To the Right of the Verb

To the Right of the Verb:

An Investigation of Clitic Doubling and Right Dislocation in three Spanish Dialects

By

Valeria A. Belloro

To the Right of the Verb: An Investigation of Clitic Doubling
and Right Dislocation in three Spanish Dialects

By Valeria A. Belloro

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

This study provides a new approach to Spanish clitic doubling, where the analysis of these constructions is based on spontaneous data considered within their broad discourse context. This type of analysis appeals to cognitive and pragmatic categories and advances an explicit comparison between the Argentine, Mexican and Spanish regional variants.

The approach presented here is based on embracing the graduality and heterogeneity that emerge from natural data, particularly in regard to a phenomenon that represents a change in progress at different stages in the different varieties. It demonstrates that the principle that best captures the conditions governing doubling constructions is the cognitive accessibility assumed of the target referent in the mind of the interlocutor. It distinguishes different subtypes of "doubling", which are taken to represent different points along a continuum of grammaticalization of discourse strategies, with the discourse relation between an anaphor and its antecedent at one end and prototypical grammatical agreement at the other. These different degrees of grammaticalization are apparent when examining the morphophonological and morphosyntactic behavior of dative and accusative clitics in each dialect, and they correlate with the subtypes of "doubling" that are most common in each one.

Finally, the author presents an analysis of these constructions in Role and Reference Grammar, a model that allows the interplay of syntax, semantics and pragmatics evidenced by clitic doubling to be naturally accounted for, and provides a formalization of these structures consistent with the diversity of naturally-occurring data.

ACKNOWLEDGEMENTS

My interest in clitic doubling structures began with my doctoral research, for which I analyzed doubling constructions in Buenos Aires Spanish. By the end of that project, it was clear to me that these innovative (or at least uncanonical) constructions were guided by pragmatic principles, and I had achieved a better understanding of the interplay between Spanish syntax and its semantic and pragmatic interfaces. And while I was satisfied with the answers the research was putting forward, at the same time I could see that the very same process was opening up new zones of uncertainty. Were all the instances of doubling in the data *really* doubling constructions, or could some of them be instances of right-dislocations? How could one decide between the two analyses while sticking to naturally occurring discourse and avoiding examples constructed *ad-hoc*? What was the state of dialectal spreading of doubling constructions in other Spanish dialects? Was it related to the degree of grammaticalization reached by the pronominal clitics? If so, was it possible to characterize an accordingly gradient array of doubling subtypes? Could this diversity be formally captured by a grammatical model, making it theoretically significant and potentially relevant for other languages with similar phenomena? I conceive of this book as a step forward in the road to achieving fully satisfactory answers to these questions.

Part of the analysis of the Argentinean data appeared as Belloro (2009 and 2012a). Belloro (2011) discusses differences between the treatment of

LIST OF ABBREVIATIONS

ABL	Ablative	LSC	Layered structure of the clause
ACC	Accusative		
ACS	Accessible	M	Masculine
ACT	Active	NP	Noun phrase
AFD	Actual focus domain	NMR	Non macrorole
AGR	Agreement	NOM	Nominative
AGX	Agreement index	NUC	Nucleus
AUX	Auxiliar	P	Partitive
CL	Clitic	PFD	Potential focus domain
COM	Command	PFV	Perfective
DAT	Dative	POCS	Post-core slot
DO	Direct object	POS	Possessive
CCP	Clitic + coreferential phrase	PRCS	Pre-core slot
		PRED	Predicate
DET	Determiner	PSA	Privileged syntactic argument
F	Feminine		
INA	Inactive	RDP	Right-detached position
ING	Ingressive	RP	Referential phrase
IO	Indirect object	RRG	Role and Reference Grammar
LDP	Left-detached position		
LS	Logical structure		

CHAPTER ONE

INTRODUCTION

Situated within the domain of research on the interface between Syntax and Pragmatics, the focus of this book is on so-called "clitic doubling constructions" in Spanish and the conditions that govern their occurrence in natural discourse.

Clitic doubling constructions are those in which a given entity is grammatically encoded within the clause by both a pronominal clitic and an independent lexical phrase coreferential with it. In Spanish, doubling involves exclusively accusative and dative arguments.

These constructions raise a series of issues that 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 the syntactic function of the other element is.

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 are the questions of whether the restrictions on doubling are syntactic, semantic or pragmatic in nature, and of how to distinguish doubling from dislocation.

Finally, given that these constructions are always grammatically optional, it is necessary to determine in which contexts they do effectively occur and to what communicative purpose.

Because of all the challenges they present, clitic doubling constructions have been a pet topic of grammatical analyses for the past few decades. The studies devoted to this issue have provided valuable insights into the complexities involved. At the same time, however, there are various alternative hypotheses coexisting in the literature, and it is often unclear

whether they result from progressive refinements of the linguists' intuitions or from grammatical differences in the dialects analyzed.

It is also the case that many of these grammatical studies have started to incorporate notions whose scope goes beyond clause-internal syntax, such as presuppositionality or topicality. These notions highlight the relevance of describing the particular discourse context where these constructions occur, which is somArgentineething that studies focusing on isolated sentences cannot do.

All of the aforementioned reasons have motivated this attempt to provide an analysis of clitic doubling constructions in Spanish based not only on naturally-occurring examples, but on examples which can be related to the broader communicative context in where they are used. Given the prominence of the Argentine dialect for the phenomenon under study, the claims presented here incorporate the analysis of spontaneous discourse produced by speakers of this variety of Spanish. However, since it is clear that not all Spanish dialects behave the same in this regard, two other varieties are considered, those of Mexico and Spain, and the data on accusative and dative doubling is compared among these three varieties, uncovering some interesting trends.

The relation between the Spanish pronominal clitics and their lexical counterparts in many respects resembles grammatical agreement relations between controllers and their targets. In fact, it has long been argued that systems of grammatical agreement evolve diachronically from anaphoric relations between incorporated pronouns and their discourse-antecedents (Givón 1976). In many languages, it is possible to show the transition from one stage to the next, as represented by forms whose behavior shares characteristics of both grammatical agreement markers and pronominal inflections (Bresnan 2001, Corbett 2003, Creissels 2001). It is proposed here that, in Spanish, this change in progress is apparent in the different degrees of grammaticalization displayed by, on the one hand, dative versus accusative clitics and, on the other, by both datives and accusatives in one dialect versus another. It is argued that these different degrees of grammaticalization, which have morphophonological and morphosyntactic manifestations, are also directly reflected in the freedom with which doubling constructions are used. In this regard, the traditional opposition between doubling and dislocation is claimed to prove insufficient, and a more fine-grained distinction is proposed, based predominantly on discourse-function.

One of the main theses of this study is that the morphosyntactic, semantic, and pragmatic variables that have been proposed so far 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. It is argued that this standpoint on the coding of accessibility allows for reconciling previous insights and for capturing the empirical data more accurately. 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 that reflect the speaker's assumptions about the addressee's state of knowledge and attention at the time of the utterance (Lambrecht 1994).

Assigning Spanish pronominal clitics 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 diachronic 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. At the same time, given that the process is to a great extent driven by discourse-pragmatics, its formalization is only possible within a theoretical model which has in place mechanisms for capturing and comparing phenomena that lie at the syntax-semantics-pragmatics interface. Role and Reference Grammar (Van Valin & LaPolla 1997, Van Valin 2005, 2008) represents such a model and is the framework that will be used to formalize the findings derived from the empirical data.

With these objectives in mind, the organization of this study is as follows. The next sections of this Introduction present a brief overview of the issues involved: the problem of distinguishing between doubling and dislocation, the ambiguous grammatical status of Spanish pronominal clitics, the asymmetries that hold between datives and accusatives, and the additional diversity posed by dialectal variation. Each of these issues is treated in the subsequent chapters of the book. Chapter 2 is a summary of the most relevant studies of clitic doubling in Spanish: those realized from a more theoretically-oriented standpoint, and those more concerned with descriptions based on corpus. It presents data from Argentine Spanish that challenge the current hypotheses and justify the need for a different perspective. Chapter 3 deals with the status of dative and accusative clitics in Argentine Spanish and argues for an approach that takes cognitive accessibility as the guiding principle accounting for the conditions that govern these constructions. Chapter 4 revises the status of pronominal clitics in Spanish, based on the morphophonological and morphosyntactic

behavior they exhibit in the Argentine, Mexican, and Spanish dialects. It is argued that they fall at different points along a continuum between inflectional affixes that serve as prototypical agreement markers and incorporated pronouns performing grammatical functions. Chapter 5 discusses the arguments put forth in the literature for distinguishing between doubling and dislocation, and advocates for a more fine-grained (and less syntactocentric) distinction of alternative constructions. Four subtypes of "doubling" constructions are distinguished: afterthoughts, antitopics, doubling-proper and pseudo-agreement. The distribution of each one in the three dialects under study is examined based on discourse-data. Finally, Chapter 6 provides the basic tenets of Role and Reference Grammar and advances a formal analysis of these constructions incorporating the categories that proved relevant in the descriptive analysis of the corpus data. The conclusions in Chapter 7 summarize the findings and proposals advanced in this work.

1.1. Doubling and Dislocation

The constructions that are the focus of this study are those involving a coreferential chain between a pronominal clitic and a postverbal lexical phrase; that is, sequences like those in (1)

(1) a. *yo **le**   quería   decir **al   señor Carrascal**…*
         I   DAT3SG   wanted   tell   to.the mister Carrascal
         'I wanted to tell Mr. Carrascal…' / 'I wanted to tell him, Mr. Carrascal…' (Spain, CORLEC)[1]

(2) b. *hace   tanto   que   no   **la**   veo   **a Susana**…*
         makes   much   that   NEG   ACC3fsg   I.see   a Susana
         'I haven't seen Susana in so long' / 'I haven't seen her in so long, Susana' (Argentina, HCBA)

As hinted at in the translations, one of the problems these sequences pose is that it is not obvious whether the lexical phrases are inside the clause (and there is a "double" representation of the syntactic objects), or right dislocated, and therefore outside the domain of grammatical relations.

The fact that the lexical phrases can be left out and the pronominal clitics are able to satisfy the valence of the verb has been at the origin of their treatment as pronominal arguments; an approach pioneered by Kayne (1975). Under Kayne's view, clitics are heads that originate in argument

position but, due to their weak phonological nature, "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, typically, in strict complementary distribution within the clause. In French, for instance, it is generally agreed upon that sequences of a clitic plus a coreferential phrase are always instances of right dislocations. Further, it is also commonly stated that there is some sort of intonation break separating both constituents, conventionally represented by a comma.

(3)   a.   *Jean-Paul l'   a   aimée /   a   aimée Simone.*
           Jean-Paul ACC.3sg AUX loved   AUX   loved Simone
           'Jean-Paul loved her' / 'Jean-Paul loved Simone.'

      b.   *Jean-Paul l'   a   aimée,   Simone.*
           Jean-Paul ACC.3sg   AUX loved   Simone
           'Jean-Paul loved her, Simone.'

      c.   **Jean-Paul l'   a   aimée Simone.*
           Jean-Paul   ACC.3sg AUX loved Simone
           'Jean-Paul loved Simone.'

The hypothesis that the same situation applies to Spanish, with coreferential phrases always 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. Jaeggli (1986) was among the first to point out that in Spanish there is no intonation break evidencing an extra-clausal position of the doubled noun phrase, which may even be followed by additional clausal material. Similar claims were made by Suñer (1988), who also was the first to argue that these chains should be analyzed as an instance of object agreement. These studies included examples such as the following:

(4)   a.   *Parece que tuvieron que llevarla a la hija del Cnel. Martínez de urgencia a los Estados Unidos*. (Jaeggli 1986: 33)
           'It seems they had to take Cnel. Martínez's daughter urgently to the US.'

  b. *Lo último que escuché, claro que **la** encontré pesada **la audición**,
     fue el reportaje*. (Suñer 1988:400)
     'The last thing I heard, although I found the show boring, was the
     interview'.

  c. *Yo **lo** voy a comprar **el diario** justo antes de subir*. (Suñer
     1988:400)
     'I'll buy the newspaper right before coming up'.

Pioneering studies such as these consolidated the general agreement
that (at least some dialects of) Spanish allow the clause-internal co-
occurrence of a pronominal clitic and a coreferential lexical phrase, i.e.
"clitic doubling", and since then there has been much work devoted to
these constructions. The question we posed at the beginning, however,
remained for the most part unanswered: If doublings and dislocations
coexist in the system, how are they to be distinguished? Do all dislocations
involve a prosodic break? Does doubling necessarily exclude it? Is this the
only criterial feature that can be invoked? How should we analyze
spontaneous data for which there is no *formal* feature distinguishing
between the two superficially similar constructions? Theory-oriented
studies typically concentrated on constructed examples illustrating some
theoretical point, with limited concern for spontaneous data and discourse
context. On the other hand, corpus-oriented studies have been blind to the
distinction between doubling and dislocation, treating all superficially
similar sequences alike. One of the challenges for this study is to search
for criteria that can be used for distinguishing different kinds of sequences
involving a clitic and a coreferential phrase as they appear in natural
discourse.

1.2. The Grammatical Status of Pronominal Clitics

The discussion of the syntactic role of the doubled lexical phrase has
been accompanied by varying positions about the grammatical status of
the clitics. The ambiguous nature of these forms, which share
characteristics of both words and affixes, has consequently allowed for
views that highlighted either their lexical or inflectional nature. Some
defended the view of clitics as lexical elements, arguing that, differently
from affixes, they can "climb" in the syntactic structure, do not affect the
stress pattern of the verb, and show nominal inflectional morphology
(Jaeggli 1986). Still, one must note these reasons do not always hold:
whereas the phenomenon of "clitic climbing" does not generally have any

correlate among affixes, 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, at least in Argentine Spanish, some enclitic cluster constructions cause the stress to shift to the final syllable (e.g. *poné* 'put' vs. *poneteló* 'put it on you'). Likewise, 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') has also been acknowledged. Finally, with respect to the issue of clitics exhibiting the inflectional morphology characteristic of nouns (i.e. such as the *–s* marker for plurals and the contrast *–a/o* for gender), it is clear that it only applies to a subset of the paradigm.

On the other hand, there have been studies that show that clitics in Romance languages, in general, tend to exhibit many characteristics likening them to inflectional affixes (Monachesi 2005). Some of these features are their rigid host selection and ordering, the gaps that occur in the clitics' "paradigm", and morphophonological idiosyncrasies like the ones mentioned above. As a consequence, and even though the morphophonological and morphosyntactic dimensions are, in principle, independent of one another, the characteristics that liken Spanish clitics to inflectional affixes have underscored their view as object agreement markers. In effect, it has been noted that the kind of feature erosion displayed by the clitics is typical of agreement systems (Franco 2000) and, for the case of datives in particular, an important body of diachronic research has shown their progressive grammaticalization (Company 1998, 2004, 2006).

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, it is well known that agreement phenomena may be 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 (Comrie 1989). This predicts that datives, which typically refer to animate referents, should enter in "agreement" relations more frequently than direct objects, which typically refer to inanimates. As summarized in the next section and discussed more extensively in Chapter 3, the asymmetry in the productivity of dative doubling vs. accusative doubling follows exactly this pattern.

1.3. Asymmetries between Dative and Accusative Doubling

It is well known that dative doubling is much less restricted than accusative doubling. This asymmetry is reflected both in their relative degree of dialectal spreading, as well as in the degree of grammatical or pragmatic restrictions that each construction is subject to. Dative doubling occurs rather productively in all Spanish varieties. Moreover, the dative clitic is in fact obligatory when in a chain with a strong pronoun (5)a, and in constructions introducing a benefactive (5)b or an experiencer (5)c, or involving inalienable possession (5)d, among other contexts.

(5)   a. *Le hablaron **a ella***.
          'They spoke to her.'

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

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

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

On the other hand, dialects seemingly differ in their preference for including the dative clitic in other grammatical contexts. For instance, in Peninsular Spanish, the clitic is assumed to be optional in relation to lexical recipients/goals (Demonte 1994, Mayer 2003), whereas in many varieties of Latin American Spanish the co-occurrence of the clitic has become the norm in these contexts as well (Bentivoglio 1978, Bogard 1992, Company 2006, Jaeggli 1981 Silva-Corvalán 1981):

(6)   a. *Di **a María** unas flores*. (Preferred in Peninsular Spanish)

      b. ***Le** di **a María** unas flores*. (Preferred in American Spanish)
          'I gave María some flowers.'

In contrast, accusative doubling is much more restricted. The only condition in which accusative doubling is obligatory across all dialects is in the context of a pronominal phrase, as in (7).

(7)   ***Lo** vi **a él***.
       'I saw him.'

In addition, some contexts where it is marginally tolerated have been noted. The *Diccionario Panhispánico de dudas* (RAE) affirms that DO-doubling is "normal" if the tonic complement is the pronoun *todo* 'everything' (8)a), as well as when it is a numeral preceded by an article and refers to an animate entity (8)b); or when it is realized as the indefinite *uno* 'one' and its referent is the current speaker (8)c). Interestingly, this source notes cases in which it has an emphatic value (8)d).

(8)   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, accusative doubling is considered ungrammatical, save for some regions of the Americas–especially Argentina–were it has been recognized to occur more productively. To date there are no studies that compare the productivity and pragmatic function of dative and accusative doubling in different dialects based on natural discourse, which is one of the tasks assumed here.

1.4. A Note on Dialectal Variation and Methodology

At this point, a more explicit comment must be made with regards to dialectal differences. As we saw above, it is generally acknowledged that there is a considerable degree of variation in these constructions, something hardly surprising considering Spanish is the primary language in over twenty countries. Despite this, the literature is sometimes 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. Furthermore, where claims are made with respect to some particular dialects, the lack of an explicit data gathering methodology or at least agreed upon labels may be source of additional misunderstanding. Consider the following example as a case in point. Jaeggli (1981) proposes that in "River Plate Spanish" accusative doubling is only possible if the coreferential phrase is marked by *a*. As we will see below, the occurrence of this form is crucial for the theoretical constructs proposed to explain doubling in this dialect. Thus, a sentence like (9) is judged ungrammatical.

(9)   *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 terms the "Porteño" dialect. Is it possible then that non *a*-marked phrases can be doubled in "Porteño" but not in "River Plate" Spanish? It may well be that these are different dialects governed by different principles and that "Porteño dialect (…) takes certain River Plate forms to an extreme" (Mayer 2003:12; for conclusions along the same lines see also Everett 1996:71 or Gutierrez-Rexach 2000:315, 331).

The plausibility of this hypothesis is reduced, however, when one notes that there is little base for considering "River Plate" and "Porteño" as two different dialects. After all, "porteño" is the term used to refer to the people of Buenos Aires, a city that sits on the shores of the "Río de la Plata", often (inexplicably) referred to as "River Plate" in the English literature. The different labels used by Suñer and Jaeggli are just a matter of taste, and it is unfortunate that they have led to erroneous assumptions.

Claims about the spreading and characteristics of accusative doubling have also been made with respect to "South American Spanish" (Lyons 1999a: 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 accusative doubling is homogeneous across all South American dialects; and on the other hand there is data suggesting that it may also occur in some North American varieties (specifically, in Mexican Spanish) as well as in Spain.

For instance, whereas Mayer (2003:16) argues, in tune with most of the literature, that "Standard Peninsular Spanish prohibits clitic doubling of direct objects, even if they are animate", Gutiérrez Rexach (2000:31) challenges this view, and observes that accusative doubling "in spoken and

colloquial Castillan Spanish (…) is quite regular". Likewise, Franco (2000:171) studies accusative doubling in the Spanish spoken in the Basque country, and Suñer (1989) provides examples of accusative doubling in the Spanish of Madrid. Once again, in these works there is disagreement with regards to the conditions under which accusative doubling is allowed, and it is unclear whether this is due to dialectal differences or analytical perspectives. Additionally, it is also unclear whether the same necessary conditions argued in each case for doubling couldn't also license right dislocations.

A similar picture repeats in the Americas. Beside the numerous analyses of Argentine Spanish, accusative doubling has been claimed to occur in the Spanish of Ecuador (Suñer 1989), Peru (Sánchez 2010), Mexico (Alarcón & Orozco 2004, Parodi & Santa Ana 2002), and Uruguay (Groppi 2002), again under seemingly diverse conditions.

This diversity makes apparent the importance of, on the one hand, unambiguously defining the dialect or dialects that serve as the object of the analysis. Unless this condition is met, it is impossible to establish whether contradictory grammaticality judgments (as for instance those regarding the obligatoriness of *a*-marking mentioned 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 varieties.

On the other hand, given that the literature has by now shown that doubling does occur in some dialects, it is vital to work with natural discourse to understand the full extent in which these constructions are actually used. Thirdly, since some of the variables that seem to condition these constructions go beyond clause-internal syntax, it is important to have available for analysis at least a considerable portion of the discourse where they appear. Finally, even though the very notion of dialect always and necessarily implies a great level of idealization, it would be desirable to be able to compare, under the same methodological approach, how different varieties behave. With these considerations in mind, the following chapters will deal with natural data, from the most part conversational, from corpora representing the Argentine, Mexican, and Peninsular varieties.

Now, even if I tried to fulfill all of these requirements, I must also note now some limitations. The corpora used in this study have been gathered, from the most part, from the cities of Buenos Aires, Mexico City, and Madrid. Sometimes I will refer to Argentine, Mexican, and Peninsular Spanish based on what is evidently partial data from speakers native to one particular city of each country. Overgeneralizations should be

minimized, but are, to an extent, unavoidable, and the only thing one can do is to explicitly acknowledge them. Secondly, there is still, to my knowledge, no annotated corpus that can be used to retrieve automatically potential cases of doubling (and that at the same time permit access to the discourse context where the utterance occurred). The gathering of the data had to be done by hand, by reading through the transcriptions of each corpus. This, together with the fact that we are dealing with structures that are not particularly frequent, limits the amount of instances available for analysis. In particular, for the data on accusatives, I recovered around 600 instances of sequences of a clitic plus a coreferential phrase in postverbal position from the corpora of transcriptions of oral interviews and conversations in the three dialects under study. Thus, those who expect the sort of quantitative analysis now common in corpus linguistics will necessarily be disappointed. I take here a compromising standpoint that is bound to leave some unsatisfied: On the other hand, I tried to fulfill what I think is the crucial need to work with natural data, in its broad discourse context, from explicitly comparable varieties, taking into consideration both the formal and functional aspects that have an impact on these constructions. At the same time, I tried to accomplish the above while incorporating the heterogeneity that comes with the analysis of changes in progress, and advancing a proposal for sorting this heterogeneous data into functionally motivated classes that can be used to categorize new data. The final challenge, relevant for those more theoretically-oriented, was to provide a grammatical formalization that is consistent with the diversity uncovered from the descriptive analysis. The broadness of these objectives always threatens to endanger the depth of some aspect of the analysis. I hope to have achieved a good balance.

---

[1] Morpheme-by-morpheme glosses are provided only when necessary for the point at hand. Coreferential structures are marked in boldface. The source of each example is provided right beside it or indicated in the text preceding it. If no source indication is provided, the example is constructed.

CHAPTER TWO

PREVIOUS ACCOUNTS

2.1. Corpus-Driven Studies

In the Introduction it was mentioned that compared to the numerous formal analyses of clitic doubling, there are only a handful of functionally-oriented analyses that study this construction in its discourse-context. One of the pioneering studies of this type is Barrenechea and Orecchia (1977), based on some of the data later published as *El Habla Culta de la Ciudad de Buenos Aires* (Barrenechea 1987). Despite the scarcity of examples provided by the authors, who take a strong quantitative approach, and the general succinctness of their descriptive interpretations, this study was one of the first to empirically analyze the productivity of dative doubling versus accusative doubling, and to show how doubling correlates with definite phrases, *a*-marked NPs, and human referents.

A more influential empirical study, this time involving Chilean Spanish, was advanced in Silva-Corvalán (1981), confirming that doubling is more frequent with datives than accusatives, and that it correlates with definiteness and specificity. Crucially, Silva-Corvalán argues that these features are the expression of an increase in discourse-topicality, which is advanced as the underlying explanatory variable. In this regard, Silva-Corvalán presents her study as supporting Givón's (1976) hypothesis linking the spreading of agreement to topicality hierarchies. Reinterpreting clitic doubling as a case of object agreement, Silva-Corvalán provides the basis for a unified account of the phenomenon. Interestingly, however, the involvement of referents that were just mentioned in the discourse context (and therefore maximally continuous, or "active") is a feature most typically exhibited by dislocations, although the difference between the two constructions is an issue that her study does not raise.

The interpretation of doubling as object agreement is also defended by Weissenrieder (1995), who focuses on dative doubling. Wiessenreder observes that 75% of her examples consist of dative arguments minimally encoded by the clitic, something expected given the correlation between

datives and highly topical entities. The cases of doubling (16%) are observed to occur wherever the topic is less continuous. There is a small percentage (4%) of structures involving a strong pronoun, associated with a shift on topical emphasis, and a similar amount (5%) of cases where the dative phrase does not trigger verb agreement (i.e. the clitic), which are observed to correlate with participants that are interpreted as topically unimportant and least often mentioned in the text. Examples of each type are presented below (Weissenrieder 1995: 173):

(10) a. **_Le_** _pide disculpas_.
        'He asks her for forgiveness.'

     b. _..._**_le_** _está haciendo bien_ **_a ella_**.
        'It is making her feel well.'

     c. **_Le_** _grita de todo_ **_a la chica_**.
        'He yells at the girl.'

     d. _Pide_ **_al ordenanza_** _un café doble_.
        'He asks the waiter for a double espresso.'

    Thus, this study follows Silva-Corvalán's in confirming the correlation between clitics and topicality, inasmuch as the frequent presence of the dative clitic in the clausal structure is viewed precisely as an effect of the sensitivity of agreement phenomena on the relative topicality of the referents involved. Note, however, that the contexts where dative doubling occurs are those where the topicality of the referent is _not_ as high.
    A somewhat different perspective is taken in Colantoni (2002). Colantoni presents an analysis of dative and accusative doubling based on a corpus of interviews with six residents of different towns in the province of Corrientes, Argentina. This is a _leísta_ dialect, and accusatives referring to humans take the _le_ form instead of _lo_. Furthermore, the clitic typically does not show number agreement, as illustrated in the following example:

(11)   _Hay muchas que no_ **_le_** _quieren amamantar_ **_a sus hijitos_**.
        'There are many of them who don't want to nurse their children.'

    Colantoni rightfully argues that the extra _bleaching_ of the clitic represented by the loss of number agreement conspires for making doubling more frequent, since a lexical phrase is often needed for correct identification of the intended referent. In effect, she proposes that one of