviewing some company executive, and describing one day in his life, as a veiled form of referring to the current socio-political situation of the country. It is in this context that one of them recalls an Irish movie he has just seen, and starts to retell its plot. The story is about a big executive, who one day goes to visit his father and discovers he has been killed by an English soldier:

(130)   Él entra a la pieza, lo ve, se da cuenta que está muerto después de un rato, y entonces--- se le ocurre mirarle el cuerpo y le encuentra unos golpes. Él empieza a averiguar--- y averigua que <u>un soldado inglés</u>--- por una canción irlandesa anti inglesa que el padre cantó--- en la taberna--- se lo armó la gresca y lo golpeó y lo pateó en el suelo y qué sé yo. Y el padre sufría del corazón y se

murió. Se puso mal--- entonces el tipo decide vengarse. Entonces parece que va a ser--- un retorno--- y efectivamente el hombre con su mujer arma un escándalo en una fiesta que hace la mujer sabiendo que murió el padre. Arma un gran escándalo y qué sé yo, lo... lo... le... lo trompea a uno de los capos de la empresa donde él está, que está... que está en la fiesta. Se va, tiene... se... se enrieda [sic] con una- - - con una mujer madura, pero irlandesa ahí en... en Liverpool donde murió el padre. Parece que vuelve, va a buscarlo <u>al soldadito</u>. Primero se tortura un poco, anda de acá para allá, no sabe muy bien si ... si se debe vengar o no. Reacciona diciendo: "No no tengo que vengarme." Después decide y va, lo busca, prepara un fierro, hace todo [sic] una cosa muy preparada- - - y finalmente- - - **lo** golpea terriblemente **al chico**.   (HC:XXXI)

'He enters the room, sees him, after a while he realizes that he [the father] is dead, and then--- he thinks of checking his body and he finds some bruises. He starts to ask around--- and finds out that <u>an English soldier</u>--- because of an anti-British song his father had sang in the tavern--- he created a fuss and beat him and kicked him on the floor and I don't know what. And the father suffered from his heart, and so he dies. He went mad--- then the guy decides to take revenge. Then it seems that it is going to be--- a payback and in effect the guy makes a scandal with his wife at a party that the wife organizes knowing that his father is dead. He makes a big fuss and I don't know what, he beats one of the bosses of the company where he works, who is… that is at the party. He goes away, he has… he becomes involved with an--- older woman, but an Irish woman there in… in Liverpool where his father had died. It seems that he goes back, he goes to look for <u>the little soldier</u>. First he tortures himself a little, he goes back and forth, he doesn't quite know if… if he must take revenge or not. Then he decides and goes, looks for him, prepares a metal bar, he does a very elaborate thing--- and finally--- he beats **the kid** terribly.'

I want to call attention to the way the speaker refers to one of the characters: it is "an English soldier" (*un soldado Inglés*) when introduced by the first time; "the little soldier" (*el soldadito*) toward the end; to finally becomine "the kid" (*el chico*) in the last clause. Note that in this last mention the NP "el chico" appears doubled by the accusative clitic.

One can speculate that, if the accessibility hypothesis is correct, and the association between clitic doubling and accessible referents has become conventionalized, then the

way is open for speakers to start exploiting the meaning associated with this form for rhetorical ends. To state it simply, I believe that in this particular context clitic doubling allows the speaker to refer to "the English soldier" as "the kid", and still assume that the hearer will assign these two (quite diverging) descriptions to the same referent. In other words, that the hearer will be able to "accommodate" the *presupposition of accessibility* carried by the accusative clitic[42].

---

[42] The same phenomenon seems to be behind the use of a doubling construction in two examples we introduced in (108) (§4.3.2), which I repeat here for convenience.

(i) –Llámame a <u>Doña María Josefa</u> -le ordena. Hace días que quiere ver**la a la cuñada del Restaurador**.
'-Get <u>Doña María Josefa</u> for me -she orders her. She has wanted to see the **Restorer's sister in law** for days.'

(ii) […] ¿qué puede extrañar que una noche hasta <u>Sarmiento</u> bailara una contradanza?
'Véan**lo al viejo**, bailando', dijo Urquiza muerto de risa.
'[…] why should it be strange that one night even <u>Sarmiento</u> danced a   contradance?
'Look at **the old guy**, dancing', said Urquiza dying of laughter.'

Building a referential link between "the Restorer's sister in law" and "Doña María Josefa"; and between "Sarmiento" and "el viejo", would be a much more difficult task if the second NP in each case were not marked by the clitic.

5 DATIVE DOUBLING IN BUENOS AIRES SPANISH

This Chapter is concerned with the conditions that determine the use of dative doubling constructions. After a brief summary of the most relevant literature on the topic, it is argued that dative doubling should be conceived of as the mirror image of accusative doubling. In the case of datives, the unmarked way of encoding them is via clitics alone, which are "doubled" by lexical phrases when there is a deviation from their expected high activation status.

5.1. INITIAL CHARACTERIZATION

In this chapter we will analyze the conditions that govern dative doubling. Analogous to our treatment to accusative doubling, I will consider as "dative doubling constructions" those instances in which a postverbal dative argument co-occurs with a dative clitic, e.g. *le*, *les* or its allomorph *se*, as in the following examples.

(131) a. Carolina **le**   dio   **a**   **Esteban** un libro de cuentos.
        Carolina DAT3S gave.3s DAT Esteban   a   book of short stories
        'Carolina gave Esteban a book of short stories.'

     b. Carolina **les**   dio   **a**   **los chicos** un libro de cuentos.
         Carolina DAT3PL gave.1s DAT the children a   book of short stories
         'Carolina gave the kids a book of short stories.'

     c. Carolina **se**   lo   dio   **a**   **los chicos**.
         Carolina DAT3 ACC3MS gave.3s DAT the children
         'Carolina gave it to the kids.'

The dative clitic inflects for number but not for gender, as illustrated in (131)a) and (131)b). The allomorph *se* occurs when the third person dative and accusative clitics co-occur. This form is invariable for both number and gender (131)c).

On the other hand, lexical dative arguments are always marked by *a*. This particle should be distinguished from the one that marks animate accusatives. In the case of datives it occurs obligatorily, regardless of the animacy of the referent:

(132) a. Carolina vio **Ø** **la película**.
        Carolina saw   the movie.
        'Carolina saw the movie.'

    b. Carolina vio **a** **la directora de la película**.
        Carolina saw ANIM the director of the movie.
        'Carolina saw the movie director.'

    c. Carolina **le** dio mucha importancia *a* **la película**.
        Carolina DAT3S gave much importance DAT the movie
        'Carolina gave a lot of importance to the movie.'

Spanish datives form a very heterogeneous class, in terms both of their morphosyntactic properties and their semantic content. As noted by Company (2006:488), all authors agree that there are, at least, two groups. There is disagreement, however, on the parameters for categorizing each one. Demonte (1994), for instance, proposes a classification based on the semantics of the predicate, distinguishing between datives associated with verbs of transfer (e.g. *dar* 'give') and datives associated with verbs of creation (e.g. *cocinar* 'cook'). Gutiérrez Ordoñez (1999), on the other hand, organizes the data based on whether the predicate is inherently associated with three argument roles or not.

It is frequently suggested that, for certain verbs of the type of *dar*, the dative argument may be encoded exclusively as a nominal, as in (133).

(133)   Carolina **Ø** dio   mucha importancia **a   la película**.
         Carolina   gave much   importance   DAT the movie
         'Carolina gave a lot of importance to the movie.'

With verbs of the second group, a third participant may optionally appear, and it may be encoded either as a prepositionally marked adjunct (134) or as a dative argument (135).

(134)   Carolina cocinó   una paella **para Esteban**.
         Carolina cooked.1S a   paella for Esteban
         'Carolina cooked a paella for Esteban.'

(135) a. Carolina **le**   cocinó   una paella **a   Esteban**.
         Carolina DAT1S cooked.1S a   paella DAT Esteban
         'Carolina cooked Esteban a paella.'

      b. Carolina **le**   cocinó   una paella.
         Carolina DAT1S cooked.1S a   paella
         'Carolina cooked Esteban a paella.'

If the target participant is marked by a preposition, the co-occurrence of the dative clitic is ungrammatical (e.g. **Le cocinó una paella para Estaban*). If it is marked as a dative argument, the co-occurrence of the clitic is considered obligatory (e.g. **Cocinó una paella a Esteban*), and the optional element is the coreferential NP.

Based on this morpho-syntactic distribution we have, then, two groups of verbs which differ in how they allow dative participants to be coded.

	Verbs like *dar* 'give'	**Verbs like *cocinar* 'cook'**
1.	CLITIC ONLY	CLITIC ONLY
2.	CLITIC DOUBLING	CLITIC DOUBLING
3.	NP-ONLY	*NP-ONLY

Table 7. *Dative encoding alternatives for* dar-*type and* cocinar-*type verbs*

The table shows that verbs like *dar* allow their datives to alternate between clitic-exclusive, clitic-doubling, and NP-exclusive encoding. Verbs like *cocinar*, on the other hand, allow their datives to alternate between clitic-exclusive and clitic-doubling encoding, banning the NP-exclusive alternative.

It has been proposed that, in cases like (133) the *a* marking the third participant does not assign dative case but should be analyzed as a preposition, akin to the one in (134) (Demonte 1995, Cuervo 2003; but see Suñer 1988, Campos 1999, and §2.1). On the other hand, it is assumed that in the cases where the NP does not occur, the argument is realized by a phonologically silent *pro*. Together, these two assumptions justify the claim that dative doubling is always obligatory. Here I take a more conservative approach, considering as proper doubling structures only those examples in which both the clitic and the dative phrase are morphosyntactically realized in the clause.

5.2. DATIVE DOUBLING AND INFORMATION STATUS

Given the different encoding alternatives allowed for each verb-type, in analyzing the function of dative doubling we have to compare doubling constructions with clitic-exclusive constructions for one type of verbs (i.e. the *cocinar*-type); and doubling

constructions with both clitic-exclusive and NP-exclusive encoding for verbs of the other type (i.e. the *dar*-type). A problem for this approach is that, as mentioned above, there is no complete agreement about which verbs belong to each group. The verb *hacer* 'make' is ascribed to one class by Demonte (op.cit. 437) and to the other by Gutiérrez Ordoñez (op.cit. 1869, 1872)[43]. *Poner* 'put' is traditionally considered a three-place verb (and even as instantiating a type of transfer); it is, however, traditionally included in the class of *cocinar* 'cook' (Demonte 1994: 437).

All this makes it difficult to decide exactly which verbs are predicted to be able to encode dative arguments in exclusive nominal form, and to devise the corresponding parameters for tracking the relevant instances in a corpus. An additional difficulty comes from the fact that different verbs of transfer included in the *dar*-group exhibit different frequencies of dative encoding of any type. Data from ADESSE[44] show that *dar* 'give' encodes a recipient in 98.8% of the clauses where it occurs, whereas for *ofrecer* 'offer' this percentage falls to 56.5%; and for *vender* 'sell' the encoding of a recipient only occurs in 23.1% of the cases.

Thus, in order to track possible occurrences of datives encoded in exclusive nominal form, I chose 12 of the most typical verbs of transfer, including transfer of knowledge, possession and information (4 of each sub-class). Where a verb returned no

---

[43] Demonte characterizes *hacer* as a predicate of creation, for which the incorporation of a dative argument determines the presence of the dative clitic. Ordoñez characterizes *hacer* like a verb with which datives may occur in exclusively nominal form (i.e. like *dar* 'give').

[44] *Alternancias de Diátesis y Esquemas Sintáctico-Semánticos del Español*. Developed at the Universidad de Vigo. http://webs.uvigo.es/adesse/enlaces.html Accessed June 16, 2007.

instance of dative encoding for the Buenos Aires dialect, I replaced them with the next best alternative. The resulting data are presented in Table 8.

VERB	ENCODING			TOTALS
	CLITIC	CL-DOUBLING	NP	
Decir 'say, tell'	52	7	0	59
Dar 'give'	43	7	1	51
Contar 'tell'	26	2	0	28
Preguntar 'ask'	15	2	0	17
Pedir 'ask for'	7	0	1	8
Enseñar 'teach'	4	0	0	4
Regalar 'give'	3	1	0	4
Pasar 'pass'	2	1	0	3
Entregar 'deliver'	1	1	0	2
Mostrar 'show'	1	1	0	2
Comprar 'buy'	0	0	1	1
Recomendar 'recommend'	0	1	0	1
TOTALS	**154**	**23**	**3**	**180**
PERCENTAGES	**85.6%**	**12.8%**	**1.7%**	**100%**

Table 8. *Percentages of encoding forms for* dar-*type verbs*

From this table we see that the most frequent type of dative encoding occurring with transfer verbs is by means of a clitic alone (154/180=85.6%), there is some incidence of clitic doubling structures (23/180=12.8%), and NP-exclusive encoding is extremely low (3/180=1.7%). These cases are presented in (136). In two of them the dative NP is non-referential ((a) and (b)). In the third example the hesitations that appear suggest that there may be processing factors involved in the non-canonical marking of the dative object.

(136) a.   Hay que dar oportunidades **a todos**.   (HX:XXI)
           'Opportunities have to be given to all.'

b. Nosotros no habíamos querido pedir el auto prestado **a nadie**.   (HC:XXII)
    'We hadn't wanted to borrow the car from anyone.'

c. Ayer a las seis de la tarde salí a comprar ....eh... eh... unas cosas ... eh... eh... u...
    unas cosas **a Alvarito**.   (HC:XXXII)
    'Yesterday at six in the afternoon I went out to buy…eh…eh… some things…
    eh…eh… s… some things for Alvarito.'

Corpus data from Mexican Spanish presented by Maldonado (2002) can be
interpreted in analogous fashion. Analyzing the phenomenon within the framework of
Cognitive Grammar, Maldonado argues that the omission of the clitic correlates with a
weaker conceptual linkage between discourse-participants. He notes that "third-person
plural-impersonal constructions [(137)a), (137)b)] and generic referents such as
institutions, groups, or masses either in subject [(137)c)] or indirect object [(137)d)]
position depict a looser subject-dative linkage. Under these circumstances, the clitic *le*
can be left out" (op. cit. p. 18).

(137) a. Dieron un día extra de asueto **a los trabajadores del Estado**.
          'They gave the State employees an extra free day.'

      b. Autorizaron un aumento del 10% **al precio de la gasolina**.
          'They authorized a 10% increase in the gas price.'

      c. Los ingenieros agrónomos manifestaron su apoyo **al aspirante priísta**.
          'The engineers demonstrated their support to the PRI candidate.'

      d. Corresponderá **a las autoridades** vigilar el caso.
          'It will be the authorities' duty to examine the case.'

What we find, then, is that NP-exclusive encoding seems to be associated with
referents extremely low in individuation. If this is the case, it comes as no surprise that

the number of NP-exclusive tokens in the Buenos Aires corpus is so small[45]. There is

extensive evidence that the dative case-role is associated with topical entities, which have

high degrees of individuation and discourse saliency (Greenberg 1974, Givón 1984,

2001). NP-exclusive encoding, then, seems to occur in cases where the dative participant

deviates maximally from this expectation[46].

We can now turn to the analysis of the alternation between clitic-exclusive and

clitic-doubling encoding. For these two encoding types we do not need to distinguish

between different verb classes. The analysis of the Buenos Aires corpus shows that clitic-

exclusive encoding is, again, the most frequent case. From a total of 780 clauses, in 76%

of the instances (593/780) there was no coreferential NP[47]. Again, these frequencies

simply confirm the long-noted correlation between datives and active referents involved

in topical chains. This referential persistence justifies an attenuated form of encoding, and

thus explains the pervasiveness of datives minimally encoded by clitic forms. In fact,

diachronic studies of Spanish have shown that clitic-exclusive encoding is, actually, the

basic and more frequent manifestation of dative arguments throughout the history of the

language (Company 2006: 493, 504).

---

[45] Analogous results from corpora analyses were obtained by Bogard (1992), who examined the encoding of dative arguments in *El habla culta de la ciudad de México* (Lope Blanch 1971). The dative arguments in this corpus show the occurrence of the dative clitic in 92% of the cases (453/491), and of a non clitic-doubled dative phrase in only 8% (38/491). Also similar is the data presented in Weissenrieder (1995), reviewed in Chapter 2 (§2.2.3).

[46] Furthermore, it has been pointed out that datives typically refer to singular definite entities, and that datives realized as indefinites or with generic reference are often unattested in corpora, a trait in which the Spanish language has been consistent for centuries (Company 2006: 503).

[47] These data were obtained looking for occurrences of the dative form *le* in the two volumes of the Barrenechea corpus. This particular clitic form is considered the unmarked manifestation of indirect objects in Spanish (Company 2006: 504).

In the Buenos Aires corpus, typical examples involve instances where the referent encoded by the clitic has been mentioned in the preceding clause, as in the sentences presented in (138).

(138) a. El tipo llega y **le** abre la puerta esta mujer.   (HC:XXX)
       'The guy arrives and this woman opens the door for him.'
       (lit. 'opens him the door')

    b. Ya he dicho muchas veces que la originalidad no tenía el valor en aquel
       entonces que se **le** atribuye hoy.   (HC:XX)
       'I have said many times that in back then originality didn't have the value
       that is attributed to it today'

    c. …estuve hablando con---Susi y **le** contaba mis angustias.   (HC:XXIV)
       '…I was talking to---Susi and I was telling her my problems.'

On the other hand, clitic-doubling constructions occur when the antecedent is farther away in the discourse context, or when the dative participant has no antecedent at all. Consider the following examples.

(139) a. ¿Pedro Páramo? Eh... escucháme, Pedro Páramo, mirá--- yo lo leí este
       año cuando fui a la facultad. Eh... es la historia de un tipo--- cuya madre al
       momento de morir--- le dice que vaya--- a un pueblo donde vive--- Pedro
       Páramo. Pedro Páramo es su padre, él es hijo de Pedro Páramo. Entonces el tipo
       le cierra los ojos **a su madre**--- y va a ese pueblo   (HC:XXX)

       '*Pedro Páramo*? Eh… listen to me, *Pedro Páramo*, look--- I read it this
       year when I was in college. Eh… it is the story of a guy--- whose mother at
       the time of death--- tells him to go--- to a town where ---Pedro
       Páramo lives. Pedro Páramo is his father, he is Pedro Páramo's son. Then the
       guy closes **his mother's** eyes (lit. 'he closes the eyes to his mother')
       --- and goes to that town.'

144

b. No. Yo me sicoanalicé y estoy de vuelta del sicoanálisis. Yo hago una sicoterapia... me he sicoanalizado muchos años con sicoanalistas de primerísima categoría, entre los cuales Emilio Rodrigué, que es de escuela inglesa; lo hice cuando él volvió de Inglaterra--- pero--- yo creo que el sicoanálisis es una escuela que tiene que ser revisada íntegramente--- es decir--- yo **le** doy muchísima importancia **a los aspectos sociales** que los sicoanalistas no dan.   (HC:XIV)

'No. I did psychoanalysis and I am done with it. I do
psychotherapy… I have been in analysis many years with first rate analysts,
among them Emilio Rodrigué, that comes from the English school;
I did it when he came back from England--- but--- I think that
psychoanalysis is a school that has to be completely revised--- I mean---
I give a lot of importance **to the social aspects**, and psychoanalysts don't.'

Dative doubling occurs when exclusive pronominal encoding would have failed at helping the hearer establish the intended reference. In the example in (139)a) this is due to the presence of competing participants (i.e. the protagonist, his mother, Pedro Páramo, the town where Pedro Páramo lives). In the example in (139)b), exclusive pronominal encoding would have failed because the referent is discourse-new, although identifiable, as signaled by the definite article that introduces it.

If dative arguments are normally definite, topical, cognitively *active* in the mind of the interlocutors, then doubling emerges as a marked construction that codifies a marked meaning: it occurs when the referent in question is *less accessible* than normally expected for a dative argument; i.e. where the clitic alone would not suffice for establishing the intended referent.

Cases of NP-exclusive encoding are simply extreme examples of a deviation from the expected pattern, as they involve referents that not only are not *accessible*, but further are not even identifiable.

5.3. COMPARING DATIVE AND ACCUSATIVE DOUBLING

The pervasiveness of the clitic in the encoding of dative arguments, as reflected by its frequency of occurrence in natural discourse, is a natural correlate of the lack of constraints that hold over dative doubling constructions. In effect, it is important to stress that, contrary to accusative doubling, dative doubling cannot be posited to be constrained by the referential features of the lexical phrase. It may involve animate or inanimate referents (139); and it is claimed to be possible with both specific and non-specific ones (Suñer 1988, cf. e.g. (25)b). Suñer (1988: 395 fn.6) has claimed that the clitic is dispreferred in chains with plural bare nouns (140)a; but it has been argued that the acceptability of the clitic improves as soon as the referent is qualified (140)b or the phrase is otherwise made phonologically heavier (140)c (Cuervo 2003: 43, (b) and (c) are her examples).

(140) a. ?**Les/Ø** donaré todos mis bienes **a museos**.
'I will donate all my possessions to museums.'

b. **Les** donaré todos mis bienes **a museos locales**.
'I will donate all my possessions to local museums.'

c. **Les** donaré todos mis bienes **a museos y bibliotecas**.
'I will donate all my possessions to museums and libraries.'

Not only there are no contexts in which the dative clitic is categorically banned, but on the contrary there are many structures in which it is obligatory (cf. §1.3.). The dative clitic is not only obligatory if the dative phrase is realized by a strong pronoun (141)a;

but also if the coreferential phrase occurs pre-verbally (141)b[48], encodes a semantic experiencer (141)c) or appears with any of the verbs which may "incorporate" a dative argument into their structure, of the *cocinar*-type reviewed above (141)d). It has been argued that the dative clitic is also obligatory if the dative is realized by a proper name (141)e) (Mayer 2003, Company 2006:491) or if it involves inalienable possession (141)f) (Fernandez Soriano 1993).

(141) a. ***Ø/Le** preguntaron **a él**.
      DAT3S asked.3PL   DAT he
      'They asked him.'

   b. **A Juan *Ø/le** trajeron   chocolates.
      DAT Juan   DAT3S brought.3PL chocolats
      'They brought chocolates to Juan.'

   c. **A Paula *Ø/le** alegró   la noticia.
      DAT Paula   DAT3S made.happy.3S the news
      'The news pleased Paula.'

   d. Piazzolla ***Ø/le**   encontró un ritmo nuevo **a la ciudad**.
      Piazzolla   DAT3S found.3S a rhythm new DAT the city
      'Piazzolla gave the city a new rhythm.' (lit. 'found it a new rhythm')

   e. ***Ø/Le** contó **a Gisela** su secreto más íntimo.
      DAT3S told.3S DAT Gisela her secret most intimate
      'She let Gisela in on her most intimate secret.'

   f. ***ø/Les** dolía la cabeza **a todos los que vieron esa película**.
      DAT3PL hurt.3S the head DAT all   the.ones saw.3PL that movie
      'All the people who saw that movie had a headache.'

---

[48] Note that contrary to DOs, for which this position is alweays marked, preverbal position is unmarked for datives associated with the role of experiencer (cf. Gutiérrez Bravo 2002).

The distributional potential of the dative clitic contrasts sharply with the restrictions on the occurrence of the accusative clitic. Consider the table below, which summarizes the conditions that force, allow, or prohibit the occurrence of a clitic given a coreferential full NP.

	DATIVE CLITIC	ACCUSATIVE CLITIC
OBLIGATORY	NP=preverbal NP=semantic experiencer NP="incorporated" dative NP=proper name NP=inalienable possession	**Never**
POSSIBLE	*Elsewhere*	NP="accessible"
BANNED	**Never**	*Elsewhere*

Table 9. *Comparing the privileges of occurrence of dative vs. accusative clitics*

Each column represents the virtual mirror-image of the other. There are several conditions that force the occurrence of the dative clitic, and none that force the occurrence of the accusative. On the other end, there are no contexts in which the dative clitic is banned[49], but many were the accusative clitic is. It is in the most restricted contexts that the accusative clitic can *occur*, and the dative clitic can be *omitted*. The middle row shows that there are no restrictions for the potential appearance of the dative clitic, whereas there are severe limitations on the possible appearance of the accusative.

---

[49] Save, arguably, when in a chain with certain plural bare nouns (cf. above).

As we saw, analyses of texts show that this distributional potential is reflected in the frequency of occurrence of each form in actual discourse.

It seems, then, that whereas from a pragmatic perspective it is justified to analyze DO-doubling constructions as the "duplication" of a nominal argument by a pronominal clitic, IO-doubling structures require the opposite approach, acknowledging the fact that the unmarked realization of dative arguments is via pronominal clitics. The markedness of each encoding type for a given case-role can be illustrated as in Figure 1, which represents the markedness of the relation between case-role and encoding form for both dative and accusative arguments. It captures that the unmarked realization of dative arguments is via clitics alone, and the unmarked realization of accusative arguments is via NPs alone (two facts that have been long recognized), whereas clitic doubling of dative and accusative arguments should be conceived of the marked form of encoding in both cases, although with respect to different "standards".
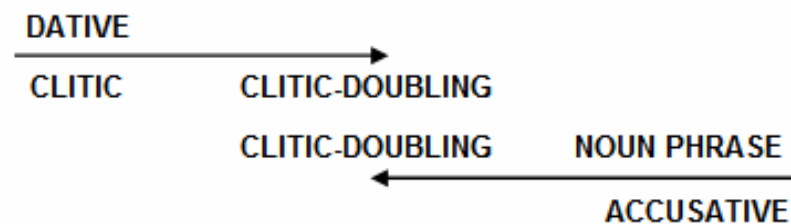

Figure 1. *Markedness of the relation between case-role and encoding form*

5.4. SUMMARY AND DISCUSSION

In this Chapter we have seen that the dative clitic seems to be one step further than the accusative clitic in the process of grammaticalization that marks the transformation of bound pronouns into grammatical agreement markers. The dative clitic still behaves as a pronoun, as it can fulfill the valence of the verb (and in fact is the most frequent morphosyntactic representation of dative arguments), but it is closer than the accusative clitic to the behavior of a pure agreement marker in that the semantic / pragmatic conditions for its occurrence are much less constrained and, further, it has become obligatory in many grammatical contexts.

In effect, the clitic may be omitted with a certain group of verbs in certain constructions, particularly with non-specific, often non-referential, participants, in order to mark, as Maldonado (2002: 18) puts it, a "looser subject-dative linkage". Given the typical discourse prominence of dative arguments, it is not surprising that tokens of this class are rarely found in actual texts. On the other hand, the discourse prominence typically associated with datives renders optional the appearance of a coreferential phrase, whose occurrence depends on the relative accessibility of the target referent: If the entity in question is the current focus of attention, as is commonly the case, the clitic alone can suffice for establishing reference. If this entity is deemed cognitively less accessible, the coreferential phrase helps to identify it.

The interpretation of dative doubling constructions as serving to mark referents of low accessibility goes in accordance with the data presented in Weissenrieder (1995) and Colantoni (2002) and summarized in Chapter 2 (§2.2.3, 2.2.4). I agree with Colantoni

that dative doubling should be interpreted as "the duplication of the clitic by a lexical expression", which is used to help identify the intended referent. I differ from her in concluding that this is the function of clitic doubling in general. On the contrary, accusative doubling has the *opposite* function, as argued in Chapter 4.

On the other hand my corpus data agree with Weissenrieder's in showing that the clitic (alone) is the most frequent form of encoding of dative arguments in Spanish. Note, however, that I am not suggesting that *doubling* is the consequence of a topical argument triggering object-verb agreement. Quite on the contrary, it is the consequence of a non-topical argument requiring its nominal expression. More precisely, whereas the greater grammaticalization of the dative clitic is a consequence of the general topicality of dative arguments, the context-specific occurrences of doubling result from a deviation from this tendency. If the general tendency were automatically translated to particular discourse contexts (i.e. if dative doubling were automatically correlated with topicality), we would find dative doubling constructions to be extremely frequent, and we have seen that in fact they are not.

From the analysis presented in Chapter 4 it is clear that accusative doubling structures are also rare. In effect, the unmarked realization of accusative arguments is via a (non-doubled) lexical phrase. This form of encoding is consistent with the fact that the unmarked position for direct objects is also the unmarked focus position; and focal elements are, further, typically associated with non-accessible referents (Lambrecht 1994: 166). In other words, there are a number of factors that conspire for making DO-doubling structures infrequent: in canonical transitive sentences the DO occurs within the focal

constituent, and the focal constituent is the best place to include brand-new referents. Since DO-doubling is only possible if the referent is identifiable, DO-doubling also ends up as a marked structure used to code a marked meaning: In this case, a relatively *accessible* focal direct object. I mentioned that dative doubling constructions are "marked" for the opposite reasons: they occur when a participant encoded with a dative case-role is less accessible than is usually the case. Thus, we can revise the schema presented in Figure 1 and reinterpret it in terms of the markedness of the relation between case-role and activation status, as in Figure 2.
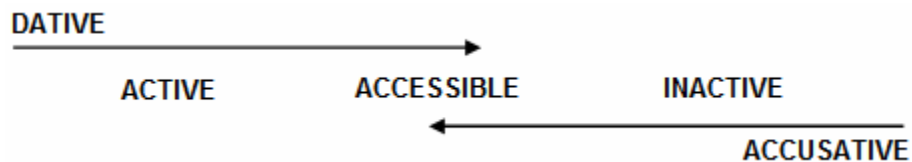
Figure 2. *Markedness of the relation between case-role and activation status*

What this Figure says is that we can conceive of dative and accusative clitic doubling constructions as working in opposite directions, but pursuing a common functional goal. I am not claiming that there is any processing implications to this, or that these "opposite directions" should be interpreted as anything else than an analytical perspective that let us capture the functional motivations that bring together what appeared as separate phenomena governed by independent factors.

In effect, we can now return to the chart introduced in Chapter 4 (Table 6), and complete it with a corresponding line capturing the relation between dative doubling and activation status. The result is presented below.

ACC.	**La** invité 'I invited her.'	**La** invité **a Mabel**. 'I invited Mabel.'	Invité **a Mabel**. 'I invited Mabel.'
FORM	CLITIC	CLITIC-DOUBLING	NP
STATUS	ACTIVE	ACCESSIBLE	INACTIVE
DAT.	**Les** dieron un día de asueto. 'They gave them a day off.'	**Les** dieron un día de asueto **a los trabajadores**. 'They gave the workers a day off.'	Dieron un día de asueto **a los trabajadores**. 'They have the workers a day off.'

Table 10. *Clitic doubling and activation levels*

6   CLITIC DOUBLING IN ROLE AND REFERENCE GRAMMAR

In Chapters 4 and 5 I presented a corpus-based study of clitic doubling constructions in Spanish. In this Chapter I will propose an analysis of these constructions within the framework of Role and Reference Grammar. First I will introduce the model, making particular reference to RRG's treatment of the interface between syntax, semantics and pragmatics, as it relates to the constructions under analysis. I will then advance a Role and Reference Grammar account of clitic doubling in Spanish that is consistent with its functional motivation, as discussed in the previous Chapters.

6.1. PRESENTATION OF THE MODEL

As mentioned in Chapter 1, Role and Reference Grammar (RRG) is a structural-functionalist model, which takes a "communication-and-cognition" perspective to linguistic analysis. One of the basic tenets of this perspective is that morphosyntactic forms cannot be analyzed without reference to their use in particular contexts. Accordingly, RRG is concerned with investigating the interaction of structure, meaning and communicative function. The following quote, from Van Valin (2005:1), summarizes the premises of this framework.

> "Language is a system of communicative social action in which grammatical structures are employed to express meaning in context. While all languages can achieve the same basic communicative ends, different languages use different linguistic means to achieve them, and an important aspect of these differences concerns the divergent ways syntax, semantics and pragmatics interact across languages."

In accordance with this conception of language, RRG is organized in terms of bidirectional mappings between semantic and syntactic representations, which are permeated by the influence of discourse pragmatics. The interface between these components is governed by a set of bidirectional rules, called "linking algorithms", which capture both the production and comprehension aspects of language use. This organization can be represented as in Figure 3 (Van Valin 2005: 2).

SYNTACTIC REPRESENTATION

Linking
Algorithm

SEMANTIC REPRESENTATION

DISCOURSE-PRAGMATICS
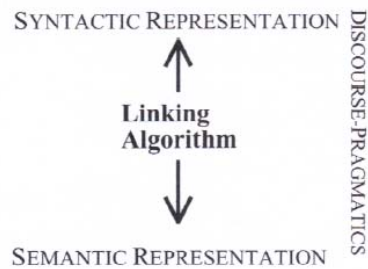
Figure 3. *Organization of Role and Reference Grammar*

I will introduce the basic concepts behind the RRG treatment of the syntactic, semantic and pragmatic representations and their interfaces, focusing on the aspects more relevant to Spanish clitic doubling constructions. There are a number of important features of the theory which I will not include in this review, as they have no direct bearing on the phenomenon under analysis. Among them, the characteristics of the layered structure of the NP or RRGs theory of juncture and nexus types. A comprehensive presentation of the theory can be found in Van Valin (2005) and Van Valin and LaPolla (1997).

6.1.1. The Layered Structure of the Clause

There are two general principles on which the RRG theory of clause structure is founded. The first dictates that the universal features of clause structure should be captured without being imposed on particular languages categories for which there is no evidence. The second principle states that where different languages have comparable structures, these structures should be represented in comparable ways.

These constraints have to be obeyed without resorting to abstract levels of syntactic representation. Being a monostratal theory, RRG considers only one level of syntactic representation: that which corresponds with the actual morphosyntactic chain.

In other words, RRG seeks to uncover universal generalizations about clause structure without biasing the analysis in favor of Indo-european categories and without positing underlying syntactic structures where differences among languages are neutralized.

In RRG the clause is conceived as a layered structure of grammatical units, termed the "layered structure of the clause" (LSC). The LSC consists of a "core" and an optional "periphery". The "core" contains the verb or other predicating element, termed "nucleus", and its direct and oblique arguments, while the periphery contains the potential adjuncts.

These units are based on the universal distinction between predicates, arguments and adjuncts, which motivate the universal syntactic units on the LSC. This is represented in the following table (Van Valin 2005: 5).

Semantic element(s)	Syntactic unit
Predicate	Nucleus
Argument in semantic representation of predicate	Core argument
Non-arguments	Periphery
Predicate + Arguments	Core
Predicate + Arguments + Non-arguments	Clause (= Core + Periphery)

Table 11. *Semantic units underlying the syntactic units of the LSC*

The following figure illustrates a preliminary approach to the layered structure of a Spanish transitive sentence ('Carolina bought the flowers in Plaza Francia').
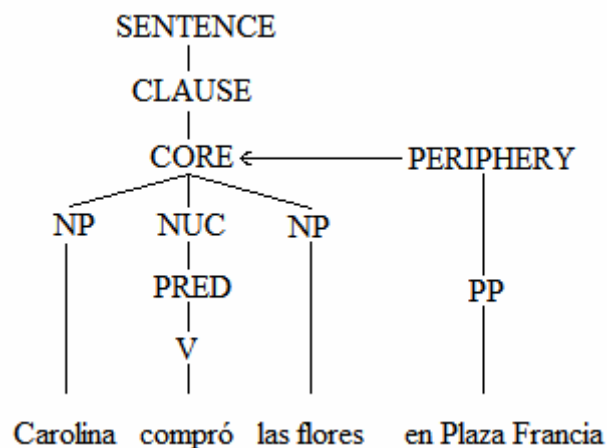
Figure 4. *The Layered Structure of the Clause (preliminary)*

Whereas these elements are semantically motivated and universal, there may be pragmatically motivated, language-specific additional elements beyond the ones represented in Figure 4. In Spanish, a "pre-core slot" (PrCS), preceding the core-internal slot typically occupied by the subject, captures the position where question words and pre-posed focal elements occur. There is no pause or intonation break separating these

157

elements from the rest of the sentence, nor there is a resumptive pronoun within the clause referring to these elements when they function as semantic arguments of the clause. These features distinguish the PrCS from the "left-detached position" (LDP). The LDP is inside the sentence but outside the clause, and it is where topicalized elements occur. Elements in the LDP are typically separated from the clause by a pause or intonation break, and when they function as semantic arguments they are co-indexed inside the clause. In languages in which these elements may also occur in sentence final position we talk, in turn, about a "post-core slot" (PoCS) and a "right-detached position" (RDP).

Each layer of the clause structure is modified by an operator projection, which captures the morphological realization of grammatical categories such as illocutionary force, tense, aspect and negation. These operators are ordered with respect to their relative scope. At one end illocutionary force modifies the clause as a whole, and it is therefore a clausal operator. At the other end, aspect is a nuclear operator modifying the event or state itself. (The complete list of operators is presented below).

The constraints on the ordering of predicates and arguments that may be operative in a language are independent of the constraints on the ordering of operators, which is sensitive to the relative scope of each one. In order to capture this divergence, Johnson (1987) proposed a formalization in which operators are represented in a different projection than predicates and arguments. This is illustrated in Figure 5, which also contains an abstract representation of a sentence with pre- and post-core slots and left- and right-detached positions.
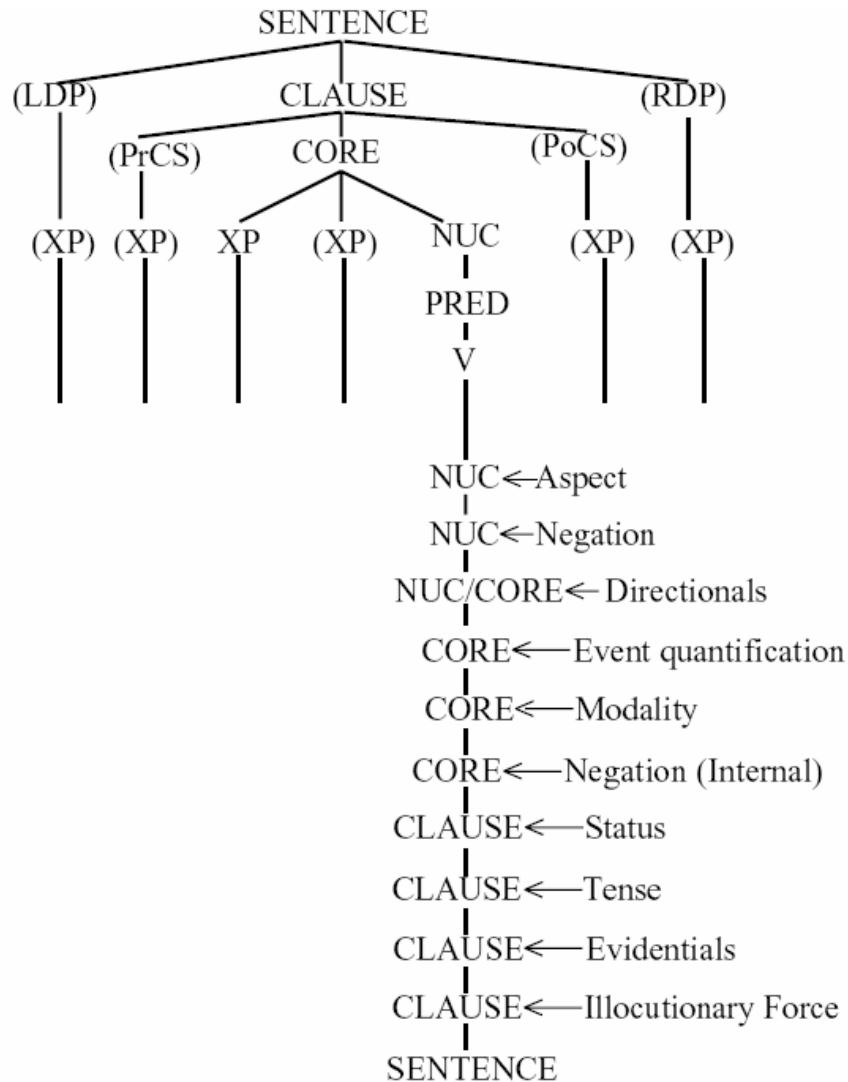
Figure 5. *The constituent and operator projection in the LSC*

In RRG, the syntactic representation of a sentence does not depend on the application of phrase-structure rules, but is determined by selectional principles which target a language-specific inventory of syntactic templates. Syntactic templates are merged to create appropriate syntactic representations of actual sentences in the language. An example of the combination of English syntactic templates is presented in Figure 6

(op.cit. 15). Spanish syntactic templates and selectional principles will be discussed in §6.2.1.

Figure 6. *Example of English syntactic templates and their combination*

6.1.2. The Semantic Representation

In RRG the semantic representation of the sentence is based on the lexical structure of the predicating element. Building on the categories proposed by Vendler (1967), RRG recognizes six basic *Aktionsart* classes, each of which has a causative counterpart.

i. STATE:
   El ladrón está asustado. ('The thief is afraid.')

i''. CAUSATIVE STATE:
   El perro asustó al ladrón. ('The dog frightened the thief.')
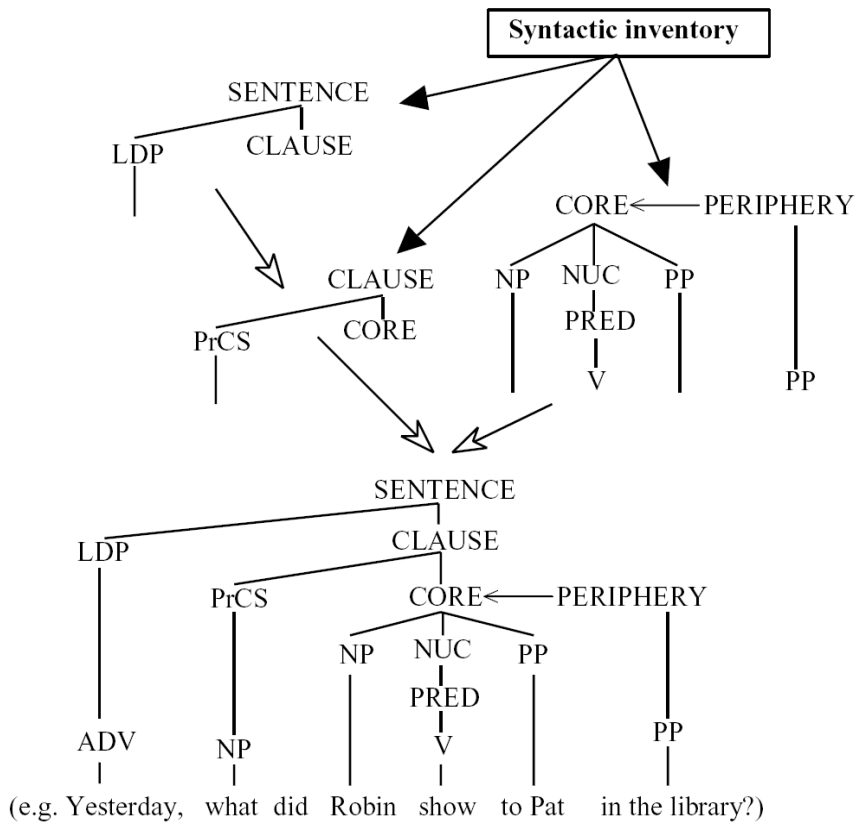
ii. ACHIEVEMENT:
   La bomba detonó. ('The bomb detonated.')

ii'. CAUSATIVE ACHIEVEMENT:
   El oficial detonó la bomba. ('The officer detonated the bomb.')

iii. SEMELFACTIVE:
   La tierra tembló. ('The earth trembled.')

iii'. CAUSATIVE SEMELFACTIVE:
   El terremoto hizo temblar la tierra. ('The earthquake make the earth tremble.')

iv. ACCOMPLISHMENT:
   El bloque de hielo se derritió. ('The ice block melted.')

iv'. CAUSATIVE ACCOMPLISHMENT:
   El sol derritió el bloque de hielo. ('The sun melted the ice block.')

v. ACTIVITY:
   Los atletas corrieron por la pista. ('The athletes run in the runway.')

v'. CAUSATIVE ACTIVITY:
   El entrenador hizo correr a los atletas. ('The coach made the athletes run.')

vi. ACTIVE ACCOMPLISHMENT:
   El niño se comió un chocolatín. ('The kid ate a chocolate bar.')

vi''. CAUSATIVE ACTIVE ACCOMPLISHMENT:
   La mamá le dio de comer un chocolatín. ('The mother fed him a chocolate bar.')[50]

Each class is formally captured by representational schemas based on Dowty

(1979). States and activities are considered basic, with the remaining representations

---

[50] Note that active accomplishments, which result from the telic use of activity verbs, are generally derived from (activity) verbs of motion, consumption and creation. Contrary to English, Spanish is scarce in lexical examples of causative active accomplishments.

161

derived from them. States are represented as bare predicates (e.g. **know'** (x, y)), and activities as containing the element **do'**; e.g. **do'** (x [**eat'** (x, y)]). Achievements are captured by an ingressive operator INGR, semelfactives by the operator SEML, and accomplishments by the operator BECOME. There is an extra operator, PURP, for "purposive", which captures the basic meaning of English *for* and Spanish *para*[51]. Causative verbs use the operator-connective CAUSE to link the causing action or event to the result state. Examples of the different representations are illustrated in Table 12.

Aktionsart class	Logical Structure	Example
STATE	**feel'** (ladrón, [**afraid'**])	*El ladrón está asustado.*
ACTIVITY	**do'** (atletas [**run'** (atletas)]	*Los atletas corrieron.*
ACHIEVEMENT	INGR **detonate'** (bomba)	*La bomba detonó*
SEMELFACTIVE	SEML **do'** (tierra [**tremble'** (tierra)])	*La tierra tembló*
ACCOMPLISHMENT	BECOME **melted'** (bloque de hielo)	*El bloque de hielo se derritió*
ACTIVE. ACCOMP.	[**do'** (niño [**eat'** (niño, chocolatín)])] & [INGR **consumed'** (chocolatín)]	*El niño se comió un chocolatín.*
CAUSATIVE STATE	[**do'** (perro, Ø)] CAUSE [**feel'** (ladrón, [**afraid'**])]	*El perro asustó al ladrón.*
CAUSATIVE ACTIVITY	[**do'** (entrenador, Ø)] CAUSE [**do'** (atletas [**run'** (atletas)])]	*El entrenador hizo corer a los atletas.*
CAUSATIVE ACHIEVEMENT	[**do'** (oficial, Ø)] CAUSE [INGR **detonate'** (bomba)]	*El oficial detonó la bomba.*
CAUSATIVE SEMELFACTIVE	[**do'** (terremoto, Ø)] CAUSE [SEML **do'** (tierra [**tremble'** (tierra)])]	*El terremoto hizo temblar la tierra.*

---

[51] PURP is the abbreviation of the logical structure **want'** (x, LS$_2$) & DO (x, [[LS$_1$ ◊ CAUSE [LS$_2$]]). Thus, the semantic representation of a sentence such as *Robin baked a cake for Sandy* is [[**do'** (Robin, Ø)] CAUSE [BECOME **baked'** (cake)] PURP [BECOME **have'** (Sandy, cake)]. See Joly (1993) and Van Valin and LaPolla (1997:382-4) for discussion.

| CAUSATIVE ACCOMPLISHMENT | [**do'** (sol, Ø)] CAUSE [BECOME **melted'** (bloque de hielo)] | *El sol derritió el bloque de hielo.* |
| CAUSATIVE ACTIVE ACCOMPLISHMENT | [**do'** (mamá, Ø)] CAUSE [[**do'** (niño [**eat'** (niño, chocolatín)])] & [INGR **consumed'** (chocolatín)]] | *La mamá le dio de comer un chocolatín al niño.* |

Table 12. *Lexical representations for* Aktionsart *classes*

The theory proposes two semantic macroroles: ACTOR and UNDERGOER, which are a generalization over more traditional thematic roles, justified by their common grammatical behavior. The relation between argument positions in the logical structure and macrorole assignment is captured in the actor-undergoer hierarchy presented in Figure 7 (op.cit. 61)[52].

ACTOR   UNDERGOER
⟶
                                    ⟵
Arg. of   1st arg. of 1st arg. of   2nd arg. of   Arg. of
DO   **do'** (x,... **pred'** (x, y)   **pred'** (x, y)   **pred'** (x)
[⟶' = increasing markedness of realization of argument as macrorole]
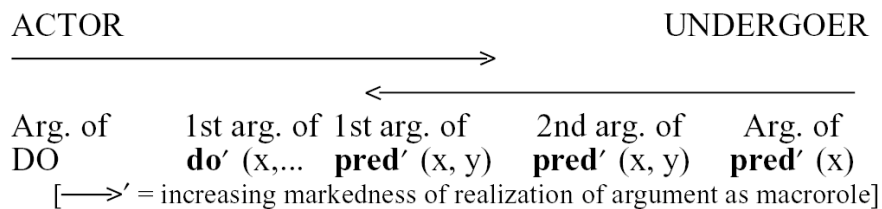
Figure 7. *The Actor-Undergoer hierarchy*

This hierarchy captures that "argument of **DO**" (agent) is the unmarked choice for actor, and "argument of **pred'**(x)" (patient) is the unmarked choice for undergoer. A marked choice for actor is possible only if the higher-ranking arguments are not present in the clause. Conversely, a marked choice for undergoer accounts for phenomena such as English dative shift, or locative alternation in Spanish (e.g. *Cargó las cajas en el*

---

[52] The linking system which determines, among other things, the assignmet of macroroles, will be introduced in §6.2

163

camión vs. *Cargó el camión con cajas*; 'He loaded the boxes in the truck' vs. 'He loaded the truck with boxes'). In any logical structure with two arguments, the leftmost will be the actor and the rightmost will be the undergoer. When macroroles occur as core arguments, they are always direct (with the exception of actors of passive constructions).

The number of macroroles a verb takes generally corresponds to the traditional notion of transitivity: one-macrorole verbs are intransitive, two-macrorole verbs are transitive. The two notions, however, should be distinguished. This is done by referring to syntactic- or S-transitivity and macrorole or M-transitivity. S-transitivity refers to the number of direct core arguments that appear in the syntax. For instance, in *Carolina gave Esteban a big surprise* there are three direct core arguments in the syntactic structure. In terms of M-transitivity, there are only two macroroles, actor and undergoer. The third participant is in this case referred to as non-macrorole core argument. It is "direct" in that sentence, and it would be "oblique" in the non-dative shift version, where it is prepositionally marked: *Carolina gave a big surprise to Esteban*.

Just as the notion of "indirect object" has no theoretical status, the same is true for the other canonical grammatical relations, i.e. "subject" and "direct object". "Direct objects" coincide with what in RRG is captured by the notion of "macrorole direct core argument". For the traditional notion of "subject" RRG uses the concept of "privileged syntactic argument" (PSA). For accusative languages such as Spanish, the PSA in active constructions corresponds to the highest-ranking direct core argument according to the following hierarchy.

Arg.ofDO>1starg.of **do´**(x...)>1st arg. of **pred´**(x,y)>2ndarg.of **pred´**(x, y)>arg.of **pred´**(x)

Figure 8. *The PSA selection hierarchy*

Contrary to other theories, in RRG case assignment rules make no reference to grammatical relations, as these have no status in theory. Instead, case assignment depends on the macrorole and direct core status of each argument. This is illustrated for Spanish in (142).

(142)   Case assignment rules for Spanish:
        a. Highest ranking core macrorole takes nominative case.
        b. Other core macrorole takes accusative case.
        c. Non-macrorole direct core arguments take dative as their default case.

For oblique arguments, the equivalent of case assignment is prepositional assignment. Here I will include only the rule assigning *para*, to which we will refer below.

(143)   Prepositional assignment rule for Spanish:
        Assign *para* to non-macrorole *y* argument in logical structure segment:…
        PURP [BECOME **pred'** (y, z)]

6.1.3. The Pragmatic Structure

Following Lambrecht (1994), RRG recognizes two aspects of discourse-pragmatics that influence sentence structure. The first aspect concerns the distribution of information in the sentence in terms of the pragmatic relations established among referents (i.e. the role of its constituents as "topic" or "focus"). The second aspect concerns the pragmatic

properties of referents, as it involves the activation status of the referents in the mind of the interlocutors, and its reflection in the forms chosen to encode them.

The distribution of information in the sentence, or its "focus structure", is the grammatical system that serves to encode the relative scope of the pragmatic assertion with respect to the pragmatic presupposition. Following Lambrecht, the pragmatic presupposition is the set of propositions lexico-grammatically evoked by an utterance, which the hearer is assumed to know, or believe, or be willing to accept by interpreting the utterance. Crucially, it contains whatever elements are topical in the syntactic structure. In effect, the topic of a sentence is that particular referential expression that is associated to an entity which is assumed as center of interest in the communicative exchange. Note that, as Lambrecht stresses, it is *not* the case that the topic should be given or presupposed. What is presupposed is that it will be involved in the proposition due to its status as center of interest[53].

The topical elements that are part of the pragmatic presupposition are opposed to focal elements related with the pragmatic assertion. The pragmatic assertion, in this theory, is the proposition that the hearer is expected to know (or believe, or accept) as a result of interpreting the utterance. Again, just as it makes no sense to talk about a topic as presupposed, it also makes no sense to talk about a focus as new. What is new (what is

---

[53] Lambrecht (op.cit.151) also points out that just as one cannot say that the topic "is presupposed", it also makes little sense to say that a referent, or a noun phrase, is "presupposed": "It is no doubt the inherent relationship between topic and pragmatic presupposition that has led to the widespread terminological habit of calling the topic of a sentence "presupposed". This habit is as misleading as that of calling a definite NP, or even its referent, "presupposed". Anything presupposed is propositional in nature (such as some shared belief or knowledge), but topic referents are for the most part not propositions but entities".

informative) is the relationship that the utterance establishes between new and old information. In this regard, focus is defined as "the semantic component of a pragmatically structured proposition whereby the assertion differs from the presupposition" (op.cit. 213).

Lambrecht posits three basic focus types, related with three basic communicative functions. "Predicate focus" in the unmarked focus structure; and its basic communicative function is to realize a comment on the topic. "Sentence focus" emerges when a new referent (or a new event or state) is introduced in the discourse context. "Argument focus" ("narrow focus" in RRG) has the function of identifying a referent.

All languages have some grammatical strategies to encode the informational status of its constituents. In other words, in each language there are conventional associations between the form of a sentence and the distribution of information in terms of topic and focus. Languages differ, however, in two respects. The first is the degree in which the focal structure may influence the syntactic structure. The second is the relative prominence languages grant to syntactic or morphological devices (in addition to intonation) in the marking of focus structure.

This can be seen more clearly through examples. The sentences below, from different languages, represent the unmarked manifestation of the three basic focus types recognized in the theory. (Examples from languages other than Spanish are originally from Lambrecht 1994:223).

(144)   Predicate focus
Q: What happened to your car?
A: My car / it broke DOWN.   English
      (La mia macchina) si é ROTTA.   Italian
      (Ma voiture) ell est en PANNE.   French
      (Kuruma wa) KOSHOO-shi-ta.   Japanese
      (El auto) se me DESCOMPUSO.   Spanish

(145)   Sentence focus
Q: What happened?
A: My CAR broke down.   English
      Mi si é rotta la MACCHINA.   Italian
      J'ai ma VOITURE qui est en PANNE.   French
      KURUMA ga KOSHOO-shi-ta.   Japanese
      Se me descompuso el AUTO.   Spanish
      El auto se me descompuso.

(146)   Narrow Focus
Q: I heard that your motorcycle broke down.
A: My CAR broke down.   English
      Si é rotta la mia MACCHINA.   Italian
      C'est ma VOITURE qui est en panne.   French
      KURUMA ga koshoo-shi-ta.   Japanese
      El AUTO se me descompuso.   Spanish
      Se me descompuso el AUTO.
      Es el AUTO el que se me descompuso.

As illustrated by these examples, in English it is possible to express the different

focal structures by changing the locus of intonational prominence; no morpho-syntactic

change in the structure of the sentence is required. In Japanese, the different focal

structures are distinguished by a combination of intonation and morphological marking.

In French and Italian there is a restriction over pre-verbal focal elements, and therefore it

is not possible in these languages to mark sentence focus or narrow focus stressing

preverbal elements. Therefore, they encode the relevant focus structures through syntactic means.

In Spanish, the unmarked focus position coincides with the last syllable of the melodic group of the clause; and thus predicate focus structures are simply marked by a neutral focal accent on the last constituent of the core. Sentence focus is realized by constructions in which the PSA is explicit and typically occurs post-verbally. As shown in (145), however, Spanish also allows for sentence-focus structures to be realized with the default SVO order (Zubizarreta 1999: 4225). Narrow focus can target any constituent in the sentence. If it is marked (if it targets an element other than the last constituent of the core in a canonical SVO sentence), it receives an emphatic accent and is interpreted contrastively. As illustrated in (146), narrow focus on the PSA can also be expressed by means of an inverted construction or by a cleft. In other words, Spanish is like English in that it is potentially possible for the same morphosyntactic string (*El auto se me descompuso*) to encodes any of the three focus types, depending on the placement of the intonational accent. It is, however, a language which also has flexible word order, and thus it is also possible to encode the distinction between the unmarked predicate focus structure and sentence and narrow focus by altering the syntactic structure, as the examples above illustrate[54].

---

[54] When there is more than one structural alternative for encoding a particular focus type (as for instance in (145) and (146)), it remains to be established which particular construction is preferred under which particular pragmatic conditions. For some recent studies on different aspects of Spanish focus structure, see for instance Gutiérrez Bravo (2002), Domínguez (2004), Mora Bustos, (2007) and Bogard (2007).

RRG captures the restriction over potential placement of focal elements with the notion of "potential focus domain" (PFD). The potential focus domain refers to the domain in the sentence where focus may fall, whereas the "actual focus domain" (AFD) targets the actual focal element in a specific construction. Thus, the facts from Italian, as those from many other languages with similar constraints on pre-verbal focal elements, can be explained in terms of restrictions on the PFD. Consider for example the representations on Figure 9.
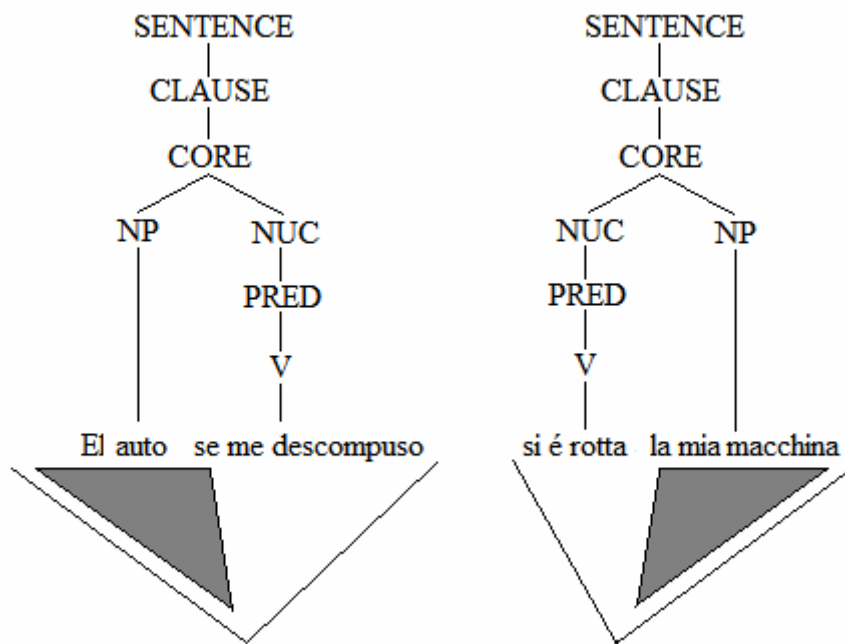

Figure 9. *Representation of the Potential and Actual Focus Domains in Spanish and Italian Narrow Focus Constructions*

In the Spanish example, the actual focus falls on the first constituent of the core, the PSA, marking a narrow focus construction. Italian does not allow preverbal elements

inside the core to be focal, and thus in Italian an inversion construction is used so that the PSA occurs post-verbally. The difference between the two languages does not lie in the actual focus domain (both languages allow PSAs to be focal), but in the *potential* domain of focus: In Italian focal elements must be post-verbal, whereas this restriction that does not hold in Spanish. As a consequence, Italian syntax adapts to the relative rigidity of its focus structure. To put if differently, the Spanish example may potentially serve to express different focus structures (the same morphosyntax could be used to encode sentence focus or predicate focus, for instance). On the contrary, the non canonical placement of the PSA in the Italian example marks it as a focal constituent.

The different strategies used by different languages to encode focus structures, and in particular the degree to which focus structure may influence syntactic structure (in particular, word order), are at the origin of the RRG typology of languages in terms of the interaction of "syntax" and "information status". There are languages where the flexibility of word order adapts to the rigidity of focus structure, such as the Italian example just reviewed. At the other end, there are languages where it is the focus structure which adapts to the rigidity of word order (e.g. English). Finally, there are also languages where both word order and focus structure are rigid, requiring more dramatic changes in the syntax to express marked focus types (e.g. French), and languages where both word order and focus structure are relatively flexible (e.g. Russian). This is captured in the following table (Van Valin 2005: 77).

	Rigid Focus Structure	Flexible Focus Structure
Rigid Syntax	French, Toba Batak	English, Toura
Flexible Syntax	Italian, Sesotho	Brazilian Portuguese, Croatian, Russian, Polish

Table 13. *Interaction of Syntax and Focus Structure*

Where should Spanish be included? Spanish has free word order, at least compared with English or French. Does it also have flexible focus structure? Crucially, the criterion used for determining the "flexibility" of focus structure concerns the potential focus domain. As discussed above, Spanish lacks the constraints present in French or Italian. Thus, even if the four categories of the chart represent idealized groups, it seems justified to place Spanish in the same group as Brazilian Portuguese, Croatian, Polish, and other languages with flexible syntax and flexible focus structure.

The following examples illustrate different focus types, using as example DO-doubling structures.

(147)　Q　¿Cómo **lo** hiciste **el trabajo**?
　　　　　'How did you do the job?'

(148)　A　-Pedro es un gran tipo. Mucho mejor que su hermano.
　　　　B　-Sí, yo **lo** quiero mucho **a Pedro**.

　　　　'A　-Pedro is a great guy. (He is) much better than his brother.
　　　　B　-Yes, I love Pedro very much.'

(149)　Q　-¿Qué pasó?
　　　　A　-Me parece que **lo** vi **a Pedro** entrando en ese negocio.

　　　　'Q　-What happened?
　　　　A　-I think I saw Pedro going into that shop.'

172

(150)   Q   -¿Qué hiciste ayer?
         A   **-Lo** vi **a Pedro**.

         'Q: -What did you do yesterday?
         A   -I saw Pedro.'

(151)   Q   -¿A quién viste ayer?
         A   **-Lo** vi **a Pedro**.

         'Q   -Who did you see yesterday?
         A   -I saw Pedro.'

(152)   Q   -¿Así que viste a Juan?
         A   -No, **lo** vi **a Pedro**.

         'Q   -So you saw Juan?
         A   -No, I saw Pedro.'

Following the typology proposed in Lambrecht (1994), in (147) and (148) the doubled constituent is topical. In (147) the focus falls in the question word, and the rest of the sentence codifies the topic under discussion. In (148) the speaker functions as a primary topic and Pedro as a secondary one, as the utterance is about the relationship between the two participants (cf. op.cit. 147ff). Examples (149) to (152) illustrate the more common situation in which the DO coincides with the focus domain. (149) is a case of sentence focus, (150) of predicate focus, and (151) and (152) illustrate narrow focus on the DO argument, completive in the first case and contrastive in the second.

At the beginning of this section we mentioned that focus structure is one of the discourse-pragmatics aspects that the theory incorporates, the other aspect concerning the activation state of discourse referents. Building on the works by Prince (1981) and Chafe (1987), among others, RRG recognizes five possible activation levels: "*active*, i.e.

actively under consideration in the discourse by means of direct mention; *accessible*, i.e.

not actively under consideration but readily recognized by the addressee due either to

knowledge of the world or occurrence in the immediate environment of the speech

situation; *inactive*, i.e. previously mentioned but not actively under consideration and not

assumed by the speaker to be recognized by the addressee; *brand new anchored*, i.e. not

previously mentioned but related to something already mentioned or accessible; and

brand new unanchored, i.e. not previously mentioned or related to anything previously

mentioned" (Van Valin 2005: 79).

The five possibilities are represented in Figure 10 (from Van Valin and LaPolla
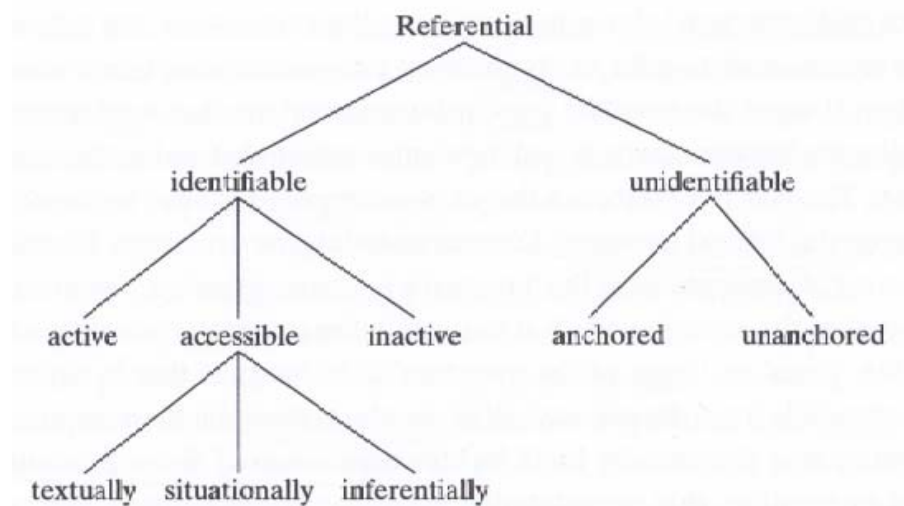
1997: 201).

Figure 10. *Cognitive states of discourse referents*

There is a natural interaction between focus structure, formal encoding, and

referent's activation. RRG proposes a markedness hierarchy that captures the interface

between pragmatic function (i.e. topic or focus) and the form of encoding. It is presented in Figure 11[55].

Markedness of Occurrence as Focus

Zero   Clitic/Bound Pro   Pro [-Stress]   Pro [+Stress]   Definite NP   Indefinite NP
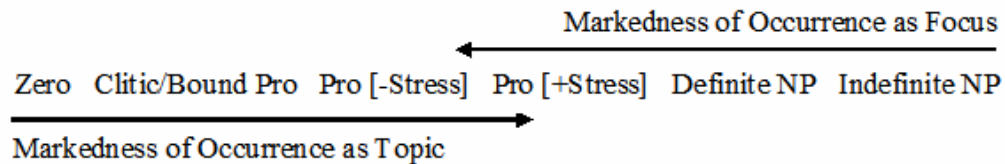
Markedness of Occurrence as Topic

Figure 11. *Pragmatic function and form of encoding*

This hierarchy captures the fact that attenuated forms are the unmarked realization of topics, whereas indefinite NPs are the unmarked realization of focal elements[56]. On the other hand, Lambrecht (1994: 165) proposes a scale of topic acceptability that captures the relation between pragmatic function and activation level, as in Figure 12.

Active > Accessible > Inactive > Brand-new anchored > Brand-new unanchored

Figure 12. *Pragmatic function and activation level*

Finally, there are works which focus on the relation between form of encoding and activation level. An example of this approach is the following "Giveness hierarchy" proposed by Gundel *et al.* (1993: 275).

---

[55] RRG's pragmatic markedness hierarchy is based on proposals from Givón (1983), Levinson (1987), Ariel (1990), Gundel *et al.* (1993) and Lambrecht (1994).

[56] The two ends are not completely symmetrical: indefinite NPs *may* occur as marked realizations of topical elements, but focal elements *cannot* be encoded by a clitic alone or by zero.

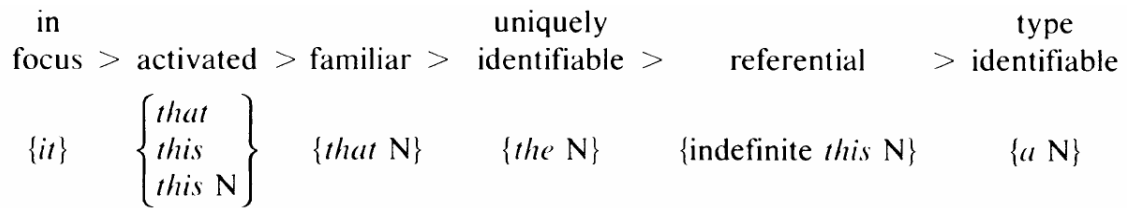| in | | uniquely | | | type |
| focus > | activated > familiar > | identifiable > | | referential > | identifiable |
| $\{it\}$ | $\begin{Bmatrix} that \\ this \\ this\ \text{N} \end{Bmatrix}$ $\{that\ \text{N}\}$ | $\{the\ \text{N}\}$ | $\{\text{indefinite } this\ \text{N}\}$ | | $\{a\ \text{N}\}$ |

Figure 13. *Form of encoding and activation level*

The pragmatic component in RRG is crucial for the analysis of a number of grammatical phenomena. It makes possible a pragmatic account of VP-like constituents (Van Valin 2005, §3.5), and it explains the interpretation of certain quantifier scope and so-called "extraction" phenomena (op.cit. §3.6, Van Valin and LaPolla 1997, §5.5; Van Valin 1995, 1998, Van Valin and LaPolla, §9.5, Van Valin 2005, §7.6) as well as constraints on pronominalization (Van Valin and LaPolla 1997, §5.6). Focus structure, derived from a dynamic representation of discourse referents akin to the one proposed in Discourse Representation Theory (Kamp and Reyle 1993; Asher 1999; von Heusinger 1999), has also being invoked for capturing phenomena such as "zero anaphora" (Van Valin 2005, §5.4.1) and "missing verbs" in Japanese (Shimojo, in press.). The evidence presented in previous chapters shows that clitic doubling constructions are also crucially sensitive to discourse pragmatic categories. In the following section I propose an analysis of Spanish clitic doubling constructions incorporating the analytic tools of Role and Reference Grammar presented here.

6.2. AN RRG ANALYSIS OF SPANISH CLITIC DOUBLING

6.2.1. The Clausal Structure of Clitic Doubling Constructions

In order to advance an RRG account of clitic doubling, the first task is to provide an appropriate layered structure for these constructions. A basic typological distinction, which may have a direct bearing on clause structure, is that between head-marking and dependent-marking languages (Nichols 1986). This typology is based on the locus of occurrence of grammatical morphemes that indicate a certain dependency relation between a head and its dependents. The dependency may be at the level of the phrase (e.g. the relation between a noun and an adjective), at the level of the clause (e.g. the relation between a predicate and its arguments), or at the level of the sentence (the relation between the main predicate and an embedded clause). We will be mainly concerned with clause-level dependencies.

As their name indicates, "head-marking" and "dependent-marking" make reference to the place where the dependency marker occurs. Languages in which dependency markers consistently appear in heads (of different levels) are "head-marking languages", whereas in "dependent-marking languages" the markers appear in the dependent.

One of the most salient characteristics of head-marking patterns at the clause level is that the verb can function as a sentence on its own. In effect, Nichols notes that "in head-marking grammatical relation, the dependent is usually an optional element of the constituent. For instance, in languages with consistently head-marked clauses, the verb itself normally constitutes a complete sentence; full NPs are included only for emphasis, focus, disambiguation, etc." (op.cit. 107). It is common for head-marking languages to

resist traditional grammatical questions based on dependent-marking syntax, since in head-marking languages the markers on the verb often display the grammaticalization of pragmatic and discourse relation, rather than strictly syntactic relations (Nichols 1986: 114)[57].

How is clausal-structure of head-marking and dependent-marking languages captured in RRG? Recall that one of the goals of RRG is to uncover the aspects of clause structure found in all human languages, providing consistent representations for comparable constructions, without imposing categories particular to one language onto a different one where these categories play no role. Finally, clause-structure commonalities are based on the actual form of each sentence, since the theory does not postulate abstract levels of syntactic representation.

In dependent-marking languages the dependents of the verb occur as NPs, and thus these NPs are core constituents in the layered structure of the clause. On the other hand, Van Valin (2005: 17ff) argues that, in head-marking languages, the markers on the verb are core constituents, whereas the optional NPs which are coreferential with them are clause constituents outside of the core.

Despite these differences, a comparable layered structure can be provided for both language-types. This is illustrated in Figure 14 (from Van Valin op.cit. 17), which represents the English and Tzotzil sentences in (153). (The Tzotzil examples are from

---

[57] The way in which dependency-markers are labeled often depends on the characterization of the language as head-marking or dependent-marking. For consistently head-marking languages, clause-level morphology is usually referred to as "cross-reference". For languages characterized as dependent-marking, the occurrence of clause-level head-marking morphology, indicating the relationship between the verb and the nominal argument it syntactically governs, is rather described as "agreement".

Aissen 1987). The structures differ in that arguments are realized by NPs (either full nominals or independent pronouns) in English (a dependent-marking language), and for pronominal affixes in Tzotzil (a head-marking language). In this last language type, coreferential NPs, if included, are extra-CORE constituents linked to the CLAUSE node.

(153) a.   He carried her.

   b.   ʔi-Ø-s-pet.
        ASP-3ABS-3ERG-carry
        'He/she carried him/her/it.'

   c.   ʔi-Ø-s-pet   lokel ʔantz   ti   tul-e.
        ASP-3ABS-3ERG-carry away woman DEF rabbit-DEF
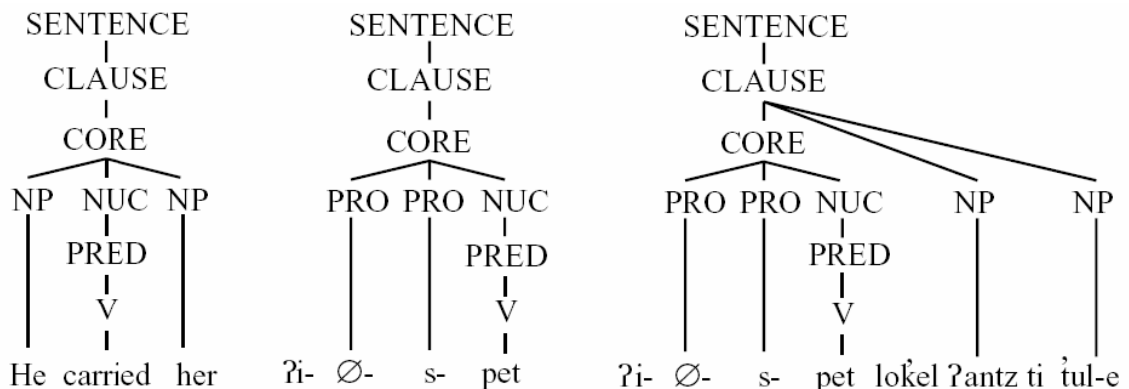   'The rabbit carried away the woman.'

Figure 14. *The Layered Structure of the Clause for head-marking and dependent-marking languages*

Head-marking languages and dependent-marking languages are the two "polar" types. There are, also, languages which mark the dependency twice, on both the head and the dependent. These are called "double-marking languages". Finally, there are languages

which show some head-marking and some dependent-marking patterns, and these are called "split-marking languages". It is typical for split-marking languages to exhibit head-marking patterns at the level of the clause, and dependent-marking patterns at lower levels.

One can argue that Spanish belongs to this last type, although this categorization is not absolutely clear-cut. Spanish displays unambiguous dependent-marking patterns at the level of the phrase (modification is marked on the dependent adjective; possession is marked on the dependent possessor). On the other hand, whether the language is considered as showing head-marking or dependent-marking pattern at the level of the clause depends, in part, on the status one grants to the affixes found on the verb[58].

In this regard, there are two extreme views: the "pure agreement view" and the "pure pronominal view" (Corbett 2003). The "pure agreement view" argues that if the verb displays person/number/gender features, the responsible affix should be considered as an agreement marker. In the version that requires agreement features to have local controllers (i.e. the clausal element that the feature agrees with), the required NP has to be inserted at some abstract level of representation. At the other end, the "pure pronominal view" argues that if an affix has pronominal features, then it must be analyzed as an incorporated pronoun, and therefore as outside the domain of agreement phenomena.

---

[58] A matter that is often also influenced by the particular tradition in which the analysis is based. As noted by Corbett, "languages in certain parts of the world are normally treated as having verb agreement and those in other parts as having pronominal affixes, and the distinction is not always one of substance. Moreover, given the gradient nature of the categories, there are many cases where the distinction may not be a reasonable one to draw" (2003: 21).

The categorization of verbal affixes is further complicated by the fact that some languages have so-called "pronominal affixes" which constitute a sort of intermediate category between clearer instances of agreement markers or pronouns. Corbett proposes a number of characteristics, typical of one or the other class, which can be used as a heuristic for distinguishing between pure agreement markers, pronominal affixes, and free pronouns[59].

For instance, pronominal affixes (but not free pronouns) are compatible with generic readings and may be used with indefinite objects. To the limited extent in which Spanish clitics co-occur with indefinite NPs or NPs receiving generic readings, they are set apart from more clearly pronoun-like forms. Spanish clitics are closer to pronouns, however, in their sensitivity to what Corbett calls "descriptive content". Prototypical agreement markers are expected to be immune to the referential properties of the controller. At the other end, there are languages in which pronouns are very limited in reference. Some languages, in turn, lie somehow in the middle: Iroquoian languages, for instance, mark grammatical patients only when they are animate, showing greater descriptive content than agreement markers (op.cit. 15), a behavior that resembles the constraints that govern the occurrence of clitics in Spanish.

Another heuristic concerns quantifier scope. It is argued that "the presence of markers with quantifiers like *every* and negative operators like *nobody* is inconsistent with an incorporated pronoun interpretation. If a language has such markers they must be grammatical agreement markers" (Corbett op.cit. 27). Since Spanish clitics may occur in

---

[59] Some of these heuristics were originally proposed in Bresnan and Mchombo (1987).

the target contexts (cf. §2.2.1; 5.1), they appear to be closer to agreement markers in this respect.

Based on data from Chicheŵa, Bresnan and Mchombo (1987) propose that only agreement markers (and not incorporated pronouns) can co-occur with questioned phrases. The rationale depends on the fact that, in Chicheŵa as in other Bantu languages, incorporated pronouns are linked via anaphoric agreement to the topic of the sentence (cf. §3.3), and therefore can not be focalized via a wh-question. We saw that in Spanish doubled question words are possible (although to a greater extent for datives than for accusatives). This suggests, in terms of Bresnan and Mchombo's criterion, that Spanish clitics may be located closer to agreement markers than to incorporated pronouns.

On the other hand, prototypical agreement marks only one argument, whereas pronominal affixes usually mark all the main argument positions. Thus, whereas the suffix that appear in Spanish finite verbs is a clearer instance of agreement, the clitics, targeting direct and indirect objects, are closer to pronominal affixes under this criterion.

To sum up, Spanish clitics share properties of both pronominal affixes and agreement markers (cf. also the discussion in §1.2), and therefore do not fit into the existing representations illustrated in Figure 14. As opposed to both typical dependent-marking and head-marking languages, in Spanish *neither* the dependent nominal nor the object marker on the verb are obligatory. Spanish differs from dependent-marking languages in that the markers on the verb suffice to create a grammatical sequence, and from head-marking languages in that the independent NPs, when they occur, cannot be treated as appositional adjuncts.

Given these facts, I propose for Spanish a clausal representation that differs both from the head-marking and the dependent-marking layered structures. In languages of this split-marking type, the argumental NPs appear as constituents of the core, whereas their pronominal features are linked to an "agreement index" node (AGX) inside the nucleus[60]. Consider the following sentence from the Buenos Aires corpus, presented along with one of its possible variants. The layered structure of both sentences is represented in Figure 15.

(154) a. El año pasado yo **la**   invité   **a**   **Mabel**.   =(83)a)
          the year past   I   ACC3FS invited.1S ANIM Mabel
          'Last year I invited Mabel.'

     b. El año pasado **Ø** **la**   invité   **Ø^{61}**.
          the year past   ACC3FS invited.1S
          'Last year I invited her.'

---

[60] The idea of an agreement index node was originally suggested to me by Dan Everett. The details of this representation were originally proposed in Belloro (2004a and 2004b), and adopted in Bentley (2006) and González Vergara (2006).

[61] The inclusion of the symbol Ø in these examples is meant to help graphically distinguishing between alternatives. It does *not* mean that I am positing an empty category of any sort. Indeed, a major claim of RRG is that there are not phonologically null elements (e.g. "PRO", "pro", "traces") in the syntax.

Figure 15. *The layered structure of the clause in "split-marking" languages*

The difference between (154)a) and (154)b) is that, in the second, both a (strong) pronoun and a nominal argument are omitted (i.e. *yo*, and *Mabel*). They are, accordingly, omitted from the layered structure of that sentence. The structure of the nucleus is the same in both clauses. I do not posit empty categories realizing the missing nominals; in the structure for (154), it is the bound forms which perform as arguments. This representation captures the functional ambiguity of Spanish verbal morphemes, while keeping the syntactic structure constant. At the same time, it captures the particularities that distinguish Spanish clause structure, as well as that of other so-called "pro-drop" or "split-marking languages", from both canonical head-marking and dependent-marking patterns. In other words, by positing this new node we can capture the fact that the verb-governed dependents can be omitted without affecting the grammaticality of the phrasal unit, while at the same time taking into account the relevance of the lexical NPs for all

sorts of grammatical phenomena that cannot be accounted for without explicit reference to their behavior.

We mentioned in §6.1.1. that, in RRG, the syntactic representation of a sentence is determined by a selectional principle, which targets a language-specific inventory of syntactic templates. The selectional principle is formulated as follow (Van Valin op.cit. 130):

(155)   Syntactic template selection principle:
         The number of syntactic slots for arguments and argument-adjuncts[62] within the core is equal to the number of distinct specified argument positions in the semantic representation of the core.

Whereas this principle is universal, there may be language-specific qualifications to it. For instance, English requires dummy subjects for verbs like *rain*, which lack any distinct specified argument position in the semantic representation. Thus, there is an English-specific qualification to the principle in (155), which states that all cores in the language have a minimum syntactic valence of 1.

Analogously, we need a Spanish-specific qualification which accounts for the missing syntactic slots in those cases where the arguments are exclusively realized morphologically. I propose the following formulation:

(156)   Spanish-specific qualification:
         Argument positions filled exclusively by feature bundles in the semantic representation of the core do not require syntactic slots in the core template.

---

[62] For a definition of the category "argument-adjunct" in RRG, see Van Valin and LaPolla (1997 §4.4.1.1, 7.3.2.2) and below (§6.2.3).

185

The principle in (155) together with the qualification in (156) guides the selection of the appropriate syntactic templates in the semantics-to-syntax linking. Differing from head-marking and dependent-marking languages, the syntactic inventory of split-marking languages contains a CORE template which includes the agreement index node[63].

6.2.2. Semantic Representation and Linking Rules

As mentioned in §6.1.2, RRG bases the semantic representation of the sentence on the lexical structure of the predicating element. Once the appropriate structure has been selected, each argument position is filled with the corresponding referential expressions, whose particular form is motivated by discourse-pragmatic considerations. We mentioned that arguments in the logical structure are linked to macroroles, according to the actor-undergoer hierarchy in Figure 7. Finally, the PSA selection hierarchy in Figure 8, and the selection principles in (155) and (156), capture the relation between macroroles and non-macrorole arguments and morphosyntactic function.

There is a series of rules that govern all these relations, which is called the "linking algorithm". The linking algorithm is in turn governed by a "completeness constraint", as stated in (157).

(157)   Completeness constraint:
         All of the arguments explicitly specified in the semantic representation of the sentence must be realized morpho-syntactically in the sentence, and all of the referring expressions in the syntactic representation of a sentence must be

---

[63]   For a specification of the inventory of Spanish syntactic templates in accordance with this proposal see González Vergara (2006).

linked to an argument position in a logical structure in the semantic representation of the sentence.

The linking algorithms are bidirectional. There are syntax-to-semantics algorithms and semantics-to-syntax algorithms, since the linking is meant to capture the production side and the comprehension side of language processing (Van Valin 2006). We will first present the linking algorithms for Spanish, and then illustrate their functioning with concrete examples.

(158)   Spanish Linking Algorithm: Semantics → Syntax:

    1. Construct the semantic representation of the sentence, based on the logical structure of the predicator and select the realization of each argument based on the activation level of its referent:

        a. If active, fill in the respective argument position with relevant bundle of pronominal features.

        b. If accessible, fill in the respective argument position with the corresponding nominal, plus its pronominal features.

        c. If inactive or non-identifiable, fill in the respective argument position with the corresponding nominal exclusively.

    2. Determine the actor and undergoer assignments, following the actor-undergoer hierarchy (Figure 7).

    3. Determine the morphosyntactic coding of the arguments:

        d. Select the PSA, based on the PSA selection hierarchy (Figure 6).

        e. Assign the XPs the appropriate case markers and/or adpositions, as indicated in (142) and (143).

    4. Select the syntactic template(s) for the sentence following the principles in (155) and (156).

5. Assign the elements in each argument positions to the appropriate slots in the syntactic template:

    a. Assign pronominal features to the AGX.

    b. Assign nominals to the appropriate positions in the clause, subject to focus structure.

       i. Assign focal elements to the last position in the core (default)

    c. Assign any [+WH] arguments to the precore slot.

(159)   Spanish Linking Algorithm: Syntax → Semantics:

1. Determine the macrorole(s) and other core argument(s) in the clause.

    a. If the verb is intransitive, then assign the privileged syntactic argument macrorole status.

    b. If the verb is transitive, determine its voice.

       i. If the voice is active, assign actor status to the PSA and undergoer status to the other direct core argument.

       ii. If the voice is passive, the undergoer is PSA.

          1. Assign actor status to the object of *por* in the periphery of the core.

          2. If there is no actor in the periphery, then replace the variable representing the highest ranking argument in the logical structure with 'Ø'.

    c. Determine the linking of the pronominal features of the predicate in the nucleus:

       i. Link the nominative features to the actor.

       ii. Link the accusative features to the undergoer.

       iii. Link the dative features to the non-macrorole.

2. Retrieve from the lexicon the logical structure of the predicate in the nucleus of the clause and with respect to it determine the actor and undergoer assignments, following actor-undergoer hierarchy Figure 7.

a. If there is a non-macrorole core argument, link it with the first argument position in the state predicate (default).

3. Link the arguments determined in step 1 with the arguments determined in step 2 until all core elements are linked.

4. If there is an element in the pre- or postcore slot,

i. Assign it the remaining unlinked argument position in the semantic representation of the sentence.

ii. If there are no unlinked argument positions in the sentence, treat the wh-word like a predicative preposition and follow the procedure in step 4, linking the wh-word to the first argument position in the logical structure.

We can now illustrate the workings of the semantics-to-syntax linking algorithm with one of the structures introduced in Figure 15. The relevant sentence is repeated below.

(160) El año pasado yo **la**   invité   **a   Mabel**.   =(154)a)
       the year past   I   ACC3FS invited.1S ANIM Mabel
       'Last year I invited Mabel.'

According to the Spanish linking algorithm presented in (158), the first step is to construct the appropriate semantic representation of the sentence. In these sentences we have an activity verb, and therefore we select [**do'** (x, [**invite'** (x, y)])]. We need to add, to this abstract representation, the elements that fill each argument position, together with their activation levels. Both referents are *accessible*: the speaker because despite being directly accessible from the speech situation is not the focus of attention; Mabel because, although new to the discourse-context, is a common acquaintance of both interlocutors. In this sentence we also have a temporal modifier. RRG considers them as predicative

elements, taking the logical structure of the core as their argument. The output of the first step in the linking is thus the following structure:

last year'[**do'**(yo[1S]$_{ACS}$,[**invite'**(yo[1S]$_{ACS}$,Mabel[3FS]$_{ACS}$)]).

The argument of **do'** (coindexed with the first argument of **invite'**) is doubly filled, by the strong pronominal *yo* and the features first person and singular number. Analogous double filling occurs with the second argument of **invite'**, were we have the noun *Mabel* and the features third person, feminine gender, singular number. The double filling of these argument positions reflects the accessible status of the corresponding referents, in accordance with the linking algorithm presented in (158). (The accessible status of the referents is expressed in the ACS subscript beside each argument position.)

After having built the appropriate semantic representation of the sentence, the second step in the linking is to determine the actor and undergoer assignments, following the actor-undergoer hierarchy in Figure 7. First argument of "**do'** (x…" is the unmarked choice for actor, and thus *yo[1S]* is assigned actor status. The second argument of **pred'** (x, y) is the unmarked choice for undergoer, and then undergoer status is assigned to *Mabel[3FS]*.

The third step is to determine the morphosyntactic coding of the arguments. According to the PSA selection hierarchy in Figure 8, *yo[1S]* is selected as PSA and assigned nominative case; whereas *Mabel[3FS]* is assigned accusative. The fourth step is to select the syntactic templates, following the principles in (155) and (156). We have two doubly filled argument positions in the logical structure, and thus we need a core template with two syntactic slots. Because of the temporal modifier, we also need a

periphery template, which is merged to the core. The result is the layered structure in Figure 16.
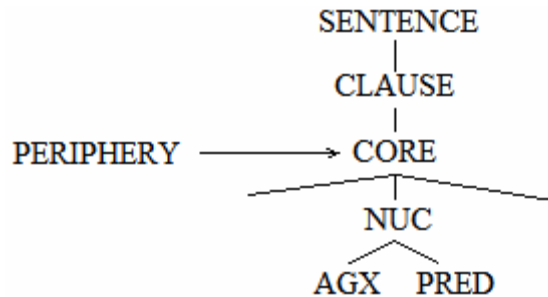
```
                              SENTENCE
                                 |
                              CLAUSE
                                 |
        PERIPHERY ────────→   CORE
                              ───┴───
                               NUC
                              ╱    ╲
                           AGX    PRED
```

Figure 16. *Syntactic template*

The fifth and final step concerns the linking of the elements in each argument position to the appropriate slots in the syntactic template. The temporal modifier is linked to the periphery, and the two NPs are linked to the two core internal positions. Due to the typical structure of Spanish predicate-focus constructions, the topical NP *yo* is linked to the first position in the core, and the focal NP *Mabel* is linked to the last core-internal position. Crucially, as proposed in §6.2.1., the pronominal features are linked to AGX, where they are morpho-phonologically realized. The features present in the actor position (i.e. first person, singular number), marked nominative, are realized as a verbal suffix *-é*, whereas the features from the undergoer (i.e. third person, feminine gender, singular number), marked accusative, are realized as a preverbal clitic *la*[64].

---

[64] In a recent article Kailuweit (in press) challenges the idea of linking the pronominal features marked nominative to the AGX. He argues that "subject agreement must be sharply separated from objet agreement" since "there is no evidence for a combination or interference of the two phenomena". Recall, however, the aforementioned tendency, in colloquial speech, to attach

As it was argued in Chapter 4, the occurrence of the accusative clitic depends on the activation status of the relevant referent. In this particular example the relevant argument position is marked as accessible, and therefore the clitic co-occurs with the NP in the clausal structure. The result of all the steps in the linking process is represented in Figure 17.
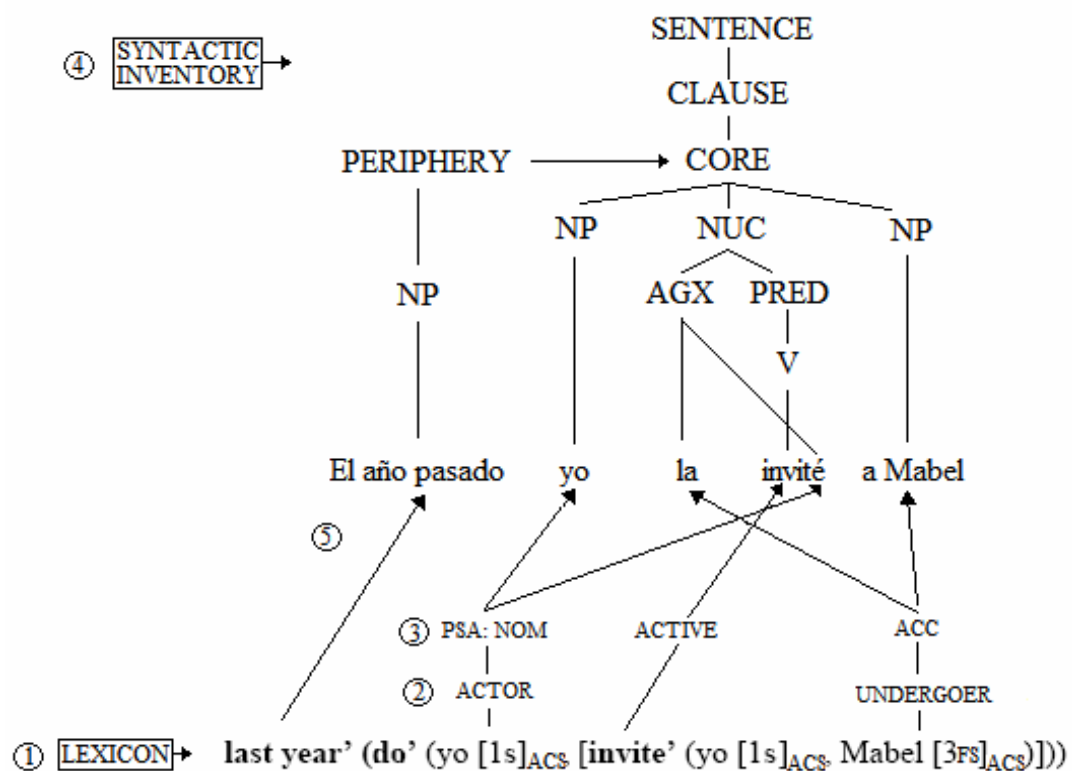

Figure 17. *Linking diagram for a DO-doubling structure*

The linking of either of the non-DO doubled versions of this sentence proceeds analogously. The two possible alternatives are presented in (161), and illustrated in Figure 18 (the NP-only version) and Figure 19 (the clitic-only version).

---

the plural "subject agreement" morpheme to the cluster formed by the verb and the clitic mentioned in §1.2.

(161) a. El año pasado yo invité **a Mabel**.
      the year past   I   invited.1S ANIM Mabel
      'Last year I invited Mabel.'

  b. El año pasado **la** invité.   =(154)b
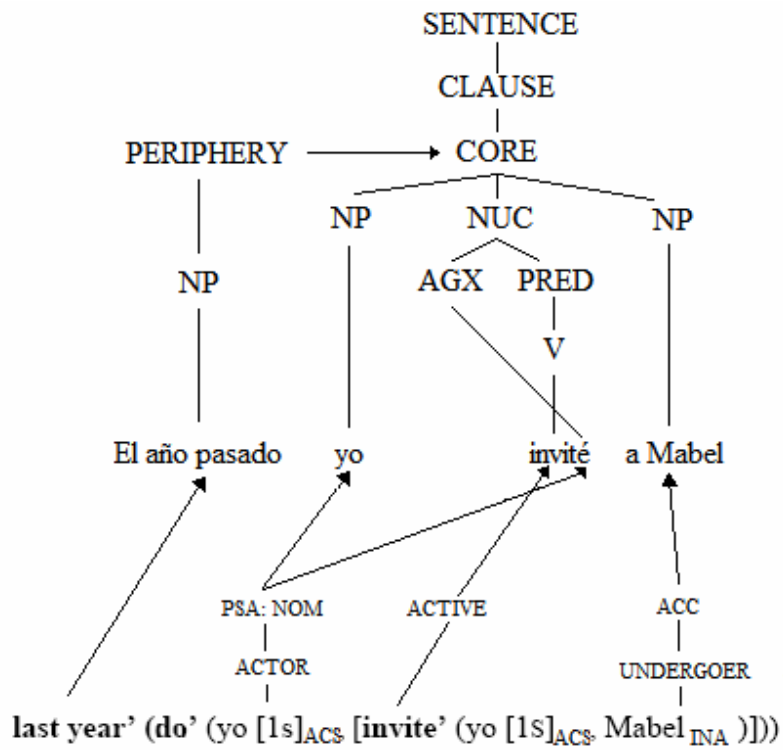   the year past   ACC3FS invited.1S
      'Last year I invited her.'

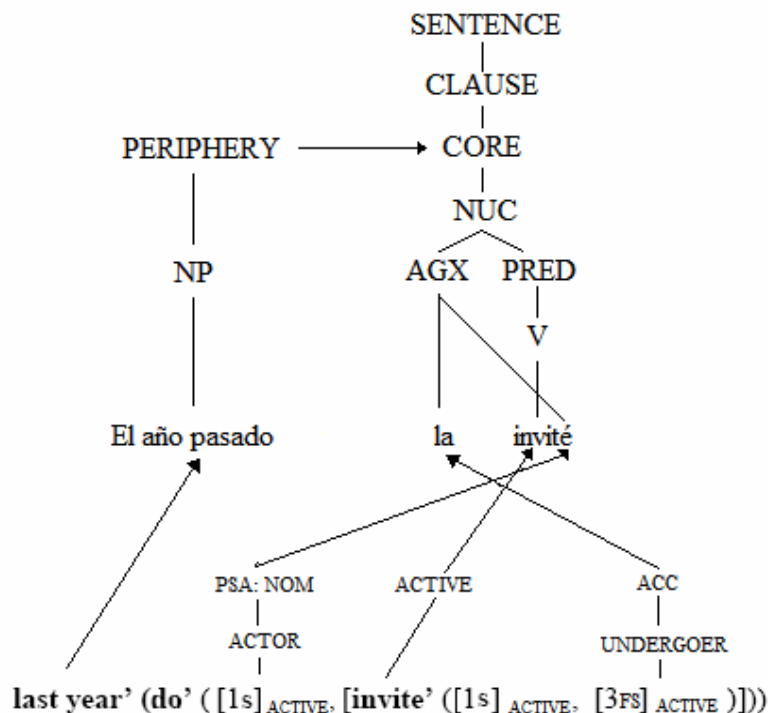Figure 18. *Linking diagram for a non DO-doubling structure (NP-only version)*

Figure 19. *Linking diagram for a non DO-doubling structure (CL-only version)*

Consider the logical structure corresponding to the sentence in (161)a) (Figure 18). Note that the argument in the second position of the **invite'** predicate, i.e. the undergoer, is marked as inactive. This accounts for the fact that the pronominal features corresponding to this position cannot be realized, and only those contained in the argument of **do'**, i.e. the actor, will appear in the morphosyntaxtic structure.

In the logical structure of (161)b), on the other hand, the undergoer argument is marked as active, and it only contains a bundle or pronominal features. Note that in this version the actor argument is marked as active as well; and therefore the relevant position in the logical structure is also filled exclusively by pronominal features. According to the qualification presented in (156), argument positions filled exclusively by feature bundles

diminish the number of syntactic slots required in the core template. Thus, we need to select a core template containing just the nucleus. The nucleus is composed of the predicate and the AGX, where the pronominal features of the actor and undergoer arguments are realized, and serve to satisfy the completeness constraint.

In terms of the realization of the accusative argument, note that the crucial structural difference between the sentences represented in Figure 18 and Figure 19, as well as between them and the DO-doubled structure in Figure 17, arises from the different activation status of the accusative argument in each case. This difference in activation status motivates the different encodings, and thus this representation allows us to capture the pragmatic correlations postulated at the end of Chapter 4.

In Chapter 5 I discussed the pragmatics of dative doubling, and we saw that the unmarked realization of dative arguments is by means of clitics alone, whereas independent NPs typically serve to introduce new referents or to disambiguate the intended target between competing participants. On the other hand, there are marked cases in which the referent is particularly low in identifiability and the clitic is omitted. Thus, even though between datives and accusatives the markedness of each encoding type is virtually the opposite, they converge in doubling constructions, which in both cases emerge as an "intermediate" point that encodes a deviation from the expected default.

We can now turn to the linking of structures containing a dative argument. We will do so based on one of the corpus examples provided by Maldonado (2002). Consider first the sentence in (162)a) and its alternatives in (b) and (c).

(162) a. Dieron   un asueto **a   los trabajadores**.   (based on (137)a))
          gave.3PL a day.off   DAT the workers
          'They gave the workers a day off.'

    b. **Les**   dieron   un asueto **a   los trabajadores**.
          DAT.3PL gave.3PL a day.off   DAT the workers
          'They gave the workers a day off.'

    c. **Les**   dieron   un asueto.
          DAT3PL gave.3PL a day.off
          'They gave them a day off.'

We need to create an appropriate logical structure for each sentence. All involve a predicate of transfer, and more precisely a causative accomplishment. Thus, the logical structure of the predicates are represented by the general structure in (163).

(163)   [**do'** (x, Ø)] CAUSE [BECOME **have'** (y, z)]

The logical structures of each of the sentences in (162) differ, however, on the relative accessibility of the dative argument, and therefore also on how the position for this argument (i.e. the first variable on the BECOME **have'** component) is filled in each case[65]. With respect to (162)a), we saw that, for dative arguments, the clitic may be

---

[65] Considering the actor participant of so-called "third person impersonal constructions" as *active* is an oversimplification. It is encoded *as if* active, and thus I will code it thus here, leaving for further studies a more insightful account of its cognitive status. The "direct object" realized as the indefinite *un asueto* is inactive in all three cases, which accounts for the fact that it will be encoded exclusively as an NP (i.e. not in a DO-doubling construction).

omitted only if the corresponding referent is particularly low on identifiability. Thus, in this case this position must be marked either inactive or brand-new. (Recall Table **10**, presented in page 153). We will code the relevant referent here as inactive, allowing the corresponding argument position to be filled exclusively by the nominal *trabajadores*. The resulting logical structure is presented in (164)a below.

In the sentences in (162)b we have a dative doubling construction. I argued that these constructions are used to encode referents relatively higher in accessibility[66]: we can thus consider the relevant argument position in this case as marked accessible. In the logical structure this position is consequently filled by both the nominal *trabajadores* and the pronominal features third person, masculine and plural, as in (164)b below.

Finally, the correlations between topicality, activeness and attenuated encoding examined in §6.1.3 justify considering the dative argument (162)c as active. The argument position in this case is then filled exclusively by the corresponding pronominal features, as in (164)c.

(164) a. [**do'** ([3PL]$_{ACT}$, Ø)] CAUSE [BECOME **have'** (trabajadores$_{INA}$, asueto$_{INA}$)]

b. [**do'** ([3PL]$_{ACT}$, Ø)] CAUSE [BECOME **have'** (trabajadores[3MPL]$_{ACS}$, asueto$_{INA}$)]

c. [**do'** ([3PL]$_{ACT}$, Ø)] CAUSE [BECOME **have'** ([3MPL]$_{ACT}$, asueto$_{INA}$)]

The structures in (164) count as the output of the first step of the semantic-to-syntax linking. The second step is to determine the actor-undergoer assignment.

---

[66] I.e., relatively higher *with respect to the higly marked NP-exclusive encoding*. As argued in Chapter 5, these constructions mark referents as relatively lower in accessibility than expected for dative arguments, if we take into account the typical association between datives, topicality and clitic-only encoding discussed in that Chapter.

Following the hierarchy in Figure 7, the first argument of **do'** (x, y) is selected as the actor macrorole, and the second argument of BECOME **have'** (y, z) as the undergoer, leaving the remaining argument as a non-macrorole. The third step in the linking involves the selection of the PSA and the assignment of case. PSA status falls on the first argument of **do'** (x, y), which is assigned nominative case. There is no lexical PSA, and thus the case features are assigned exclusively to the pronominal features. Following the case assignment rules in (142), the highest ranking core macrorole (i.e. the actor) receives nominative case, the other macrorole receives accusative, and the non-macrorole receives dative.

The fourth step in the linking entails the selection of the appropriate syntactic templates. The three sentences contain three specified argument positions their semantic representation. In (164)a) and (164)b), however, one of these argument positions is filled *exclusively* by feature bundles, and thus for these structures we select core templates with only two syntactic slots. The structure in (164)c) contains two argument positions filled exclusively by feature bundles, and therefore for this structure we select a core template with just one syntactic slot.

The final step in the linking implies the assignment of the elements in each of the argument positions to the appropriate slots in the syntactic template: the pronominal features to the AGX, and the nominals to the NP nodes. The result of the linking for the three structures is presented below.
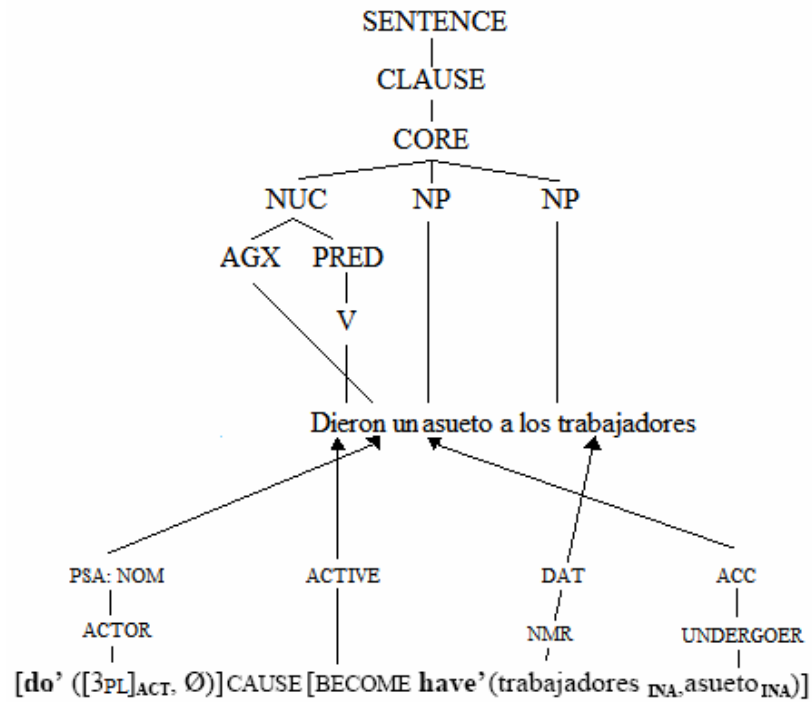
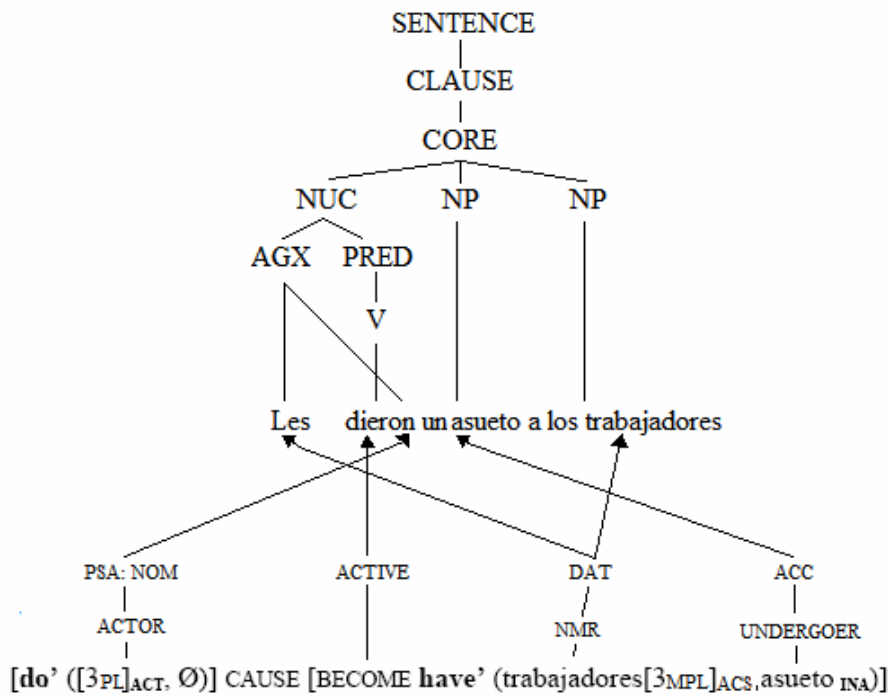Figure 20. *Linking diagram for non IO-doubling structure (NP-only version)*

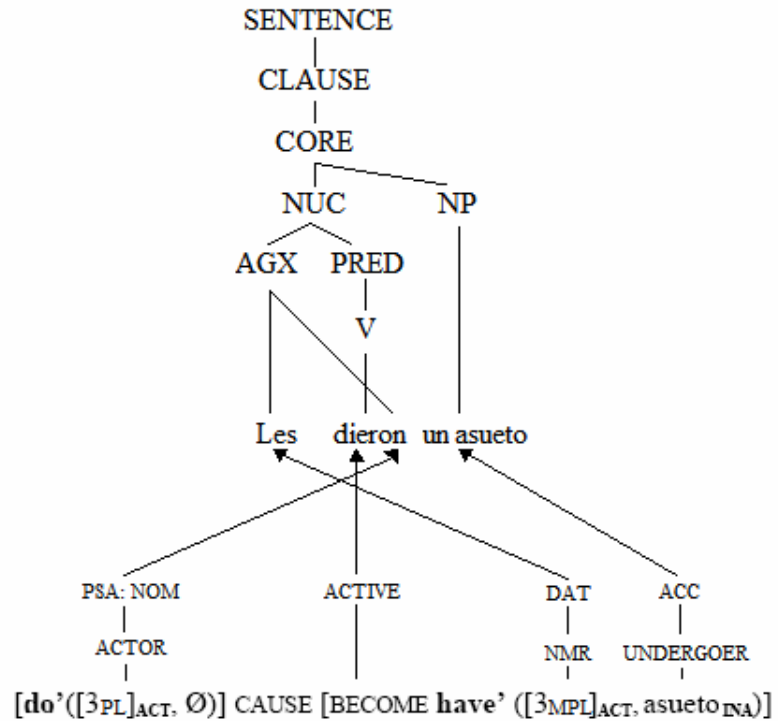Figure 21. *Linking diagram for an IO-doubling structure*

SENTENCE
|
CLAUSE
|
CORE
|
NUC   NP
|
AGX   PRED
|
V

Les   dieron un asueto

PSA: NOM   ACTIVE   DAT   ACC
| |
ACTOR   NMR   UNDERGOER
| |
[**do'**([3PL]ACT, Ø)] CAUSE [BECOME **have'** ([3MPL]ACT, asueto INA)]

Figure 22. *Linking diagram for a non IO-doubling structure (CL-only version)*

6.2.3. A brief note on "incorporated" datives

In Chapter 5 we saw that datives come in two varieties, depending on whether they are required by the verb or not. Let us call them "required" datives and "incorporated" datives, just to facilitate the discussion. According to the literature, the two groups differ in that "required" datives are the only ones that can be encoded exclusively as NPs. The linking example we illustrated step-by-step above, thus, represents the most complex case, because it includes an encoding alternative which is only available for a particular group of verbs. The linking of "incorporated" datives is simpler, since the pronominal

features of the relevant argument are always realized (i.e. there is no NP-exclusive encoding).

We should address, however, how this "incorporation" of extra participants can be captured in RRG. The addition of an argument not required by the structure of the predicate entails a change in the semantic representation of the sentence, and in fact Van Valin (2005:59) notes that, cross-linguistically, one of the reasons why logical structures are expanded is "to specify the source, path (or means) and goal (or recipient) with motion active accomplishment predications and transfer predications". Here I will briefly discuss the cases of *comprar* 'buy' and *cocinar* 'cook'[67].

In the literature on the subject, *comprar* is often considered a three-place verb; that is, a verb requiring a dative as part of its valence (Demonte 1994, Gutiérrez Ordoñez 1999). This is probably based on the fact that *comprar* encodes events of transfer of possession; and we mentioned that predicates of transfer were categorized as requiring a third participant. Accordingly, *comprar* is assumed to be able to occur with dative arguments encoded exclusively in lexical form; i.e. it is assumed to behave like *dar* (165).

(165)   Carolina **Ø/le** compró/dio un libro **a Esteban**.
            'Carolina bought/gave Esteban a book.'

Comprar displays, however, two features that distinguish it from *dar*, and likens its datives to the "incorporated" class. The first is that it does not require the third participant

---

[67] A complete characterization of Spanish datives is beyond the scope of this work. The references discussed in Chapter 5 represent the main proposals that have been put forth in the literature, and can be consulted for a more comprehensive analysis of the issues involved.

to occur in the syntactic structure (166). The second, that it accepts alternative encodings for the (intended) recipient, which can appear either marked by *a* or by *para*, an alternative that is not available for *dar* (167).

(166) a.   Carolina compró un libro.
     'Carolina bought a book.'

   b. *Carolina dio un libro.
     'Carolina gave a book.'

(167) a.   Carolina compró un libro **para Esteban**.
     'Carolina bought a book for Esteban.'

   b. *Carolina dio un libro **para Esteban**.
     'Carolina gave a book for Esteban.'

Within the RRG framework, the alternative encodings of the (intended) recipient that may occur with *comprar* can be derived from the different logical structures of the alternative construction, as illustrated in the examples below.

(168)   Carolina compró un libro.
     'Carolina bought a book.'

     [**do'** (Carolina, Ø)] CAUSE [BECOME **have'** (Carolina, libro)]

(169)   Carolina compró un libro **para Esteban**.
     'Carolina bought a book for Esteban.'

     [[**do'** (Carolina, Ø)] CAUSE [BECOME **have'** (Carolina, libro)]]
                    & PURP [BECOME **have'** (Esteban, libro)]

(170)   Carolina **le** compró un libro **a Esteban**.
         'Carolina bought Esteban a book.'

[[**do'** (Carolina, Ø)] CAUSE [BECOME **have'** (Carolina, libro)]]
                                      & CAUSE [BECOME **have'** (Esteban, libro)]

The difference between (168) and (169) is that, in the second, the extra "purposive" component[68] captures the addition to the structure of an intended recipient (i.e. a "beneficiary"). Since this participant (i.e. *Esteban*) occurs in the PURP component, the conditions for preposition assignment in (143) are met, and this argument is marked by *para*. And since there is a shared argument in the CAUSE and the PURP components (namely, *libro*), the prepositionally marked argument (i.e. *Esteban*) is linked as an oblique "argument-adjunct" to the core[69]. (Van Valin and LaPolla 1997: 383).

In the structure in (170), on the other hand, there is also an extra component, but it is marked by the CAUSE operator. Therefore, the case assignment rules presented in (142) determine that it receives dative case; it is thus marked by *a* and linked to the core as a *direct argument*. Since the pronominal features of an argument position assigned dative case are linked to the AGX, they are also, by default, realized as a dative clitic. Recall that only direct arguments (i.e. marked nominative, accusative and dative) are linked to the AGX. If the constituent is linked as an oblique it cannot serve as controller for verb agreement, which captures the fact that the dative clitic is blocked in cases like (169). Note that, in the RRG system of semantic decomposition, the assumed variation in the

---

[68] Recall that PURP is the abbreviation of the purposive component expressed by the logical structure **want'** (x, LS_2) & DO (x, [[LS_1 ◊ CAUSE [LS_2]]) (cf. § 6.1.2, p.162).
[69] The phrase introduced by *para* in these examples is treated as an "argument-adjunct" because of the shared argument (*libro*) in the logical structures of the two components (Van Valin and LaPolla 1997, §4.4.1, 7.3.2.2).

affectedness of the third participant in each type of structure (i.e. the claim that *Esteban* is more affected in (170) than in (169)) can be derived from the semantic components that form the logical representation of each sentence, and that determines whether this participant will occur as a direct (more affected) or as an oblique (less affected) constituent.

We can now turn to the examination of *cocinar* 'cook'. *Cocinar* belongs to the class of "creation" predicates, according to Demonte's classification (1994). There is no claim in the literature that this verb may require a dative argument, or that, if there is a dative in the structure, this argument may be encoded exclusively as (non-doubled) lexical phrase. In (171) there is a representation of this predicate in a canonical "transitive" structure. In (172) the third participant (the intended recipient of the *paella*) is introduced under the PURP logical structure. The conditions for preposition assignment are met, and this participant is thus marked by *para* and linked to the core as an oblique argument-adjunct. In (173), in turn, the third participant is introduced by a CAUSE operator, assigned dative case, and linked to the core as a direct argument. The linking of this participant as a direct argument licenses its pronominal features to be realized by the verb, something that was blocked in the previous case.

(171)   Carolina cocinó una paella.
        'Carolina cooked a paella.'

        [**do'** (Carolina, [**cook'** (Carolina, paella)])] & [INGR **exist'** (paella)]

(172)   Carolina cocinó una paella **para Esteban**.
'Carolina cooked a paella for Esteban'

[[**do'** (Carolina, [**cook'** (Carolina, paella)])] & [INGR **exist'** (paella)]]
PURP [BECOME **have'** (Esteban, paella)]

(173)   Carolina **le** cocinó una paella **a Esteban**.
'Carolina cooked Esteban a paella.'

[[**do'** (Carolina, [**cook'** (Carolina, paella)])] & [INGR **exist'** (paella)]]
CAUSE [BECOME **have'** (Esteban, paella)]

Comparing the examples of *dar* 'give', *comprar* 'buy' and *cocinar* 'cook' we see then that, in RRG terms, the presence of a "recipient" dative argument results in three different patterns, depending on how the event(s) of transfer are encoded in the semantic structure. I repeat the relevant examples below.

(174)   Carolina **le** dio un libro **a Esteban**.
'Carolina gave Esteban a book.'
[**do'** (Carolina, Ø)] CAUSE [BECOME **have'** (Esteban, libro)]

(175)   Carolina **le** compró un libro **a Esteban**.
'Carolina bought Esteban a book.'
[**do'** (Carolina, Ø)] CAUSE [BECOME **have'** (Carolina, libro)]
CAUSE [BECOME **have'** (Esteban, libro)]

(176)   Carolina **le** cocinó una paella **a Esteban**.
'Carolina cooked Esteban a paella.'
[**do'** (Carolina, [**cook'** (Carolina, paella)])] & [INGR **exist'** (paella)]
CAUSE [BECOME **have'** (Esteban, paella)]

In the example with *dar*, the "recipient" is part of the basic logical structure associated with the predicate. On the other hand, the "clitic doubled" structures with *comprar* and *cocinar* entail the addition of extra causative component in the logical

structure to represent the "recipient". (Note further that there is no possible causal component that could be added to the representation of *dar*).

To end this section, I would like to address two arguments commonly made in the literature in relation with dative doubling constructions. The first is that these constructions are analogous to "dative shift" constructions in languages like English (Demonte 1995, Cuervo 2003, among others). From an RRG perspective, dative shift involves a marked choice for undergoer, targeting the "recipient" rather than the "theme". A reflection of this is the possibility of having a passive construction where the marked undergoer functions as PSA (177). This is not an option in Spanish, where passive PSAs are always the lowest ranking argument according to the PSA selection hierarchy (178)[70].

(177) a. Some governments$_{[A]}$ grant special protection$_{[U]}$ to the elderly$_{[NMR]}$.
 b. Some governments$_{[A]}$ grant the elderly$_{[U]}$ special protection$_{[NMR]}$.

 a'. Special protection$_{[PSA]}$ is granted by some governments to the elderly.
 b'. The elderly$_{[PSA]}$ are granted special protection by some governments.

(178) a. Algunos gobiernos$_{[A]}$ otorgan una protección especial$_{[U]}$ a los ancianos$_{[NMR]}$.
 b. Algunos gobiernos$_{[A]}$ **les** otorgan una protección especial$_{[U]}$ **a los ancianos**.

 a'. Una protección especial$_{[PSA]}$ es otorgada a los ancianos por algunos gobiernos.
 b'.*Los ancianos$_{[PSA]}$ son otorgados una protección especial por algunos gobiernos.

---

[70] Spanish allows marked undergoer selection for some verbs with locative alternation, as mentioned in p. 206 and illustrated below. Note that the locative alternation does allow for either non-actor argument to function as PSA in a passive construction, differently from what happens in dative contructions, such as those in (177) and (178).
(i) Juan cargó las cajas en el camión.   (i)' Las cajas fueron cargadas en el camión.
  'Juan loaded the boxes in the truck.'   'The boxes were loaded in the truck.'
(ii) Juan cargó el camión con las cajas.   (ii)' El camión fue cargado con las cajas.
  'Juan loaded the truck with the boxes.'   'The truck was loaded with the boxes.'

An RRG analysis of dative doubling constructions as an instance of dative shift would have to account for the fact that the doubled arguments do not behave as undergoers, and that they are marked by dative case, reserved in Spanish for non-macrorole core arguments. One of the reasons why this analysis is tempting is that it captures the greater degree of affectedness over the third participant (marked by *a*), as compared with those cases in which the third participant is encoded as the object of a preposition (i.e. marked by *para*). For languages such as English, which allow marked choices for undergoer with *give*-type verbs, the greater degree of affectedness of the "recipient" is a result of its macrorole status in "dative shift" constructions. For languages such as Spanish, which do not allow a marked choice for undergoer with these verbs, however, the RRG framework still makes it possible to capture these potential semantic contrasts in terms of the distinction between direct arguments and oblique argument-adjuncts, without the need of assimilating Spanish "dative doubling" to English "dative shift".

This brings us to the second argument commonly found in the literature, which is that the dative clitic is an applicative morpheme, with the function of licensing the occurrence of additional core elements (e.g. Ibáñez 2002, Cuervo 2003). Traditional analyses of applicative constructions posit that they entail a marked choice of "direct object" and, in RRG terms, a marked undergoer assignment. In Spanish there does not seem to be strong evidence of the promotion of a non-argument to undergoer status, nor for the "demotion" of the canonical undergoer. There is, however, evidence that non-objects may be promoted to *direct core argument* status, as discussed above. In effect, the

Spanish facts resemble those reported from Kinyarwanda, where non-arguments become direct core arguments but not undergoers; or, in its morphology, those reported from Sama, where "incorporated" beneficiaries trigger the addition of a bound form on the verb (see Van Valin 2005: 121 for discussion and examples). In other words, the occurrence of the dative clitic with *a*-marked recipients can be understood not as a cause but as a consequence of greater affectedness over these participants, as compared with prepositionally marked ones. In other words, it does not license the occurrence of a dative NP (which results from the addition of another component to the logical structure plus the application of the corresponding linking rules), but follows from the tighter semantic / pragmatic linkage that this addition reflects.

6.3. SUMMARY AND DISCUSSION

In this Chapter I presented a Role and Reference analysis of DO- and IO-doubling structures in Spanish. As mentioned in the introduction to this Chapter, Role and Reference Grammar pays particular attention to the potential effect of discourse-pragmatics on morpho-syntactic phenomena. In effect, it considers that "[f]rom the lexicon to the syntactic inventory and to the different steps in the linking algorithms that relates them, discourse-pragmatics can influence every aspect of grammar, and a major claim of RRG is that an important source of cross-linguistic variation is the differences in this interaction" (Van Valin 2005: 182). The following diagram represents the different levels at which the "cognitive model of context" may interact with the different steps in the linking rules (from op.cit. 182):
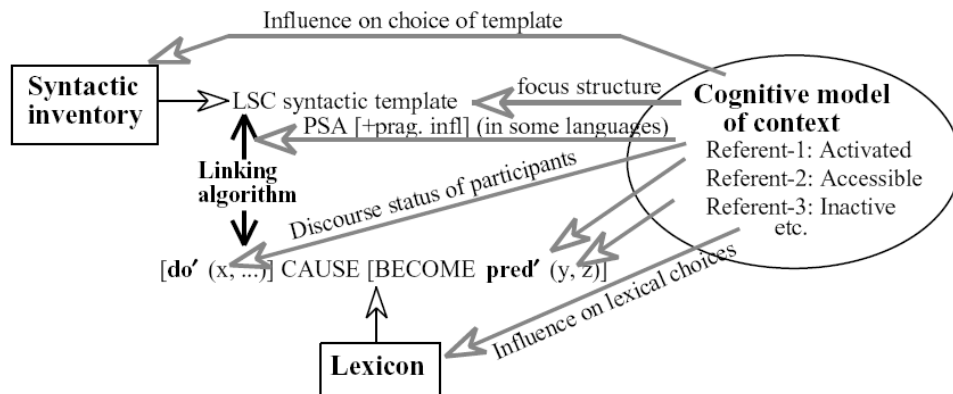
Figure 23. *The interaction of discourse-pragmatics and grammar in RRG*

The theory is explicit with respect to the potential effects of pragmatic information on the morpho-syntactic structure (Van Valin and LaPolla 1997, Chapter 5; Van Valin 2005, Chapter 3); and it incorporates this information in the linking algorithms, as represented in the diagram above. In this regard, this model is particularly well suited for capturing phenomena such as Spanish clitic doubling constructions, heavily affected, as argued in this dissertation, by the pragmatic status of the discourse participants. Thus, in this Chapter I proposed Spanish-specific linking rules that account for the realization of the core participants, and a new node in the layered structure where pronominal features are realized. I suggested that the layered structure of the Spanish clause should be different from the one proposed for typical head-marking and dependent-marking languages, inasmuch as it should reflect the status of the language as split-marking (Nichols 1986). These additions to the general machinery provided by the theory, I suggest, allow us to capture the particularities of Spanish clitics in a way that is consistent with the findings that result from the analysis of corpus data.

7   Conclusions

The main goal of this dissertation has been to provide a functionally-oriented analysis of Spanish clitic doubling based on data from naturally occurring texts, and to advance a grammatical analysis of these constructions informed by the empirical findings. After presenting a general overview of Spanish clitic doubling constructions and the challenges they pose, I discussed the main proposals that have been advanced for explaining them, both from formal and functional perspectives.

Theoretical approaches have proposed that DO-doubling is more constrained than IO-doubling due to an extra semantic constraint holding over doubled DOs; namely, that DOs must be specific / presupposed. These studies suggest that since doubled IOs do not require any similar semantic value to be "matched", IO doubling is less constrained. On the other hand, corpus-based studies have related clitic doubling with topicality, arguing, in general terms, that IO-doubling is more widespread than DO-doubling because of the inherent greater topicality of dative arguments. Thus, DO-doubling is limited to (or at least typical with) topical participants, in the sense of participant which have been previously referred to in the text.

In fact, similar features (specificity, topicality) have been claimed to be at play in determining the possibility of doubling constructions across Spanish dialects, as well as in other languages of the same and different families. I reviewed claims about clitic doubling (in particular DO-doubling) constructions in the Spanish of Spain, Chile, Perú, Ecuador, Uruguay, and México, showing that even though Argentinean Spanish is the most paradigmatic "doubling" dialect, these structures are by no means limited to it.

I then examined more carefully the predictions made by the different hypotheses and evaluated them with respect to a corpus of interactional speech. We saw that in fact none of them could successfully account for the data, and I proposed that, in fact, a clearer picture of the principles involved in clitic doubling constructions arises from an analysis of the particular discourse context where these constructions occurr. I suggested that DO-doubling marks referents as more active than expected for lexically realized DOs in focus position: specifically, as "accessible". On the other hand, I proposed that IO-doubling should also be analyzed as a deviation from the canonical realization of datives, which are typically encoded as clitics alone. Under this perspective, then, IO-doubling constructions also serve to mark referents which deviate from the expectation; namely those which are less active than typically anticipated for a dative argument (arguably, those which can be considered "accessible" or "inactive" under Chafe's classification).

Spanish clitics have an ambiguous grammatical nature, sharing characteristics with both pronouns and agreement markers, as discussed in Chapter 6. In that Chapter I argued that, without denying this ambiguity, it is nevertheless possible to incorporate clitics (and clitic doubling constructions) into a grammatical analysis consistent with the general characteristics of the language and, at the same time, the pragmatic function that doubling constructions are argued to fulfill. A proposal for such an analysis was advanced within the framework of Role and Reference Grammar.

I conceive of the hypotheses defended here as a refinement on the ones currently proposed in the literature; and as widely compatible, I believe, with the general insights behind them. At the same time, however, one of the goals of this study has been to start

bridging the gap that often exists (as far as the study of clitic doubling is concerned) between theoretical and corpus-based proposals, by presenting a formal account of these constructions based on the categories uncovered by naturally-occurring data.

There is, however, much work that still needs to be done. Among the most relevant tasks, two deserve special attention. First, there is the question of distinguishing between "clitic doubling" and so-called "clitic right-dislocation" constructions. As mentioned in Chapters 2 and 3, there is a series of grammatical tests that have been proposed for determining whether a language allows "clitic doubling", at least in some contexts, or not. However, these tests are of limited use for deciding over particular tokens encountered in natural discourse, and this difficulty is stressed by the alleged optionality, in some languages, of disruptions in the intonational contour between phrases typically considered as "dislocated" and the rest of the clause. In this regard, the different pragmatic meaning of "doubled" vs. "dislocated" phrases can serve as a basis for deciding between the two possible analyses in specific contexts, but there is more research that needs to be done for distinguishing among the pragmatic meanings of "dislocated" constituents of different sorts (cf. Lambrecht 1984, Grosz and Ziv 1997, Averintseva-Klinsch 2006, among others).

On the other hand, there is also the need for a more precise measurement of the exact level of cognitive accessibility a referent encoded as direct object or indirect object must have, according to the hypothesis defended here, in order to be involved in a doubling construction. An answer to this question requires an experimental protocol

devised for testing the different levels of activation of "doubled" vs. "non-doubled" direct

and indirect objects in Spanish, and this is a task I leave open for the future.

References

Abbott, B. (1993). "Referentiality, specificity, strenght & individual concepts." <u>West Coast Conference on formal Linguistics</u> 12.

Abbott, B. (1995). "Some remarks on specificity." <u>Linguistic Inquiry</u> 26: 341-347.

Abbott, B. (2000). "Presuppositions as nonassertions." <u>Journal of Pragmatics</u> 32: 1429-1437.

Abbott, B. (2005). Definiteness and Indefiniteness. <u>Handbook of Pragmatics</u>. L. R. H. a. G. Ward. Oxford, Blackwell.

Abbott, B. (2006). Definite and Indefinite. <u>Encyclopeadia of Language and Linguistics, 2nd edition</u>. K. Brown. Oxford, Elsevier. 3: 392-399.

Aissen, J. (1987). <u>Tzotzil clause structure</u>. Dordrecht, Reidel.

Aissen, J. (2000). Differential object marking: Iconicity vs. economy, University of California at Santa Cruz.

Alarcón Neve, L. J. and W. Orozco Lecona (2004). <u>El uso del pronombre clítico de objeto directo en el discurso narrativo de los niños de 5to. grado de primaria.</u> Séptimo Encuentro Internacional de Lingüística en el Noroeste. , Hermosillo, Sonora, Unison.

Alcina Caudet, M. A. (1999). Las expresiones referenciales. Estudio semántico del sintagma nominal, Universitat de Valencia. Doctoral Dissertation.

Allen, K. (1983). Anaphora, Cataphora, and Topic Focusing: Function of the Object Prefix in Swahili. <u>Current Approaches to African Linguistics</u>. I. Dihoff. Dordrecht, Foris.

Anagnostopoulou, E. (1994). Clitic Dependencies in Modern Greek, Universität Salzburg. Ph.D. thesis.

Anagnostopoulou, E. and A. Giannakidou (1995). <u>Clitics and Prominence, of Why Specificity is not Enough</u>. Papers from the 31st. Regional Meeting of the Chicago Linguistic Society. Parasession on Clitics, Chicago, University of Chicago.

Aoun, J. (1981). The Formal Nature of Anapgoric Relations, MIT. Ph.D dissertation.

Ariel, M. (1988). "Referring and accessibility." <u>Journal of Linguistics</u> 24: 65-87.

Ariel, M. (1990). Accessing noun-phrase antecedents. London and New York, Routledge.

Asher, N. (1999). Discourse and the focus/background distinction. Focus: linguistic, cognitive and computational perspectives. P. Bosch and R. van der Sandt. Cambridge, Cambridge University Press.

Auger, J. (1994). Pronominal Clitics in Quebec Colloquial French: A Morphological Analysis, University of Pennsylvania. Ph.D. dissertation.

Averintseva-Klisch, M. (2006). German Right Dislocation and Afterthought in Discourse (submitted).

Barrenechea, A. M. (1987). El habla culta de la ciudad de Buenos Aires. Buenos Aires, Universidad Nacional de Buenos Aires.

Barrenechea, A. M. and T. Orecchia (1977). La duplicación de objetos directos e indirectos en el español hablado en Buenos Aires. Estudios sobre el español hablado en las principales ciudades de América. J. M. Lope Blanch. México, Universidad Nacional Autónoma de México: 351-381.

Belletti, A. (1999). Italian/Romance clitics: Structure and derivation. Clitics in the Languages of Europe. H. van Riemsdijk. Berlin, Mouton de Gruyter.

Belletti, A. (2005). "Extended doubling and the VP periphery." Probus(17): 1-35.

Belloro, V. (2004a). A Role and Reference Grammar Account of Third-person Clitic Clusters in Spanish. Buffalo, University at Buffalo. M.A. thesis.

Belloro, V. (2004b). A Role and Reference Grammar Approach to Spanish Clitics. Proceedings from the 2004 International Role and Reference Grammar Conference. B. Nolan. Dublin, Institute of Technology Blanchardstown.

Belloro, V. (2006). "What's this clitic doing in my sentence?" VI International Role and Reference Grammar Conference. University of Leipzig.

Belloro, V. (2007). "The pragmatics of dative doubling in Spanish". 2007 International Conference on Role and Reference Grammar. México, DF.

Beninca, P. (1988). L'ordine degli elementi nella grase e le costruzione marcate. Grande Gramatica Italiana di Consultazione. L. Renzi. Bologna, Il Mulino. 1: 153-194.

Bentivoglio, P. (1978). Formación de clíticos: análisis sobre el habla culta de Caracas. Corrientes actuales en la dialoctología del Caribe Hipánico. H. López Morales. Puerto Rico, Editorial Universitaria: 13-29.

Bentivoglio, P. (1983). Topic Continuity and Discontinuity in Discourse: A Study of Spoken Latin-American Spanish. Amsterdam and Philadelphia, John Benjamins.

Bentivoglio, P. (1993). Full NPs in Spoken Spanish: A Discourse Profile. Linguistic Symposium on Romance Languages. W. Ashby, M. Mithun, G. Perissinotto and E. Raposo. Amsterdam and Philadelphia, John Benjamins.

Bentley, D. (2006). Split intransitivity in Italian. Berlin, Mounton de Gruyter.

Birner, B. a. G. W. (1994). Uniqueness, familiarity, and the definite article in English. Berkeley Linguistic Society 20, Berkeley, CA.

Bleam, T. (1999). Leísta Spanish and the Syntax of Clitic Doubling. Delaware, University of Delaware. Unpublished Ph.D. thesis.

Bogard, S. (1992). El estatus del clítico de complemento indirecto en español. Reflexiones lingüísticas y literarias. R. Barriga Villanueva and J. García Fajardo. México, El Colegio de México. 1. Lingüística: 171-186.

Bogard, S. (2007a). Objeto directo preverbal en español: función y expresión de la construcción inversa. I Encuentro de estudios de gramática del español. Semántica, pragmática y prosodia: Reflejos en el orden de palabras en español. Centro de Estudios Lingüísticos y Literarios, El Colegio de México, México, DF.

Bogard, S. (2007b). Spanish Atelic Sentences Complemented by a Bare Noun Phrase: Two Macroroles or Just One? 2007 International Conference on Role and Reference Grammar, UNAM, México DF.

Bolinger, D. (1977). There. Meaning and Form. London, Longman: 90-123.

Bonet, E. (1991). Morphology After Syntax: Pronominal Clitics in Romace, MIT. Ph.D. dissertation.

Borer, H. (1984). Parametric Syntax. Dordrecht, Foris.

Bresnan, J. (2001). Lexical-Functional Syntax. Oxford, Blackwell.

Bresnan, J. and L. Moshi (1990). "Object asymmetries in comparatie Bantu syntax." Linguistic Inquiry(21): 147-185.

Bresnan, J. and S Mchombo (1987). "Topic, pronoun, and agreement in Chichewa." Language(63): 741-782.

Campos, H. (1999). Transitividad e intransitividad. <u>Gramática descriptiva de la lengua española</u>. I. Bosque and V. Demonte. Madrid, Espasa-Calpe. 2: 1519-1574.

Capone, A. (2001). Theories of presuppositions and presuppositonal clitics. <u>Proceedings of the Bielefeld Bi-dialogue workshop</u>. P. Kuenhlein, Rieser and H. Zeevat.

Cardinaletti, A. (2002). "Against Optional and null clitics. Right dislocation vs. marginalization." <u>Studia Linguistica</u> 56(1): 29-57.

Carroll, S. (1982). Redoublement et dislocation en français populaire. <u>La syntaxe comparée du français standard et populaire: approches formelle et fonctionnelle</u>. C. Lefebvre. Québec, Éditeur officiel du Québec. 1.

Casson, R. (1983). "Schemata in Cognitive Anthropology." <u>Annual Review of Anthropology</u> 12: 429-462.

Ceccetto, C. (1999). "A Comparative Analysis of Left and Right Dislocation in Romance." <u>Studia Linguistica</u> 53(1): 40-67.

Chafe, W. (1976). Giveness, Contrastiveness, Definiteness, Subjects, Topics, and Point of View. <u>Subject and Topic</u>. C. Li. New York, Academic Press: 25-55.

Chafe, W. (1987). Cognitive Constraints on Information Flow. <u>Coherence and Grounding in Discourse</u>. R. Tomlin. Amsterdam, John Benjamins: 21-51.

Chafe, W. (1994). <u>Discourse, Consciousness and Time</u>. Chicago, The University of Chicago Press.

Chomsky, N. (1981). <u>Lectures on Government and Binding</u>. Dordrecht, Foris.

Christophersen, P. (1939). <u>The articles: A study of their theory and use in English</u>. Copenhagen, Munksgaard.

Clark, H. H. and S. E. Haviland (1977). Comprehension and the Given-New Contract. <u>Discourse processes: advances in research and theory.</u> . R. Freedle. Norwood, NJ, Ablex Publishing. 1. Discourse Production and Comprehension: 1-40.

Colantoni, L. (2002). Clitic Doubling, Null Objects and Clitic Climbing in the Spanish of Corrientes. <u>From Words to Discourse: Trends in Spanish Semantics and Pragmatics</u>. J. Gutiérrez Rexach. Amsterdam, ELSEVIER: 321-336.

Cole, P. (1978). On the Origins of Referential Opacity. <u>Syntax and Semantics. Volume 9. Pragmatics</u>. P. Cole, New York: Academic Press: 1-22.

Company Company, C. (2006). El Objeto Indirecto. <u>Sintaxis histórica de la lengua española</u>. C. C. C. (Directora). México DF, Universidad Nacional Autónoma de México, Fondo de Cultura Económica. 1. La frase verbal: 477-572.

Comrie, B. (1989). <u>Language universals and linguistic typology</u>. Chicago, University of Chicago.

Corbett, G. (2003). Agreement: The range of the phenomenon and the principles of the Surrey database of agreement. <u>To appear in Agreement: A Typological Perspective. Special number of Transactions of the Philological Society.</u> .

Creissels, D. (2001). A typology of subject marker and object marker systems in African languages. <u>International Symposium: Typology of African Languages</u>. Köln.

Croft, W. (1988). Agreement vs. Case Marking and Direct Objects. <u>Agreement in Natural Language</u>. M. a. F. Barlow, Ch. Standfor, Center for the Study of Language and Information: 159-180.

Cuervo, M. C. (2003). Datives at Large. Cambridge, MA, MIT. Ph.D. thesis.

Culicover, P. a. M., L. (1988). "On Exploring the Limits of Syntax." <u>Syntax and Semantics</u> 29. The limits of Syntax.

de Miguel, M. E. (1997). <u>La amante del Restaurador</u>. Buenos Aires, Planeta Argentina.

Demonte, V. (1994). La ditransitividad en español. <u>Gramática del español</u>. V. Demonte. México, El Colegio de México: 431-470.

Demonte, V. (1995). "Dative alternation in Spansih." <u>Probus</u> 7: 5-30.

Demuth, K. and M. Johnson (1989). "Interaction Between Discourse Functions and Agreement in Setawana." <u>Journal of African Languages and Linguistics</u> 11: 22-35.

Diesing, M. (1992). <u>Indefinites</u>. Cambridge, MA, The MIT Press.

Domínguez, L. (2004). Mapping Focus: The Syntax and Prosody of Focus in Spanish, Boston University. Ph.D. dissertation.

Donellan, K. (1966). "Reference and Definite Descriptions." <u>Philosophical Review</u> 75: 281-304.

Donohue, M. (1999). <u>A grammar or Tukang Besi</u>. Berlin, Mouton de Gruyter.

Dowty, D. (1979). <u>Word meaning and Montague Grammar</u>. Dordrecht, Foris.

Dryer, M. (1996). "Focus, pragmatic presupposition, and activated propositions." <u>Journal of Pragmatics</u>(26): 475-523.

Du Bois, J. (1980). Beyond Definiteness: The Trace of Identity in Discourse. <u>The Pear Storie. Cognitive, Cultural, and Liguistic Aspects of Narative Production</u>. W. Chafe. Norwood, New Jersey, ABLEX Publishing Corporation: 203-274.

Duranti, A. and E. Ochs (1979). Left-dislocation in Italian Conversation. <u>Syntax and Semantics. </u>. T. Givón. 12: Discourse and Syntax: 377-416.

Enç, M. (1991). "The semantics of specificity." <u>Linguistic Inquiry</u>(22): 1-25.

Erteschik-Shir, N. (1973). On the nature of island constraints. Cambridge, MA, MIT. Ph.D. dissertation.

Farkas, D. (1995). Specificity and Scope. <u>Actes du Premier Colloque Langes & Grammaire</u>. L. Nash and G. Tsoulas. Paris: 119-137.

Fernández Soriano, O. (1989). Rección y ligamiento en el español: Aspectos del parámetro de sujeto nulo, Universidad Autónoma de Madrid. Tesis doctoral.

Fernández Soriano, O. (1993). Los pronombres átonos en la teoría gramatica. Repaso y balance. <u>Los pronombres átonos</u>. O. Fernández Soriano. Madrid, Taurus.

Fillmore, C. (1982). Frame semantics. <u>Linguistics in the Morning Calm</u>. T. L. S. o. Korea. Seoul, Hanshin: 111-137.

Fodor, J. and I. Sag (1982). "Referential and Quantificational Indefinites." <u>Linguistics and Philosophy</u>(5): 355-398.

Franco, J. (1993). On object agreement in Spanish. Los Angeles, University of Southern California. Ph.D. thesis.

Franco, J. (2000). Agreement as a Continuum. <u>Clitic Phenomena in European Languages</u>. F. a. d. D. Beukema, M. Amsterdam, John Benjamins.

Franco, J. and E. Mejías-Bikandi (1997). "Overt and Covert Raising to Spec AGR." <u>Linguistic Analysis</u>(27): 79-107.

Franco, J. and E. Mejías-Bikandi (1999). The Presuppositionality Condition and Spanish Clitic-Doubled Objects. <u>Formal Perspectives on Romance Linguistics</u>. J.-M. Authier, Bullock, Barbara and Reed, Lisa. Amsterdam/Philadelphia, John Benjamins: 107-119.

Fraurud, K. (1996). Cognitive Ontology and NP form. <u>Reference and Referent Accessibility</u>. T. a. G. Fretheim, J. Amsterdam/Philadelphia, John Benjamins: 65-87.

García-Miguel, J. (1991). "La duplicación del objeto directo e indirecto com concordancia." <u>Verba</u>(18): 375-410.

Gierling, D. (1997). Clitic doubling, specificity and focus in Romanian. <u>Clitics, pronouns and movement</u>. J. Black and V. Motapanyane. Amsterdam, Philadephia, John Benjamins.

Givón, T. (1976). Topic, Pronoun, and Grammatical Agreement. <u>Subject and Topic</u>. C. Li. New York, Academic Press.

Givón, T. (1978). Definiteness and Referentiality. <u>Universals of Human Language</u>. J. Greenberg. Stanford, Stanford University Press.

Givón, T. (1983). <u>Topic Continuity in Discourse. A Quantitative Cross-Language Study</u>. Amsterdam/Philadelphia, John Benjamins.

Givón, T. (1984). Direct object and dative shifting: The semantics and pragmatics of case. <u>Objects</u>. F. Plank. New York, Academic Press.

Givón, T. (2001). <u>Syntax</u>. Amsterdam and Philadelphia, John Benjamins.

González Vergara, C. (2006). Las construcciones no reflexivas con 'se'. Una propuesta desde la Gramática del Papel y la Referencia. <u>Facultad de Filología</u>. Madrid, Universidad Complutense de Madrid. Tesis doctoral.

Greenberg, J. (1974). <u>The relation of frequency to semantic feature in a case language (Russian)</u>. Stanford, Stanford University.

Groppi, M. (2002). <u>Observando las duplicaciones</u>. XIII Congreso Internacional, Asociación de Lingüística y Filología de América Latina, Costa Rica, Universidad de Costa Rica.

Grosz, B. and Y Ziv (1997). Centering, Global Focus, and Right Dislocation. <u>Centering in Discourse</u>. Walker, M. A. Joshi and E. Prince, eds., Oxford University Press.

Gundel, J., N. Hedberg, et al. (1993). "Cognitive Status and the Form of Referring Expressions in Discourse." <u>Language</u> 69(2): 274-307.

Gundel, J., N. Hedberg, et al. (2001). "Definite Descriptions and Cognitive Status in English." <u>Journal of English Language and Linguistics</u>.

Gutiérrez Bravo, R. (2002a). <u>Prominence Scales and Unmarked Word Order in Spanish</u>. Proceedings of the Workshop "Semantic and Syntactic Aspects of Specificity in Romance Languages", Konstanz.

Gutiérrez Bravo, R. (2002b). Structural Markedness and Syntactic Structure: A Study of Word Order and the Left Periphery in Mexican Spanish. Ph.D. dissertation, Department of Linguistics. University of California at Santa Cruz.

Gutiérrez Ordoñez, S. (1999). Los dativos. <u>Gramática descriptiva de la lengua española</u>. I. Bosque and V. Demonte. Madrid, Espasa-Calpe. 2: 1855-1930.

Gutiérrez Rexach, J. (2000). "The Formal Semantics of Clitic Doubling." <u>Journal of Semantics</u>(16): 315-380.

Haspelmath, M. (2004). "Explaining the ditransitive person-role constraint: a usage-based account." <u>Constructions</u>(2).

Hawkins, J. (1978). <u>Definiteness and indefiniteness: a study in reference and grammaticality prediction.</u> Atlantic Highlands, NJ, Humanities Press.

Heim, I. (1988). <u>The Semantics of Definite and Indefinite Noun Phrases</u>. New York and London, Garland.

Hualde, J. (1989). Double Object Constructions in KiRimi. <u>Current Approaches to African Linguistics</u>. R. Botne and P. Newman. Dordrecht, Foris.

Hurtado, A. (1984). "On the Properties of LF." <u>Cornell Working Papers in Linguistics</u>(5).

Ibañez Cerda, S. (2002). <u>Introduciendo participants en la estructura argumental: el caso del aplicativo 'le' del español</u>. XIII Congreso Internacional de la Asociación de Lingüística y Filología de América Latina., Costa Rica, Universidad de Costa Rica.

Ibañez Cerda, S. (2006). El papel del clítico 'le' en las construcciones de duplicación de dativo. Evidencia diacrónica. <u>Jornadas Filológicas 2006</u>. Instituto de Investigaciones Filológicas, UNAM.

Ioup, G. (1977). "Specificity and the Interpretation of Quantifiers." Linguistics and Philosophy 1(2): 233-245.

Jackendoff, R. (1972). Semantic Interpretation in Generative Grammar. Cambridge, MA, MIT Press.

Jaeggli, O. (1981). Topics in Romance syntax. Dordrecht, Foris.

Jaeggli, O. (1986). Three issues in the theory of clitics: Case, doubled NPs, and extraction. . Syntax and Semantics. The Syntax of Pronominal Clitics. H. Borer. New York, Academic Press. 19.

Jolly, J. (1993). Prepositional analysis within the framework of Role and Reference Grammar. New York, Peter Lang.

Kagan, O. (2006). Specificity as Speaker Identifiability. The Ninth Symposium on Logic and Language, Budapest, Research Institute for Linguistics, Eötvös Loránd University.

Kaiser, G. (1994). More about INFL-ection; The acquisition of clitic pronouns in French. Bilingual First Language Acquisition. French and German Grammatical Development . J. Meisel. Amsterdam, John Benjamins.

Kailuweit, R. (in press). 'Floating plurals', Prodrop and Agreement-an optimality-based RRG approach. Proceedings from the 2006 Rol and Reference Gramar Conference. R. D. Van Valin, John Benjamins.

Kallulli, D. (2000). Direct Object Clitic Doubling in Albanian and Greek. Clitic Phenomena in European Languages. F. a. d. D. Beukema, M. Amsterdam, Philadelphia, John Benjamins.

Kamp, H. and U. Reyle (1993). From discourse to logic. Hingham, MA, Kluwer.

Karttunen, L. (1976). Discourse Referents. Syntax and Semantics. J. McCawley. New York, Academic Press. 7: Notes from the Linguistic Undeground.

Karttunen, L. and S. Peters (1977). Requiem for Presupposition. 3rd. Annual Meeting of the Berkeley Linguistic Society, Berkeley, University of California.

Kawachi, K. (2007). Topics in the Semantics of Siddhama. Linguistics. Buffalo, Univeristy at Buffalo. Ph.D. dissertation.

Kayne, R. (1975). French Syntax: The Transformational Cycle. Cambridge, MIT Press.

Keyne, R. (1994). <u>The Antisymmetry of Syntax</u>. Cambridge, Mass., MIT Press.

Kiss, K. (1988). "Identificational versus information focus." <u>Language</u> 74(2): 245-273.

Klamer, M. (2004). Kambera. <u>The Austronesian Languages of Asia and Madagascar</u>. S. a. H. Adelaar, N. London, Routledge.

Laca, B. (2006). El Objeto Directo. La marcación preposicional. <u>Sintaxis histórica de la lengua española. Primera Parte: La frase verbal</u>. C. C. Company. México, Fondo de Cultura Económica, UNAM. 1: 421-475.

Laidig, W. and C. Laidig (1990). "Larike pronouns: duals and trials in a Central Moluccan languages." <u>Oceanic Linguistics</u> 29: 87-109.

Laka, L. (1993). The structure of inflection: A case study in X syntax. <u>Generative Studies in Basque Linguistics.</u> . J. Hualde and J. Ortiz de Urbina. Amsterdam, John Benjamins: 21-27.

Lakoff, G. (1987). <u>Women, Fire, and Dangerous Things: What Categories Reveal about the Mind</u>. Chicago, University of Chicago Press.

Lambrecht, K. (1994). <u>Information structure and sentence form. A theory of topic, focus, and the mental representations of discourse referents</u>. Cambridge, CUP.

Leonetti, M. (1990). <u>El artículo y la referencia</u>. Madrid, Taurus Universitaria.

Leonetti, M. (1999). El artículo. <u>Gramática descriptiva de la lengua española</u>. V. Demonte and I. Bosque. Madrid, Espasa-Calpe: 787-890.

Leonetti, M. (2000). The Asymmetries between the definite article and demonstratives: A procedural account. <u>7th International Pragmatics Conference</u>. Budapest.

Leonetti, M. (2003). <u>Specificity and Object Marking: the Case of Spanish *a*</u>. Semantic and Syntactic Aspects of Specificity in Romance Languages, Konstanz, Universität Konstanz.

Leonetti, M. (2007) Clitics do not encode specificity. In Kaiser, G. and M. Leonetti (eds.) Proceedings of the Workshop "Definiteness, Specificity and Animacy in Ibero-Romance Languages". Arbeitspapier – Fachbereich Sprachwissenschaft der Universität Konstanz.

Levinson, S. (1987). "Pragmatics and the grammar of anaphora: a partial pragmatic reduction of binding and control phenomena." <u>Journal of Linguistics</u>(23): 379-434.

Lewis, D. (1979). "Scorekeeping in a language game." Journal of Philosophical Logic 8: 339-359.

Llorente, A. and J. Mondéjar (1972). "La conjugación objetiva en las lenguas románicas." Prohemio 3: 5-27.

Lope Blanch, J. M. (1971). El habla culta de la ciudad de México. Materiales para su estudio. México, UNAM.

Lope Blanch, J. M. (1977). Prólogo. Estudios sobre el español hablado en las principales ciudades de América. J. M. Lope Blanch. México, Universidad Nacional Autónoma de México.

Lyons, C. (1999b). Definiteness. Concise Encyclopedia of Grammatical Categories. K. Brown and J. Miller. Amsterdam, Elsevier: 125-131.

Lyons, C. (1999a). Definiteness. Cambridge, Cambridge University Press.

Maldonado, R. (2002). "Objective and subjective datives." Cognitive Linguistics 13(1): 1-65.

Mandler, J. M. (1984). Stories, Scripts, and Scenes: Aspects of Schema Theory. Hillsdale, NJ, Lawrence Erlbaum.

Mayer, E. (2003). Clitic Doubling in Limeño. A Case Study in LFG. Linguistics, Australian National University. M.A. thesis.

Miller, P. (1992). Clitics and Constituents in Phrase Structure Grammar. New York, Garland.

Milsark, G. (1977). "Toward an Explanation of Certain Peculiarities of the Existential Construction in English." Linguistic Analysis(3): 1-29.

Monachesi, P. (2003). The verbal complex in Romance: a case study in grammatical interfaces.

Mora Bustos, A. (2007). Orden de constituyentes, foco y entonación. I Encuentro de estudios de gramática del español. Semántica, pragmática y prosodia: Reflejos en el orden de palabras en español. Centro de Estudios Lingüísticos y Literarios, El Colegio de México, México DF.

Nash, J. (1992). Aspects of Ruwund Grammar, University of Illinois at Urbana-Champaign. Ph.D. dissertation.

Nerlich, B. and D. Clarke (2000). "Semantic fields and frames: Historial explorations of the interface between language, action, and cognition." Journal of Pragmatics (32): 125-150.

Nichols, J. (1986). "Head-marking and dependent-marking grammar." Language 62: 56-119.

Oh, C. K. and D. E. Dinneen (1979). Presupposition. New York, San Francisco, London, Academic Press.

Ouali, H. (2006). Unifying Agreement Relations: A Minimalist Analysis of Berber. Linguistics, University of Michigan.

Parodi, C. and O. Santa Ana (2002). Clíticos de Objeto Directo: Análisis sociolingüístico en el marco minimalista. XIII Congreso International, Asociación de Lingüística y Filología de América Latina, Costa Rica, Universidad de Costa Rica.

Partee, B. (1972). Opacity, coreference, and pronouns. Semantics of Natural Language. D. Davidson and G. Harman. Dordrecht, Reidel.

Payne, D. (1985). Aspects of the grammar of Yagua: A typological perspective. Los Angeles, UCLA. Ph.D. dissertation.

Payne, T. E. (1992). The Twin Stories. Participant Coding in Yague Narrative. Berkeley, Los Angeles, Oxford, University of California Press.

Pensado, C. E. (1995). El compremento directo preposicional. Madrid, Visor Libros.

Perlmutter, D. (1971). Deep and surface constraints in syntax. New York, Holt, Rinehart and Winston Inc.

Prince, E. (1981). Toward a taxonomy of given-new information. Radical Pragmatics. P. Cole. New York, Academic Press: 223-256.

Prince, E. (1992). The ZPG letter: subjects, definiteness, and information-status. Discourse description: diverse analyses of a fund raising text. S. Thompson and W. Mann. Philadelphia and Amsterdam, John Benjamins: 295-325.

Puig, M. (1988). El beso de la mujer araña. Barcelona, Ediciones 8, Biblioteca de Bolsillo.

Quilis, A., M. Cantarero, et al. (1985). Los pronombres le, la, lo y sus plurales en la lengua española hablada en Madrid. Madrid, Consejo Superior de Investigaciones Científicas.

RAE (2005). Diccionario Panhispánico de dudas [//buscon.rae.es/dpdI/].

Rivas, A. (1977). A Theory of Clitics, MIT.

Rivero, M. L. (1977). "Specificity and Existence: A Reply." Language 53(1): 70-85.

Roberts, J. (2005). Ra in Persian and Information Structure.

Sanchez, L. (no date). Clitic doubling and the checking of focus. New Brunswick, NJ, Rutgers University, Department of Spanish and Portuguese.

Schank, R. and R. Abelson (1977). Scripts, plans, goals and understanding: . Hillsdale, NJ, Lawrence Erlbaum.

Schick, I. P. (2000). Clitic Doubling Constructions in Balkan-Slavic Languages. Amsterdam, Philadephia, John Benjamins.

Seidl, A. and A. Dimitriadis (1997). The Discourse Function of Object Marking in Swahili. Chicago Linguistic Society. The Main Session.

Shimojo, M. (in press). How missing is the missing verb? The verb-less numeral quantifier construction in Japanese. Proceedings from the 2006 International Conference on Role and Reference Grammar. R. D. Van Valin.

Silva-Corvalán, C. (1981). "The Diffusion of Object-Verb Agreement in Spanish." Papers in Romance(3): 163-176.

Silva-Corvalán, C. (1984). "Topicalización y pragmática en español." Revista Española de Lingüística 14(1).

Sportiche, D. (1995). Clitic Constructions. Phrase Structure and the Lexicon. I Zaring and J. Rooryck. Dordercht, Kluwer Academic Publishers: 213-276.

Steriade, D. (1980). Clitic-Doubling in Romania Top WH Constructions and the Analysis of Topicalization. Papers from the 16th Annual Meeting, Chicago Linguistic Society.

Strawson, P. F. (1950). "On referring." Mind(59): 320-344.

Strozer, J. R. (1977). "Clitics in Spanish." Dissertation Abstracts International 37: 6451A.

Suñer, M. (1999a). Clitic-Doubling of Strong Pronouns in Spanish: An Instance of Object Shift. <u>Grammatical analyses in Basque and Romance Linguistics: Papers in honor of Mario Saltarelli.</u> . J. Franco, A. Landa and J. Martin. Amsterdam, John Benjamins: 233-256.

Suñer, M. (1999b). Dialectal variation and clitic-doubled direct objects. <u>Studies in Romance Linguistics</u>. C. a. D. Kirschner, Janet. Amsterdam/Philadelphia, John Benjamins. 60: 377-395.

Suñer, M. (2000). El traslado de objetos: reflejos semánticos en la sintaxs del español y el germánico con algún detalle sobre el griego. <u>V Encuentro Internacional de Lingüística en el Noroeste</u>. Z. B. A. Estrada Fernández, Isabel. Hermosillo, Editorial Unison. 1: 17-38.

Suñer, M. (1988). "The role of agreement in clitic-doubled constructions. ." <u>Natural Language and Linguistic Theory</u>(6): 391-434.

Tannen, D. and C. Wallat (1987). "Interactive Frames and Knowledge Schemas in Interaction: Examples from a Medical Examination/Interview." <u>Social Psychology Quarterly</u> 50 (2. Special Issue: Language and Social Interaction): 205-216.

Torrego, E. (1995). On the Nature of Clitic Doubling. <u>Evolution and Revolution in Linguistic Theory: Studies in Honor of C.P. Otero</u>. H. Campos and P. Kempchinsky. Washington, DC, Georgetown University Press.

Torrego, E. (1999). El complemento directo preposicional. <u>Gramática descriptiva de la lengua española</u>. V. Demonte and I. Bosque. Madrid, Espasa-Calpe. 2: 1779-1805.

Uriagereka, J. (1995). "Aspects of the syntax of clitic placement in Western Romance." <u>Linguistic Inquiry</u>(26): 79-123.

Vallduví, E. (2001). La dislocació del verb en catalá. <u>Actes del Nove Colloqui d'Estudis Catalans a Nord-America</u>. M. Bover and M. Vidal-Tibbits. Barcelona, Abadia de Montserrat: 467-494.

Van Valin, R. D., Ed. (1993). <u>Advances in Role and Reference Grammar</u>. Amsterdam/Philadelphia, John Benjamin.

Van Valin, R. D. (no date). Summary of RRG.

Van Valin, R. D. (1995). Toward a functionalist account of so-called 'extraction constraints'. <u>Complex structures: A functionalis perspective</u>. B. e. a. Divriendt. Berlin, Mouton de Gruyter: 29-60.

Van Valin, R. D. (1998). The acquisition of WH-questions and the mechanisms of language aquisition. <u>The new psychology of language: cognitive and functional approaches to language structure</u>. M. Tomasello. Mahwah, NJ, LEA: 221-249.

Van Valin, R. D. (2005). <u>Exploring the Syntax-Semantics Interface</u>. Cambridge, Cambridge University Press.

Van Valin, R. D. (2006). On the relationship between syntactic theory and models of language processing. <u>Semantic role universals and argument linking: Theoretical, typological and psycho-neurolinguistic perspectives</u>. Bornkessel, Ina et al. Berlin, Mouton de Gruyter: 263-302.

Van Valin, R. D. and R. LaPolla, J. (1997). <u>Syntax. Structure, meaning and function</u>. Cambridge, Cambridge University Press.

Vendler, Z. (1967). <u>Linguistics in philosophy</u>. Ithaca, Cornell University Press.

von Heusinger, K. (1999). Intonation and information structure, University of Konstanz. Habilitationschrift.

von Heusinger, K. (2002). "Specificity and Definiteness in Sentence and Discourse Structure." <u>Journal of Semantics</u> 19: 245-274.

Wald, B. (1979). The development of the Swahili object marker: A study of the interaction of syntac and discourse. <u>Discourse and Syntax</u>. T. Givón. New York, Academic Press: 505-524.

Wanner, D. (1987). "Clitic Pronouns in Italian: A Linguistic Guide." <u>Italica</u> 64(3): 410-442.

Ward, G. and B. Birner (2004). Information Structure and Non-canonical Syntax. <u>The Handbook of Pragmatics</u>. L. R. Horn and G. Ward. Malden, Mass., Blackwell Publishing Ltd.: 153-174.

Weissenrieder, M. (1995). "Indirect Object Doubling: Saying Things Twice in Spanish." <u>Hispania</u>(78): 169-177.

Woolford, E. (1999). "Animacy Hierarchy Effects on Object Agreement." <u>New Dimensions in African Linguistics and Languages (Trends in African Linguistics 3)</u>: 203-216.

Woolford, E. (2001). Conditions on Object Agreement in Ruwundu (Bantu). <u>The UMass Volume on Indigenous Languages</u>. E. Benedicto. Amherst, MA.

Zubizarreta, M. (1999). Las funciones informativas: Tema y Foco. <u>Gramática descriptiva de la lengua española</u>. I. Bosque and V. Demonte. Madrid, Espasa-Calpe. 3.

Zwicky, A. and G. Pullum (1983). "Cliticization vs. Inflection: English n't." <u>Language</u> 59(3): 502-513.