OPTIONAL AGREEMENT AND GRAMMATICAL FUNCTIONS: A CORPUS STUDY OF DATIVE CLITIC DOUBLING IN SPANISH

by
Roberto Aranovich
Licenciado en Letras, Universidad de Buenos Aires, 1997
Magister en Lingüística, Universidad Nacional del Comahue, 2002

Submitted to the Graduate Faculty of
the Kenneth P. Dietrich School of
Arts and Sciences in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

University of Pittsburgh

2011

UNIVERSITY OF PITTSBURGH
KENNETH P. DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Roberto Aranovich

It was defended on
February 7th 2011

and approved by

Lori Levin, Associate Research Professor, Carnegie Mellon University

Alan Juffs, Associate Professor, University of Pittsburgh

Yasuhiro Shirai, Professor, University of Pittsburgh

Erin O'Rourke, Assistant Professor, University of Alabama at Tuscaloosa

Roberto Mayoral Hernández, Assistant Professor, University of Alabama at Birmingham

Thesis Advisor: Lori Levin, Associate Research Professor, Carnegie Mellon University

Thesis co-Advisor: Alan Juffs, Associate Professor, University of Pittsburgh

OPTIONAL AGREEMENT AND GRAMMATICAL FUNCTIONS:

A CORPUS STUDY OF DATIVE CLITIC DOUBLING IN SPANISH

Roberto Aranovich, PhD

University of Pittsburgh, 2011

Spanish ditransitive constructions are characterized by the optionality of dative clitic doubling (DCLD), the co-occurrence of an unstressed dative pronoun with a co-referential indirect object (IO).

This fact has not received a satisfactory account in the literature, which has largely overlooked the optionality of the phenomenon or tried to reduce it to syntactic or lexical considerations (Strozer, 1976; Demonte, 1995). Our goal is to describe and explain the distribution of Dative Clitic Doubling in ditransitive sentences, as well as to study the implications of this phenomenon to the overall grammar of Spanish, in particular its interaction with word order.

We argue that the optionality of DCLD is an instance of optional object agreement, a widespread phenomenon in the languages of the world (Comrie, 1989; Woolford, 1999), which is favored by the pragmatic salience of the IO (high degree of animacy and givenness). We also argue that the distribution of DCLD is independent of word order, a claim that follows from the fact that Spanish encodes grammatical functions through agreement rather than word order. We support our claims with the results of a quantitative study of ditransitive sentences.

The study of Spanish ditransitive constructions is complemented by a quantitative study of another dative construction in Spanish, the possessive construction. The conclusion of this comparison is that dative case is favored by pragmatic prominence across different construction types.

From a cross-linguistic perspective, the dissertation compares Spanish DCLD and English dative-shift, two constructions that have been considered analogous in the literature (Demonte, 1995). In this respect, our conclusion is that the two constructions are essentially different as a result of an important typological difference between Spanish and English: Spanish is a Direct/Indirect Object language and English is a Primary/Secondary Object language (Dryer, 1986; Raúl Aranovich, 2007).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

This dissertation was possible thanks to the funding and support provided by the Department of Linguistics, University of Pittsburgh. I want to express my appreciation for the efforts made by the members of the committee in order to help me improve this dissertation: Lori Levin, Alan Juffs, Yasuhiro Shirai, Erin O'Rourke, and Roberto Mayoral Hernández. Al remaining errors are mine. In particular, I should thank Lori for supervising my dissertation through Skype since my relocation in Eau Claire, Wisconsin. Alan's help has been very important, especially in the final stages of revising this dissertation. Scott Kiesling, as graduate student advisor, also helped to make this dissertation possible. Pascual José Masullo has been my advisor for many years at Universidad Nacional del Comahue and Pittsburgh. After returning to Argentina, he was not able to be in the committee, but he has made an important contribution to my training in Linguistics and to this dissertation. I also want to thank Marcela Depiante for her long lasting encouragement and support, as well as for providing comments on an earlier version of this dissertation. Lynsey Wolter also provided comments on a draft. I want to thank my colleagues at the Department of Foreign Languages, University of Wisconsin, Eau Claire, for covering my classes while I was in Pittsburgh defending this dissertation.

This dissertation is dedicated to my wife and daughters, Mija, Ariana, and Ilana.

List of abbreviations and symbols used in this dissertation: *(grammatically unacceptable sentence), % (variably acceptable sentence), 1 (1st person), 2 (2nd person), 3 (3rd person), MASC (masculine gender), FEM (feminine gender), SG (singular number), PL (plural number), NOM (nominative case), ACC (accusative case), DAT (dative case) DO (direct object), IO (indirect object), OBL (oblique), CLD (clitic doubling) DCLD (dative clitic doubling), ACLD (accusative clitic doubling), CLLD (clitic left dislocation, CLRD (clitic right dislocation), th (theme), rec (recipient). The use of the term "grammatical acceptability" follows Lyons (1968) definition of grammaticality as the part of the acceptability of utterances which can be accounted for in terms of the rules of grammar.

1.0 INTRODUCTION

Spanish ditransitive constructions are characterized by the optionality of dative clitic doubling (DCLD), the co-occurrence of an unstressed dative pronoun with a co-referential indirect object (IO) (a Noun Phrase introduced by the preposition *a* fulfilling the thematic relation of recipient), as shown below.

(1) *Juan (le) envió un libro a María.*
Juan DAT.3SG sent a book to María
'Juan sent a book to María.'

This fact, to our best knowledge, has not received a satisfactory account in the literature, which has largely overlooked the optionality of the phenomenon or tried to reduce it to syntactic or lexical considerations (Strozer, 1976; Demonte, 1995). Our goal is to describe and explain the distribution of Dative Clitic Doubling, as well as to study the implications of this phenomenon to the overall grammar of Spanish, for instance the syntactic representation of ditransitive sentences.

DCLD in Spanish is a phenomenon that resists a complete description in terms of deterministic statements. A deterministic statement provides a set of necessary and sufficient conditions (Evans, 2007) for the application of a categorical rule, which applies 100 percent of the time (Labov, 1969; Walker, 2010). Categorical rules can only partially account for the distribution of DCLD in Spanish. For instance, if the IO is a pronoun, DCLD is categorically required:

(2) **(Le) di el libro a ella.*
DAT.3SG gave the book to her
'I gave the book to her.'

However, leaving aside a few contexts, to be discussed in Chapter 2, categorical rules cannot successfully describe the distribution of DCLD. Instead, the occurrence or not of DCLD seems to be in most contexts in free variation, a situation that can be described by an optional rule (Labov, 1969; Walker, 2010). For instance, DCLD can occur with animate or inanimate IOs (examples from *Corpus del Español*, http://corpusdelespanol.org/):

(3) a. Animate IO (*Amaranta)*:
En cierta ocasión ***le*** *envió* ***a Amaranta*** *un papelito desde la cárcel, pidiéndole el favor de bordar una docena de pañuelos de batista con las iniciales de su padre.*
'Once he sent Amaranta a piece of paper from prison, asking the favor to embroider a dozen handkerchiefs with the initials of his father.'

b. Inanimate IO *(la investigación,* 'research'):
*Sí, tengo una carga académica reducida para que pueda dedicar****le*** *más tiempo* ***a la investigación.***
'Yes, I have a reduced teaching load so I can spend more time on research.'
(*Corpus del Español*)

DCLD can occur with definite or indefinite IOs:

(4) a. Definite IO (*el estudio,* 'study'):
Todos los días ***le*** *dedica unas pocas de horas* ***al estudio****.*
'Every day he spends a few hours studying.'

b. Indefinite IO (*un cliente*, 'a client'):
Afirmó que nunca ***le*** *aconsejaría* ***a un cliente*** *que inventara información.*
'He said he would never advise a client to invent information.'
(*Corpus del Español*)

DCLD can occur with verbs that denote a change of possession or with verbs that denote a change of location:

(5) a. Change of possession (*dar*, 'to give'):
*Y hoy día entonces uno le pide al papá para dar****le*** *el regalo* ***a la mamá*** *y a la mamá para dar****le al papá****.*
'And today then one asks the dad in order to give the gift to the mom and to the mom in order to give it to the dad.'

b. Change of location (*arrojar*, 'to throw'):
Sin detenerse ***le*** *arrojó* ***al recepcionista*** *la llave de su cuarto y le hizo un gesto de guardar silencio al muchacho ya que empezaba a decirle que una señorita lo esperaba en el lobby.*
'Without stopping he threw his room key to the receptionist and made a silent gesture to the boy as he began to tell him that a lady was waiting for him in the lobby.'
(*Corpus del Español*)

DCLD can occur when the recipient precedes the theme or when the theme precedes the recipient:

(6) a. Theme precedes recipient:
La costumbre de la época era que los hombres les dedicaran [$_{TH}$ *boleros*] [$_{REC}$ *a sus novias*] *en las serenatas.*
'The custom of the time was that men dedicate boleros to their girlfriends in/during the serenade.'

b. Recipient precedes theme:

... el general Castilla vio que era necesario darle [$_{REC}$ *a Lima*] [$_{TH}$ *un mercado espacioso*]*, ...*

'… General Castilla found it necessary to give Lima a spacious market, ...'

(*Corpus del Español*)

The study of DCLD in Spanish is relevant in the current debate in the field of Linguistics, which has been increasingly focused during the last years on optional phenomena, also described using the term *gradient* (Bod et al., 2003; Gries, 2003; Fanselow et al., 2006; Aarts, 2007; Tagliamonte, 2011).

A possibly comparable optional phenomenon is English dative-shift (Arnold et al., 2000; Wasow, 2002; Bresnan et al., 2007). Dative-shift in English consists of the promotion of the recipient of a ditransitive verb from oblique (OBL) to object, resulting in a "double object construction". The alternative, without dative-shift is a construction in which the recipient is an oblique introduced by the preposition *to* ("dative prepositional construction"). The alternation between the dative prepositional construction and the double object construction is called the "dative alternation":

(7) a. Dative prepositional construction:

John gave a book to Mary.

b. Double object construction:

John gave Mary a book.

Given the optionality of DCLD, we resort to usage data gathered from corpora, which was subject to a quantitative study. The design of our study has been inspired by recent work on English dative-shift (Arnold et al., 2000; Wasow, 2002; Bresnan et al., 2007) which considers the interaction of multiple factors as predictors of speakers' choice of construction.

The studies of English dative-shift mentioned above arrive at the conclusion that speakers' choices cannot be predicted by a single factor. Instead, multiple factors, such as animacy, definiteness, givenness, lexical semantics, and grammatical complexity, are independently significant as predictors. The results of our study of DCLD in Spanish show interesting similarities with those obtained for English dative-shift, since factors such as animacy and givenness are significant in Spanish DCLD as well.

Although animacy and givenness are predictors of DCLD, our results also show interesting differences with respect to English dative-shift. Grammatical complexity, measured either by the length of words or the grammatical category of the theme and recipient, is not a predictor of DCLD, as opposed to what happens with English dative-shift. This fact, in our opinion, can be linked to the role of grammatical complexity as a predictor of word order (Hawkins, 1994, 2004) combined with the dissociation of word order and grammatical functions in Spanish. The association between grammatical complexity and dative-shift in English, we will argue, is not direct, but a side effect of the fact that English resorts to word order as a mechanism of overt coding of grammatical functions.

Regarding word order, our data lead to a discussion of the order of theme and recipient in ditransitive constructions, which can display either the order theme/recipient or recipient/theme independently of the occurrence of DCLD, as shown below.

(8) a. *María (le) envió un libro a Juan.*
María DAT.3SG sent a book to Juan
'María sent a book to Juan.'

b. *María (le) envió a Juan un libro.*
María DAT.3SG sent to Juan a book
' María sent a book to Juan.'

In this respect, our study shows that DCLD and order of theme and recipient are not associated variables and that they are constrained by different factors; animacy and givenness, in addition to dialectal and stylistic variation, in the case of DCLD, and mainly grammatical complexity in the case of the order of theme and recipient.

In order to complete the overview of dative constructions in Spanish, DCLD of recipients is compared with dative encoding of other thematic roles. DCLD is not only possible, but also required, in a set of Spanish constructions, in which the dative argument fulfills a set of thematic roles, including experiencer, beneficiary, and possessor, as shown below:

(9) a. Experiencer dative:
*A Juan *(le) gusta el helado.*
to Juan DAT.3SG likes the ice cream
'Juan likes ice cream.'

b. Benefactive dative[1]:
*Juan *(le) preparó la cena a María.*
Juan DAT.3SG prepared the dinner to María
'Juan fixed dinner for María.'

1 See discussion about the acceptability of this example in 2.1.2.3.

c. Possessive dative:
*Juan *(le) cortó el pelo a su amigo.*
Juan DAT.3SG cut the hair to his friend
'Juan cut his friend's hair.'

In order to compare the distribution of DCLD in ditransitive and non-ditransitive constructions, we conducted a quantitative study of one of the non-ditransitive dative constructions: the possessive construction. The results of that study show that, in spite of formal differences, a unified account of dative case in Spanish is possible, since the factors that constrain it overlap across different construction types.

1.1 THEORETICAL ASSUMPTIONS

Before describing the phenomena under study, we consider it necessary to make explicit the descriptive tools to be used. Throughout the dissertation, we will make use of a set of assumptions about the structure of human language. Our basic assumptions are related to the nature and inventory of grammatical functions and thematic roles, as well as the principles that regulate the mapping between them, especially when there are multiple possible mappings of thematic roles to grammatical functions (argument alternations).

1.1.1 Argument alternations

We assume that DCLD in Spanish is a mechanism of overt coding of grammatical functions and that it is involved in argument alternations, at least when it is required or preferred (experiencer,

benefactive, and possessive IOs). In Chapter 2 we will discuss if there is an argument alternation when DCLD is optional (recipient IOs).

An argument alternation takes place when, for a given predicate, there are two possible alignments of thematic roles to grammatical functions. For instance, the beneficiary argument of some predicates can be linked to an IO or an oblique (OBL) (a prepositional phrase that cannot undergo clitic doubling, see discussion in Chapter 2).

(10) a. Beneficiary → IO:

Juan le hizo un café a María.
Juan DAT.3SG made a coffee to María
'Juan made a coffee for María.'

b. Beneficiary → OBL:
Juan hizo un café para María.
Juan made a coffee for María
'Juan made a coffee for María.'

The alternations in the mapping between thematic roles and grammatical functions constitutes what has been called the "linking problem" (Maling, 2001).[2]

1.1.2 Grammatical Functions

Syntactic theories differ in the status of grammatical functions as primitives or derived categories. For instance, Chomsky (1986) claims that grammatical functions are derivative concepts that can be reduced to syntactic configurations. However, we will assume that

2 See Butt (2006) for an overview of Linking Theory.

grammatical functions are primitives of syntactic analysis, which cannot be further reduced to other properties, such as syntactic configuration, case, or agreement (Bresnan, 2001, Givón, 1984, 2001, Culicover and Jackendoff, 2005).

Grammatical functions can be characterized by taking into account two different sets of properties: overt coding properties and behavioral properties (Keenan, 1976; Givón, 1984, 2001). Overt coding properties include word order, agreement, and nominal morphology, in addition to case marking prepositions. Behavioral properties consist of the ability of a grammatical function to participate in specific constructions or processes, such as passivization or secondary predication. The inventory of grammatical functions that will be used in our descriptions include subject (SUBJ), direct object (DO), indirect object (IO) (Alsina, 1996), primary object (PO), secondary object (SO) (Dryer, 1986), and oblique (OBL).

Although the definition of SUBJ is far from uncontroversial,[3] our discussion will be focused on the differences between DO, IO, PO, SO, and OBL because these are the relevant grammatical functions necessary to describe the distribution of dative constructions.

It is a fact that, as far as nominative-accusative languages are concerned, SUBJ and DO enjoy a privileged position in the inventory of grammatical functions. From the point of view of linking, SUBJ and DO have been described as the only grammatical functions able to fulfill an unrestricted set of thematic roles (Levin, 1985), while other grammatical functions typically are able to fulfill a specific thematic role or a particular set of them. The SUBJ grammatical function can be occupied by arguments fulfilling diverse thematic roles, such as agent, patient, experiencer, instrument, etc., as shown below:

3 See Falk (2006) for a discussion of the nature of the SUBJ grammatical function.

(11) a. Agent SUBJ:
Juan rompió la ventana.
Juan broke the window
'Juan broke the window.'

b. Patient SUBJ:
La ventana se rompió.
the window SE broke
'The window broke.'

c. Experiencer SUBJ:
Juan se asustó.
Juan SE frightened
'Juan got frightened.'

d. Instrument SUBJ:
La piedra rompió la ventana.
the stone broke the window
'The stone broke the window.'

Oblique grammatical functions, on the other hand, are introduced by prepositions, at least in Spanish and English. Those prepositions restrict the possible thematic roles fulfilled by OBL arguments. An OBL introduced by the preposition *en* ('in') can indicate spatial or temporal location or manner. The semantics of the preposition restricts the possible thematic roles fulfilled by its complement:

(12) a. Spatial location OBL

Juan estudió en París.

Juan studied in Paris

'Juan studied in Paris.'

b. Temporal location OBL

Juan estudió en el verano.

Juan studied in the summer

'Juan studied in in the summer.'

c. Manner OBL:

Juan estudió en serio.

Juan studied in seriousness

'Juan studied seriously.'

SUBJ and DO are also considered, in the functionalist tradition, to be the grammaticalization of the main and secondary topics of a sentence (Givón, 1984, 2001). In addition to these semantic and pragmatic properties, SUBJ and DO display a set of formal properties that set them apart from other grammatical functions. SUBJ and DO are more likely to be overtly marked and to display salient behavioral properties. Because of all of these special properties, SUBJ and DO have been singled out in different grammatical frameworks as being special grammatical functions, also called "core", as opposed to "oblique", grammatical functions. The distinction has been stated in terms of "structural" vs. "inherent" case in Government and Binding (Chomsky, 1981, 1986). Nominative and accusative are considered "structural" and oblique cases are considered "inherent". Structural case is independent of thematic role assignment and inherent case is contingent to thematic role assignment. In Lexical Functional Grammar (Bresnan, 2001),

SUBJ and OBJ are defined as semantically unrestricted grammatical functions, in opposition to OBLs, which are characterized as semantically restricted (Levin, 1985).

The distinguishing formal property of SUBJs in Spanish is that they trigger agreement in person and number with the verb. From the point of view of overt coding, Spanish DOs display Differential Object Marking (DOM) (Bosson, 1986, Aissen, 2003). DOM is the tendency for languages with overt grammatical function marking of direct objects to optionally mark them according to their prominence. Spanish DOs are marked by the preposition *a* (personal *a*) when they are high in animacy and definiteness, and left unmarked when they are low in those scales.

(13) Contrast in definiteness:

a. *Vi *(a) la estudiante.*
saw to the student
'I saw the student.'

b. *Vi (a) una estudiante.*[4]
saw to a student
'I saw a student.'

(14) Contrast in animacy:

a. *Vi *(a) la estudiante.*
saw to the student
'I saw the student.'

b. *Vi (*a) la casa.*
saw to the house
'I saw the house.'

4 The indefinite NP *una estudiante* takes personal *a* if it is interpreted as a specific indefinite but it does not take if it is interpreted as a non-specific indefinite. See discussion in Chapter 4.

From the behavioral point of view, the promotion of a DOs to SUBJ of a passive sentence usually produces an acceptable sentence:

(15) a. *El ejército destruyó la ciudad.*
the army destroyed the city
'The army destroyed the city.'

b. *La ciudad fue destruida (por el ejército).*
the city was destroyed by the army
'The city was destroyed (by the army).'

OBLs, on the other hand, are the prototypical *non-core* grammatical function. OBLs are always introduced by a preposition, as opposed to objects, and the preposition restricts the semantic interpretation of the argument. From the formal point of view, in addition to being introduced by a preposition, OBLs do not present salient overt marking nor salient behavioral properties. Obliques cannot agree with the verb or undergo clitic doubling. Some examples of OBLs in Spanish are the phrases *con un cuchillo* (instrumental), *en Europa* (locative)*, hacia su casa* (goal) in the following sentences:

(16) a. *María cortó el pan con un cuchillo.*
María cut the bread with a knife
'María cut the bread with a knife.'

b. *Juan vive en Europa.*
Juan lives in Europe
'Juan lives in Europe.'

c. *María caminó hacia su casa.*
María walked towards her house
'María walked towards her house.'

The status of IOs in syntactic theory is controversial, to the extent that in some frameworks, such as Lexical Functional Grammar (Bresnan, 2001), they are not considered a grammatical function of their own[5]. The dismissal of the IO as a core grammatical function is due to the fact that in English it is circumscribed to a specific set of thematic roles (recipient or beneficiary). Therefore, it is semantically restricted and becomes a candidate to be demoted from the inventory of core grammatical functions. In addition, in English, it lacks overt coding or behavioral core properties. However, the situation is quite different in Spanish and other Romance languages, as pointed out by Masullo (1992), Alsina (1996), and Givón (1984, 2001). Masullo notices that IOs can fulfill as wide a range of thematic roles as SUBJs or DOs do. Also, there is evidence that IOs have behavioral properties of core grammatical functions. The discussion of the status of IOs in Spanish is one of the main issues of this dissertation and it will be addressed in Chapter 2.

PO and SO are grammatical functions which are restricted to the so-called "primary object languages", such as English. Dryer (1986) proposes that Accusative languages can be classified in two types depending on the way they treat their objects. Direct Object/Indirect Object languages (DO/IO) give the same treatment to the themes of ditransitive (double object) and transitive constructions. Primary Object/Secondary Object languages (PO/SO) give the same treatment to the recipient of a ditransitive (double object) and the theme of a transitive construction. Under this approach, English lacks DOs, and the theme argument of a transitive

5 See Alsina (1996) for a discussion of the status of IOs in Romance languages in the framework of LFG.

clause should be called PO instead. In English, POs of transitive sentences and POs of ditransitive (double object) sentences have the common behavioral property of becoming the subject of a passive sentence:

(17) a. *The army* (SUBJ) *destroyed the city* (PO).

b. *The city* (SUBJ) *was destroyed (by the army)* (OBL).

(18) a. *John* (SUBJ) *gave Mary* (PO) *a book* (SO).

b. *Mary* (SUBJ) *was given a book (by John).*

On the other hand, in a DO/IO language like Spanish, the recipient argument of a ditransitive sentence cannot become the subject of the passive. It is the theme argument of the ditransitive (DO) that becomes the subject of the passive instead (Raúl Aranovich, 2007):

(19) a. *María* (SUBJ) *(le) envió el libro* (DO) *a Juan* (IO).

María DAT.3SG sent the book to Juan
'María sent the book to Juan.'

b. *El libro* (SUBJ) *le fue enviado a Juan (por María).*
the book DAT.3SG was sent to Juan by María
'The book was sent to María (by John).'[6]

6 This sentence is possible in English as the passive of *John sent the book to María*, which is not a ditransitive. This sentence has only one object (*the book)* and an oblique (*to María*). The PO of this sentence is *the book*, which is therefore allowed to become the subject of the passive.

c. ***Juan*** (SUBJ) ***fue enviado el libro (por María).***[7]
Juan was sent the book by María
'Juan was sent the book (by María).'

1.1.3 Thematic roles

Thematic roles are the semantic labels used in order to designate the participants in the situations denoted by predicates. However, defining the inventory of thematic roles is a challenging task, because it is not always clear which semantic distinctions are grammatically relevant. A possible approach is to resort to individual, predicate-specific thematic roles. In this approach, a verb like "build" has two arguments, a "builder" and a "built". Although this approach is descriptively accurate, it seems to be missing a possible generalization.

The linking of thematic roles to grammatical functions is chaotic at times, but it is also quite predictable under some circumstances. Those predictable patterns can be captured by resorting to thematic role "types", like agent and patient. But it is not easy to determine the inventory of thematic role types. According to Ackerman and Moore (2001, 18-9):[8]

> Most commonly, the characterizations and definitions of these roles have been quite impressionistic and opportunistic: there has been little success in providing compelling, independently motivated characterizations of these roles, as well as few convincing arguments for delimiting a specific inventory of necessary roles which figure in linguistic explanation.

7 This sentence is acceptable in Japanese, a language that is also considered a Direct/Indirect Object language (Yasuhiro Shirai, personal communication).
8 A point that has been made earlier by Jackendoff (1987).

As Dowty (1991) has pointed out, it is almost impossible to draw a clear boundary between thematic role types. An example is the distinction between "agent", "author", and "instrument" (Frawley, 1992: 203). Frawley defines agent as "the deliberate, potent, active instigator of the predicate: the primary involved doer". The author, on the other hand, "has all the characteristics of an agent, but is not the direct cause of the act". And instrument is defined as "the means by which a predicate is carried out". Frawley provides the following examples of sentences with agent, author, and instrument subjects.

(20) a. ***The boy*** *dried the clothes with the blow drier* (agent).

b. ***The truck*** *ran over the rose bushes with its back tire* (author).

c. ***The rock*** *broke the window* (instrument).

Frawley presents evidence that these thematic distinctions are grammatically relevant in Russian, among other languages. However, as Dowty pointed out, by adopting finer distinctions, linguists can miss bigger generalizations about which argument will be subject and which argument will be object that can be stated by making reference to more general thematic roles (for instance, agent, defined as the argument that initiates an event, independently of volition).

Dowty (1991) rejects the notion of thematic role types as discrete categories with clear boundaries and proposes instead two cluster-concepts, Proto-Agent and Proto-Patient. Each of these prototypical categories is characterized by a set entailments, which are associated with the meaning of each verb. This approach allows for arguments to bear either of the two Proto-Roles

to different degrees, depending on the number of entailments provided by each predicate.

Dowty (1991: 572) provides the following lists of Proto-Agent and Proto-Patient entailments:

Contributing properties for the Agent Proto-Role:

a. volitional involvement in the event or state
b. sentience (and/or perception)
c. causing an event or change of state in another participant
d. movement (relative to the position of another participant)
(e. exists independently of the event named by the verb)[9]

Contributing properties for the Patient Proto-Role:

a. undergoes change of state
b. incremental theme
c. causally affected by another participant
d. stationary relative to movement of another participant
(e. does not exist independently of the event, or not at all)

Notice that Dowty's formulation captures the similarities and differences between "agent", "author", and "instrument" in the examples discussed above. Only "agents" are volitional and sentient, but the three roles cause an event or change of state in another participant and move relative to the position of another participant. "Agent" would be the most prototypical Proto-Agent, since it presents all of the Proto-Agent entailments and no Proto-Patient entailment whatsoever. "Author" and "instrument" present a subset of Proto-Agent entailments, but they can also present some Proto-Patient ones. "Instruments", in particular, can be causally affected by

9 Dowty lists this entailment in brackets because it could be attributed to the discourse status of subjecthood rather than to Proto-Role entailments. The same comment applies to Proto-Patient property e.

another participant. The relevance of Proto-Role theory for linking theory will be discussed in the next section.

Although we will assume Dowty's Proto-Role theory, we will use thematic role types for descriptive purposes, in particular while describing the phenomenon of DCLD in Chapter 2. The inventory and definitions of the thematic role types to be used are the following (definitions from Culicover, 2009, unless otherwise stated):

a. Experiencer: Participant which is in a perceptual or cognitive state. Example:

(21) ***A Juan*** *le gusta el helado.*
to Juan DAT.3S likes the ice cream
'Juan likes ice cream.'

b. Recipient: Participant that comes into possession of something. Example:

(22) *Juan le vendió un libro* ***a María.***
Juan DAT.3SG sold a book to María
'Juan sold a book to María.'

c. Benefactive: Participant for whose benefit the action is performed (Givón, 1984, 2001). Notice that recipient could be considered a type of benefactive. In general, benefactive is used to refer to entities which benefit from an event but without necessarily becoming the possessor of something. The definition of the boundary between recipient and benefactive underscores the advantages of the Proto-Roles approach. Example:

(23) *Juan le preparó la cena **a María.***
Juan DAT.3SG prepared the dinner to María
'Juan fixed dinner for María.'

d. Possessive: Participant which is in possession of something. This thematic role occurs both in the nominal and the verbal domain. Examples:

(24) a. Nominal argument:
*El pelo **de Juan.***
the hair of Juan
'Juan's hair.'

b. Verbal argument:
*Le cortó el pelo **a Juan.***
DAT.3SG cut the hair to Juan
'He/she cut Juan's hair.'

Notice that, as verbal argument, possessive is also difficult to distinguish from benefactive.

e. Theme: Participant which is in a location or state, or changes location or state. Example:

(25) *Juan le vendió **un libro** a María.*
Juan DAT.3SG sold a book to María
'Juan sold a book to María.'

f. Goal: Participant that denotes the final location or state of a change. Example:

(26) *Juan caminó* ***hacia su casa.***
Juan walked towards his house
'Juan walked towards his house.'

g. Location: Participant that denotes the location of a theme. Example:

(27) *Juan vive* ***en Europa*** .
Juan lives in Europe
'Juan lives in Europe.'

1.1.4 Linking theory

The definitions and inventories of grammatical functions and thematic roles play an essential role in linking. Moreover, the assumptions about grammatical functions and thematic roles are essential to the characterization of the linking mechanism itself.[10] For instance, Baker (1988) claims that the mapping of thematic roles to grammatical functions is completely uniform, as long as grammatical functions are considered as derivative notions and restated in terms of syntactic configurations at an underlying level of representation. In Lexical Functional Grammar (Bresnan, 2001) linking is regulated by linking rules that are constrained by hierarchies of grammatical functions and thematic roles. In this way, agent, the highest thematic role in the thematic hierarchy, is linked to SUBJ, the highest grammatical function. The problem in either case is that whatever linking principles are postulated, it is necessary for them to make reference

10 See Butt (2006) for an overview.

to thematic relations, which, as has been previously discussed are themselves controversial in linguistic theory.

An alternative approach to linking, formulated as the linking of Proto-Roles to grammatical functions, has been proposed by Dowty (1991). Dowty's approach is expressed by the Argument Selection Principle (ASP), as stated below:[11]

> ARGUMENT SELECTION PRINCIPLE: In predicates with grammatical subject and object, the argument for which the predicate entails the greatest number of Proto-Agent properties will be lexicalized as the subject of the predicate; the argument having the greatest number of Proto-Patient entailments will be lexicalized as the direct object (Dowty, 1991: 576).

This principle predicts that, for instance, agent, author, and instrument type thematic roles will be lexicalized as subject, as long as there is no other argument with a greatest number of Proto-Agent properties.

The linking patterns of psychological predicates illustrate the advantages of the prototypical approach to thematic roles. The linking of thematic roles to grammatical functions in psychological predicates such as *fear* and *frighten* is problematic from the point of view of linking because they display opposed linking patterns.[12] The thematic roles which are generally assumed to participate in these predicates are experiencer, previously defined, and stimulus, understood as the source or cause of the perceptual or cognitive state experienced by the experiencer. In the case of *fear*, the experiencer is linked to SUBJ and the stimulus to PO. The

11 This formulation is valid for nominative-accusative languages, but not for ergative-absolutive languages.
12 Chomsky (1965) discusses the syntax of psychological predicates.

opposed happens in the case of *frighten*, since the stimulus is linked to SUBJ and the experiencer to PO, as shown below:

(28) a. *John fears the storms.*

b. *The storms frighten John.*

In order to provide a unified analysis of these predicates it is necessary to claim that either the experiencer in both cases occupies the same syntactic position at an underlying level of syntactic representation, following Baker (1988), or that different sets of thematic roles are involved. If the thematic roles are not the same, any linking mechanism will be expected to treat them differently. The first move requires the adoption of additional assumptions about syntactic representations, including underlying levels of representations, derivations, and empty categories, among others. The second move requires the postulation of additional thematic relations. A third option, which we will not discuss here, would be to assume a decompositional theory of predicates (Pinker, 1989; Jackendoff, 1990a).

These problems are avoided if the thematic relations are defined in terms of Proto-Roles. The alternative linking can be understood as a result of the fact that *frighten* entails that its stimulus causes the event or change of state in another participant, while *fear* does not present that entailment. On the other hand, *frighten*, but not *fear*, entails that its experiencer is causally affected by another participant. As a result, the experiencer has more Proto-Agent entailments than the stimulus in the case of *fear*, but the opposite happens in the case of *frighten*, a situation that predicts the linking patterns observed in the data without resort to neither additional syntactic

machinery, nor additional thematic relations.

The linking patterns in ditransitive predicates, such as *give*, which are at the core of our inquiry, are not addressed by the Argument Selection, which only makes reference to predicates that have two (agent and patient), but not three arguments (agent, patient or theme, goal or beneficiary). The linking in ditransitive predicates is addressed by Corollary 2 of the ASP:

> COROLLARY 2: With a three-place predicate, the non-subject argument having the greater number of entailed Proto-Patient properties will be lexicalized as the direct object and the non-subject argument having fewer entailed Proto-Patient properties will be lexicalized as an oblique or prepositional object (and if two non-subject arguments have approximately equal numbers of entailed Proto-Patient properties, either or both may be lexicalized as direct object) (Dowty, 1991: 576).

Going back to the English dative alternation, Corollary 2 of the ASP predicts that in a ditransitive sentence there is a competition between the theme and the recipient for the object position. This view is compatible with the results of the study of English dative-shift performed by Bresnan et al. (2007) which shows that pragmatically relevant grammatical properties of both the theme and the recipient (animacy, definiteness, givenness) are significant predictors of dative-shift. We will deal in more detail with the issue of lexical semantics in section 3.2.

1.2 OVERVIEW OF THE DISSERTATION

Chapter 1 has presented the topic and theoretical background of the dissertation. Chapter 2 describes in detail the phenomena and discusses the grammatical status of dative clitics and dative arguments. We argue that dative clitics are agreement markers and that IO is a core grammatical function in Spanish. Chapter 3 presents the different approaches that have been taken in the literature about English dative-shift and Spanish DCLD. Our conclusion is that DCLD in Spanish cannot be accounted for in purely lexical or syntactic terms. Chapter 4 describes our quantitative study of DCLD in dative constructions and its results, which show that DCLD in Spanish is an optional phenomenon constrained by multiple factors. The results also show that DCLD and order of theme and recipient are not associated variables. Chapter 5 describes our quantitative study of possessive constructions. The results show important similarities in the distribution of dative case in ditransitive and non-ditransitive constructions. Chapter 6 presents the general conclusions of the dissertation.

2.0 DATIVE CLITIC DOUBLING AND GRAMMATICAL FUNCTIONS

The purpose of this chapter is to present the basic data about dative clitic doubling in Spanish and discuss the status of dative clitics and IOs. Regarding dative clitics in doubling constructions, we will argue that they behave as inflectional affixes (agreement markers) rather than as clitic pronouns. Being agreement markers, it is natural to assume that dative clitics are part of the system of overt coding of grammatical functions. Regarding IOs, we will argue that in Spanish they display the properties of a core grammatical function. The chapter also discusses the functional load of dative clitics in ditransitive and non-ditransitive constructions.

2.1 CLITIC DOUBLING

Clitic Doubling (CLD) is a widely studied feature of Spanish grammar (Strozer, 1976; Suñer, 1988; Andrews, 1990; Halpern, 2001; Anagnostopoulou, 2006). It consists of the co-occurrence of an accusative or dative unstressed pronoun phonologically attached to the verb with a co-referential Noun Phrase (NP).

(1) a. Accusative Clitic Doubling (ACLD):
Yo lo vi a Juan.
I ACC.3SG.MASC saw to Juan
'I saw Juan.'

b. Dative Clitic Doubling (DCLD):

Yo le di un libro a Juan.

I DAT.3SG gave a book to Juan

'I gave a book to Juan.'

Example a. shows the doubling of a Direct Object (DO) by the 3rd person, singular, masculine, accusative clitic *lo.*[13] Example b. shows the doubling of an Indirect Object (IO) by the 3rd person, singular, dative clitic *le*.

Spanish clitic pronouns are inflected for person, number, and case, as well as gender in the 3rd person accusative. There is case syncretism in the 1st and 2nd persons; therefore the opposition between accusative and dative is morphologically marked only in the 3rd person. Clitic and antecedent agree in person, number, gender, and case (Suñer, 1988). The following chart shows the paradigm of Spanish accusative and dative clitic pronouns.

Table 2.1: Unstressed set of Spanish pronouns (clitics)

	1sg	2sg	3sg	1pl	2pl	3pl
Acc.			*lo/la*			*los/las*
	me	*te*		*nos*	*os*[14]	
Dat.			*le*			*les*

13 This example is acceptable in some varieties of American Spanish, but not in Peninsular Spanish.

14 *Os* as a 2nd person plural pronoun is used in Peninsular Spanish, but has been lost in Latin American Spanish, which uses the 3rd person plural forms (*los, las, les*) instead. *Os* in Peninsular Spanish alternates with the 3rd person forms depending on the degree of formality of the situation.

2.1.1 Accusative clitic doubling

Accusative clitic doubling is required in Spanish when the DO is pronominal. This fact is not subject to dialectal variation:

(2) **(La) vi a ella.*
ACC.3SG.FEM saw to her
'I saw her.'

However, there is an important degree of dialectal variation when the DO is not a pronoun. Peninsular Spanish disallows accusative clitic doubling in those cases, independently of animacy or other considerations:

(3) a. *(*La) vi a María.* (Peninsular Spanish)
ACC.3SG.FEM saw to María
'I saw María.'

b. *(*La) vi la casa.* (Peninsular Spanish)
ACC.3SG.FEM saw the house
'I saw the house.'

However, some varieties of Latin American Spanish do allow for optional accusative clitic doubling of proper nouns (Limeño Spanish; Mayer, 2006; Rioplatense Spanish; Suñer, 1988) and specific common nouns (Rioplatense Spanish; Suñer, 1988).

(4) a. *(La) vi a María.* (Limeño Spanish,

ACC.3SG.FEM saw to María Rioplatense Spanish)

'I saw María.'

b. *(La) vi la casa.* (Rioplatense Spanish)

ACC.3SG.FEM saw the house

'I saw the house.'

Another context that requires the use of the clitic is Clitic Left or Right-dislocation (CLLD, CLRD). This fact is not subject to dialectal variation:

(5) a. *A María, *(la) vi ayer.*

to María ACC.3SG.FEM saw yesterday

'María, I saw her yesterday.'

b. **(La) vi ayer, a María.*

ACC.3SG.FEM saw yesterday to María

'I saw her yesterday, María'

(6) a. *La casa, *(la) vi ayer*

the house ACC.3SG.FEM saw yesterday

'The house, I saw it yesterday.'

b. **(La) vi ayer, la casa.*

ACC.3SG.FEM saw yesterday the house

'I saw it yesterday, the house.'

The use of the clitic in the case of CLLD or CLRD, however, is not considered an instance of doubling in the relevant sense because the DO in the core sentence is the clitic itself (Bresnan,

2001) or a null pronoun (Suñer, 1988) and the dislocated NP is in the periphery of the sentence. Therefore, the clitic acts as a resumptive pronoun, as in the case of English Left Dislocation (*Mary, I saw her yesterday*). Since CLLD and CLRD are not cases of doubling in the relevant sense, they will not be considered in this study.

2.1.2 Dative clitic doubling (DCLD)

Dative clitic doubling (DCLD) is not supposed to display the high degree of dialectal variation observed in accusative clitic doubling (Fernández Soriano, 1999), although it has been noticed that it is more common in the Americas than in Spain (Becerra Bascuñán, 2006). Dative clitic doubling is required when the IO is pronominal. Also, dative clitics act as resumptive pronouns in cases of CLLD or CLRD. These facts are not subject to dialectal variation:

(7) **(Le) di el libro a ella.*

DAT.3SG gave the book to her
'I gave the book to her.'

(8) a. *A María, *(le) di el libro ayer.*
to María DAT.3SG gave the book yesterday
'María, I gave the book to her yesterday.'

b. **(Le) di el libro ayer, a María.*
DAT.3SG gave the book yesterday to María
'I saw her yesterday, María.'

Regarding clitic doubling of non-pronominal IOs, the distribution varies among predicate types. Doubling is optional if the IO is the recipient argument of ditransitive predicate like *enviar* ('to

send') or *lanzar* ('to throw'). But doubling is required if the IO is the experiencer argument of a psychological predicate like *gustar* ('to like'), which is syntactically intransitive. *Gustar* ('to like') is considered intransitive in Spanish because it cannot take a DO. Doubling is also required when the IO is a benefactive or a possessor.

(9) a. Recipient dative:
(Le) envié el libro a Juan.
DAT.3SG sent the book to Juan
'I sent the book to Juan.'

b. Experiencer dative:
*A Juan *(le) gusta este helado.*
to Juan DAT.3SG likes this ice cream
'Juan likes this ice cream.'

c. Benefactive dative:
*Juan *(le) preparó la cena a María.*[15]
Juan DAT.3SG prepared the dinner to María
'Juan fixed dinner for María.'

d. Possessive dative:[16]
*Juan *(le) cortó el pelo a su amigo.*
Juan DAT.3ST cut the hair to his friend
'Juan cut his friend's hair.'

The following subsections provide more detail about the linking patterns in each type of dative construction.

15 See discussion about the acceptability of this example in 2.1.2.3.
16 In Chapter 5 it will be shown that not only inalienable possessors undergo dativization.

2.1.2.1 Recipient datives (Masullo, 1992; Fernández Ordoñez, 1999; Delbecque and Lamiroy, 2006) Recipient datives occur as dependents of ditransitive predicates such as *dar* 'to give', *obsequiar* 'to present', *entregar* 'to hand', *aclarar* 'to explain', *asegurar* 'to assure', *confesar* 'to confess', *acercar* 'to bring close(r)', *lanzar* 'to throw', *llevar* 'to bring', *ofrecer* 'to offer', *conferir* 'to confer', *dedicar* 'to dedicate', etc.

The most intriguing property of DCLD in the context of all the recipient datives is its optionality:

(10) *Juan (le) envió una carta a María.*
Juan DAT.3SG sent a letter to María
'Juan sent a letter to María.'

Another property of the construction is that the order of DO and IO is optional, both DO/IO and IO/DO orders are possible:

(11) a. *Juan (le) envió una carta a María.*
Juan DAT.3SG sent a letter to María
'Juan sent a letter to María.'

b. *Juan (le) envió a María una carta.*
Juan DAT.3SG sent a María a letter
'Juan sent a letter to María.'

This word order alternation has been analyzed as a phenomenon analogous to English dative-shift (Demonte, 1995). As was mentioned in Chapter 1, English dative-shift is an alternation of grammatical functions which allows for two possible linking patterns. In the prepositional dative

construction (*John gave the book to Mary*), the theme is linked to Primary Object (PO) and the recipient to Oblique (OBL). In the double object construction, the recipient is linked to PO and the theme to Secondary Object (SO).[17] This alignment of thematic roles to grammatical functions can be confirmed by looking at the behavior of these constructions under passivization:

(12) English prepositional dative construction:

a. *John gave the book* (PO) *to Mary* (OBL).

b. *The book was given to Mary (by John).*

c. **To Mary was given the book (by John).*

d. **Mary was given the book to (by John).*

(13) English double object construction:

a. *John gave Mary* (PO) *the book* (SO).

b. *Mary was given the book (by John)*.

c. **The book was given Mary (by John)*.[18]

Notice that in the English prepositional dative construction the theme is the PO, as shown by the fact that it becomes the subject of the passive. On the other hand, in the English double object construction, the recipient is the PO, as shown by the same test. The syntactic effects of English

17 It has been claimed that the double object construction and the dative prepositional construction do not have the same meaning (Pinker, 1989). See discussion in Chapter 3. The status of English as a Primary/Secondary Object language has been discussed in Chapter 1.

18 This construction is acceptable with some verbs in British English.

dative-shift are represented in the following diagram, which shows the linking of thematic roles to grammatical functions in both versions of the alternation:[19]

(14) a. English prepositional dative construction:

Thematic roles: *give*<	agent,	theme,	recipient>
	|	|	|
Grammatical functions:	SUBJ	PO	OBL

b. English double object construction:

Thematic roles: *give*<	agent,	theme,	recipient>
	|	|	|
Grammatical functions:	SUBJ	PO	SO

However, opposite to what happens in English, recipients cannot become the subject of a passive in Spanish, irrespective of the occurrence of DCLD or the order of theme and recipient:

(15) a. *Juan (le) envió el libro a María / a María el libro.*
Juan DAT.3SG sent the book to María / María the book
'Juan sent the book to María/María the book.'

b. *El libro (le) fue enviado a María (por Juan).*
the book DAT.3SG was sent to María by Juan
'The book was sent to María (by Juan).'

c. **María fue enviada el libro (por Juan).*
María was sent the book by Juan
'Mary was sent the book (by Juan).'

19 We will not address in this dissertation the issue of how lexical entries should look like.

In **María fue enviada el libro (por Juan)*, *María* has become subject of the passive, as shown by the passive participle feminine agreement. This sentence is unacceptable independently of the order of constituents in the active counterpart. This fact shows that the theme is always the DO in Spanish ditransitive constructions and constitutes evidence of an important typological difference between English and Spanish: the former is a Primary Object/Secondary Object language but the latter is a Direct Object/Indirect Object language, as has been discussed in Chapter 1. Therefore, DCLD and dative-shift do not seem to be completely analogous. The difference would be that in Spanish the alternation does not affect the grammatical coding of the theme, which is always a DO, but only the coding of the recipient, which can be overtly marked as IO by the clitic or not.

Linking patterns in ditransitives According to Masullo (1992), the occurrence of DCLD in Spanish is the morphological expression of an argument alternation in which the recipient is linked to either OBL or IO. The recipient would be OBL when there is no DCLD and IO when DCLD takes place, as shown below:

(16) a. Ditransitive with DCLD:

Thematic roles: *dar*<	agent,	theme,	recipient>
	|	|	|
Grammatical functions:	SUBJ	DO	IO

b. Ditransitive without DCLD:

Thematic roles: *dar*<	agent,	theme,	recipient>
	|	|	|
Grammatical functions:	SUBJ	DO	OBL

Masullo's analysis of Spanish dative constructions is based on the idea that the distribution of DCLD in Spanish is the same as that of applicative markers in languages that have that type of morphology, as in the case of Bantu languages (Mchombo, 2001). An applicative marker is an affix that is attached to the verb when an argument alternation consisting of the promotion of an oblique grammatical relation to core grammatical status (the so-called "applied object") and the demotion of the object to secondary object status takes place.

However, the evidence for the existence in Spanish of the argument alternation depicted above is not conclusive. The most compelling argument is that assuming that DCLD of IOs is required would provide a unified account of DCLD in ditransitive and non-ditransitive constructions (experiencer, benefactive, and possessive datives). In the following sections we will describe those non-ditransitive dative constructions and we will return to the discussion of the existence of an argument alternation in the ditransitives later in the chapter.

2.1.2.2 Experiencer datives (Masullo, 1992; Fernández Ordoñez, 1999; Ackerman and Moore, 2001; Delbecque and Lamiroy, 2006) Experiencer datives occur as IOs of a set of psychological verbs (*asustar* 'to frighten', *molestar* 'to bother', *preocupar* 'to preoccupate', etc.). These verbs take two arguments, an experiencer and a stimulus, which are realized as IO and subject (SUBJ), respectively. In this context, DCLD is required:

(17) *A Juan *(le) asustan los ruidos.*
to Juan DAT.3SG frighten the noises
'The noises frighten Juan.'

This pattern alternates with another one, in which the experiencer is accusative instead of dative (the stimulus is the SUBJ in both cases):

(18) *Los ruidos (lo) asustan a Juan.*
the noises ACC frighten to Juan
'The noises frighten Juan.'

Therefore, there are two possible alignments of thematic roles to grammatical functions. The stimulus is SUBJ in both cases, but the experiencer alternates as IO/DO (see below).

(19) a. Dative experiencer (*A Juan le asustan los ruidos*)

Thematic roles: *asustar*<	experiencer,	stimulus>
	|	|
Grammatical functions:	IO	SUBJ

b. Accusative experiencer (*Los ruidos lo asustan a Juan*)

Thematic roles: *asustar*<	stimulus	experiencer >
	|	|
Grammatical functions:	SUBJ	DO

2.1.2.3 Benefactive datives (Masullo, 1992; Fernández Ordoñez, 1999; Delbecque and Lamiroy, 2006) In Spanish, benefactives can occur either as OBLs or IOs. As obliques, benefactives are introduced by the preposition *para* ('for') and do not allow DCLD. As IOs, benefactives are introduced by *a* (“to”) and require DCLD:

(20) a. *Juan (*le$_i$) preparó la comida para María$_i$.*
Juan DAT.3SG prepared the food for María
'Juan prepared the meal for María.'

b. *Juan *(le$_i$) preparó la comida a María$_i$.*
Juan DAT.3SG prepared the food for María
'Juan prepared the meal for María.'

These judgments, which coincide with ours, are from Masullo (1992) and Fernández Ordoñez (1999). However, according to Roberto Mayoral Hernández (personal communication) *Juan preparó la comida a María* is acceptable without DCLD. It is possible that DCLD in this context, which is considered required in the literature, is actually subject to variation. However, for the purposes of this discussion, we will assume that DCLD in this context is the preferred option.

In this alternation, the benefactive can be realized as either OBL or IO.

(21) a. Dative benefactive (*Juan le preparó la comida a María.*)

Thematic roles: *preparar*<	agent,	patient,	benefactive>
	|	|	|
Grammatical functions:	SUBJ	DO	IO

b. Oblique benefactive (*Juan preparó la comida para María.*)

Thematic roles: *preparar*<	agent,	patient,	benefactive>
	|	|	|
Grammatical functions:	SUBJ	DO	OBL

2.1.2.4 Possessive datives (Masullo, 1992; Fernández Ordoñez, 1999; Delbecque and Lamiroy, 2006) Possessives[20] in Spanish can be realized as an embedded genitive NP or, in certain contexts, as a syntactically independent IO. Dative clitic doubling is not allowed in the former, but required in the latter:

(22) a. *Juan (*le$_i$) cortó [$_{NP}$ el pelo **de** María$_i$]*.
Juan DAT.3SG cut the hair of María
'Juan cut María's hair.'

b. *Juan *(le$_i$) cortó [$_{NP}$ el pelo] [$_{PP}$ **a** María$_i$]*.
Juan DAT.3SG cut the hair to María
'Juan cut María's hair.'

This alternation involves two different alignments of thematic roles to grammatical functions. One of them contains a genitive possessor embedded within the DO. In the other one, the possessor is a dependent[21] of the verb (IO).

(23) a. Dative possessive (*Juan le cortó el pelo a María* 'Juan cut María's hair')

Thematic roles: *preparar*<	agent,	patient,	possessive>
	|	|	|
Grammatical functions:	SUBJ	DO	IO

20 The role of the distinction between alienable and inalienable possession regarding dativization is discussed in Chapter 5.

21 We use *dependent* in a broad sense as any unit standing in a relation of *dependency* ("Any relation in which one element is taken to imply the other" according to Matthews, 2007). In this use, *dependent* includes both complements and modifiers.

b. Genitive possessive (*Juan cortó el pelo de María* 'Juan cut María's hair'.)

Thematic roles: *preparar*<	agent,	patient	<possessive>>
	|	|	|
Grammatical functions:	SUBJ	[DO	[GENITIVE]]

The fact that the possessor is an independent constituent in the dative construction, but not in the genitive one, is confirmed by the fact that the dative can be detached from the possessed noun, for instances by left-dislocation. That option is not available in the case of the genitive construction:

(24) a. *A Juan, María le pintó la casa.*
to Juan María DAT.3ST painted the house
'Juan, María painted the house for him.'

b. **De María, Juan pintó la casa.*
of María Juan painted the house

2.1.3 Discussion

Masullo (1991) provides the insight that dative clitic doubling is associated with argument alternations (dativization of experiencers, benefactives, and possessors). Masullo extends this analysis to ditransitive verbs, claiming that dative clitic doubling of recipients is actually required. The appearance of optionality would be the result of the existence of two superficially similar but underlyingly different constructions. The confounding factor, according to Masullo, would be the ambiguity of Spanish *a,* which can be either a semantically vacuous prepositional case marker or a goal preposition. The use of *a* as a goal preposition is exemplified below:

(25) *Juan fue a Buenos Aires.*
Juan went to Buenos Aires
'Juan went to Buenos Aires.'

The use of *a* (Latin *ad* > *a*) as a case marker developed in early stages of Romance (Pensado, 1995). Latin *ad* governed dative case and its use as an (objective) case marker spread from dative to accusative. Consider the example of optional dative clitic doubling in a ditransitive construction shown below:

(26) *Yo (le) di un libro a Juan.*
I DAT.3SG gave a book to Juan
'I gave a book to Juan.'

Masullo claims that if there is clitic doubling *a Juan* is an actual IO, but it is an OBL if doubling does not occur. The goal of this analysis is to provide a unified account of dative clitic doubling, eliminating optionality from the grammar. In our opinion, a unified account should be preferred if it is able to account for the whole range of data. However, this analysis presents empirical problems. Consider the behavior of obliques with respect to left-dislocation (see below).

(27) a. *Juan (*le) preparó la comida para María.*
Juan DAT.3SG prepared the food for María
'Juan prepared the food for María.'

b. *Para María, Juan (*le) preparó la comida.*
for María Juan DAT.3SG prepared the food
' María, Juan prepared the food for her.'

The previous example shows that when an oblique (*para María* 'for Mary') undergoes left-dislocation, DCLD is not allowed. Therefore, if the recipient of a ditransitive verb is an oblique, it should be able to undergo left-dislocation without DCLD. But this prediction does not hold. If the recipient of a ditransitive verb is left-dislocated, DCLD is required (see below).

(28) a. *Juan envió un libro a María.*
Juan DAT.3SG sent a book to María
'Juan sent a book to María.'

b. *A María, Juan *(le) envió un libro.*
to María Juan DAT.3SG sent a book
' María, Juan sent a book to her.'

The only way to account for the previous example under the assumption that *a María* is an oblique would be to stipulate that obliques cannot be left-dislocated in Spanish, contrary to fact.

Another empirical problem for this analysis is that, as has been repeatedly pointed out, DCLD is obligatory if the recipient is a pronoun (see below).

(29) *Juan *(le) envió un libro a ella.*
Juan DAT.3SG sent a book to her
'Juan sent a book to her.'

Therefore, in this sentence, *a ella* ('to her') can only be an IO, but not an oblique, because obliques do not allow for DCLD. Therefore, this analysis predicts that the unacceptable sentence **Yo di un libro a ella* should be possible as an alternative to the acceptable *Yo le di un libro a ella*. The difference would be that in one case the recipient would be an OBL and in the other

case it would be an IO. The only way out of this problem would be to stipulate that pronouns cannot fulfill oblique grammatical functions, contrary to fact, as shown below.

(30) *Juan preparó la comida para ella.*
Juan DAT.3SG prepared the food for her
'Juan prepared the food for her.'

Our conclusion is that the analysis of DCLD in ditransitive constructions does not provide evidence that an argument alternation has taken place. Instead, we will assume that optional DCLD in ditransitive sentences is an instance of optional overt coding of the grammatical function IO. Optional coding, by means of optional agreement marking or optional case marking is a well attested phenomenon in natural language (Comrie, 1989; Woolford, 1999). Spanish is an example of a language that has optional case marking of DOs, by means of personal *a.* Also, accusative clitic doubling has been analyzed as an instance of optional agreement of the verb with the DO (Suñer, 1988).

In the reminder of the chapter we will discuss the status of dative clitics as optional agreement markers and the status of the IO as a core grammatical function in Spanish.

2.2 ON THE GRAMMATICAL STATUS OF DATIVE CLITICS

The discussion of the grammatical status of pronominal clitics has focused on characterizing them as either pronominal arguments or agreement markers (Givón, 1976; Bresnan and Mchombo, 1987; Suñer, 1988; Bresnan, 2001). A pronominal argument (or incorporated

pronoun) is a bound morpheme that saturates an argument of the verb. A pronominal argument behaves syntactically as a morphologically independent NP. Therefore, pronominal arguments are in complementary distribution with lexical NPs fulfilling the same grammatical function, with the exception of dislocated topics (Bresnan, 2001: 145). An agreement marker, on the other hand, does not saturate an argument of the verb and co-occurs with a co-referential NP specified for the same agreement features (person, number, and gender). An uncontroversial example of an agreement marker in Spanish is the subject agreement affix (*-n*, '3PL' in the example below).

(31) *Los estudiantes camin-a-n por el campus.*
the students walk-TV-3PL through the campus
'The students walk through campus.'

The subject agreement marker does not saturate an argument of the verb. Its function is to cross-reference the verb with one of its arguments, the subject. On the other hand, Spanish stressed pronouns are morphologically and syntactically independent constituents which do saturate an argument of the verb in the same fashion as a full lexical NP. See example below, where the stressed pronoun fulfills the subject grammatical function (*ellos*, 'NOM.3PL.MASC'):

(32) *Ellos camin-a-n por el campus.*
NOM.3PL.MASC walk-TV-3PL through the campus
'They walk through campus.'

If clitics are pronominal arguments, they are supposed to behave syntactically as stressed pronouns, the only difference being morphophonological in nature (clitic pronouns are morphophonologically attached to a host). If clitics are agreement markers, they are supposed to behave as subject agreement. If pronominal clitics behave like stressed pronouns, in clitic doubling constructions the associated NP would be a left or right dislocated constituent (Jaeggli, 1982). However, the analysis of dative clitics as pronominal arguments is problematic, as will be discussed below. The empirical criteria that can be used in order to support the analysis of Spanish dative pronominal clitics as agreement markers are word order, focus, interrogation, and referential properties. Our discussion follows the criteria applied by Bresnan and Mchombo (1987) to the discussion of Chichewa agreement system, also discussed in Bresnan (2001).

The claim that doubled clitics are agreement markers is further supported by evidence of their behavior as inflectional affixes rather than clitics. We discuss six criteria to distinguish affixes from clitics: degree of selection with respect to the host, rigid ordering, coordination, arbitrary gaps, idiosyncrasies, and verb left-detachment (Zwicky and Pullum, 1983; Monachesi, 1998, 2005).

Our conclusion is that doubled clitics behave as agreement markers, a type of inflectional affix. This conclusion is relevant in the context of the dissertation because being agreement markers, it is expected that clitics would be part of the system of overt marking of grammatical functions.

2.2.1 Pronominal clitics as agreement markers

In this section we will discuss the empirical criteria that support the analysis of Spanish dative pronominal clitics as agreement markers: word order, focus, interrogation, and referential properties

2.2.1.1 Word order and dislocation Agreement markers do not saturate an argument of the verb. Therefore, they freely co-occur with co-referential arguments in the sentence. This is the case of subject agreement in Spanish, as shown above. Stressed pronouns, on the other hand, cannot co-occur with a co-referential argument in the sentence, unless it is a dislocated topic:

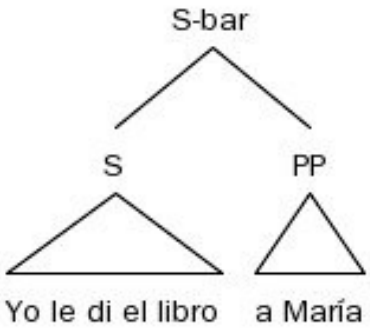
(33) a. Non-dislocated constituent doubled by a stressed pronoun:
**Los estudiantes$_i$ ellos$_i$ camin-a-n por el campus.*
the students NOM.3PL.MASC walk-TV-3PL through the campus
*'The students they walk through campus.'

a. Dislocated constituent doubled by a stressed pronoun*:*
Los estudiantes, ellos camin-a-n por el campus.
the students NOM.3PL.MASC walk-TV-3PL through the campus
'The students, they walk through campus.'

The analysis of clitics as pronominal arguments implies that the clitic itself is an argument of the verb, while the associated NP is a topic occurring in a dislocated, peripheral position, either to the right or to the left of the core sentence (Jaeggli, 1982; Aoun, 1985). In this analysis, along the lines of the Pronominal Argument Hypothesis (Jelinek, 1984), the clitic performs the same function of a stressed pronoun, referring anaphorically or cataphorically to a dislocated topic. In

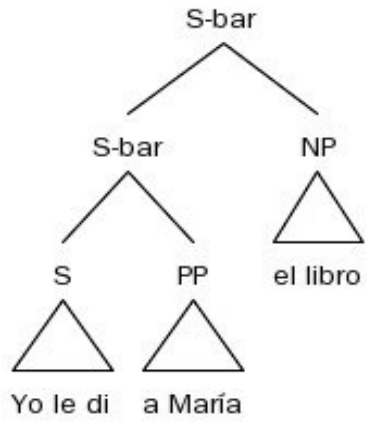
this approach, the syntactic structure of a sentence like: *Yo le di el libro a María* ('I gave a book to Mary') would be the following:

(34)



Regarding the possible positions of recipient NPs in the sentence, it has been already pointed out that both the order theme/recipient or recipient/theme are attested. So, in addition to *Yo le di el libro a María*, it is also possible to say *Yo le di a María el libro*. If the order is recipient/theme, the recipient NP can be a right-dislocated topic only if the theme NP is right-dislocated as well, as shown below.

(35)



Since Spanish allows for topic stacking (Zagona, 2002: 221), both to the left and the right of the core sentence, this analysis is plausible. However, it is not viable because both left and right-dislocation of accusative objects require accusative clitic doubling, which does not take place in this sentence. For the topic-stacking analysis to be viable, the sentence should be *Yo se lo di a María el libro*, where *lo* is the accusative clitic and *le* is substituted by *se* as the result of morphologically conditioned allomorphy. This example shows that dative clitics can double arguments of the verb which are not dislocated. Therefore, dative clitics behave like subject agreement markers rather than like stressed pronouns in this respect.

2.2.1.2 Focus and interrogation Another difference in the behavior of the subject agreement affix and stressed pronouns is that the former, but not the latter, can co-occur with a co-referential interrogative or focused constituent. See examples below:

(36) a. Interrogative subject:

¿Quiénes camin-a-n por el campus?
who walk-TV-3PL through the campus
'Who walks through campus?'

b. Focused subject:
LOS ESTUDIANTES camin-a-n por el campus (no los visitantes).
the students walk-TV-3PL through the campus (not the visitors)
'It is the students that walk through campus (not the visitors).'

But stressed pronouns cannot co-occur with co-referential interrogative or focused constituents:

(37) a. Interrogative subject:
**¿Quiénes ellos camin-a-n por el campus?*
who NOM.3PL.MASC walk-TV-3PL through the campus

b. Focused subject:
**LOS ESTUDIANTES ellos camin-a-n por el campus(no los visitantes).*
the students NOM.3PL.MASC walk-TV-3PL through the campus (not the visitors)

These facts occur because stressed pronouns can only co-occur with co-referential dislocated constituents (topics). But interrogative and focused constituents cannot be dislocated because they express new information instead of given information. The examples below show that interrogative and focused constituents cannot be dislocated (the dislocation is indicated by a separation of the constituent from the sentence by means of a comma).

(38) Interrogative subject:
¿Quiénes, camin-a-n por el campus?
who walk-TV-3PL through the campus

Focused subject:
**LOS ESTUDIANTES, camin-a-n por el campus(no los visitantes).*
the students walk-TV-3PL through the campus (not the visitors)

The examples below show that interrogative and focused IOs can undergo DCLD:

(39) a. Interrogative IO:
¿A quién le diste el libro?
To who DAT.3SG the book
'Who did you give the book to?'

b. Focused IO:
A MARIA le di el libro, no a Juan.
to María DAT.3SG the book not to Juan
'It was Mary that I gave the book to, not John.'

These facts show that dative clitics behave like subject agreement affixes rather than as stressed pronouns with respect to interrogative and focused constituents.

2.2.1.3 Emphatic pronouns Spanish subject agreement can be either anaphoric or grammatical, but not emphatic. It is grammatical when it matches the person and number features of the subject, in which case it constitutes the grammatical encoding of the subject grammatical function:

(40) *Juan fue a la librería.*
Juan went to the bookstore
'Juan went to the bookstore.'

It is anaphoric when the subject has been dropped and its reference is recovered from context:

(41) *Allí, compró los libros que necesitaba.*
there bought the books that needed
'There, he bought the books he needed.'

On the other hand, the use of Spanish stressed pronouns is emphatic. Spanish stressed pronouns are always redundant because the information they provide is also provided by either the subject agreement affix or a pronominal clitic (clitic doubling of stressed pronouns is always required). In the following example, the use of the stressed pronoun would be given an emphatic interpretation.

(42) *Allí, él (no María) compró los libros que necesitaba.*
there NOM.3SG.MASC (not María) bought the books that needed
'There, HE (not María) bought the books he needed.'

Dative clitics, as well as accusative ones, behave analogously to subject agreement markers. They can either match the person, number, and gender features of the IO or be interpreted anaphorically:

(43) a. Grammatical agreement:

Juan le envió el libro a María.
Juan DAT.3SG sent the book to María
'Juan sent the book to María.'

b. Anaphoric agreement:
María necesitaba un libro. Juan se lo envió.
María needed a book. Juan SE ACC.3SG.MASC sent
'María needed a book. Juan sent it to her.'

But clitic pronouns, either accusative or dative, cannot be emphatic in Spanish. It is necessary to use a stressed pronoun in order to get an emphatic interpretation of the IO:

(44) *María necesitaba un libro. Juan se lo envió a ella (no a su hermana).*
María needed a book. Juan SE ACC.3SG.MASC to her (not to her sister)
'María needed a book. Juan sent it to HER (not to her sister).'

These facts show that dative (and accusative) clitics behave like subject agreement markers rather than stressed object pronouns with respect to their interpretation.

2.2.2 Pronominal clitics as inflectional affixes

Further evidence for the status of doubled dative clitics as agreement markers is provided by the fact that they behave as inflectional affixes according to the criteria proposed by Zwicky and Pullum (1983) in their analysis of English contracted auxiliaries and contracted negation. Zwicky and Pullum arrive at the conclusion that English contracted auxiliaries (*'s, 've*) are clitics but English contracted negation (*n't*) is an inflectional affix. The criteria they use are degree of

selection with respect to the host, arbitrary lexical gaps, phonological idiosyncrasies, semantic idiosyncrasies, syntactic restrictions, and restriction on the combination of clitics and affixes. The evidence, although not conclusive, is compatible with the claim that pronominal clitics behave as inflectional affixes.

Monachesi (1998, 2005), applied those criteria to the analysis of Italian and Romance clitic pronouns. As would be expected, given the similarities between Italian and Spanish clitics, Monachesi's results can be easily extended to Spanish. The comparison emphasizes the contrast between the expected behavior of pronominal clitics if they were syntactic objects (words), governed in their distribution by phrase structure rules, and their observed behavior as morphological objects (affixes), governed by morphological rules instead. The criteria, as stated by Monachesi, are the following:

2.2.2.1 Degree of selection with respect to the host Affixes, but not clitics, exhibit a high degree of selection with respect to the lexical category of the stem that acts as host. Zwicky and Pullum show that English *'s* and *'ve* can attach to different lexical categories (preposition, verb, adjective, adverb, in addition to pronoun). The following examples are from Zwicky and Pullum, 1983: 504):

(45) a. *The person I was talking to's going to be angry with me*. [preposition]

b. *The ball you hit's just broken my dining room window.* [verb]

c. *Any answer not entirely right's going to be marked as an error.* [adjective]

d. *The drive home tonight's been really easy.* [adverb]

Inflectional affixes, on the other hand, are selective with respect to their host. Zwicky and Pullum mention the examples of 'noun plural', 'verb past', and 'adjective superlative' in English, which attach only to a specific lexical category. In this respect, Spanish pronominal clitics, as well as Italian and Romance pronominal clitics in general,[22] attach exclusively to verbs. Moreover, like inflectional affixes, pronominal clitics do not affect the lexical category of the host (Monachesi, 2005: 44).

2.2.2.2 Rigid ordering This criterion is introduced by Monachesi in her discussion of Italian clitics. Spanish and Romance clitics appear in a rigid order. The order of pronominal clitics in Spanish follows this template (Perlmutter, 1971; Zagona, 2002): *se* > 2nd > 1st > 3rd dative > 3rd accusative

22 Romanian pronominal clitics are an exception. They can combine with negation, complementizers, nouns, and interrogative words (Dobrovie-Sorin, 1994; Monachesi, 2005).

(46) *se* > 2nd:

a. *Juan se te escapó.*

Juan SE 2.SG escaped

'Juan escaped from you.'

b. **Juan te se escapó.*

Juan 2.SG SE escaped

(47) 2nd > 1st:

a. ***Te me*** *enojaste.*

2.SG 1.SG got.angry

'You got angry at me.'

b. ***Me te*** *enojaste.*

2.SG 1.SG got.angry

(48) 1st > 3rd:

a. ***Me lo*** *compraste.*

1.SG ACC.3SG.MASC bought

'You bought it for me.'

b. ***Lo me*** *compraste.*

ACC.3SG.MASC 1.SG bought

(49) 3rd dative > 3rd accusative:

a. ***Se lo*** *dieron a Juan.*

SE ACC.3SG.MASC gave to Juan

'They gave it to Juan.'

b. ***Lo se*** *dieron a Juan.*

ACC.3SG SE gave to Juan

According to Monachesi (1998, 2005), the templatic distribution of Romance pronominal clitics makes it possible to analyze their placement as the result of the same type of morphological processes that account for the placement of inflectional affixes in morphologically rich languages (agglutinating languages).

This rigid order is observed whether the clitic cluster is a proclitic, as in the previous example, or an enclitic, as in the following one.

(50) a. *Quieren dár-**se-lo** a Juan.*
want give-DAT.3SG-ACC.3SG.MASC to Juan
'They want to give it to Juan.'

b. **Quieren dár-**lo-se** a Juan.*
want give-ACC.3SG.MASC-SE to Juan

The occurrence of proclisis or enclisis follows strict rules, which can be subject to dialectal variation. In standard varieties of Spanish, proclisis is the norm, except in infinitivals, gerunds, and imperatives. This variability in the direction of attachment is not a common property of inflectional affixes. Monachesi (2005: 63-4) makes reference to the existence of mobile position inflectional affixes in Arabic (Fontana, 1993), Huave (Noyer, 1994), and Afar (Fulmer, 1990). However, this is not the usual behavior of inflectional affixes and more research is necessary in this respect.

2.2.2.3 Coordination The criterion of coordination is introduced by Monachesi in her discussion of Italian clitics. Spanish pronominal clitics, as well as the Italian ones, cannot have wide scope over coordination. In this respect, pronominal clitics behave like any other inflectional affix rather than a syntactic constituent.

A syntactic constituent has wide scope over coordination, as in the case of the DO *el libro* ('the book') in the following sentence:

(51) a. *Juan leyó y entendió el libro.*
Juan read and understood the book
'Juan read and understood the book.'

That is not the case for inflectional affixes. An inflectional affix has scope only over the word it is attached to. The following example shows that the past tense/person agreement portmanteau affix cannot have scope over coordination:

(52) a. **Juan leyó y entend- el libro.*
Juan read.3sg.PAST and understand (root) the book

b. **Juan le- y entendió el libro.*
Juan read (stem) and understood the book
'Juan read and understood the book.'

Pronominal clitics behave like the past tense/person agreement inflectional affix rather than like a syntactic constituent. The comparison is not straightforward because verbal roots in Spanish are bound and cannot occur without either finite or non-finite inflection, but it suggests that pronominal clitics behave like affixes regarding coordination:

(53) a. *El libro, Juan lo leyó y lo entendió.*
the book Juan ACC.3SG.MASC read and ACC.3SG.MASC understood
'(The book,) Juan read it and understood it.'

b. **El libro, Juan lo leyó y entendió.*[23]
the book Juan ACC.3SG.MASC read and understood
'(The book,) Juan read and understood it.'

c. **(El libro,) Juan leyó y lo entendió.*
the book Juan read and ACC.3SG.MASC understood
'(The book,) Juan read and understood it.'

This fact is independent of enclisis or proclisis:

(54) a. *(El libro,) Juan quería leerlo y entenderlo.*
the book Juan wanted read-ACC.3SG.MASC and understand-ACC.3SG.MASC
'(The book,) Juan wanted to read it and understand it.'

b. **(El libro,) Juan quería leerlo y entender.*
the book Juan wanted read-ACC.3SG.MASC and understand
'(The book,) Juan wanted to read and understand it.'

23 Acceptable with the interpretation that Juan understood something by reading the book, not that he understood the book itself.

c. **(El libro,) Juan quería leer y entenderlo.*
the book Juan wanted read and understand-ACC.3SG.MASC
'(The book,) Juan wanted to read and understand it.'

2.2.2.4 Arbitrary gaps Arbitrary gaps in a set of possible combinations is a characteristic of inflectional paradigms (Zwicky and Pullum, 1983: 504). Zwicky and Pullum notice that there are no arbitrary gaps in the set of host-clitic combinations involving *s′* and *′ve*. But inflectional paradigms can display arbitrary gaps. Zwicky and Pullum mention the verb *stride*, which lacks a past participle.

Monachesi cites many Italian verbs which lack verbal forms: *involvere* ('to wrap') lacks past tense forms; *solere* ('to be used') and *sapere* ('to know') lack the present participle form; *incombere* ('to impend'), *splendere* ('to shine'), *prudere* ('to itch'), *urgere* ('to be necessary'), and *vigere* ('to be in use') lack the past participle form.

An arbitrary gap in the combination of pronominal clitics in Italian and Spanish disallows the combination of a 1st or 2nd person accusative with a dative clitic. In those cases, the dative clitic should be replaced by a stressed third person pronoun (Monachesi, 1998):

(55) a. **Juan le me presentó.*
JUAN DAT.3SG 1SG introduced

b. **Juan me le presentó.*
JUAN 1SG DAT.3SG introduced

c. *Juan me presentó a él/ella.*
JUAN 1SG introduced to him/her
'Juan introduced me to him/her.'

2.2.2.5 Morphophonological idiosyncrasies Morphophonological idiosyncrasies are characteristic of affixed words (Zwicky and Pullum, 1983: 504). According to Zwicky and Pullum, no morphophonological idiosyncrasies exist within clitic groups containing *'s* and *'ve*. But morphophonological idiosyncrasies are common in inflected words. Some English examples provided by Zwicky and Pullum are *oxen* and *dice* for the plural affix, *slept* and *went* for the past tense affix, and *best* and *worst* for the superlative.

Monachesi provides examples of morphophonological idiosyncrasies that affect Italian clitics. For instance, if the third person dative feminine *le* precedes a clitic beginning with *l-* or *n-*, it is replaced by the masculine form *gli*:

(56) a. **Le le ho date.*
DAT ACC have given

b. *Gli le ho date.*
DAT ACC have given
'I have given them to her/him.'
(Monachesi, 2005: 55)

An Spanish example of a morphophonological idiosyncrasy affecting pronominal clitics is the morphologically conditioned allomorphy *le/se*. When the 3rd person dative clitic precedes the 3rd person accusative clitic, *le* is replaced by *se*:

(57) a. **Le lo dio a María.*
DAT.3SG ACC.3SG.MASC gave to María
'he/she gave it to María.'

b. *Se lo dio a María.*
SE ACC.3SG.MASC gave to María
'he/she gave it to María.'

2.2.2.6 Verb left-detachment According to Zwicky and Pullum (1983: 504), syntactic rules can affect affixed words but not clitic groups. In English, they claim, "no syntactic operations exist which treat a word combined with one of the clitics *'s* or *'ve* as a unit. […] But inflected nouns, verbs, adjectives, and adverbs are of course regularly treated as units by syntactic operations".

In Romance there is a syntactic rule that affects pronominal clitic-verb complexes. That rule, which cannot affect VPs, is infinitival left-detachment (Monachesi, 1998: 313). It consists of the doubling of a matrix verb in the indicative by a left-detached infinitive:[24]

(58) a. *Dárselo, no se lo di.*
give-SE-ACC.3SG.MASC not SE ACC.3SG.MASC gave
'As for giving it to him/her, I didn't give it to him/her.'

b. ** Dárselo a ella, no se lo di.*
give-SE-ACC.3SG.MASC to NOM.3SG.FEM not SE ACC.3SG.MASC gave
'As for giving the book to *María,* I didn't give it to her.'

24 Unfortunately, the rule of verb-left-detachment is limited to infinitivals; finite verbs cannot be left-detached in the same way: **Se lo di, no se lo di* ('As for give it to him/her, I didn't give it to him/her').

Since the pronominal clitic-verb complex can be affected by this syntactic rule, we should conclude that the status of the complex is that of an inflected word, not that of a phrasal constituent.

2.2.3 Grammaticalization

Spanish pronominal clitics behave, according to most of the tests, as agreement markers, a type of inflectional affix. The affixal behavior of pronominal clitics can be understood as the result of a process of grammaticalization, as has been claimed by Givón (1976, 2001).

Givón claims that Spanish clitics have become agreement markers as the result of a diachronic process of reanalysis. The process consists of three steps. The first one is the use of unstressed pronouns as anaphors referentially dependent on a left-dislocated topic (*topic-shift*, TS). This strategy would be used by speakers in order to change the current discourse topic by reintroducing a previous one.

In the second step, clitics are used as cataphors referentially dependent on a right-dislocated topic (*after-thought topic shift*, AS). The use of after-thought topic shifting produces a reinforcement of the referent of the cataphoric pronoun. The overuse of this construction would have resulted in the reanalysis of right-dislocated objects as plain objects, and clitics as agreement markers. As a result of this process, defined as the *demarking* of a marked or semi-marked structure, clitics become grammaticalized as agreement markers, as shown below.

(59)

TS ("marked")		AT ("semi-marked")		NEUTRAL ("demarked")
the man, I saw him	=>	I saw him, the man	=>	I saw-him the man

(Givón, 1976: 157)

Givón (1976) further discusses the possible functional load of agreement, suggesting that agreement markers can become associated with argument alternations: "If a language develops a 'complete' subject, accusative, and dative agreement, one may consider the possibility that the agreement system can then become, for the speaker, a way of signaling the syntactic type ('transitivity') of verbs" (1976: 168). The idea that clitics can express lexical information (transitivity type) could explain why dative clitic doubling is required or preferred when there is an argument alternation which introduces a dative argument (experiencer, benefactive, and possessive datives).

The claim that Spanish clitics can fulfill this function is not new, since it has been already made with respect to the reflexive clitic in the causative alternation. Verbs which participate in this alternation have two possible argument structures, a causative one with an agent as subject and a patient as object, and an unaccusative one with a patient as subject and no expression of the agent. A typical example of a verb that participates in the causative alternation is *break:*

(60) a. *John broke the window.*

b. *The window broke.*

The causative alternation is morphologically unmarked in English. But in Spanish, it is marked by the occurrence of the reflexive clitic *se* in the unaccusative form, as shown below:

(61) a. *Juan rompió la ventana.*
Juan broke the window
'Juan broke the window.'

b. *La ventana se rompió.*
the window SE broke
'The window broke.'

The grammaticalization of the reflexive clitic as an unaccusativity marker explains its use as part of the citation form of the verb (romperse vs. romper[25]). This process is widely acknowledged in the literature (Mendikoetxea, 1999).

Therefore, it seems that dative clitic doubling in Spanish is not a completely unified phenomenon. The functional load of the dative clitics is different in ditransitive and non-ditransitive constructions. As the result of grammaticalization, it is required in those contexts in which an argument alternation has taken place, but it does not express an argument alternation in ditransitive constructions. The phenomenon of DCLD in ditransitive constructions should be understood in terms of optional agreement instead and its functional load is to (optionally) overtly mark those recipients which are pragmatically salient. The hypothesis that pragmatic salience is the trigger of agreement is based on the typological observation that objects which are pragmatically salient are more likely to trigger agreement (Comrie, 1989; Croft, 1988; Woolford, 1999). Croft (1988) expresses this fact in the following terms (italics in the original):

25 Notice that Spanish non-finite verbal forms have enclisis instead of proclisis.

> Agreement -i.e., person-based agreement, also called 'cross-reference' or 'indexing'- indexes the *important or salient* arguments. This concept is a pragmatic one: salience is a relationship between the speaker and a referent in the described situation -that is, the speaker's attitude or point of view towards the referent- rather than a relation between two entities in the described situation itself. Salience correlates with being high on the case, animacy, and definiteness hierarchies, since the most salient entities are those most closely involved in the described event, closest in nature to the speaker, and most easily identifiable. The natural correlation predicts that where the presence vs. absence of agreement is grammaticalized, it will always align itself with high animacy, high definiteness, and core grammatical relations (Croft, 1988: 168).

The role of pragmatic salience as a trigger of agreement will be demonstrated by the results of our quantitative study of ditransitive sentences (Chapter 4).

2.3 ON THE GRAMMATICAL STATUS OF INDIRECT OBJECTS

In this section, we will discuss the status of Spanish IOs as a core or oblique grammatical function. We will use three criteria to argue that IOs in Spanish are core grammatical functions: lack of semantic restrictions, overt coding properties, and behavioral properties. For descriptive purposes, we understand that a syntactic constituent is an IO if it is a Noun Phrase introduced by the preposition a which optionally agrees in person and number with the verb by means of a dative clitic.

As was discussed in Chapter 1, there is a split between core and non-core grammatical functions. Core grammatical functions have been claimed to be semantically unrestricted, to present a distinctive set of properties, from the point of view of both overt marking and

behavioral patterns, and to be the grammaticalization of the main and secondary topics of a sentence.

2.3.1 Semantic restrictions

As has been observed by Masullo (1992), IOs in Spanish can fulfill a wide range of thematic relations, including recipient, experiencer, benefactive, and possessive. Therefore, it should not be considered a semantically restricted case. Although IOs are introduced by the preposition *a* ('to'), their semantic interpretation is not restricted by it. The preposition *a*, when introducing an IO, should be considered semantically vacuous, unlike when it is used to introduce a goal OBL, as in *Juan fue a la escuela* ('Juan went to school'). Goal prepositional phrases, as well as other OBL phrases, cannot undergo DCLD:

(62) a. **Juan le fue a la escuela.* (*a*-OBL)
Juan DAT.3SG went to the school
'Juan went to school.'

b. **Juan le preparó la comida para María.* (*para*-OBL)
Juan DAT.3SG the food for María
'Juan prepared the food for María.'

Moreover, IOs and OBLs introduced by *a* are different grammatical functions, to the extent that they can co-occur in the same sentence. In grammatical frameworks that rely on the concept of grammatical function, there is the assumption that each argument is associated with a

grammatical function and that no grammatical function occurs more than once in a sentence (Principle of Function-Argument Biuniqueness in Lexical Functional Grammar, Bresnan, 2001: 311). The following example shows the co-occurrence of an IO and a goal OBL:

(63) a. *Juan le envió el libro a María a su casa.*
Juan DAT.3SG sent the book to María to her house
'Juan sent the book to María to her house.'

The possibility of this co-occurrence suggests that the IO and the goal OBL are two different grammatical functions.

2.3.2 Overt marking properties

The most salient overt marking property of IOs is the fact that they optionally agree with the verb in person and number by means of the dative clitic phonologically attached to it. An additional overt marking property is the preposition *a* itself, which in Spanish functions as a prepositional case marker that introduces not only IOs, but also DOs when they are animate and specific.

2.3.3 Behavioral properties

Regarding their syntactic behavior, it has been pointed out (Masullo, 1992; Alsina, 1996; Givón, 2001) that IOs in Spanish and Romance display properties usually associated with core rather than oblique grammatical functions. Experiencer IOs, as well as experiencer SUBJs, can control the understood subject of a non-finite clause (examples adapted from Givón, 2001):

(64) a. Experiencer SUBJ controller

Juan quiere caminar por el bosque.
Juan wants walk through the forest
'Juan wants to walk through the forest.'

b. Experiencer IO controller
A Juan le gusta caminar por el bosque.
to Juan DAT.3SG walk through the forest
'Juan likes to walk through the forest.'

In addition, IOs as well as DOs can be the target of secondary predication, another behavioral property which is exclusive of core grammatical functions (examples adapted from Catalan, Alsina, 1996):

(65) a. DO-oriented secondary predication
(La) quieren retratar a María disfrazada de vestal.
ACC.3SG.FEM want portray to María disguised of vestal
'They want to portray María disguised as a vestal.'

b. IO-oriented secondary predication
(Le) quieren hacer un retrato a María disfrazada de vestal.
ACC.3SG.FEM want make a portrait to María disguised of vestal
'They want to make a portrait of María disguised as vestal.'

This behavioral evidence is not directly relevant because it does not apply to the IOs of the predicates under study. However, this review of properties of Spanish IOs strongly suggests that IOs should be considered core rather than oblique grammatical functions. The IO grammatical

function is not semantically restricted in Spanish. Its overt coding is analogous to that of DOs: it is the target of agreement with the verb by means of the dative clitic and is introduced by the same prepositional case marker that is required by salient DOs. In addition, IOs display, in some contexts, behavioral properties normally attributed to core grammatical functions.

2.4 CONCLUSIONS

In this chapter we have discussed the distribution of DCLD and the status of dative clitics and IOs in Spanish. We have arrived at the conclusion that dative clitics are agreement markers with different functional loads in ditransitive and non-ditransitive constructions. In ditransitive constructions, they optionally mark the recipient as IO. Our hypothesis is that pragmatically salient IOs are overtly marked as dative, an issue that will be discussed in Chapter 4. In non-ditransitive constructions, the dative clitic signals that an argument alternation has taken place. The argument alternations that take place in non-ditransitive constructions consist of the promotion of an argument to IO, which is a core grammatical function in Spanish.

3.0 APPROACHES TO DCLD AND DATIVE-SHIFT

In this chapter, we present a review of the hypotheses that have been posed in the literature about the distribution of DCLD in Spanish ditransitive constructions. Since the study of DCLD has been to a large extent influenced by the study of English dative-shift, we will also discuss here the literature on that topic. The distribution of DCLD in Spanish has been analyzed in syntactic (Demonte, 1995; Bleam, 2003; Cuervo, 2003) or lexico-semantic terms (Strozer, 1976; Romero Morales, 2008). We will argue that DCLD is not reducible to a particular syntactic structure or to lexical semantics.

3.1 SYNTACTIC APPROACH

Demonte (1995) is the first of a series of analyses of DCLD in ditransitive constructions that claims that Spanish ditransitive sentences with and without DCLD have different syntactic structure. Demonte assumes that Spanish DCLD is a phenomenon analogous to English dative-shift and claims that a version of Larson's (1988) "VP-shells" analysis can account for the facts in Spanish. Larson's analysis is formulated in the framework of the Principles and Parameters Theory, (P&P, Chomsky, 1981; Chomsky and Lasnik, 1993). Our trees do not attempt to replicate the details of Larson's analysis but to display the structural facts which are relevant to our discussion.

Syntactically, dative-shift in English, as well as other languages, consists of the promotion of the recipient of a ditransitive verb from OBL to Primary Object (PO), resulting in a "double object construction". The alternative, without dative-shift, is a construction in which the recipient is an OBL introduced by the preposition *to* ("dative prepositional construction"). The alternation between the dative prepositional construction and the double object construction is called the "dative alternation":

(1) a. Dative prepositional construction:
John gave a book to Mary.

b. Double object construction:
John gave Mary a book.

Notice that the term "dative" is used here as a semantic rather than a grammatical term, since it refers to the recipient, independently of its grammatical properties. This use of the term is radically different from its use in Spanish grammars, where "dative" refers to a grammatical function, that of Indirect Object (IO) (see Chapter 2).

The fact that the recipient is the PO in the English double object construction follows from its adjacency to the verb and is further confirmed by its ability to become the subject of the sentence under passivization:

(2) a. *John gave a book to Mary.*

b. *Mary was given a book (by John).*

In the dative prepositional construction, it is the theme, instead of the recipient, which is adjacent to the verb and the theme is also the argument that becomes the subject of the sentence under passivization:

(3) a. *John gave a book to Mary.*

b. *A book was given to Mary.*

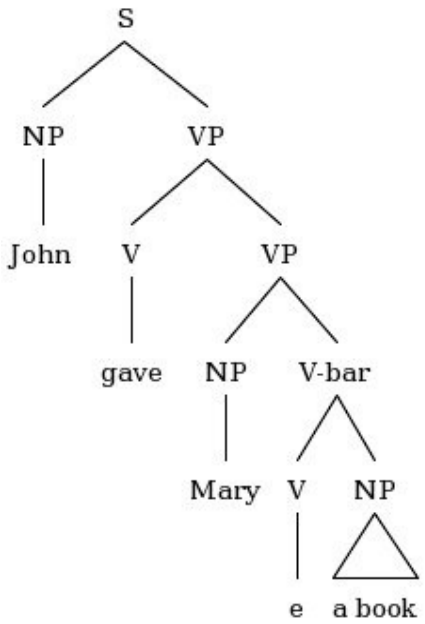
Larson (1988) claims that at an underlying level of syntactic structure, the double object construction lacks the preposition *to*, which is responsible for assigning case to the recipient in the dative prepositional construction. As a result, in order to avoid having an argument devoid of grammatical marking or "abstract case", the recipient is displaced to DO position. That position is, linearly speaking, to the right and adjacent to the verb, a fact that is supposed to account for the recipient/theme order in the double object construction. That position, according to Larson, is not only to the left of the theme, but it also asymmetrically c-commands it ("c-command" is a short form of "constituent-command", a type of the more general syntactic notion of "command"). Command, a structural relation between nodes in a tree, is defined in the following terms:

> In the most general case, a node A in one branch 'commands' a node B in another branch if *(a)* both are dominated by the same node C and *(b)* C is the lowest branching node that dominates A (Matthews, 2007: 64)

Symmetrical c-command holds between daughters of the same node, while asymmetric c-commands holds between a node and all the material dominated by its mother, including, vacuously, itself.

Larson's claim that the recipient asymmetrically c-commands the theme in the double object construction is incompatible with a ternary branching structure. Notice that in a ternary branching structure, there is mutual symmetric c-command of the theme and the recipient instead of asymmetric c-command of the theme by the recipient. In order to obtain asymmetric c-command, the recipient should occupy a higher position in a binary-branching tree, as shown in the tree below. The binary-branching structure proposed by Larson includes an "iteration" of the VP ("VP-shell") and an "empty category" as the head of the lower VP, which we will not discuss here:

(4) Double object construction:



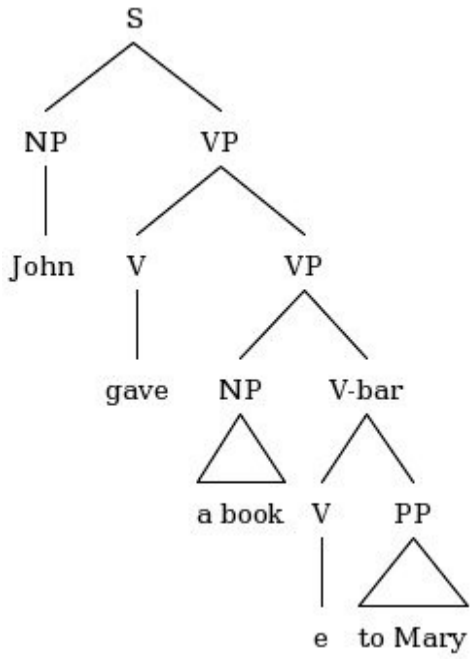
Notice that this analysis of the double object construction makes the claim that the recipient and the theme are a single syntactic constituent ([$_{VP}$ *Mary a book*]), since they are both dominated by a single node, the lower VP. Notice, however, that *Mary* and *a book* seem to be independent constituents, since they can be questioned independently:

(4) a. *What did you give Mary?*

b. *Who did you give a book?*

Moreover, Larson proposes that in the dative prepositional construction the opposite situation holds regarding c-command: the theme c-commands the recipient in a configuration structurally analogous to the one depicted below:

(5) Dative prepositional construction:



Larson made that claim based on evidence regarding anaphoric reflexives and other referential dependencies. Larson discusses the following contrast in the double object construction:[26]

(6) a. *I showed John himself (in the mirror).*

b. **I showed himself John (in the mirror).*

In the first example, the theme is a reflexive referentially dependent on the recipient. In the second example, the recipient is a reflexive which cannot be interpreted as referentially dependent on the theme and lacks an antecedent, rendering the sentence unacceptable. The opposite occurs in the dative prepositional construction:

(7) a. *I showed John to himself (in the mirror).*

b. **I showed himself to John (in the mirror).*

Larson says that these contrasts can be explained under the assumption that antecedents should asymmetrically c-command a reflexive in order to obtain a referential dependency ("Principle A of Binding Theory", Chomsky and Lasnik, 1993). Therefore, "recipients" should c-command "themes" in the double object construction and "themes" should c-command "recipients" in the dative prepositional construction.

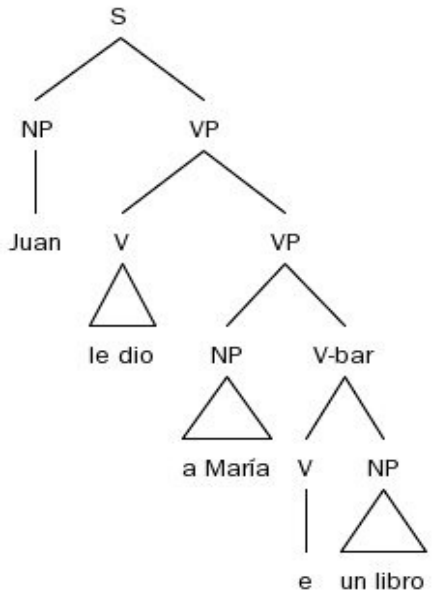
Notice, however, that Larson violates Occam's razor by proposing additional structure in order to explain phenomena that have received alternative explanations which do not require additional assumptions. Barss and Lasnik (1986) propose an explanation of the reflexivization

26 Originally from Barss and Lasnik (1986).

patterns adding "precedence" to c-command as a requisite on reflexivization[27]. Linear order (precedence), as opposite to the structure postulated by Larson, is directly observable. Therefore, it is safe to assume that speakers have access to it. Another approach that does not require additional syntactic machinery is the one taken in *Lexical Functional Grammar* (LFG) (Bresnan, 2001), in which constraints on reflexivization are stated in terms of a hierarchy of grammatical functions instead of syntactic configurations.

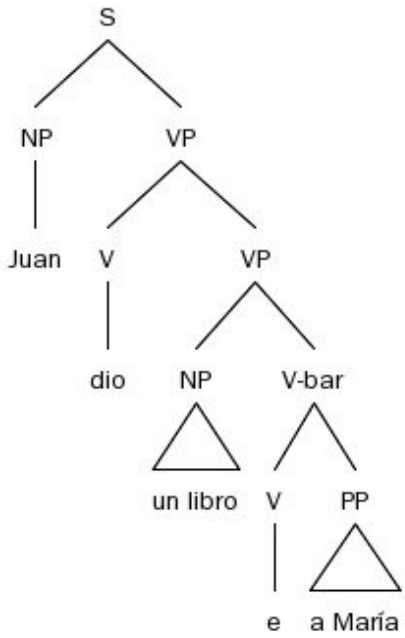
According to Demonte (1995), Spanish ditransitive sentences with DCLD have the structure proposed by Larson for the double object construction, and Spanish ditransitive sentences without DCLD have the structure proposed by Larson for the dative prepositional construction:

(8) a. <u>DCLD construction (adapted from Demonte, 1995)</u>:



27 See also Jackendoff (1990b), Culicover and Jackendoff (2005), and Culicover (2009) for a critical assessment of Larson's analysis and Croft (2001: 41-7) for criticism on both Larson and Jackendoff.

b. Ditransive construction without DCLD (adapted from Demonte, 1995):



Like Larson, Demonte supports her claim using evidence related to the behavior of anaphoric reflexives and other referential dependencies. As in the case of Larson, Demonte's argument relies on a configurational theory of referential dependencies (P&P binding theory). However, it is not clear if the data regarding Spanish reflexives are analogous to those of English. The Spanish reflexive equivalent to *himself/herself* is the expression *a sí mismo/misma*, which is introduced by *a*. The antecedent of the reflexive, if human, will be also introduced by *a*. Since *a* in either case is not a lexical preposition, but an object marker, a ditransitive construction with both theme and recipient introduced by *a* is difficult to process and hardly acceptable, independently of the occurrence of DCLD or word order.

(9) a. ?(*Le) mostr*é a Juan a sí mismo (en el espejo).

b. ?(*Le) mostré a sí mismo a Juan (en el espejo).*

Therefore, the contrast that would indicate that the recipient c-commands the theme in DCLD constructions is difficult to judge.

In addition, there are other differences between the English double object construction and Spanish DCLD: the behavior of the constructions under passivization and word order facts.

3.1.1 DCLD and the passive

As was previously stated, in the double object construction, the recipient becomes the SUBJ of the passive. But this is not the case in Spanish ditransitive constructions, since only the theme can become the SUBJ of the passive, independently of DCLD:

(10) a. *Juan (le) envió el libro a María.*
Juan DAT.3SG sent the book to María
'Juan sent the book to María.'

b. **María fue enviada el libro.*
María was sent the book
'María was sent the book.'

c. *El libro le fue enviado a María.*
the book DAT.3SG was sent to María
'The book was sent to María.'

Therefore, the recipient in DCLD constructions does not occupy a syntactic slot or fulfill the same grammatical function as the patient of a transitive sentence, as would be expected under Demonte's analysis.

3.1.2 Word order

As has been previously mentioned, the order of theme and recipient in Spanish ditransitive constructions is optional, independently of DCLD. This is not the case in the double object construction, in which the recipient is right-adjacent to the verb. This fact is expected if we take into account that word order is the main mechanism of overt coding of grammatical functions in English. Spanish, on the other hand, resorts to agreement as the main mechanism of overt coding of grammatical functions, a fact that results in a more flexible word order.

In spite of the flexibility of word order in Spanish, Demonte (1995) claims that the order of theme and recipient is not free in sentences without DCLD. In those sentences, according to Demonte, the order theme/recipient is the only fully acceptable one:

> Even if it is accepted that Spanish is a free word-order language, constraints on the arrangement of sentence constituents have to be acknowledged. In Goal structures, where the clitic can be absent, the unmarked order is V DO IO. The order V IO DO ranges from being felt as stylistically marked to having an ungrammatical flavor. What some speakers say is that in this second case the structure 'asks for a clitic' (Demonte, 1995: 20).

Demonte provides the following judgments, which are the expected ones under her analysis:

(11) a. *Di el libro a María. / Entregué las llaves al dueño.*
gave the book to María / delivered the keyw to.the landlord
'I gave the book to María.'/ 'I delivered the keys to the landlord.'

b. *%Di a María el libro / ?Entregué al dueño las llaves*
gave to María the book / delivered to.the landlord the keys
'I gave María the book.'/ 'I delivered the landlord the keys.'

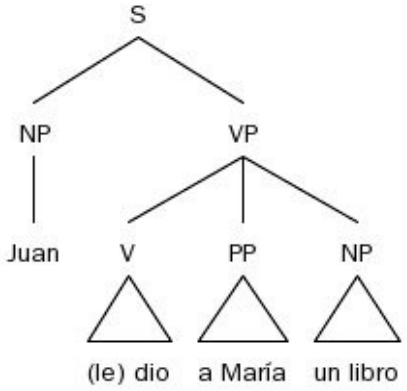
This state of affairs, in Demonte's account, follows from her analysis of Spanish ditransitive constructions based on Larson's analysis of the English double object and dative prepositional constructions. In Chapter 4 we will present the results of our corpus study of Spanish ditransitive constructions, which shows that the opposite situation holds: theme-recipient word order is possible with DCLD, as in the following corpus example:

(12) *... el senador priísta Eduardo Andrade Sánchez ya le envió una carta a Lott ...*
'... Senator Eduardo Andrade Sánchez (PRI) has already sent a letter to Lott ...'
(Corpus del Español)

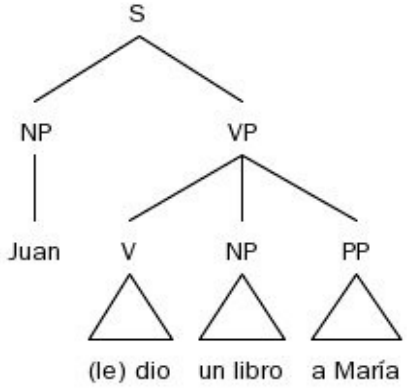
3.1.3 An alternative to VP-shells: flat VP

An alternative analysis of ditransitive VPs, without VP-shells, consists of a flat, ternary branching VP, in which the theme and the recipient are daughters of the same projection (Jackendoff, 1990b; Culicover and Jackendoff, 2005). The trees below show the structure of Spanish ditransitive VPs following Culicover and Jackendoff's (2005) proposal.

(13) a. Flat VP (recipient/theme order):



b. Flat VP (theme/recipient order):



The evidence in favor of a flat VP is not conclusive, but the burden of proof should be on the more complex, not the simpler analysis (the more complex analysis being the one that postulates the existence of a more complex structure which is not necessary in order to account for basic constituency facts). In the case of Spanish, the flat VP analysis depicted above is compatible with the idea that the grammar does not determine the order of theme and recipient. Both orders are available and the occurrence of one or another order would be determined by functional, not syntactic, constraints.

3.2 LEXICO-SEMANTIC APPROACH

The occurrence of DCLD in Spanish ditransitive constructions has been analyzed in the literature as a manifestation of lexico-semantic properties of the predicates involved (Strozer, 1976; Romero Morales, 2008). This interpretation follows a long tradition of analyses of the English dative alternation which associates the dative prepositional construction (*John gave a book to Mary)* with the notion of "change of location" and the double object construction (*John gave Mary a book)* with the notion of "change of possession" (Green, 1974, Oehrle, 1976, Pinker, 1984, 1989), a hypothesis that has been called by Bresnan et al. (2007) the "Meaning-to-Structure Mapping Hypothesis" (MSMH).

According to Pinker (1984, 1989), there are two types of constraints on English dative-shift: a lexico-semantic one and a morphological one. The lexico-semantic constraint limits the double object construction to verbs that mean or can be interpreted as 'X causing Y to have Z' (Pinker, 1989:110). Pinker mentions the following as typical examples of verbs that occur in the double object construction:

(14) *give, pass, hand, sell, pay, trade, lend, loan, serve, feed.*

Some verbs of change of location can be reinterpreted to denote a change of possession and are allowed in the double object construction. This is the case of the verbs of instantaneous imparting of force causing ballistic motion. Pinker provides the following examples:

(15) *Lafleur throws / tosses / flips / slaps / kicks / pokes / flings / blasts him the puck; he shoots, he scores!*

But, also according to Pinker, verbs of continuous imparting of force causing accompanied motion cannot be reinterpreted to denote a change of possession and are not allowed in the double object construction. Pinker provides the following examples:

(16) **I carried / pulled / pushed / schlepped / lifted / lowered / hauled John the box.*

Pinker also mentions other verb classes that do and do not participate in the double object construction (dativizable verbs and non-dativizable verbs):

(17) Dativizable verbs:

a. Verbs of future having: *offer, promise, bequeath, leave, refer, forward, allocate, guarantee, allot, assign, advance, award, reserve, grant.*

b. Verbs of future not having: *cost, spare, envy, begrudge, bet, refuse, ask, save, charge, fine, forgive, ?deny.*

c. Verbs of communication: *tell, show, ask, teach, pose, write, spin, read, quote, cite.*

d. Verbs of obtaining: *get, buy, find, steal, order, win, earn, grab.*

(18) Non-dativizable verbs:

a. Verbs of manner of speaking: **John shouted / screamed / murmured / whispered / shrieked / yodeled / yelled / bellowed / grunted / barked / Bill the news.*

b. Verbs of choosing: **I chose / picked / selected / favored / indicated / preferred / designated her a dress.*

The morphological constraint on the double object construction is that most Latinate verbs do not accept it (Pinker, 1989: 119). For instance *donate* and *contribute* do not accept the double object construction, in spite of denoting a change of possession event. Taking into account both the lexico-semantic and morphological constraints, dative-shift in English seems to be quite sensitive to lexical constraints.

DCLD in Spanish is not constrained by any morphological constraint. Of course, all Spanish verbs are Latinate. Also, DCLD occurs freely with some of the verbs which are not dativizable in English, according to Pinker, for instance, verbs of manner of speaking, as shown below (examples from *Corpus del Español*):

(19) *Gritar* ('to shout'):
Malvina ***les gritó*** *a los nativos que arrastraran el vehículo hasta junto a la ventana del despacho.*
'Malvina shouted to drag the vehicle up next to the office window to the natives.'

(20) *Susurrar* ('to whisper'):
*Llegué antes de la hora convenida y l****e susurré*** *a Matilde en la cocina que me sentía sucio y curtido por la sal y el sol, ...*
'I arrived before the appointed time and whispered that I felt dirty and weathered by salt and sun to Matilda in the kitchen, ...'

Romero Morales (2008) claims that lexical semantics is the main factor that constrains the distribution of DCLD in Spanish. If this is the case, DCLD would not be supposed to occur when the event denoted by the verb cannot be interpreted in terms of "change of possession". In order to support this claim, Romero Morales (2008) presents the following judgment:

(21) a. *Ramón y Cajal entregó su vida a la ciencia.*
Ramón y Cajal gave his life to the science
'Ramón y Cajal gave his life to science'

b. **Ramón y Cajal le entregó su vida a la ciencia*
Ramón y Cajal DAT.3SG gave his life to the science

According to Romero Morales, DCLD is not allowed in this context because the recipient (*la ciencia*, 'science') cannot own the theme (*su vida*, 'his life'). But, in spite of this claim, DCLD is found in usage data in contexts analogous to that of the example. The following are some examples found on the internet:

(22) a. *Hace 14 años partió un excepcional venezolano que **le entregó** su vida a las artes.*
'Fourteen years ago an exceptional Venezuelan that gave his life to the arts left.'
(http://elobservador.rctv.net/Noticias/VerNoticia.aspxNoticiaId=270779&Tipo=34)

b. ... *una persona que **le entregó** su vida a la patria...*
'... a person who dedicated his life to his homeland...'
(http://www.elpilon.com.co/inicio/le-dieron-ultimo-adios-a-policia-muerto-en-atentado- de-la-guerrilla-en-caqueta/)

c. *Mimí González **le entregó** su vida a la danza.*
'Mimí González gave her life to dance.'
(http://www.revistabellasartes.com/videos/)

As in the case of the example provided by Romero Morales, these examples have an abstract recipient which cannot own the theme (*la ciencia*, 'science'; *la patria* 'the homeland'; *la danza*, 'dance'). The availability of these examples poses a challenge for the claim that lexical semantics is the main constraint on the distribution of DCLD in ditransitive constructions.

3.3 MULTIPLE FACTORS APPROACH

Wasow (2002)[28] presents an analysis of word order alternations in English in which he discusses the interaction of grammatical weight and information structure as predictors of word order in contexts of optionality. As Wasow points out, although English is a language with a relatively fixed word order, it allows for optionality in some constructions, the dative alternation being one example. Wasow shows, by means of a corpus study and a psycholinguistic experiment, that both grammatical weight and information structure are independently significant as predictors of word order.

This line of inquiry was followed by Bresnan et al. (2007), who present a corpus study of the dative alternation which shows that a wide set of factors, not only grammatical weight and information structure, are significant predictors of the dative alternation (animacy, definiteness, givenness, pronominality, and weight of the theme and recipient, and lexical semantics of the verb). The results of this work have improved our understanding of the empirical facts associated with the English dative alternation (details of these studies will be discussed in Chapter 4). However, Bresnan et al. (2007) do not provide an account of the way in which each of the factors that have been found to be statistically significant contribute to the dative alternation. In particular, Bresnan et al. (2007) do not discuss the fact that in the dative alternation there are two different levels of analysis which are intimately intertwined: the assignment of grammatical functions and the determination of word order. In the dative prepositional construction, the theme is linked to Primary Object and the recipient is linked to Oblique. In the double object

28 The resarch presented by Wasow (2002) has been previously presented in Arnold et al. (2000).

construction, the recipient is linked to Primary Object and the theme is linked to Secondary Object. The Primary Object precedes the Secondary Object. But the use of word order as a mechanism of overt marking of grammatical functions is a language specific characteristic of English and other languages. Therefore, what does it mean that a specific feature is a significant predictor of the dative alternation? Does it mean that it is a predictor of the choice of grammatical functions or that it is a predictor of word order? The two dimensions of linguistic structure are logically independent. The study of DCLD in Spanish, in which overt grammatical coding and word order are independent, can provide an insight into this problem.

3.4 CONCLUSIONS

In this chapter we have reviewed the analyses of dative-shift and DCLD in the literature based on syntactic and lexico-semantic considerations and shown that they do not seem to account for the Spanish data. It is our understanding that the problems with these analyses result from the wrong assumption that dative-shift and DCLD are analogous phenomena, an assumption that we have discussed and refuted in Chapter 2.

In Chapter 4 we present our quantitative study of DCLD in ditransitive constructions, which has been modeled after the studies of the English dative alternation discussed in this section, but with the goal of telling apart which factors constrain grammatical marking (DCLD) and which factors constrain constituent order.

4.0 QUANTITATIVE STUDY OF DITRANSITIVE SENTENCES

4.1 HYPOTHESES

In Chapter 3 we have discussed the syntactic and lexico-semantic approaches to DCLD in Spanish ditransitive sentences and pointed out they are not able to account for the whole range of data. We also discussed, in Chapter 1, the possibility of analyzing DCLD as a variable instead of a categorical rule (Labov, 1969; Walker, 2010). Variable rules describe a linguistic process and state what is the probability that the rule will apply. These rules are useful when it is not possible to make a deterministic statement about the distribution of linguistic forms (Walker, 2010). A deterministic statement consists of a list of necessary and sufficient conditions for the application of a rule (Evans, 2007).

An important property of variable rules is that, if a large number of observations is taken into account, their application is not random. In this way, it is possible to make quantitative generalizations about the likelihood of application of a rule (Walker, 2010). The purpose of our quantitative study is to make those quantitative generalizations. Our study has been inspired by studies on English dative-shift (Arnold et al., 2002; Wasow, 2002; Bresnan et al., 2007).[29] These studies arrive at the conclusion that speakers' choices cannot be predicted by a single factor. Instead, multiple factors, such as animacy, definiteness, givenness, lexical semantics, and grammatical complexity, are independently significant as predictors.

29 In these studies, the phenomenon of dative-shift is described using the term *gradient*, which we understand is equivalent to saying that is is governed by a *variable rule*.

Our hypothesis, which will be empirically tested by means of the quantitative studies described in this chapter, is that the occurrence DCLD in Spanish should also be described as a variable rule constrained by multiple factors, at least in the contexts under study (main hypothesis). However, as a result of typological differences between Spanish and English, we expect the factors that constrain DCLD and dative-shift to overlap only partially.

The typological differences which we consider relevant are:

a. The fact that word order is the main mechanism of overt grammatical coding in English, while Spanish resorts mainly to agreement in order to fulfill that function (Secondary hypothesis #1: occurrence of DCLD and order of theme and recipient are dissociated phenomena).

b. The assumption that dative-shift consists of the promotion of the recipient, which becomes the PO, and the demotion of the theme, which becomes the SO. DCLD, on the other hand, only affects the grammatical marking of the recipient, which is overtly marked as dative by means of agreement with the clitic (Dryer, 1986; Raúl Aranovich, 2007) (Secondary hypothesis #2: the occurrence of DCLD is constrained by properties of the recipient but is dissociated from the properties of the theme).

The criteria to support or falsify the hypotheses are the following:

a. Main hypothesis: in the contexts under study DCLD is an optional phenomenon constrained by multiple factors.

This hypothesis would be falsified if the results of the study show that there is an independent variable, or set of independent variables, that triggers the occurrence of DCLD nearly a hundred percent of the time. In that case, the occurrence of DCLD would be the result of a categorical rather than a variable rule. This hypothesis would also be falsified if the results show that the

occurrence of DCLD is a random phenomenon, which is not statistically associated with any independent variable, or set of independent variables. In that case, the occurrence of DCLD would not be the result of a variable rule, but a random phenomenon.

b. Secondary hypothesis #1: the occurrence of DCLD and the order of theme and recipient are dissociated phenomena. This hypothesis would be falsified if there is a statistically significant association between the occurrence of DCLD and the order of theme and recipient. This hypothesis would be confirmed otherwise.

c. Secondary hypothesis #2: the occurrence of DCLD is constrained by properties of the recipient and is dissociated from properties of the theme. This hypothesis would be falsified if there is a statistically significant association between the occurrence of DCLD and any independent variable, or set of independent variables, that describe properties of the theme (animacy, definiteness, givenness, grammatical category). This hypothesis would be confirmed otherwise.

The statistical analysis will consist of the application of logistic regression in order to determine the significance of a set of independent variables while controlling for the possible effect of correlated variables (Bresnan et al., 2007). This chapter will describe the quantitative studies performed in order to check these hypotheses and determine which factors constrain the occurrence of DCLD in Spanish. We will present a description of the data, the dependent and independent variables, the statistical tests, and their results.

4.2 THE DATA

The data used for our quantitative study of DCLD in ditransitive constructions consist of 1008 sentences of the relevant types, namely ditransitive sentences with and without DCLD. The sample contains both finite and non-finite examples. The following are examples from our sample, extracted from *Corpus del Español* (Davies, 2002-, www.corpusdelespanol.org):

(1) Ditransitive sentences with DCLD:

a. *En cierta ocasión* ***le envió [a Amaranta] [un papelito]*** *desde la cárcel, pidiéndole el favor de bordar una docena de pañuelos de batista con las iniciales de su padre.*
'Once he sent Amaranta a piece of paper from prison, asking the favor to embroider a dozen handkerchiefs with the initials of his father.'

b. *Según se dice, Starr* ***le envió [una citación judicial] [a Clinton].***
'Reportedly, Starr sent a subpoena to Clinton.'

c. *Unos empleados del departamento l****e enviaron recientemente [una carta] [al alcalde de Dade, Alex Penelas]****, en la que mencionan presuntas acciones indebidas.*
'Some employees of the department recently sent a letter to Dade Mayor Alex Penelas, in which they mention alleged wrongdoing.'

(2) Ditransitive sentences without DCLD:

a. *Ladislaw Borovitz, subido en su tarima de director,* ***indicó [a los músicos] [que se levantasen].***
'Borovitz Ladislaw, standing on his stage of director, indicated to the musicians to stand up.'

b. *Esa mañana la Linch* ***envió [a la familia] [unos adornos para las niñas].***
'This morning Linch sent some decorations to the family for the girls.'

c. ***Envió [un telegrama] [a sus padres y familiares que vivían en su pueblo natal de Las Cuchillas, en la provincia de Corrientes].***

'He sent a telegram to their parents and relatives that live in his hometown of Las Cuchillas, in the province of Corrientes.'

As stated, the sentences were gathered from *Corpus del Español* (Davies, 2002-), a 100+ million word corpus of Spanish which allows users to perform advanced searches based on parts of speech, lemmas, collocations, synonyms, and frequency, in different time periods and genres. *Corpus del Español* contains data from the 13th to the 20th Century. However, since our study is synchronic, we limited our search to contemporary Spanish only (20th Century). *Corpus del Español* contains spoken and written data. The oral data consists of interviews, the written data belongs to three different genres: fiction (literature), news, and academic. We included in our study data from the four genres. The distribution of the examples according to genre is the following: oral (N = 94), fiction (N = 191), news (N = 460), academic (N = 263). *Corpus del Español* contains data from both Spain and the Americas. We included in our study data from Spain, Argentina, Bolivia, Chile, Colombia, Costa Rica, Cuba, Dominican Repuatblic, Guatemala, Honduras, Mexico, Peru, Puerto Rico, USA, and Venezuela. A total of 429 examples are from Spain and a total of 441 from the Americas. We were not able to determine the geographic origin of 138 examples.

The first step in order to build the sample was to select representative verbs belonging to the different classes of Spanish ditransitive verbs (Delbecque and Lamiroy, 1996). The verbs used in the study are the following:[30]

30 We adopted Delbecque and Lamiroy (1996) classification in this study although we consider that further research on the lexical semantics of dative taking verbs in Spanish is necessary.

Table 4.1: Verbs in the sample of ditransitive sentences

VERB	Count	Percent
aconsejar('advice')	16	1.59
adjudicar('assign')	17	1.69
aportar('provide')	52	5.16
arrojar('throw')	10	0.99
conferir('confer')	53	5.26
dar('give')	86	8.53
dedicar('dedicate')	157	15.58
donar ('donate')	20	1.98
entregar('deliver')	50	4.96
enviar('send')	140	13.89
indicar('indicate')	37	3.67
lanzar('release')	28	2.78
ofrecer('offer')	232	23.02
proporcionar('provide'	32	3.17
regalar ('present')	37	3.67
transferir ('transfer')	41	4.07
TOTAL	1008	100

The search in *Corpus del Español* was conducted using the lemma of each verb, which consists of the infinitive of the verb between square brackets (for example, "[enviar]"). The output of the search consists of a list of every concordance found in the corpus for every inflected form of the verb. For instance, the output for *enviar* ('to send') contains 2,392 concordances. The concordances are presented in groups defined by inflected form. A concordance for the inflected form *envió* ('sent') is:

(3) *en ese barrio. [60] La acompañó hasta su casa y al día siguiente* ***le envió*** *una orquídea lila con una tarjetita: «Quisiera volverla a ver».*
'in that neighborhood. He accompanied her to her house and the next day sent her a purple orchid with a card: "I would like to see you again".'

For each concordance, it is also possible to retrieve an expanded context and information about the source (date, title, author, source). For this concordance, the expanded context and source information is the following:

Source information:

Date	
Title	El peldaño gris
Author	Gayoso Manzur, Milia (1962-)
Source	http://www.cervantesvirtual.com/FichaObra.html?Ref=5605

Expanded context:

de las blusas y revisó sus polleras, todas oscuras y gastadas, todas muy con pinta de empleada pobre y vieja. A Rogelio lo conoció en el colectivo cuando volvía del trabajo en la biblioteca del Centro Cultural, él le dio el asiento y después cuando se desocupó el de al lado se sentó allí, hizo un comentario sobre el calor, la invasión de moscas y al final se bajó en la misma cuadra que ella con el pretexto de buscar una dirección en ese barrio. [60] La acompañó hasta su casa y al día siguiente **le envió una orquídea** lila con una tarjetita: « Quisiera volverla a ver ». « Cosas así ocurrían antes, treinta años atrás o se ve en las películas », dijo Georgina, divertida y emocionada, porque era la primera vez en cincuenta años que alguien le regalaba una simple flor, y encima i una orquídea! Después la fue a buscar al trabajo y hablaron de sus soledades, del primer matrimonio de él, de su esposa muerta, de los hijos ya casados, de sus cinco nietos; y hablaron

Figure 4.1: Example of expanded context in *Corpus del español*

Given that there is no searchable Spanish tree-bank available, it was not possible to determine automatically which sentences belong to the ditransitive pattern and it was necessary to review the corpus manually.

The output of a search is presented in pages of 100 concordances each. In order to gather the examples, we reviewed every other page of the output for each inflected form of each lemma. For instance, there were 515 concordances for the inflected form *envió*, divided in 6 pages. We reviewed pages 1, 3, and 5 (concordances 1 to 100, 201 to 300, and 401 to 500). The same procedure was applied to the output for every verb, with the exception of *dar* ('to give'), for which a convenience sample was selected.

After being reviewed, the examples were included in the sample if they met the following criteria:

a. The context of the sentence needs to allow for the optionality of DCLD. In other words DCLD should not be required in that context. The exclusion of contexts where DCLD is categorical rather than optional is an important part of the definition of a variable context according to the variationist methodology (Walker, 2010). DCLD is required if the recipient is pronominal (a stressed pronoun), as in the following example, which would be unacceptable without DCLD. We assume that the fact that stressed pronouns are always doubled follow from the fact that stressed pronouns are characterized by a high degree of pragmatic salience:

(4) *... él **le sugirió a ella** orillarse y platicar un ratito ...*
'… he suggested that she pull over and talk a while...'

DCLD is also required if it is left or right-dislocated, as in the following examples, also excluded from the sample:

(5) a. Left-dislocation:
A María, Juan le dio el libro.
to María Juan DAT.3SG gave the book
'María, Juan gave her the book.'

b. Right-dislocation:
Juan le dio el libro, a María.
Juan DAT.3ST gave the book to María
'Juan gave the bookto her, to María.'

In order to determine if there is dislocation, the criterion used was whether or not there is a comma. In the two previous examples, there is a comma separating the dislocated constituent from the core sentence.

b. Examples in which the recipient is not lexically expressed were not included either (*Juan ofreció un café*, 'Juan offered coffee'). Those examples were not included because they are not instances of doubling and they are not useful in order to study the order of theme and recipient.

c. Examples in which the theme is not lexically expressed were not included either (*Juan se lo ofreció*, 'Juan offered it to him/her/them'). Those examples were not included because they are not instances of doubling and they are not useful in order to study the order of theme and recipient.

d. Examples in which there is a spatial goal rather than a recipient (spatial goals are not IOs but OBLs and cannot undergo DCCL) were excluded. For instance, the following example was excluded (the locative goal is *a Francia*, 'to France'):

(6) *... pero ocurrió lo siguiente: que mi padre envió a Francia a estudiar a mis dos hermanas mayores...*
'... but the following happened: my father sent my two older sisters to France to study …'

e. In order to be included, the examples should be declarative sentences in the active voice. This

criterion is justified by the fact that interrogative and passive sentences display a marked order of constituents.

4.3 THE VARIABLES

The examples were annotated for a set of independent variables and compiled in a Calc spreadsheet.[30] As was previously stated, we performed two different quantitative studies with different dependent variables: occurrence of DCLD for Quantitative Study #1 and order of theme and recipient for Quantitative Study #2. The independent variables, which are discussed in detail below, were the same in both studies (animacy, definiteness, givenness, and grammatical category of theme and recipient, relative length of theme and recipient, and lexical semantics of the verb). The following sections contain an individual discussion of the independent variables. Given the important level of interaction among these properties, the discussion will overlap on some occasions.

4.3.1 Animacy

Animacy is a property to which natural languages have an important degree of sensitivity (Comrie, 1989). The linguistic relevance of animacy has been interpreted by Kuno (1976) in terms of the "empathy principle":

31 http://www.openoffice.org/

> Human beings tend to select as topics entities with whom they empathize, first of all themselves, then the person they are speaking to, then other human beings, then non-human animate beings, and finally the inanimate world. Therefore, morphosyntactic expressions whose function is to refer to topical entities tend to refer to entities that speakers empathize with (Kuno, 1976).

We assume that DCLD can be described as a morphosyntactic expression that refers to topical entities. Therefore, we would expect that recipients high in animacy would be more likely to undergo DCLD.

The ways in which animacy affects the formal structure of natural language are diverse, including the distribution of split ergativity (Comrie, 1989) and differential object marking (DOM, Bossong, 1991; Aissen, 1993). DOM is the tendency for languages with overt grammatical function marking of direct objects to optionally mark them according to their prominence in a scale of either animacy or definiteness (or both). The manifestation of DOM in Spanish is the distribution of personal *a*, which is required with human DOs:

(7) a. *Lo están buscando *(a) él.*
ACC.3SG.MASC are looking.for to NOM.3SG.MASC
'They are looking for him.'

b. *Están buscando *(a) Juan.*
are looking.for to Juan
'They are looking for Juan.'

c. *Están buscando *(a) su hermano.*
are looking.for to his/her brother
'They are looking for his/her brother.'

Regarding object agreement, the grammatical function that we assume is fulfilled by DCLD, the relevance of animacy has been discussed by Woolford (1999), which describes the effects of animacy on different African languages, including Bantu.

The main values of the animacy hierarchy are *human* > *animate* > *inanimate* (Comrie, 1989; Aissen, 2003). However, there are finer distinctions that have been proposed in the literature, such as the introduction of the category "organization", which refers to temporally stable entities constituted by a collective of humans with a collective purpose (Yamamoto, 1999; Garretson, 2004; Zaenen at al., 2004). The determination of the animacy of NPs can be quite complex because of the ambiguity (and/or vagueness) of linguistic expressions, as well as the widespread use of metaphor and metonymy in human language (Zaenen et at., 2004). For instance, the word *congreso* ('congress') can refer to an inanimate object (building in which the congress of a specific country functions), to an organization (congress as a collective of humans with a specific purpose), or to a set of humans (members of congress as a set of humans), etc.

For the purposes of coding, we will use the traditional animacy scale, with the addition of the category "organization". The resulting scale is *human* > *organization* > *animate* > *inanimate*.

In order to conduct the annotation we follow the coding practices established by the project Optimal Typology of Determiner Phrases,[31] henceforth OTDP, described in Garretson (2004). The animacy hierarchy used by OTDP is different from the one we are using here, but

32 http://npcorpus.bu.edu/documentation/index.html.

largely analogous. It contains three tiers, which can be broken up into two or more categories. The tiers are “top”, “middle”, and “bottom”. The top tier consists of the category “human” only. The middle tier consists of the categories “animal” and “organization”. The bottom tier consists of the categories “concrete inanimate”, “non-concrete inanimate”, “place”, and “time”. Therefore, the top tier is equivalent to “human”, as in the hierarchy we are using. The middle tier groups two categories that we are keeping separated (organization and animate). The bottom tier is equivalent to “inanimate” in our scale, although it contains finer distinctions that we are not going to use for the purposes of coding. The OTDP coding practices provide important insights in order to distinguish some of the categories, especially “human” from “organization”. In the following paragraphs we will discuss the definition of each value of animacy, as well as provide examples of each one. The definitions are based on those of OTDP.

4.3.1.1 Human “Human” is the animacy tag used for entities which are singular human beings or collectives of humans, as far as they do not constitute an “organization” (see below). A human entity can be the referent of a proper name (*Amaranta*), a kinship term (*tu padre*, 'your father'), a common noun (*campesino*, 'farmer'), or a pronoun (our sample does not contain personal pronouns as recipients or themes, but it contains demonstratives which in some cases refer to humans). Our sample also contains many examples of titles (*el director de la policía*, 'the chief of police'). We also adopt the criteria, discussed in Garretson (2004), that entities that behave as human, such as gods, belong to the category “human” (for instance: *Afrodita*, 'Aphrodite').

4.3.1.2 Animate “Animate” is the animacy tag used for entities which are non-human animals of any kind. All of the examples of this category in our sample are common nouns (*las demás*

abejas, 'the other bees'; *los depredadores*, 'the predators', *el alce*, 'the elk'), but proper nouns, kinship terms, and pronouns can also be used to refer to them.

4.3.1.3 Organization "Organization" is the animacy tag used for entities which are groups of humans seen as collective bodies characterized by a group identity. In order to identify an "organization", the following set of properties, described as an implicational hierarchy[32], is provided in Garretson (2004). The properties in the set are treated as binary features, as shown below:

+/- chartered/official

+/- temporally stable

+/- collective voice/purpose

+/- collective action

+/- collective

The following are some examples of NPs from our sample that usually refer to organizations: *el equipo colombiano* ('the Colombian team'), *el gobierno indonesio* ('the Indonesian government'), *la Organización de las Naciones Unidas* ('the United Nations').

The cutting point for a set of humans to be considered an "organization" is whether or not the entity has collective voice or purpose. Any set of humans is a collective and can have collective action without having a collective voice or purpose. An example of this distinction

33 This set of properties is represented below as an implicational hierarchy, in which any entity which has a property will also have all of the properties to its right:
chartered/official > temporally stable > collective voice/purpose > collective action > collective
An entity which is [+chartered/official] will also have all of the other properties.

from our sample is the NP *sus colegas independientes de Cuba* ('her independent colleagues from Cuba'), which refers to the collective of independent journalists who live and work in Cuba. In spite of being able to have collective action, such a collective does not necessarily have a collective voice or purpose. Therefore, we will consider it a collection of human entities and tagged it as "human". The NP appears in this context:

(8) [+collective action, - collective voice/purpose] → tagged as "human" instead of "organization":
La reportera de Univisión de Texas, Teresa Farfani, aprovechó la oportunidad para enviar un cálido saludo de los periodistas hispanos de los Estados Unidos ***a sus colegas independientes de Cuba****, ...*
'Texas Univision reporter Teresa Farfani took advantage of the opportunity to send warm greetings from the Hispanic journalists in the United States to Cuba's independent colleagues, ...'

Another example of a collective of humans without a collective voice or purpose is the referent of the NP *sus aficionados* ('its fans') in the following context:

(9) [+collective action, - collective voice/purpose] → tagged as "human" instead of "organization":
El Parc Lescure en esta ciudad, con espacio para 35,200 espectadores, se tapizó de azul y blanco y la Selección Argentina le regaló ***a sus aficionados*** *otra victoria.*
'The Parc Lescure in this city, with room for 35.200 spectators, turned blue and white and the Argentinean Team gave their fans another victory.'

We have taken the opposite decision, to consider a collective of humans as an "organization" in the case of the NP *las autoridades argentinas* ('the Argentinean authorities'), which we assume have a collective voice or purpose (in addition to being temporally stable and being official). The NP appears in the following context:

(10) [+ collective voice/purpose] → tagged as "organization":
En 1827 ofreció sus servicios ***a las autoridades argentinas*** *con motivo de la guerra con Brasil ...*
'In 1827 he offered his services to the Argentine authorities during the war with Brazil ...'

The following is another example in which a NP makes reference to the set of individual members of an organization (*los nueve miembros de un organismo colegiado, el Consejo Nacional de Gobierno*, 'the nine members of a collegial body, the National Governing Council') which we understand as a reference to the members as a collective entity, the organization itself. Notice that the head of the NP is a plural human entity. The NP appears in the following context:

(11) [+ collective voice/purpose] → tagged as "organization":
Dos años después, la reforma constitucional promovida por el presidente abolió la presidencia y transfirió el poder ejecutivo ***a los nueve miembros de un organismo colegiado, el Consejo Nacional de Gobierno***.
'Two years later, the constitutional reform promoted by the president abolished the presidency and transferred power to the nine members of a collegial body, the National Governing Council.'

4.3.1.4 Inanimate "Inanimate" is the animacy tag used for entities which are not alive and are not a collective of live entities. In Garretson (2004), inanimates constitute the "bottom" tier of the hierarchy and are subdivided in four sub-categories ("concrete", "non-concrete", "place", "time"). Although we will not use those sub-categories for the purposes of coding, they are useful in order to describe the range of NPs that belong to the category "inanimate". The criterion used in Garretson (2004) in order to differentiate "concrete" from "non-concrete" inanimates is whether or not the referent of the NP can be perceived with one of the five senses. Some examples of NPs from our sample that usually refer to "concrete" inanimates are *un*

papelito ('a small piece of paper'), *una carta* ('a letter'), *un cigarrillo* ('a cigarette'), but notice that *una carta* would not be concrete if it refers to the content of the letter instead of the letter itself. Also, some NPs which can refer to concrete inanimates can also refer to organizations in certain contexts. For instance, the NP *el templo* ('the temple'), which can refer to a building, which is an inanimate and concrete object, can also refer to the organization that functions in that building as the result of metonymy. That happens in the following context, in which *el templo* was tagged as "organization":

(12) Inanimate noun tagged as "organization":
Cuando Úrsula dispuso la reanudación de la misa dominical, Pietro Crespi le regaló ***al templo*** *un armonio alemán, ...*
'When Ursula ordered the resumption of Sunday Mass, Pietro Crespi gave the temple a German harmonium, ...'

A NP that usually refers to a non-concrete inanimate entity can also refer to a human as the result of metonymy. In the following context, a body part (*la cabeza de uno de los celadores*, 'the head of one of the guards') is used to refer to the person. The NP was tagged as human:

(13) Body part tagged as "human" (metonymy):
El alcalde Cuevas entregó a los periodistas ***la cabeza de uno de los celadores****, a quien acusó de haber facilitado la fuga a cambio de diez mil pesos.*
'Mayor Cuevas gave the journalists the head of one of the guards, whom he accused of having facilitated the escape in exchange for ten thousand pesos.'

Some examples of NPs that usually refer to non-concrete inanimates in our sample are *importantes derechos económicos* ('important economic rights'), *la jurisdicción eclesiástica* ('the ecclesiastic jurisdiction'), *el poder* ('the power').

The category "place" is used in Garretson (2004) to tag NPs that refer to entities characterized as "a stationary area in the surface of the planet (or above it) that is the potential location of a human". Notice that NPs that potentially refer to locations can be actually referring to other kind of entities. For instance, the NP *Honduras* in the following context is not used to refer to a location, but as the recipient of a donation. As such, it cannot be categorized as a "inanimate", but it was tagged instead as "organization":

(14) Locative expression tagged as "organization":
Japón dona más de L. 300 millones ***a Honduras*** *para varios proyectos.*
'Japan donates over 300 million L. to Honduras for various projects.'

The same consideration applies to the NP *el Perú* in the following context:

(15) Locative expression tagged as "organization":
... la firma MAPO Mig ofreció ***al Perú*** *una nueva versión de esos cazas al precio unitario de US$ 24 millones...*
'… the firm MAPO Mig offered to Peru a new version of these fighters at a price of $ 24 million each...'

An example of an NP that refers to a "place" and was tagged as "inanimate" is *los territorios al este del río Uruguay* ('the territories to the east of the Uruguay River') in the following context:[33]

34 It can also be considered that the referent of the NP is actually a "concrete" entity. In any case, we would tag it as "inanimate".

(16) Locative expression tagged as "inanimate":
Cuando el Tratado dio al Reino de Portugal ***los territorios al este del río Uruguay,*** *los indios se sintieron traicionados.*
'When the Treaty gave the Kingdom of Portugal the territories to the east of the Uruguay River, the Indians felt betrayed.'

The last sub-category of "inanimate" used in Garretson (2004) is "time", which is used to tag NPs that refer to periods of time. We have not found examples of this type as recipients or themes in our sample.

4.3.1.5 Coordination A special case, discussed in Garretson (2004), arises when two NPs with different value for animacy are coordinated. An example from our sample is the NP *el Consejo Técnico y la opinión pública en general* ("the Technical Council and the public opinion in general"). The first conjunct would be tagged as organization and the second as human, understanding that it makes reference to the public opinion as a set of humans). But the NP as a whole is the recipient of a communication verb and it needs a single animacy tag. In cases like this one, we have decided to use the tag that applies to the first conjunct, which in this case is "organization". The decision is to some extent arbitrary, but it is motivated by the fact that the first conjunct is closer to the verb and the optional dative clitic.

4.3.1.6 Descriptive statistics The distribution of animacy in our sample shows a clear difference between themes and recipients: 94.25% of themes are inanimate (N = 950). The percentage of human, organization, and animate themes combined is 1.39% (N = 14). 4.37% (N = 44) of the themes consist of embedded sentences in the position of complement ocommunication verbs like *indicar* ('to indicate'), and therefore are unmarked for animacy. See table below:

Table 4.2: Animacy of themes in the sample

Animacy of th.	Count	Percent
human	8	0.79
organization	5	0.5
animate	1	0.1
inanimate	950	94.25
unmarked	44	4.37
N = 1008		

In the case of recipients, the distribution is different. The most common value of animacy for recipients is "human" (49.70%, N = 501), but there are also high percentages of "organization" (23.41%, N = 236) and "inanimate" (25.69%, N = 259). Only 1.19% (N = 12) of the recipients are "animate". See table below:

Table 4.3: Animacy of recipients in the sample

Animacy of rec.	Count	Percent
human	501	49.7
organization	236	23.41
animate	12	1.19
inanimate	259	25.69
N = 1008		

The predominance of inanimate themes and human recipients has been previously noticed in the literature (Delbecque and Lamiroy, 1996).

4.3.2 Definiteness

There is no agreement among linguists about the definition of the semantic and/or pragmatic content of definiteness (Lyons, 1999; Abbot, 2006, 2010). There are two main approaches, one that understands definiteness in terms of "uniqueness" and another that understands it in terms of "familiarity". The definition in terms of "uniqueness" has its first antecedent in Russell's (1905) analysis of definite descriptions, based on the idea that definite NPs have only one potential referent. According to Russell, the sentence *a man runs away* means that there exists something which is both a man and runs away, as shown in the logical form below:

(17) *A man runs away.*
$\exists$ x [man(x) & run away(x)]

The logical form of *the man run away*, on the other hand, contains a statement that enforces the unique reading of the NP in the form of an implication of uniqueness. The logical form is the following:

(18) *The man runs away.*
$\exists$ x [man(x) & $\forall$ y[man(y) $\rightarrow$ y = x] & run away(x)]

The definition in terms of "familiarity" is based on the idea that the referent of a definite NP is already known by the listener or that the listener should be at least able to identify it using the information contained in the description (Abbot, 2006). This definition overlaps, to a considerable degree with the discussion of the cognitive status of Noun Phrases (NPs) in terms of

givenness. This fact has been noticed by Abbot (2006), who states that "uniqueness of applicability of the descriptive content, as explicated in Russell's analysis, is a strictly semantic property while the assumption of familiarity to the addressee is discourse-pragmatic in nature" (Abbott, 2004: 135-6). The analysis of definiteness in terms of familiarity implies that definiteness is a grammaticalization of the cognitive status of NPs. The cognitive status of NPs will be discussed in a separate section.

Independently of the nature of definiteness we have decided to use the acceptability of a NP as the "focus NP" of an existential sentence as a test to decide whether to tag it as definite or indefinite. Existential sentences, in English, are built using *there* and the verb *to be*, followed by a NP (the focus NP) and a predicative constituent, usually a PP (Abbot, 2010). These sentences allow only for indefinite, but not definite, NPs following the verb *to be* (definiteness effect)*:*

(19) a. *There is a man in the room.*

b. *#There is the man in the room.*

In Spanish, existential sentences are constructed with the impersonal verb *haber* in sentence initial position, which will always be in the 3rd person singular form (*hay*, in the present indicative) followed by a NP and an optional predicative constituent. The definiteness effect in Spanish follows the same pattern as in English; indefinite, but not definite, NPs are allowed in existential sentences:

(20) a. *Hay un hombre en la sala.*
there.is a man in the room
'There is a man in the room.'

b. *#Hay el hombre en la sala.*
there.is the man in the room
'There is the man in the room.'

The status of this test as a criterion to decide on the definiteness of a NP is controversial. If definiteness is defined in terms of uniqueness, universally quantified NPs would not be definite. Milsark (1974) claims that the test does not distinguish definite from indefinite, but "strong" from "weak" NPs, where "weak" NPs are non-quantificational and "strong" NPs are quantificational. According to Milsark, the existential construction introduces an existential quantifier, which is incompatible with already quantified "strong" NPs.

The existential construction provides an environment in which indefinite NPs are allowed and definite NPs are not. Independently of the ultimate nature of the definiteness effect, the acceptability of a NP as the focus NP of an existential sentence is a test widely used in order to distinguish definite from indefinite NPs. For instance, it is the main criterion for definiteness used by the project Optimal Typology of Determiner Phrases (Garretson, 2004; http://npcorpus.bu.edu/documentation/index.html).

The features to be used in order to tag the examples are: "indefinite", "specific" (meaning specific indefinite, to be discussed below), and "definite".

4.3.2.1 Definite NPs We will consider definite the NP types which are not allowed as the focus NP of an existential sentence (NPs with definite determiners, demonstrative determiners,

possessive determiners, and universal quantifiers, in addition to NPs consisting of a demonstrative or personal pronoun or a proper name):

a. NPs with definite determiners (Spanish *el*, *la*, *los*, *las* 'the') are not allowed in existential sentences, as has been shown above. The following is an example from our sample of a definite recipient with a definite determiner:

(21) Definite recipient (definite determiner):
Se sabe que envió un sobre ***a la casa matriz****, dentro había un listado de pedidos y el dinero que, seguramente previó, ya no podría llevar personalmente.*
'It is known that he sent an envelope to the headquarters containing a list of orders and money he surely foresaw he wouldn't be able to take personally.'

Notice that some NPs with definite determiners have a generic interpretation, in which case they refer to a whole class or kind (Krifka et al., 1995). These NPs are also considered definite, assuming that reference to a class is an instance of reference to a unique entity (Krifka et al., 1995). An argument in favor of treating generic NPs as definite is that they show the same definiteness effect as non-generic definite NPs:

(22) *#Hay las jirafas en África.*
there.is the giraffes in Africa
'There are the giraffes in Africa.'

b. NPs with demonstrative determiners (Spanish *este*, *esta* 'this', *estos*, *estas* 'these', *ese*, *esa* 'that', *esos*, *esas* 'those', *aquel*, *aquella* 'that over there', *aquellos*, *aquellas* 'those over there'), which are not allowed as the focus NP of an existential sentence:

(23) *#Hay ese hombre en la sala.*
there.is that man in the room
'There is that man in the room.'

The following is an example from our sample of a definite theme with a demonstrative determiner:

(24) Definite theme (demonstrative determiner):
Originaria de la antigua Grecia, la academia era un jardín público a las afueras de Atenas, cuyo propietario era Academo, habitante del Ática, que donó ***estos jardines*** *al pueblo ateniense.*
'Originating in ancient Greece, the academy was a public garden on the outskirts of Athens, owned by Akademos, a resident of Attica, who donated these gardens to the Athenian people.'

c. NPs with possessive determiners (Spanish *mi*, *mis* 'my', *tu*, *tus* 'your', *su* 'his, her', *nuestro*, *nuestra*, *nuestros*, *nuestras* 'our', *vuestro*, *vuestra*, *vuestros*, *vuestras* 'your', *sus* 'his, her, their'), which are not allowed in existential sentences:

(25) *#Hay su amigo en la sala.*
there.is his/her friend in the room
'There is his/her friend in the room.'

The following is an example from our sample of a definite recipient with a possessive determiner:

(26) Definite recipient (possessive determiner):
el 31 de julio de este mismo año le indica ***a su padre*** *que ha entregado el concierto hace cuatro meses y que el duque no le ha pagado todavía.*
'On July 31 this year, he told his father that he has given the concert four months ago and that the Duke has not yet paid him.'

d. NPs with universal quantifiers (Spanish *todo*, *toda* 'all', *cada* 'each'), which are not allowed in existential sentences:

(27) *#Hay cada amigo en la sala.*
there.is each friend in the room
'There is each friend in the room.'

The following is an example from our sample of a definite theme with a universal quantifier:

(28) Definite theme (universal quantifier):
Divagando sobre la impresión que estaría causando en los inesperados oyentes, dedicó ***cada acorde*** *al Espíritu Santo, al que exultaba por formar parte de la nunca bien comprendida Santísima Trinidad.*
'Wandering about the impression that he would be causing on the unexpected listeners, he devoted each line to the Holy Spirit, which he exulted for being part of the never well understood Holy Trinity'.

e. Demonstrative pronouns (Spanish *éste*, *ésta* 'this', *éstos*, *éstas* 'these', *ése*, *ésa* 'that', *ésos*, *ésas* 'those', *aquél*, *aquélla* 'that over there', *aquéllos*, *aquéllas* 'those over there', in addition to the neuter forms *esto*, *eso*, *aquello*), which are not allowed in existential sentences:

(29) # *Hay éste en la sala.*
there.is this in the room
'There is this in the room.'

The following is an example from our sample of a definite recipient expressed by a demonstrative pronoun, notice that DCLD is not required in the case of demonstrative pronouns, as opposed to what happens with personal pronouns:

(30) Definite recipient (demonstrative pronoun):
En 1812 asistió a una serie de conferencias impartidas por el químico Humphry Davy y envió ***a éste*** *las notas que tomó en esas conferencias junto con una petición de empleo.*
'In 1812 he attended a series of lectures by the chemist Humphry Davy and sent him notes taken at these conferences along with a request for employment.'

f. Personal pronouns are definite, but were not included the sample because DCLD of pronominal recipients is not optional, but required (*a mí* 'to me', *a ti* 'to you', *a usted* 'to you', *a él* 'to him', *a ella* 'to her', *a nosotros*, *a nosotras* 'to us', *a vosotros*, *a vosotras*, *a ustedes* 'to you', *a ellos*, *a ellas* 'to them')

g. Proper names, either first, last, or full names, which are definite are not allowed in existential sentences:

(31) # *Hay Juan en la sala.*
there.is Juan in the room
'There is Juan in the room.'

The following is an example from our sample of a definite recipient expressed by a proper noun:

(32) Definite recipient (proper noun):
O mejor dicho obsequiarle el salvoconducto ***a Víctor****, porque Ilse ya lo tenía.*
'Or rather, giving the pass to Victor, because Ilse already had it.'

4.3.2.2 Indefinite NPs The referents of indefinite NPs are assumed not to possess the semantic and or pragmatic properties associated with definite NPs. From the semantic point of view, the referent of an indefinite is not assumed to be "unique" because indefinite NPs can refer to any member of a class. In Russell's (1905) analysis, an indefinite NP is equivalent to an existential quantifier. From the discourse-pragmatic point of view, the referent of an indefinite NP is not assumed to be "familiar" to the listener (Ariel, 1988).

Indefinite NPs are felicitous in existential sentences. The following types of NPs are assumed to be indefinite:

a. NPs with indefinites determiners (Spanish, *un*, *una*, *unos*, *unas* 'a'). This category includes any NP with an indefinite determiner, except from partitive NPs and NPs modified by a relative clause in the indicative mood (see 4.3.2.3). NPs with indefinite determiners are allowed in existential sentences:

(33) *Hay un hombre en la sala.*
there.is a man in the room
'There is a man in the room.'

The following is an example from our sample of an indefinite theme:

(34) Indefinite theme (indefinite determiner):
Al cabo de dos horas, cuando la conversación empezaba a languidecer, Amparo aprovechó un descuido de Amaranta y le entregó ***una carta*** *a Rebeca.*
'After two hours, when the conversation began to languish, Amparo took advantage of the fact that Amaranta was not paying attention and handed a letter to Rebeca.'

b. NPs with existential quantifiers (Spanish *algún*, *alguna*, *algunos*, *algunas* 'some') are allowed in existential sentences:

(35) *Hay algunos hombres en la sala.*
there.is some men in the room
'There are some men in the room.'

The following is an example from our sample of an indefinite theme with an existential quantifier:

(36) Indefinite theme (existential quantifier):
Las epístolas de Pablo y Hechos ofrecen al lector ***algunos datos acerca de la vida de estas primitivas comunidades cristianas y sobre su relación con las culturas hegemónicas.***
'The epistles of Paul and Acts give the reader some facts about the life of these early Christian communities and their relationship with hegemonic cultures.'

c. Bare NPs, either singular mass nouns or plural count nouns, are allowed in existential sentences:

(37) a. *Hay agua en la piscina.*
there.is water in the pool
'There is water in the swimming pool.'

b. *Hay niños en la vereda.*
there.is children in the sidewalk
'There are children in the sidewalk.'

The following is an example from our sample of a bare NP theme (plural count noun):

(38) Indefinite theme (bare NP):
Súbitamente las bailarinas se levantaron para formar sucesivamente procesiones a los altares de Cibeles, las teorías de las Vestales mantenían el fuego sagrado, los desfiles de cortesanas bajo los pórticos dedicaban ***ofrendas*** *a Afrodita, aparecían los cortejos de las Panateneas.*
'Suddenly the dancers turn up to form processions to the altar of Cybele, the theories of Vestal held the sacred fire, the processions of courtesans under the portals dedicated offerings to Aphrodite, the Panathenaic processions appeared.'

4.3.2.3 Specificity In addition to the distinction between definite and indefinite, we tagged the indefinite NPs as either specific indefinites or non-specific indefinites. Therefore, the definiteness scale we used was: definite > specific (indefinite) > (non-specific) indefinite.

The referent of the specific indefinite NP is a particular individual known by the speaker which has not been previously introduced in discourse (Leonetti, 1999; Abbot, 2006, 2010). Rivero (1975) shows that specific indefinite NPs are modified by indicative relative clauses and

non-specific indefinite NPs are modified by subjunctive relative clauses (see examples below). Notice also that specific indefinites require personal *a* but non-specific indefinites do not allow it. These tests were used in order to tag the examples:

(39) a. Specific indefinite → modified by indicative relative clause and personal *a*:
Estoy buscando ****(a) un estudiante que habla francés***.
am looking to a student that speaks.INDICATIVE French
'I'm looking for a (particular) student that speaks French.'

b. Non-specific indefinite → modified by subjunctive relative clause and no personal *a*:
Estoy buscando ***(*a) un estudiante que hable francés.***
am looking to a student that speaks.SUBJUNCTIVE French
'I'm looking for (any) student that speaks French.'

The following is an example from our sample of a specific indefinite theme (*un documento que detalla los productos sobre los cuales espera obtener preferencias arancelarias* 'a document that details the products on which tariff preferences are expected'). The specific reading is obtained because the verb of the relative clause is in the indicative mood.

(40) Specific indefinite theme (modified by indicative relative clause):
Explicó que el proceso se encuentra bastante avanzado: hace algunas semanas el Gobierno de Chile envió al cubano ***un documento que detalla los productos sobre los cuales espera obtener preferencias arancelarias.***
'He explained that the process is well under way: a few weeks ago the Government of Chile sent to the Cuban government a document that details the products on which tariff preferences are expected.'

Partitive NPs (like *uno de los estudiantes* “one of the students”) are also specific, as is shown by the modification by an indicative relative clause and the use of personal *a*:

(41) Partitive NP (specific):
Busco a uno de los estudiantes que sabe francés.
look.for to one of the students that knows French
'I'm looking for one of the students that knows French.'

The following is an example from our sample of a specific indefinite recipient (*a uno de los discípulos* 'to one of the disciples'). The specific interpretation is associated with the partitive construction.

(42) Specific indefinite recipient (partitive):
Pietro Crespi se puso pálido, le entregó la cítara ***a uno de los discípulos****, y dio la clase por terminada..*
'Pietro Crespi turned pale, handed the sitar to one of the disciples, and considered the class finished.'

4.3.2.4 Descriptive statistics The descriptive statistics show that themes in our sample are predominantly indefinite (60.12%, N = 606). 33.63% of the themes are definite (N = 339), 1.88% are specific indefinite (N = 19), and 4.37% are embedded clauses unmarked for definiteness (N = 44):

Table 4.4: Definiteness of themes in the sample

Definiteness of th.	Count	Percent
Definite	339	33.63
Indefinite	606	60.12
Null	44	4.37
Specific	19	1.88
N = 1008		

On the other hand, recipients are overwhelmingly definite (92.86%, N = 936). Only 5.95% of the recipients are indefinite (N = 60) and 1.19% are specific indefinite (N = 12).

Table 4.5: Definiteness of recipients in the sample

Definiteness of rec.	Count	Percent
Definite	936	92.86
Indefinite	60	5.95
Specific	12	1.19
N = 1008		

4.3.3 Cognitive status

According to Gundel et al. (1993), the cognitive status of the referent of a NP can be defined as the addressee's knowledge of that referent and his/her attention state in the context in which the NP is used. The addressee has knowledge of the referent if he/she already has a mental representation of it.

In order to code for cognitive status, we will use the givenness hierarchy proposed by Gundel et al. (1993). This hierarchy is characterized by the combination of two types of

information: linguistic and extralinguistic contextual information, on the one hand, and NP form, on the other hand. Linguistic contextual information consists of the information that has been previously introduced in discourse. Extralinguistic contextual information consists of the actual physical context of the interaction, including the participants of the linguistic interaction. NP form refers to the different linguistic forms that a NP can present: a pronoun, a definite NP, an indefinite NP, as has been defined in the section on definiteness (4.1.3.2). To an important extent, Gundel et al.'s approach to cognitive status adopts the point of view that definiteness is coding "familiarity".

The role of linguistic and extralinguistic information has been the subject of a long tradition of studies which has its starting point in the Prague School of Linguistics (Firbas, 1966; Toman, 1995). Firbas made the distinction between given information (the "topic") and new information (the "focus") and studied its role in word order.

Prince (1981) presented a detailed taxonomy of given and new information, which distinguishes three degrees of givenness: new information, inferrable information, and evoked information, which are further classified into sub-types. Prince refers to the representation of referents in the addressee's memory as cards in a file. New entities can be either "brand-new" or "unused". A brand-new entity is one that is introduced in discourse for the first time and the addressee does not have a previous mental representation of it (an existing card) in memory. (Prince's example: *a beautiful dress* in *I bought a beautiful dress*). An unused entity is one that is introduced for the first time in discourse, but the addressee already has a mental representation of it or an existing card in memory (Prince's example: *Rotten Rizzo* in *Rotten Rizzo can't have a*

third term). The use of the proper noun without additional information presupposes that the addressee has a representation of its referent.

An inferrable entity is one that the addressee can infer from the information contained in the card of a previously introduced referent (Prince's example: *the stupid clerk* in *I went to the post office and the stupid clerk couldn't find a* stamp). The addressee can infer from his/her representation of "post office" that there would be a clerk in it.

An evoked entity is one that is already represented by a card in the memory of the addressee either by having been introduced in discourse ("textually evoked") or by being physically salient in the context of discourse ("situationally evoked"). Prince's example of a textually evoked referent is *the sweet lady* in *Susie went to visit her grandmother and the sweet lady was making Peking Duck* and her example of a situationally evoked referent is *me* in *Lucky me just stepped in something*.

Abbot (2010) notices that the categories proposed by Prince usually correspond to different NP types. The brand-new example consists of an indefinite NP. The unused example consists of a proper noun. The evoked referents are introduced by a definite NP and a pronoun. But Prince (1981) does not elaborate on the relation between givenness and NP type. That relation has been investigated by Ariel (1988, 1990) and Gundel et al. (1993).[34]

According to Ariel (1988, 1990), the cognitive status of the referent of a NP is a measure of the mutual knowledge of a referent by the speaker and listener, which can be established on the basis of any of the following criteria:

35 See also Hawkins (1978).

a. Community membership mutual knowledge: knowledge that speakers are assumed to possess as members of a community (analogous to the concept of "unused" referent in Prince, 1981). Typically, entities in this category will be expressed by proper nouns and definite descriptions (*Joan Smith*, *the president*, etc.).

b. Physical co-presence mutual knowledge: knowledge of referents which are physically present in the context of the utterance (analogous to the concept of "situationally evoked" referent in Prince, 1981). Typically, entities in this category will be represented by demonstrative pronouns or NPs with demonstrative determiners (*that*, *that hat*, etc.).

c. Linguistic co-presence mutual knowledge: knowledge of referents which have been previously introduced in discourse (analogous to the concept of "textually evoked" referent in Prince, 1981). Typically, entities in this category will be represented by personal pronouns or zero anaphora (*she*, *herself*, etc.).

Ariel claims that human languages provide speakers with means to code the accessibility of referents in a consistent way by using accessibility markers, organized in an "accessibility hierarchy". Accessibility markers consist of the use of pronouns, proper names, demonstratives, definite descriptions, etc. The highest level of accessibility in the scale is marked by the use of zero anaphora, while the lowest is marked by the use of a full name plus a definite description.

Ariel (1988, 1990) has conducted corpus studies showing that in English the use of accessibility markers is associated with the location of the antecedent of an anaphoric expression

in text (expressions marked as highly accessible have close antecedents, expressions marked by low accessibility markers have distant antecedents, mid accessibility expressions perform in between).

> Accessibility hierarchy (Ariel, 1988, 1990)
> Joan Smith, the president > Joan Smith > The president > Smith > Joan > That/this hat we bought last year > That hat > This hat > That > This > SHE > she > herself > Ø
> (Full name + definite description > full name > definite description > last name > first name > demonstrative determiner + NP + relative clause + demonstrative determiner + NP > demonstrative pronoun > pronoun > null pronoun)

Gundel et al. (1993), as well as Ariel, claim that there is a relation between the grammatical form of an NP and its cognitive status. But they claim that the cognitive statuses are not "mutually exclusive", but "implicationally related (by definition), such that each status entails (and is therefore included by) all lower statuses, but not vice versa" (Gundel et al., 1993: 276). They present their "givenness hierarchy" shown below, which should be understood as an implicational hierarchy:

> Givenness hierarchy (Gundel et al., 1993)
> in focus > activated > familiar > uniquely identifiable > referential > type identifiable
> (it > {that, this, this N} > that N > the N > (indefinite) this N > a N)

The values in Gundel et al.'s Givenness hierarchy are discussed below, followed by a discussion of the methodology used to annotate the sample and examples of annotation:

a. In focus NPs are at the current center of attention, including the topic of the preceding sentence. This status is necessary for the use of pronouns or zero anaphora. Our sample does not contain pronominal themes or recipients.

b. Activated/familiar: For a referent to be activated, the addressee needs to be represented in current short-term memory. This status is necessary for the use of a demonstrative pronoun (Spanish *éste*, *ésta* 'this', *éstos*, *éstas* 'these', *ése*, *ésa* 'that', *ésos*, *ésas* 'those', *aquél*, *aquélla* 'that over there', *aquéllos*, *aquéllas* 'those over there', in addition to the neuter forms *esto*, *eso*, *aquello*) or a proximal demonstrative determiner (Spanish *este*, *esta* 'this', *estos*, *estas* 'these').

For a referent to be familiar, the addressee needs to have a representation of it in memory, either short-term or long-term memory. This status is necessary for the use of a distal demonstrative determiner (Spanish *ese*, *esa* 'that', *esos*, *esas* 'those', *aquel*, *aquella* 'that over there', *aquellos*, *aquellas* 'those over there'). We have collapsed the two categories ("activated" and "familiar") into one, which we will call "activated". The reason to collapse these values is that there are extremely few examples in the sample with the value "familiar", which is conceptually closer to "activated" than to any other value.

c. Uniquely identifiable: For a referent to be uniquely identifiable the addressee should be able to identify the referent using the information provided in the description. This status is necessary for the use of definite NPs (see 4.3.2.1).

d. Referential: For a referent to be referential, the speaker needs to have the intention to refer to a particular object. This status is necessary for the use of a specific indefinite NP (see 4.3.2.3).

e. Type identifiable: For a referent to be type identifiable, the addressee needs to be able to access a representation of the type of object described by the NP. This status is necessary for the use of indefinite NPs (see 4.3.2.2).

The main difference between Gundel et al. and Ariel's approaches is that in Ariel's approach an expression can have a single cognitive status, independently of contextual considerations. For instance, a definite description (definite NP) will always be considered as having low accessibility. In Gundel et al.'s approach, on the other hand, a definite description will have "at least" the status of "uniquely identifiable". On the one hand, being "uniquely identifiable" entails that the expression is also "referential" and "type identifiable". On the other hand, a definite description can also be "familiar", "activated", or "in focus", if any of those statuses applies to the NP as the result of contextual considerations. The status "in focus" corresponds to the sentence topic(s); the status "activated" corresponds to referents that are represented in memory. Therefore, if a definite description is the current topic of the sentence, its cognitive status would be "in focus". A definite description that has been introduced in discourse but which is not the current topic would be considered to be "activated".

In order to code NPs for cognitive status, we will make use of Gundel's et al. (1993) givenness Hierarchy. The features to be used are "in focus", "activated", "uniquely identifiable", "referential", and "type identifiable".

The coding was done in two steps. In the first step, NPs were coded taking into account only their grammatical form (indefinite NP → "type identifiable", specific NP → "referential", definite NP or proper noun → "uniquely identifiable", demonstrative + NP or demonstrative pronoun → "activated").

In the second step, the context for each sentence provided by *Corpus del Español* was manually checked in order to determine if there was an available antecedent. If an antecedent was found in the previous sentence, the status of the expression was promoted to "in focus". If an antecedent was found beyond the previous sentence, the status of the expression was promoted to "active". Otherwise, the status of the expression was kept unchanged. The following examples show the contexts in which definite NPs were promoted from "uniquely identifiable" to "in focus" and "activated":

(43) Definite NP promoted to "in focus"

NP: *Clinton*

Lewinsky llegó a la oficina de su abogado una hora antes de que se anunciara el arreglo. Parecía preocupada cuando atravesó la multitud de reporteros, y no salió con sus abogados Cacheris, Jacob Stein y Nathaniel Speights cuando éstos anunciaron el acuerdo. Este se anunció mientras el abogado personal de Clinton, David Kendall, seguía negociando con Starr sobre cuándo y cómo ***Clinton*** *respondería a sus preguntas.* ***Según se dice, Starr le envió una citación judicial a Clinton****.*

'Lewinsky arrived at his lawyer's office one hour before the settlement was announced. She looked worried when she went through the crowd of reporters, and did not leave with her lawyers Cacheris, Jacob Stein and Nathaniel Speights when they announced the agreement. This was announced as Clinton's personal lawyer, David Kendall, was negotiating with Star on when and how **Clinton** would respond to his questions. **Reportedly, Starr sent a subpoena to Clinton**.'

(44) Definite NP promoted to “activated”

NP: *el campesino* 'the peasant'

Los curiosos y los vendedores se apartaron, expectantes. Los bueyes protestaron con fuertes mugidos al transponer el cordón y la madera crujió peligrosamente, pero la vieja no dejó de rezar un solo instante. - ¡Opa Barcino! ¡Fuerza Hovero! ¡Huuuuuuuuuuurararara...! - gritó ***el campesino****, mientras hacia tintinear las argollas en la punta de la picada. Los curiosos le dedicaron un fuerte aplauso y el Zorro Gris esbozó su primera sonrisa de la mañana, mientras los autos comenzaban a moverse.* ***Martín se trepó a la carreta y le indicó al campesino por donde debía dirigirse para llegar a la plaza.***

'The onlookers and the vendors went away, waiting. Oxen protested with strong lowing to transpose the cord and the wood creaked dangerously, but the old woman did not stop praying for a moment. - *¡Opa Barcino! ¡Fuerza Hovero! ¡Huuuuuuuuuuurararara...!* - shouted **the peasant,** while the rings at the tip of the bite were clinking. Devoted onlookers applauded loudly and the policeman outlined his first smile of the morning, as the cars started moving. **Martin climbed into the wagon and gave the peasant directions to reach the square**.'

The previous examples showed the “promotion” of a definite NP to “in focus” and “activated” statuses. However, if a definite NP does not have an antecedent in the immediate context, there is no promotion and the NP is annotated as “uniquely identifiable”, as in the following example:

(45) Definite NP annotated as "uniquely identifiable":

NP: su ayudante Johannes Kepler

Sin embargo, Copérnico estaba convencido de que las órbitas planetarias eran circulares, por lo que su sistema requería unas elaboraciones casi tan complicadas como el sistema de Tolomeo al que pretendía sustituir (véase Sistema de Copérnico). El astrónomo danés Tycho Brahe adoptó una fórmula de compromiso entre los sistemas de Copérnico y Tolomeo; según él, los planetas giraban en torno al Sol, mientras que el Sol giraba alrededor de la Tierra.Brahe era un gran observador y realizó una serie de medidas increíblemente precisas. ***Esto proporcionó a su ayudante Johannes Kepler los datos para atacar al sistema de Tolomeo*** *y enunciar tres leyes que se ajustaban a una teoría heliocéntrica modificada.*

'However, Copernicus was convinced that the planetary orbits were circular, so that their system required a working almost as complicated as the Ptolemaic system which it was designed to replace (see Copernican System). The Danish astronomer Tycho Brahe adopted a compromise between the systems of Copernicus and Ptolemy; in his view, the planets revolved around the Sun while the Sun revolved around the Earth. Brahe was a great observer and made a series of incredibly precise measurements. **This gave his assistant Johannes Kepler the data to attack the Ptolemaic system** and formulate three laws that were consistent with a modified heliocentric theory.'

4.3.3.1 Descriptive statistics The descriptive statistics show that themes in our sample are overwhelmingly new information. The values in the givenness hierarchy lower than "activated" account for a combined 90.08% of the themes. Only 9.92% of the themes are "activated" or "in focus". See table below:

Table 4.6: Cognitive status of themes in the sample

Cognitive status of th.	Count	Percent
Activated	91	9.03
In focus	9	0.89
Null	42	4.17
Referential	14	1.39
Type identifiable	624	61.9
Uniquely identifiable	228	22.62
N = 1008		

The same type of statistic shows that among recipients, the proportion of "given" information is much higher. 38.19% of the recipients are either "activated" or "in focus". The remaining 61.81% constitute "new" information. See table below:

Table 4.7: Cognitive status of recipients in the sample

Cognitive status of rec.	Count	Percent
Activated	361	35.81
In focus	24	2.38
Referential	8	0.79
Type identifiable	124	12.3
Uniquely identifiable	491	48.71
N = 1008		

4.3.4 Weight and complexity

The relevance of the relative weight or complexity of constituents for word order has been a topic in the linguistic literature since the formulation of the "Principle of end weight" by Behagel

(1909). The principle states that syntactic constituents are linearly ordered in a way in which shorter constituents precede longer constituents.

Hawkins (1994, 2004) has proposed a functional explanation for the effects of weight in word order. He claims that more complex constituents have a tendency to occur later in the sentence that simpler ones because the order of increasing complexity, at least in head-initial languages, results in a more efficient psycholinguistic processing.

Hawkins (1994) discusses different metrics of complexity, based on either grammatical category (S, PP, NP), number of syntactic nodes, ratio of terminal to non-terminal nodes, and number of words. He concludes that all of them produce roughly the same results. Wasow (2002) arrives to the same conclusion.

We will use two metrics of complexity: relative length of theme and recipient, measured in number of words and grammatical category of theme and recipient. The first criterion is the simplest one and has the advantage that it makes an explicit comparison of the complexity of both constituents. For the purposes of counting, a word will be considered a sequence of characters between blank spaces. The second criterion, grammatical category, provides a rough estimation of syntactic complexity understood in number of syntactic nodes (see below).

The relative length of theme and recipient will be annotated using the values described in Wasow (2002). These values are:

"recipient>>theme": the recipient is 4 or more words longer than the theme

"recipient>theme": the recipient is 1 to 3 words longer than the theme

"theme=recipient": the theme and the recipient have the same length.

"theme>recipient": the theme is 1 to 3 words longer than the recipient.

"theme>>recipient": the theme is 4 or more words longer than the recipient

The grammatical category of theme and recipient will be annotated using two features: "phrasal" and "sentential". NPs and PPs will be annotated as phrasal. Embedded clauses and complex-NPs, which contain an embedded sentence, will be annotated as sentential. These features try to capture a clear distinction in the level of syntactic complexity, since sentential constituents contain at least two additional layers of structure than NPs/PPs, the VP and S nodes.

The following sentences exemplify the annotation that will be used for relative length and grammatical category:

(46) Theme>recipient, phrasal theme and phrasal recipient:
Según se dice, Starr le envió una citación judicial a Clinton.
'Reportedly, Starr sent a subpoena to Clinton.'
Theme: *una citación judicial.* ('a subpoena.')
Recipient: *a Clinton.* ('to Clinton.')
Annotation:

Theme length	Recipient length	Relative weight	Theme category	Recipient category
3	2	theme>recipient	Phrasal	Phrasal

(47) Theme>>recipient, sentential theme, phrasal recipient:

Martín se trepó a la carreta y le indicó al campesino por donde debía dirigirse para llegar a la plaza.

'Martin climbed into the wagon and told the peasant how to reach the square.'

Theme: *por donde debía dirigirse para llegar a la plaza.* ('how to reach the square.')

Recipient: *al campesino.* ('to the peasant.')

Annotation:

Theme length	Recipient length	Relative weight	Theme category	Recipient category
9	2	theme>>recipient	Sentential	Phrasal

4.3.4.1 Descriptive statistics The following table shows the distribution of the independent variable "relative length of theme and recipient" in the sample:

Table 4.8: Relative length of theme and recipient in the sample

Relative length	Count	Percent
rec>>th	221	21.92
rec>th	291	28.87
th=rec	112	11.11
th>>rec	196	19.44
th>rec	188	18.65
N = 1008		

The following table shows the distribution of the independent variable "category of theme" in the sample:

Table 4.9: Category of theme in the sample

Category of th.	Count	Percent
Phrasal	902	89.48
Sentential	106	10.52
N = 1008		

The following table shows the distribution of the independent variable "category of recipient" in the sample:

Table 4.10: Category of recipients in the sample

Category of rec.	Count	Percent
Phrasal	923	91.57
Sentential	85	8.43
N = 1008		

4.3.5 Lexical semantics

The occurrence of DCLD in Spanish ditransitive constructions has been analyzed in the literature as a phenomenon restricted by the lexical semantic properties of the predicates involved (Strozer, 1976; Romero Morales, 2008). This interpretation follows a tradition of analysis of the English dative alternation which associates the dative prepositional construction (*John gave a book to Mary)* with the notion of "change of location" and the double object construction (*John gave Mary a book)* with the notion of "change of possession" (Green, 1974; Oehrle, 1976; Pinker, 1984, 1989), a hypothesis that has been labeled by Bresnan et al. (2007) the "Meaning-to-Structure Mapping Hypothesis" (MSMH).

Regarding DCLD in Spanish, it has been claimed that it is restricted to contexts in which the recipient can be interpreted as becoming the possessor of the theme (Demonte, 1995; Romero Morales, 2008). However, in Chapter 3 we have shown examples of DCLD found on the internet in which it is difficult to interpret the recipient as a possessor (entities like *las artes*, 'the arts'; *la patria*, 'the homeland'; and *la danza*,'dance'; are not likely to be possessors) and it is also difficult to interpret the recipient as a possessed entity (*su vida*, 'his/her life'). Moreover, the whole situation cannot be described as an instance of transfer. The examples are repeated below:

(48) DCLD without possessive interpretation of the recipient:

a. *Hace 14 años partió un excepcional venezolano que* ***le entregó*** *su vida a las artes.*
'Fourteen years ago an exceptional Venezuelan that gave his life to the arts left.'
(http://elobservador.rctv.net/Noticias/VerNoticia.aspxNoticiaId=270779&Tipo=34)

b. ... *una persona que* ***le entregó*** *su vida a la patria...*
"... a person who dedicated his life to his homeland fulfilling his duty..."
(http://www.elpilon.com.co/inicio/le-dieron-ultimo-adios-a-policia-muerto-en-atentado- de-la-guerrilla-en-caqueta/)

c. *Mimí González* ***le entregó*** *su vida a la danza.*
'Mimí González gave her life to dance'
(http://www.revistabellasartes.com/videos/)

The examples suggest that the role of lexical semantics in the distribution of DCLD is less straightforward than has been assumed in the literature. The purpose of this study is to provide a more precise picture of the interaction of lexical semantics and DCLD.

In order to annotate the examples, we took as a reference the semantic classification of ditransitive verbs using four schemata proposed by Delbecque and Lamiroy (1996). These

schema are “material transfer”, “verbal and perceptual transfer”, “physical motion”, and “abstract motion”. In the notation used by Delbecque and Lamiroy, N_0, N_1, *and* N_2 represent the subject, the direct object (accusative), and the indirect object (dative), respectively.

4.3.5.1 Material transfer schema The meaning of this class constitutes the basic transfer schema:

N_0 *makes* N_1 *enter the domain of* N_2 (the subject makes the direct object enter the domain of the indirect object). Verbs that typically participate in this schema are: *dar* 'to give', *obsequiar* 'to present', *entregar* 'to hand', etc. The notion of “entering the domain of N_2” means that N_2 is in control of N_1 but does not necessarily own it. According to Delbecque and Lamiroy, the transfer scheme can be metonymical or metaphorically extended in order to allow for abstract entities as the object of transfer. The following are examples of material transfer situations from our sample with a concrete (*las jaulas con los pájaros*, 'the cages with the birds') and abstract object of transfer (*la ciudadanía*, 'citizenship') :

(49) a. Transfer of a concrete object:
Vendí la casa de mi padre, pero antes regalé ***las jaulas con los pájaros*** *a la vecina de al lado, la que siempre los cuidaba.*
'I sold my father's house, but before I gave as a present the cages with the birds to the neighbor next door, who always took care of them.'

b. Transfer of an abstract object:
Ninguno hablaba de antecedentes, dijo Peñaloza en relación a todos los informes que llegaron al juzgado y que fueron tenidos en cuenta para darle ***la ciudadanía*** *a Al Kassar.*
'None of them was about background, said Peñaloza in relation to all the reports that arrived to the tribunal and that were taken into account in order to give citizenship to Al Kassar.'

Other typical objects of transfer are money, extensions of land, and information. These types of object of transfer occupy an intermediate position in a continuum that goes from concrete to abstract objects of transfer. See the examples below:

(50) a. Transfer of money:
El presidente Carlos Menem sobrevoló ayer la zona sur de Corrientes, afectada por las inundaciones, y posteriormente entregó al interventor federal de esa provincia, Francisco Durañona y Vedia, ***un subsidio por 2.600.000 pesos****.*
'President Carlos Menem flew yesterday over the south of (the province of) Corrientes, affected by the floods, and then gave the federal comptroller of the province, Francisco Durañona y Vedia, a grant of 2,600,000 pesos.'

b. Transfer of an extension of land:
Cuando el Tratado dio al Reino de Portugal ***los territorios al este del río Uruguay,*** *los indios se sintieron traicionados.*
'When the Treaty gave the Kingdom of Portugal the territories to the east of the Uruguay River, the Indians felt betrayed.'

c. Transfer of information:
También tuve que regar las plantas y darle ***todas las indicaciones*** *a la vecina, ...*
'I also had to water the plants and give complete directions to the neighbor, ...'

4.3.5.2 Verbal and perceptual transfer schema This schema constitutes an extension of the transfer scheme to the verbal domain: *N_0 makes N_1 enter the perceptual domain of N_2* (the subject makes the direct object enter the perceptual domain of the indirect object). Verbs that typically participate in this schema are: *aclarar* 'to explain', *asegurar* 'to assure', *confesar* 'to confess', etc. In the case of the verbal and perceptual transfer schema, there is no transfer of a material or abstract object, but of propositional or perceptual content expressed by the theme,

usually an embedded noun clause. Notice that this schema is different from instances of "transfer of information", discussed in the previous section. The following is an example of verbal transfer from our sample:

(51) Verbal transfer:
Fue así como Clara le informó a su madre que se iba a vivir con Paul.
'That was how Clara told (lit.: inform) his mother that he was going to live with Paul.'

4.3.5.3 Physical motion schema This schema constitutes an extension of the transfer schema to the motion domain: *N_0 makes N_1 move so as to bring N_1 in the realm of N_2* (the subject makes the direct object move so as to bring it in the realm of the indirect object). Verbs that typically participate in this schema are: *acercar* 'to bring close(r)', *lanzar* 'to throw', *llevar* 'to bring', etc. The physical motion schema should be interpreted in locative terms. Therefore, in this schema the direct object physically moves towards the indirect object, but the indirect object does not become in control of the direct object. The following is an example from our sample:

(52) Physical motion:
Sin detenerse le arrojó al recepcionista la llave de su cuarto y le hizo un gesto de guardar silencio al muchacho ya que empezaba a decirle que una señorita lo esperaba en el lobby.
'Without stopping he threw his room key to the receptionist and made a silent gesture to the boy as he began to tell him that a lady was waiting for him in the lobby.'

4.3.5.4 Abstract motion schema In this schema, the relation between N_1 and N_2 is not explicitly stated, but implied: *N_0 makes N_1 suitable for entering the realm of N_2* (the subject makes the direct object suitable for entering the realm of the indirect object). Verbs that typically participate

in this schema are: *ofrecer* 'to offer', *conferir* 'to confer', *dedicar* 'to dedicate', etc. A predicate of abstract motion denotes a situation in which there is neither transfer nor physical motion. An abstract relation or association is established between the theme and the recipient. Notice that this schema is different from instances of "transfer of an abstract object" discussed above. In the following example from our sample, the relation is that the purpose of playing the boleros is to ingratiate the girlfriends of the men:

(53) Abstract motion:
La costumbre de la época era que los hombres les dedicaran boleros a sus novias en las serenatas.
'The custom of the time was that men dedicate boleros to their girlfriends in/during the serenade.'

The following example, also tagged as "abstract motion", means that some changes regarding the guitar took place in the XIX Century. Notice that the label "abstract motion" is being used in this case as a default label:

(54) Abstract motion:
Como en los demás instrumentos, el siglo XIX aportó a la guitarra sus innovaciones.
'As with other instruments, the nineteenth century brought innovations to the guitar.'

4.3.5.5 Multiple meanings for a single verb A verb can participate in different schemata depending on the context of the sentence, which means that a verb can have more than one sense. The actual sense of a verb in the context of a sentence is not always easy to determine. Below we will show some examples of contexts in which the same verb displays different senses.

The verb *aportar* ('to contribute, to provide, to bring') can have a physical motion or an abstract motion sense, as can be seen in the following examples:

(55) a. Physical motion:
El río Yeniséi aporta 623 km3 de agua al año al océano Glacial Ártico
'The Yenisei River provides 623 km3 of water annually to the Arctic Ocean'

b. Abstract motion:
El uruguayo José Enrique Rodó aportó nuevas dimensiones artísticas al ensayo con su obra Ariel (1900), ...
'The Uruguayan José Enrique Rodó brought new artistic dimensions to the essay with his work Ariel (1900), ...'

The verb *dar* ('to give') can have an abstract motion or a material transfer sense, as the following examples show:

(56) a. Material transfer:
Y hoy día entonces uno le pide al papá para darle el regalo a la mamá y a la mamá para darle al papá...
'And today then one asks the dad in order to give the gift to the mom and to the mom in order to give it to the dad...'

b. Abstract motion:
Todos esos autores consiguieron darle a la literatura canaria una altura que no había tenido hasta entonces.
'All these authors were able to give the Canarian literature a height that it had not had before.'

The verb *transferir* ('to transfer') can have a material transfer or a physical motion sense, as can be seen in the following examples:

(57) a. Material transfer:

En 1966, la Administración Central transfirió la propiedad del inmueble a la Diputación y el Ayuntamiento de Valencia.

'In 1966, the central government transferred ownership of the property to the province and the city of Valencia.'

b. Physical motion:

Hitler canceló la operación debido a que los estadounidenses y británicos habían arribado a Sicilia y era preciso transferir divisiones a esta zona.

'Hitler canceled the operation because the Americans and British had arrived in Sicily and it was necessary to transfer divisions to this area.'

As the previous examples show, the same predicate can be ascribed to different verbal senses in different contexts. The purpose of the annotation is to determine the sense of the predicate in context rather than the general meaning of the verb out of context.

4.3.5.6 Descriptive statistics The following table shows the distribution of lexical semantic types in the sample:

Table 4.11: Lexical semantics of verbs in the sample (sense in context)

Lexical semantics	Count	Percent
Material transfer	382	37.9
Verbal transfer	53	5.26
Physical motion	32	3.17
Abstract motion	541	53.67
N = 1008		

4.3.6 Extra-linguistic variables

The purpose of this dissertation is to study the linguistic factors that constrain the occurrence of DCLD and the order of theme and recipient. However, we are aware that these dependent variables can be certainly affected by extra-linguistic variables as well. In particular, we are interested in the possible effect of dialectal and stylistic variation. That is why we have included as independent variables "region" and "medium". The purpose of the inclusion these variables is to control whether or not they are playing a role rather than to use them as explanatory devices.

4.3.6.1 Region In order to assess the possible influence of dialectal variation, sentences will be annotated for region using three possible values: "Spain", "Americas", and "unknown". Although the existence of important degrees of dialectal variation with respect to dative clitic doubling has not been reported in the literature, Becerra Bascuñán (2006) claims that dative clitic doubling is more productive in Latin American than in Castillian Spanish.

4.3.6.2 Medium of production In order to assess the possible influence of stylistic variation, sentences will be annotated for medium of production. The examples in the corpus belong to four genres, one spoken ("oral") and three written ("fiction", "academic", and "news"). We will tag the medium of production using the features "spoken"and "written".

4.3.7 Summary of independent variables

a. Animacy of theme and recipient: human, organization, animate, inanimate.

b. Definiteness of theme and recipient: definite, specific indefinite, indefinite.

c. Givenness of theme and recipient: in focus, active, uniquely identifiable, referential, type identifiable.

d. Relative length of theme and recipient: recipient>>theme, recipient>theme, theme=recipient, theme>recipient, theme>>recipient.

e. Category of theme and recipient: Phrasal, Sentential

f. Lexical semantics of the verb: material transfer, verbal and conceptual transfer, physical motion, abstract motion.

g. Region: Americas, Spain, unknown.

h. Medium: spoken, written.

4.4 THE METHOD

The statistical significance of each independent variable as a predictor of the occurrence of DCLD and the order of theme and recipient can be tested separately using the Chi-square test. However, that methodology would miss the possible interaction among independent variables. As has been observed by Bresnan et al. (2007) in their study of the English dative alternation, independent variables such as animacy, definiteness, givennness, and grammatical weight are frequently intertwined. Definite NPs are more likely to convey given information than indefinite NPs. NPs conveying given information are more likely to be short, because their referents can be identified without much information. Therefore, we will resort to multivariate analysis (logistic regression), a statistical method that "assesses the individual relative contribution of each factor to the observed variation when all factors are considered simultaneously" (Walker,

2010: 38). In order to run the logistic regression we will use GoldVarbX (http://individual.utoronto.ca/tagliamonte/goldvarb.htm).

To determine which independent variables (factor groups) are statistically significant, GoldVarbX performs a step-up/step-down procedure. Before the step-up procedure, the program determines the *input*, which is the overall probability that the rule will apply (for instance, that DCLD will occur). Once the *input* is obtained, each step of the step-up procedure will add a factor group (independent variable) and determine whether or not there is an improvement in the statistical model. If there is an improvement, the factor group is kept in the model, otherwise it is not. The combination of factors that produces the highest improvement in the model is considered the best run of the step-up procedure. The step-down procedure begins considering all of the factors together and subtracting one by one. If there is an improvement in the model after subtracting a factor group, that factor group will be dropped from the model. The step-down procedure also produces a best run, which should include the same factor groups than the best run produced by the step-up procedure.

The method has been chosen in order to avoid the problem of correlated variables and their interaction. However, the results can be misleading under different circumstances. The reliability of the results depend on having a representative sample[35] and applying consistent annotation criteria. A possible way of testing the accuracy of the results is to compare them with similar corpus studies, already performed or to be performed in the future. In this respect, it is a good sign that our results are analogous to those of Nishida (2010) (see 4.5.1).

36 A representative sample should be random, which means that each member of the population has equal chances of being part of the sample, and have an appropriate size (Woods et al., 1986). In our case, the population is defined as the sentences contained in *Corpus del Español*. The way in which our sample was selected is described in 4.2.

4.4.1 Simplification of variables

Before running the logistic regression for DCLD and order of theme and recipient, we made some adjustments in the independent variables and their values. The independent variables which constitute hierarchies or scales (animacy, definiteness, and givenness of theme and recipient) were reduced to binary values. The values of animacy were simplified to "human" vs. "non-human", the values of definiteness to "definite/specific" vs. "indefinite", and the values of givenness to "activated" (in focus and activated) vs. "non-activated" (uniquely identifiable, referential, and type identifiable). This simplification was done because of practical reasons: there were not enough tokens in some cells to perform statistics with GoldVarbX. The independent variable "animacy of theme" was excluded from the logistic regression altogether because its distribution, nearly categorically "inanimate", was not suitable to perform statistics.

4.5 THE RESULTS

The purpose of this section is to present and discuss the results of the quantitative study of ditransitive sentences. This section is divided in three sub-sections: 4.5.1 presents the overall results of the studies, 4.5.2 presents the results of the study on DCLD, 4.5.3 presents the result of the study on the order of theme and recipient.

4.5.1 Overall results

The overall results of our quantitative study of ditransitive sentences confirm the optionality of both phenomena under study, DCLD and the order of theme and recipient. Crucially, the overall results also confirm that DCLD and the order of theme and recipient are not associated.

DCLD occurs in only 19.05% of the tokens (N = 192) vs. 80.95% (N = 816) of the tokens without doubling, as shown in the table below:

Table 4.12: Occurrence of DCLD in the sample

DCLD	Count	Percent
Doubling	192	19.05
No doubling	816	80.95
N = 1008		

Our results are analogous to those of the Nishida (2010) corpus study of Spanish ditransitive sentences. Nishida studied the influence of definiteness and animacy of the recipient and lexical semantics of the verb on the occurrence of DCLD and the influence of information structure status and grammatical weight on the order of theme and recipient. Her conclusions are that definiteness favors DCLD and that relative weight is the main predictor of the order of theme and recipient. Nishida considered a database of 943 tokens gathered from *Corpus de referencia del español actua*l (CREA, Real Academia Española, http://www.rae.es). Nishida's database include sentences with the verbs *entregar* ('hand'), *ofrecer* ('offer'), *dar* ('give'), *otorgar* ('grant'), and *enviar* ('send'), only in the past tense. Nishida does not explicitly state why she used only examples in the past tense, but we assumed that it was in order to restrict the number of

factors potentially constraining the dependent variables. Her statistics show: 29% of tokens with DCLD and 71% of tokens without it. Our results, as well as Nishida's, suggest that DCLD is much less widespread than has been assumed in the literature.[36] The notion that DCLD is predominant or required in Spanish ditransitive predicates has been expressed, among others, by Givón (1976): 'Let us consider first object agreement in Spanish. For the neutral pattern with a verb taking both a dative and accusative object, such as "give", dative agreement is obligatory and accusative agreement unacceptable' (Givón, 1976: 161).

Our results also support the optionality of the other phenomenon under study, the order of theme and recipient. The order theme/recipient occurs in a 66.67% (N = 672) of the cases and the order recipient/theme takes place in the remaining 33.33% (N = 336) of the data.

Table 4.13 : Order of theme and recipient order in the sample

Order	Count	Percent
rec/th	336	33.33
th/rec	672	66.67
N = 1008		

As in the case of DCLD, our results are analogous to those of Nishida (2010). Her results for order of theme and recipient are 75.08% for theme/recipient (N = 708) and 24.92% for recipient/theme (N = 235).

4.5.1.1 DCLD and word order The first problem to solve in order to formulate the design of our study is whether or not DCLD and order of theme and recipient are associated variables or not. A

37 See also Company Company (2006).

configurational analysis of DCLD, like the one in Demonte (1995), would predict that an association exists because the occurrence of DCLD is explicitely linked to word order. Our analysis, in which DCLD and word order are independent phenomena, predicts that the two variables are not associated. This prediction is supported by the results. The cross-tabulation of occurrence of DCLD and order of theme and recipient indicates that there is no statistically significant relationship between word order and doubling, although there is a slight tendency towards the recipient/theme order when doubling occurs. The order recipient/theme occurs in 35.94% of the tokens with DCLD vs.32.72% of the tokens without it. The relative higher frequency of recipient/theme order in the tokens with DCLD is not statistically significant according to the result of the Chi-Square test (p = 0.395). Moreover, notice that the order recipient/theme is less frequent not only without, but also with DCLD. The results displayed on Table 4.14 indicate that there is no relationship between DCLD and order of theme and recipient.

Table 4.14: Occurrence of DCLD and order of theme and recipient (cross-tabulation)

Rows: CLITIC Columns: ORDER

	rec/th	th/rec	All
doubling	69	123	192
	35.94	64.06	100.00
no doubling	267	549	816
	32.72	67.28	100.00
All	336	672	1008
	33.33	66.67	100.00

Cell Contents: Count
% of Row

Pearson Chi-Square = 0.724, DF = 1, p = 0.395

Given these preliminary results, we will consider DCLD and order of theme and recipient as two variables that are not associated and treat them as separate dependent variables subject to separate studies. The study of the dependent variable DCLD has as its goal the determination of the factors that intervene in the overt coding of the IO with dative case. The study of the dependent variable order of theme and recipient has as its goal the study of the factors that constrain the order of constituents in ditransitive constructions.

4.5.2 Study #1: occurrence of DCLD

The logistic regression analysis performed using GoldVarbX found four factor groups (independent variables) as statistically significant predictors of DCLD. These variables were, in order of rank: region, medium, animacy of recipient, and givenness of recipient. The ranking of the variables as predictors is based on the range of the weight of the factors within the factor group. The weight of a factor indicates the probability of application (occurrence) of the dependent variable. The application value in this case is the occurrence of DCLD, which means that if DCLD occurs the rule is considered to apply. A weight higher than 0.5 indicates that the individual factor favors the application, while a weight lower than 0.5 disfavors it. The range of a factor group is obtained by subtracting the largest factor weight from the smallest factor weight in each factor group (Walker, 2010:41). The third column in the table indicates the percentage of applications (DCLD) in the factor (for instance, 25.10% of the [+human] recipients undergo DCLD). The fourth column indicates the number of tokens in the factor (for instance, 501 recipients are [+human]). The table also indicates the total N (= 1008) and the input value (= 0.143), which is the value of the corrected mean (overall probability that the rule will apply,

which in this case is the occurrence of DCLD, the application value). See table 4.15 below, which shows the results of the logistic regression.

Table 4.15: Occurrence of DCLD (logistic regression)

Total N: 1008, Input: 0.143, Application value: occurrence of DCLD

Factor Group	Factor Weight	% of rule application within factor	N within factor
Region			
Americas	0.68	29.30	441
Spain	0.27	5.80	429
unknown	0.67	27.50	138
Range:	*41*		
Medium			
written	0.47	17.00	914
spoken	0.74	39.40	94
Range:	*27*		
Animacy of rec.			
human	0.59	25.10	501
non-human	0.41	13.00	507
Range:	*18*		
givenness of rec.			
activated	0.57	24.70	385
non-activated	0.45	15.60	623
Range:	*12*		

Factor groups not selected: order of theme and recipient, relative length of theme and recipient, definiteness of theme, givenness of theme, category of theme, definiteness of recipient, category of recipient, lexical semantics.

The results show that the predictors of DCLD in ditransitive sentences belong to two classes: extra-linguistic factors (region and medium) and features of the recipient that contribute to its discourse salience (animacy, givenness). The extra-linguistic factors are the ones that have the higher weight, a fact that suggests that dialectal and stylistic variation play an important role in the occurrence of DCLD.

The results show that the occurrence of DCLD is more likely in the spoken language. It also shows that DCLD is more likely in the Americas than in Spain. Regarding linguistic factors, the table shows that the occurrence of DCLD is more likely if the recipient is [+ human] and [+activated]. Notice that no feature of the theme is a significant predictor of DCLD, as opposed to what happens in the case of English dative-shift (Bresnan et al., 2007). This discussion strengthens the claim that in the English dative alternation there is a competition between the theme and the recipient for the Primary Object position, a conflict that does not arise in Spanish. DCLD is a mechanism of overt coding of the recipient and it seems to be influenced by the semantically and pragmatically relevant properties of the recipient, not the theme.

4.5.2.1 Region The highest ranked predictor of DCLD is the region (range = 41). 29.25% of the examples from the Americas have DCLD but only 5.83% of the examples from Spain do. This distribution is significant according to the Chi-square test (Pearson Chi-Square = 84.853, DF = 2, $p \leq 0.001$).

Table 4.16: Region and occurrence of DCLD (cross-tabulation)

Rows: REGION Columns: CLITIC

	doubling	no doubling	All
Americas	129	312	441
	29.25	70.75	100.00
Spain	25	404	429
	5.83	94.17	100.00
unknown	38	100	138
	27.54	72.46	100.00
All	192	816	1008
	19.05	80.95	100.00

Cell Contents: Count
% of Row

Pearson Chi-Square = 84.853, DF = 2, p ≤ 0.001

4.5.2.2 Medium The medium was also found significant by the logistic regression. 39.36% of the spoken language examples undergo DCLD. The percentage goes down in the written language to 16.96%. According to the Chi-square test, the distribution of this independent variable is significant (Pearson Chi-Square = 27.744, DF = 1, $p \leq 0.001$).

Table 4.17: Medium and occurrence of DCLD (cross-tabulation)

Rows: MEDIUM	Columns: CLITIC		
	doubling	no doubling	All
spoken	37	57	94
	39.36	60.64	100.00
written	155	759	914
	16.96	83.04	100.00
All	192	816	1008
	19.05	80.95	100.00

Cell Contents: Count
% of Row

Pearson Chi-Square = 27.744, DF = 1, p ≤ 0.001

4.5.2.3 Animacy of recipient The feature [+human] is singled out by the logistic regression as a predictor of DCLD, which means that a [+human] recipient is more likely to be doubled. Table 4.18 shows that DCLD is more frequent when the recipient argument is [+human]: 25.15% of the human recipients undergo DCLD, but only 13.02% of the non-human recipients do. According to the results of the Chi-Square test, the association of DCLD and animacy of the recipient as significant (Pearson Chi-Square = 24.053, DF = 1, $p \leq 0.001$).

Table 4.18: Animacy of recipient and occurrence of DCLD (cross-tabulation)

Rows: AN-REC Columns: CLITIC

	dat	none	All
human	126	375	501
	25.15	74.85	100.00
non human	66	441	507
	13.02	86.98	100.00
All	192	816	1008
	19.05	80.95	100.00

Cell Contents: Count
% of Row

Pearson Chi-Square = 24.053, DF = 1, p ≤ 0.001

4.5.2.4 Givenness of recipient The feature [+activated] was also singled out by the logistic regression as a predictor of DCLD. Table 4.19 shows that DCLD is more frequent when the recipient argument is [+activated]: 24.68% of the activated recipients undergo DCLD. This proportion goes down to 15.57% if the recipient is not activated. According to the results of the Chi-Square test, the association of DCLD and animacy of the recipient as significant (Pearson Chi-Square = 12.795, DF = 1, $p \leq 0.001$).

Table 4.19: Givenness of recipient and occurrence of DCLD (cross-tabulation)

Rows: GIVE-REC Columns: CLITIC

	dat	none	All
activated	95	290	385
	24.68	75.32	100.00
non activated	97	526	623
	15.57	84.43	100.00
All	192	816	1008
	19.05	80.95	100.00

Cell Contents: Count
% of Row

Pearson Chi-Square = 12.795, DF = 1, p ≤ 0.001

4.5.2.5 Variables not selected by the logistic regression The logistic regression did not select three types of variables as predictors of DCLD: order of theme and recipient and measures of complexity of theme and recipient (relative length, grammatical category of theme and recipient), pragmatically relevant features of the theme (definiteness, givenness), and lexical semantics of the verb.

These results are expected given our hypotheses, listed below:

a. in the contexts under study DCLD is an optional phenomenon constrained by multiple factors (dialectal and stylistic factors, pragmatically relevant properties of the recipient).

b. DCLD is independent of word order: measures of complexity, which are relevant for word order, do not influence DCLD. The results of the Chi-square test also show that there is no statistically significant association between the measures of complexity and DCLD. DCLD is quite stable independently of the relative length of theme and recipient (rec>>th: 16.29%, rec>th:

19.24%, th=rec: 28.57%, th>>rec: 16.33%, th>>rec: 19.15%). The only value of relative length that shows a noticeable increment of DCLD is th=rec. This effect may be the result of the fact that if the theme and the recipient have the same length, it means that they are both short, and shorter recipients are more likely given information ("activated"). Nevertheless, the higher percentage of DCLD when the theme and the recipient have the same length is not statistically significant, as shown by the result of the Chi-square test (Pearson Chi-Square = 8.628, DF = 4, p = 0.071) in addition to the fact that the variable was not selected by the logistic regression. See table below:

Table 4.20: Relative length of theme and recipient and occurrence of DCLD (cross-tabulation)

Rows: LENGTH Columns: CLITIC

	doubling	no doubling	All
rec>>th	36	185	221
	16.29	83.71	100.00
rec>th	56	235	291
	19.24	80.76	100.00
th=rec	32	80	112
	28.57	71.43	100.00
th>>rec	32	164	196
	16.33	83.67	100.00
th>rec	36	152	188
	19.15	80.85	100.00
All	192	816	1008
	19.05	80.95	100.00

Cell Contents: Count
% of Row

Pearson Chi-Square = 8.628, DF = 4, p = 0.071

There is a higher percentage of DCLD when the theme is sentential (an embedded clause or a complex-NP) (25.47%) than when it is phrasal (NP) (18.29%), but this distribution is not statistically significant according to the Chi-square test (Pearson Chi-Square = 3.170, DF = 1, p = 0.075), in addition to the fact that the variable was not selected by the logistic regression. See table below:

Table 4.21: Category of theme and occurrence of DCLD (cross-tabulation)

Rows: CAT-TH Columns: CLITIC

	doubling	no doubling	All
phrasal	165	737	902
	18.29	81.71	100.00
sentential	27	79	106
	25.47	74.53	100.00
All	192	816	1008
	19.05	80.95	100.00

Cell Contents: Count
% of Row

Pearson Chi-Square = 3.170, DF = 1, p = 0.075

There is a higher percentage of DCLD when the recipient is phrasal (19.39%) than when it is sentential (15.29%), but this distribution is not statistically significant according to the Chi-square test (Pearson Chi-Square = 0.848, DF = 1, p = 0.357), in addition to the fact that the variable was not selected by the logistic regression. See table below:

Table 4.22: Category of recipient and occurrence of DCLD (cross-tabulation)

Rows: CAT-REC Columns: CLITIC

	doubling	no doubling	All
phrasal	179	744	923
	19.39	80.61	100.00
sentential	13	72	85
	15.29	84.71	100.00
All	192	816	1008
	19.05	80.95	100.00

Cell Contents: Count
% of Row

Pearson Chi-Square = 0.848, DF = 1, p = 0.357

c. DCLD is only involved in the overt coding of the recipient: DCLD is not involved in the overt coding of the theme. Since there is no competition of theme and recipient for the PO slot, the pragmatically relevant properties of the theme are not predictors of DCLD. The definiteness of the theme is not a predictor of DCLD. If the theme is definite or specific, there is a 20.15% of DCLD vs. a 17.04% if it is indefinite. This distribution is not statistically significant (Pearson Chi-Square = 1.452, DF = 1, p = 0.228). See table below:

Table 4.23: Definiteness of theme and occurrence of DCLD (cross-tabulation)

Rows: definiteness of theme Columns: CLITIC

	doubling	no doubling	All
-definite/specific	131	519	650
	20.15	79.85	100.00
+definite/specific	61	297	358
	17.04	82.96	100.00
All	192	816	1008
	19.05	80.95	100.00

Cell Contents: Count
% of Row

Pearson Chi-Square = 1.452, DF = 1, p = 0.228

The cognitive status of the theme is not a predictor of DCLD either. If the theme is [+activated], there is 17.00% of DCLD vs. 19.27% if it is [-activated]. Again, this distribution is not statistically significant (Pearson Chi-Square = 0.302, DF = 1, p = 0.583). See table below:

Table 4.24: Givenness of theme and occurrence of DCLD (cross-tabulation)

Rows: GIVE-TH Columns: CLITIC

	doubling	no doubling	All
-activated	175	733	908
	19.27	80.73	100.00
+activated	17	83	100
	17.00	83.00	100.00
All	192	816	1008
	19.05	80.95	100.00

Cell Contents: Count
% of Row

Pearson Chi-Square = 0.302, DF = 1, p = 0.583

4.5.2.6 Lexical semantics The logistic regression did not find the lexical semantics of the verb as a predictor of DCLD. The data show that DCLD is more common with verbal transfer (32.08%) than with abstract motion (17.74%), material transfer (19.37%), and physical motion verbs (15.63%). In spite of that tendency, the distribution is not significant according to the Chi-square test (Pearson Chi-Square = 6.698, DF = 3, p = 0.082). See table below:

Table 4.25: Lexical semantics and occurrence of DCLD (cross-tabulation)

Rows: SENSE Columns: CLITIC

	doubling	no doubling	All
abstract motion	96	445	541
	17.74	82.26	100.00
material transfer	74	308	382
	19.37	80.63	100.00
physical motion	5	27	32
	15.63	84.38	100.00
verbal transfer	17	36	53
	32.08	67.92	100.00
All	192	816	1008
	19.05	80.95	100.00

Cell Contents: Count
% of Row

Pearson Chi-Square = 6.698, DF = 3, p = 0.082

These result seems to contradict the widely held assumption that lexical semantics is, at least, one of the factors that constrains DCLD. However, there are some caveats about this conclusion:

a. Spatial goals, which cannot be interpreted as recipients, cannot undergo DCLD:

(58) **María le tiró el paquete al tacho de basura.*
María DAT.3SG threw the package to the can of garbage
'María threw away the package.' (lit.: 'María threw the package to the trash can.')

Therefore, this type of examples were not included in the sample. The results of the quantitative study do not contradict this accepted fact about DCLD. However, the results seem to show that DCLD is not significantly more likely when the situation described by the sentence can be considered an instance of transfer rather than when a ditransitive verb has an abstract sense. See the following examples:

(59) a. Material transfer:
*Uno de ellos **le entregó** a Vittorio un sobre abultado y le recriminó la falta de profesionalidad,...*
'One of them gave Vittorio a thick envelope and berated his lack of professionalism, ...'

b. Abstract motion:
Sí, tengo una carga académica reducida para que pueda dedicarle más tiempo a la investigación.
'Yes, I have a reduced teaching load so I can spend more time on research.'

b. The result depends to a large extent on the coding scheme and the coding itself. A previous run of the logistic regression, in which lexical semantics was coded with a different scheme, selected lexical semantics as a predictor of DCLD. On that occasion, lexical semantics was coded not taking into account the sense of the verb in context, but the most typical sense of the verb instead. For instance, all instances of *dar* ('to give') were coded as instances of material transfer. The following examples show two uses of the verb *dar*, one that was considered this time an instance of material transfer and one that was considered this time an instance of abstract motion. In the previous run, the two of them were tagged as "material transfer":

(60) a. Material transfer:

Y hoy día entonces uno le pide al papá para darle el regalo a la mamá y a la mamá para darle al papá...

'And today then one asks the dad in order to give the gift to the mom and to the mom in order to give it to the dad...'

b. Abstract motion:

Todos esos autores consiguieron darle a la literatura canaria una altura que no había tenido hasta entonces.

'All these authors were able to give the Canarian literature a height that it had not had before.'

c. The coding scheme that we used defines "material transfer" in a broad sense (transfer of concrete objects, money, land, information, and abstract objects). However, the distribution of DCLD varies considerably depending on the nature of the object of transfer. DCLD occurs in a 27.91% of the cases if the object of transfer is concrete, 23.81% if it is money, 20.55% if the object is abstract, 14.79% if it is information, and 0% if it is an extension of land. See the table below:

Table 4.26: Object type and occurrence of DCLD (cross-tabulation)
(transfer verbs only)

Rows: object Columns: clitic

	doubling	no doubling	All
abstract	15	58	73
	20.55	79.45	100.00
concrete	24	62	86
	27.91	72.09	100.00
information	25	144	169
	14.79	85.21	100.00
land	0	12	12
	0.00	100.00	100.00
money	10	32	42
	23.81	76.19	100.00
All	74	308	382
	19.37	80.63	100.00

Cell Contents: Count
% of Row

But this fact does not seem to have affected the results of the logistic regression. If the transfer class is split into two classes, one including transfer of concrete objects and money, and another one including the other types of objects, the result of the logistic regression does not change. Therefore, we kept the coding scheme with a broad definition of material transfer.

4.5.3 Study #2: order of theme and recipient

The logistic regression analysis performed using GoldVarbX found five factor groups (independent variables) as statistically significant predictors of the order of theme and recipient. These variables were, in order of rank, relative length of theme and recipient, lexical semantics, category of theme, category of recipient, and animacy of recipient.

Table 4.27: Order of theme and recipient (logistic regression)

Total N: 1008, Input: 0.263, Application value: order recipient/theme

Factor Group	Factor Weight	% of rule application within factor	N within factor
Relative length of th. and rec.			
theme>>recipient	0.97	93.90	196
theme>recipient	0.75	51.10	188
theme=recipient	0.49	25.90	112
recipient>theme	0.20	7.60	291
recipient>>theme	0.10	2.30	221
Range:		*87*	
Lexical semantics			
verbal transfer	0.94	96.20	53
material transfer	0.44	31.40	382
abstract motion	0.48	29.80	541
physical motion	0.30	12.50	32
Range		*64*	
Category of th.			
Phrasal	0.45	25.80	902
Sentential	0.86	97.20	106
Range:		*41*	
Category of rec.			
Phrasal	0.54	35.90	923
Sentential	0.14	5.90	85
Range:		*40*	
Animacy of rec.			
human	0.57	40.10	501
non-human	0.43	26.60	507
Range:		*14*	

Factor groups not selected: doubling, definiteness of theme, givenness of theme, definiteness of recipient, givenness of recipient, region, medium.

The results show that the main predictors of the order of theme and recipient are the different measures of grammatical weight, which were not significant predictors of DCLD, in addition to lexical semantics. However, the lexical semantic class which favors the recipient/theme order is verbal transfer, which is characterized by having an almost categorical recipient/theme order, probably as the result of the fact that the theme is usually an embedded clause, which is heavier and more complex than most NPs. Animacy of recipient also has an effect, although weaker than that of the complexity measures. These results provide further support for our claim that DCLD and order of theme and recipient are not associated and are subject to different sets of constraints.

4.5.3.1 Relative length of theme and recipient The relative weight, measured in number of words, of the theme and recipient is the highest ranked predictor of the order of theme and recipient. Notice that relative length is not only significant when the difference in length is four words or more, but it is also significant when the difference is between one to three words. The data also show a bias towards the order theme/recipient when the length of the theme and recipient are equal or when the theme is up to three words longer. The distribution of the variable is found significant by the Chi-square test (Pearson Chi-Square = 535.686, DF = 4, $p \leq 0.001$). See table below:

Table 4.28: Relative length and order of theme and recipient (cross-tabulation)

Rows: LENGTH Columns: ORDER

	rec/th	th/rec	All
rec>>th	5	216	221
	2.26	97.74	100.00
rec>th	22	269	291
	7.56	92.44	100.00
th=rec	29	83	112
	25.89	74.11	100.00
th>>rec	184	12	196
	93.88	6.12	100.00
th>rec	96	92	188
	51.06	48.94	100.00
All	336	672	1008
	33.33	66.67	100.00

Cell Contents: Count
% of Row

Pearson Chi-Square = 535.686, DF = 4, p ≤ 0.001

4.5.3.2 Lexical semantics The second ranked factor group in the logistic regression is lexical semantics. This result seems to be influenced by the fact that with verbs of verbal transfer the order recipient/theme is almost categorical (96.23%). The theme of verbal transfer verbs is usually an embedded clause, which is more complex than a NP, favoring the recipient/theme order. See table below:

Table 4.29: Lexical semantics and order of theme and recipient (cross-tabulation)

Rows: SENSE Columns: ORDER

	rec/th	th/rec	All
abstract motion	161	380	541
	29.76	70.24	100.00
material transfer	120	262	382
	31.41	68.59	100.00
physical motion	4	28	32
	12.50	87.50	100.00
verbal transfer	51	2	53
	96.23	3.77	100.00
All	336	672	1008
	33.33	66.67	100.00

Cell Contents: Count
% of Row

Pearson Chi-Square = 104.332, DF = 3, p ≤ 0.001

Notice in the tables below that verbal transfer verbs have a high percentage of theme longer than recipient (83.02%) and a high percentage of sentential themes (81.13%).

Table 4.30: Lexical semantics and relative length of theme and recipient (cross-tabulation)

Rows: SENSE Columns: LENGTH

	rec>>th	rec>th	th=rec	th>>rec	th>rec	All
abstract motion	137	156	50	89	109	541
	25.32	28.84	9.24	16.45	20.15	100.00
material transfer	78	116	54	74	60	382
	20.42	30.37	14.14	19.37	15.71	100.00
physical motion	5	16	3	1	7	32
	15.63	50.00	9.38	3.13	21.88	100.00
verbal transfer	1	3	5	32	12	53
	1.89	5.66	9.43	60.38	22.64	100.00
All	221	291	112	196	188	1008
	21.92	28.87	11.11	19.44	18.65	100.00

Cell Contents: Count
% of Row

Table 4.31: Lexical semantics and category of theme (cross-tabulation)

Rows: SENSE Columns: CAT-TH

	phrasal	sentential	All
abstract motion	510	31	541
	94.27	5.73	100.00
material transfer	350	32	382
	91.62	8.38	100.00
physical motion	32	0	32
	100.00	0.00	100.00
verbal transfer	10	43	53
	18.87	81.13	100.00
All	902	106	1008
	89.48	10.52	100.00

Cell Contents: Count
% of Row

4.5.3.3 Category of recipient and theme The grammatical category of the recipient and the theme are also found significant by the logistic regression. Sentential complements of ditransitive verbs and Complex-NPs are both longer on average and structurally more complex than phrasal constituents. 97.17% of the sentences with a sentential theme have recipient/theme order, but that percentage goes down to 25.83% if the theme is a NP. Only 5.88% of the sentences with a sentential recipient have recipient/theme order, but that percentage goes up to 35.86% if the recipient is phrasal. See tables below:

Table 4.32: Category of theme and order of theme and recipient (cross-tabulation)

Rows: CAT-TH Columns: ORDER

	rec/th	th/rec	All
phrasal	233	669	902
	25.83	74.17	100.00
sentential	103	3	106
	97.17	2.83	100.00
All	336	672	1008
	33.33	66.67	100.00

Cell Contents: Count
% of Row

Pearson Chi-Square = 217.225, DF = 1, p ≤ 0.001

Table 4.33: Category of recipient and order of theme and recipient (cross-tabulation)

Rows: CAT-REC Columns: ORDER

	rec/th	th/rec	All
phrasal	331	592	923
	35.86	64.14	100.00
sentential	5	80	85
	5.88	94.12	100.00
All	336	672	1008
	33.33	66.67	100.00

Cell Contents: Count
% of Row

Pearson Chi-Square = 31.478, DF = 1, p ≤ 0.001

4.5.3.4 Animacy of recipient In addition to grammatical complexity, animacy of the recipient is also found significant by the logistic regression. A human recipient is more likely to precede the theme than a recipient lower in the animacy scale. 40.12% of the sentences with a human recipient have recipient/theme order, but this percentage goes down to 26.33% if the recipient is not-human. See table below:

Table 4.34: Animacy of recipient and order of theme and recipient (cross-tabulation)

Rows: AN-REC Columns: ORDER

	rec/th	th/rec	All
human	201	300	501
	40.12	59.88	100.00
non-human	135	372	507
	26.63	73.37	100.00
All	336	672	1008
	33.33	66.67	100.00

Cell Contents: Count
% of Row

Pearson Chi-Square = 20.644, DF = 1, p ≤ 0.001

4.5.3.5 Variables not selected by the logistic regression The logistic regression did not select as predictors of the order of theme and recipient the occurrence of DCLD, definiteness of theme and recipient, givenness of theme and recipient, region, and medium. These results are compatible with our hypothesis that the occurrence of DCLD and the order of theme and recipient are dissociated phenomena that are regulated by different sets of constraints. The results suggest that the order of theme and recipient is regulated mainly by grammatical complexity, rather than pragmatic salience. The only measure of pragmatic salience which is significant is animacy of recipient.

Our initial assumption about the order of theme and recipient was that it is regulated by functional rather than formal factors. However, within the functionalist literature there is a debate about which factors constitute the main constraint on word order: pragmatic factors (Firbas, 1964, 1966; Givón, 1983, 1988) or grammatical complexity (Hawkins, 1994, 2004). We consider this an empirical issue. In the case of the order of theme and recipient in Spanish the

results of this study suggest that grammatical complexity is the main factor constraining word order.

Definiteness of theme and recipient The following tables show the cross-tabulation of definiteness of theme and definiteness of recipient and order of theme and recipient. The Chi-square test shows that neither variable is significant for order of theme and recipient. See tables below:

Table 4.35: Definiteness of theme and order of theme and recipient (cross-tabulation)

Rows: DEF-TH Columns: ORDER

	rec/th	th/rec	All
-definite/specific	210	440	650
	32.31	67.69	100.00
+definite/specific	126	232	358
	35.20	64.80	100.00
All	336	672	1008
	33.33	66.67	100.00

Cell Contents: Count
% of Row

Pearson Chi-Square = 0.866, DF = 1, p = 0.352

Table 4.36: Definiteness of recipient and order of theme and recipient (cross-tabulation)

Rows: DEF-REC Columns: ORDER

	rec/th	th/rec	All
definite/specific	314	634	948
	33.12	66.88	100.00
indefinite	22	38	60
	36.67	63.33	100.00
All	336	672	1008
	33.33	66.67	100.00

Cell Contents: Count
% of Row

Pearson Chi-Square = 0.319, DF = 1, p = 0.572

Givenness of theme and recipient The following tables show the cross-tabulation of givenness of theme and givenness of recipient and order of theme and recipient. The Chi-square test shows that both variables are significant for order of theme and recipient if they are considered individually. If the theme is [+activated], the percentage of recipient/theme order is only 16%. This percentage goes up to 35.24% if the theme is [-activated] (Pearson Chi-Square = 15.009, DF = 1, $p \leq 0.001$). See table below:

Table 4.37: Givenness of theme and order of theme and recipient (cross-tabulation)

Rows: GIVE-TH Columns: ORDER

	rec/th	th/rec	All
-activated	320	588	908
	35.24	64.76	100.00
+activated	16	84	100
	16.00	84.00	100.00
All	336	672	1008
	33.33	66.67	100.00

Cell Contents: Count
% of Row

Pearson Chi-Square = 15.009, DF = 1, p ≤ 0.001

If the recipient is [+activated], the percentage of recipient/theme order is only 37.40%. This percentage goes down to 30.82% if the recipient is [-activated] (Pearson Chi-Square = 4.642, DF = 1, p = 0.031). See table below:

Table 4.38: Givenness of recipient and order of theme and recipient (cross-tabulation)

Rows: GIVE-REC Columns: ORDER

	rec/th	th/rec	All
-activated	192	431	623
	30.82	69.18	100.00
+activated	144	241	385
	37.40	62.60	100.00
All	336	672	1008
	33.33	66.67	100.00

Cell Contents: Count
% of Row

Pearson Chi-Square = 4.642, DF = 1, p = 0.031

The fact that givenness of theme and recipient are significant according to the Chi-square test but not selected as a predictor of order of theme and recipient by the logistic regression may be an indication that there is an interaction of givenness with other properties, such as grammatical complexity. In general, given referents are expressed in discourse by shorter constituents. The fact that givenness is not selected by the logistic regression suggests that grammatical complexity, instead of givenness, is the factor that carries explanatory value.

Region and medium Extra-linguistic factors, such as region and medium, are not selected by the logistic regression as predictors of the order of theme and recipient. The percentage of recipient/theme order is very stable: 32.65% in the Americas, 34.27% in Spain (unknown: 32.61%). This distribution is not significant according to the Chi-square test (Pearson Chi-Square = 0.292, DF = 2, p = 0.864). This result is expected under the assumption that the order of theme and recipient is constrained by a psycholinguistic principle as Hawkins (1994, 2004) claims (in head initial languages, increasing order of complexity facilitates processing). See table below:

Table 4.39: Region and order of theme and recipient (cross-tabulation)

Rows: REGION	Columns: ORDER		
	rec/th	th/rec	All
Americas	144	297	441
	32.65	67.35	100.00
Spain	147	282	429
	34.27	65.73	100.00
unknown	45	93	138
	32.61	67.39	100.00
All	336	672	1008
	33.33	66.67	100.00

Cell Contents: Count
% of Row

Pearson Chi-Square = 0.292, DF = 2, p = 0.864

Regarding medium, there is a higher percentage of recipient/theme order in written (34.14%) than in spoken language (25.53%). This distribution, however, is not significant according to the Chi-square test (Pearson Chi-Square = 2.839, DF = 1, p = 0.092). See table below:

Table 4.40: Medium and order of theme and recipient (cross-tabulation)

Rows: MEDIUM	Columns: ORDER		
	rec/th	th/rec	All
spoken	24	70	94
	25.53	74.47	100.00
written	312	602	914
	34.14	65.86	100.00
All	336	672	1008
	33.33	66.67	100.00

Cell Contents: Count
% of Row

Pearson Chi-Square = 2.839, DF = 1, p = 0.092

Notice that the lack of significance of extra-linguistic variables for the order of theme and recipient contrasts with its significance for the occurrence of DCLD. This contrast underscores the fact that the order of theme and recipient and the occurrence of DCLD are constrained by different sets of factors.

4.6 CONCLUSIONS

The preliminary results of our corpus study of ditransitive constructions ratifies the optionality of both the occurrence of DCLD and the relative order of the theme and recipient, as well as the lack of association of these variables, which we proceed to study as dependent variables in two different studies. This result confirms our secondary hypothesis #1 (the occurrence of DCLD and order of theme and recipient are dissociated phenomena).

The results of study #1, on DCLD, show that the phenomenon is conditioned by dialectal and stylistic factors (region, medium) as well as properties relative to the pragmatic saliency of the recipient (animacy, givenness). No independent variable, or set of independent variables, triggers the occurrence of DCLD nearly a hundred percent of the time. This result confirms our main hypothesis (in the contexts under study DCLD should be described as a optional phenomenon constrained by multiple factors). This result shows that the phenomenon is optional and cannot be reduced to syntactic or lexical considerations.

The occurrence of DCLD is not influenced by properties of the theme, a result that confirms our secondary hypothesis #2 (the occurrence of DCLD is constrained by properties of the recipient but is dissociated from the properties of the theme). This result is expected under the assumption that in Spanish, as opposed to English, there is no competition of theme and recipient for acquiring core grammatical status. DCLD is only concerned with the overt grammatical marking of the recipient as dative and the features of the theme do not seem to affect it.

The results of study #2, on the order of theme and recipient, show that the phenomenon is mainly conditioned by grammatical complexity. This result suggests that the order of theme and recipient is conditioned by functional constraints. Therefore the word order facts do not provide evidence for a hierarchical structure of ditransitive VPs in Spanish along the lines of the VP-shells analysis proposed by Demonte (1995), among others.

Overall, the results confirm our main hypothesis that in the contexts under study DCLD is as an optional phenomenon constrained by multiple factors.

5.0 QUANTITATIVE STUDY OF POSSESSIVE CONSTRUCTIONS

5.1 HYPOTHESIS

As was discussed in Chapter 2, DCLD is optional in ditransitive constructions (recipient datives) and required or preferred in non-ditransitive constructions (experiencer, beneficiary and possessive datives). Also in Chapter 2, we arrived at the conclusion that the different behavior of DCLD in ditransitive and non-ditransitive constructions is due to the fact that they perform a different functional load. In non-ditransitive constructions, an argument alternation has taken place and the dative clitic is signaling the transitivity type of the verb. In ditransitive constructions, DCLD is an optional agreement marker which overtly marks pragmatically salient recipients. In ditransitive constructions, no argument alternation takes place; the alignment of thematic relations to grammatical functions is not affected.

In spite of this different functional load, we expect the factors that favor the occurrence of dative clitics in ditransitive and non-ditransitive constructions to overlap because pragmatic salience has been found to favor both the occurrence of object agreement (Comrie, 1989; Croft, 1988; Woolford, 1999) and the promotion of arguments to core grammatical functions (Lagendre, Raymond, and Smolensky, 1993). This is our secondary hypothesis #3: pragmatic salience favors the expression of an argument as dative when there is an alternation between dative and a less prominent grammatical function.

In order to test this hypothesis, we have conducted a quantitative study of possessive constructions in which the possessor can be realized as a dependent of the head of the DO (genitive) or a core argument of the verb (dative). This argument alternation is shown in the example below, in which *María* is the possessor:

(1) a. Genitive possessor (*Juan cortó el pelo **de** María* 'Juan cut María's hair')

Thematic roles: *preparar*<	agent,	patient	<possessor>>
	|	|	|
Grammatical functions:	SUBJ	[DO	[GENITIVE]]

b. Dative possessor (*Juan **le** cortó el pelo **a** María* 'Juan cut María's hair')

Thematic roles: *preparar*<	agent,	patient,	possessor>
	|	|	|
Grammatical functions:	SUBJ	DO	IO

This hypothesis would be falsified if the results of the study show that pragmatically relevant grammatical properties of the possessor (animacy, definiteness, givenness) are not statistically associated with its expression as IO (dative) instead of genitive. The hypothesis would be confirmed otherwise.

This chapter will describe the quantitative studies performed in order to check this hypothesis. We will present a description of the data, the dependent and independent variables, the statistical tests, and their results.

5.2 THE DATA

The data used for our quantitative study of possessive ditransitive constructions consist of 412 sentences. The sentences are representative of the argument alternation exemplified above, in which a possessor can be realized as either an argument of the verb (dative) or a dependent of the possessed noun (genitive). The following are examples extracted from the sample:

(2) Genitive possessor:

a. *Julián acaricia los cabellos* ***de Anudila.***
'Julian caresses Anudila's hair'

b. *Yo cierro la puerta* ***de mi dormitorio*** *y voy a la ventana.*
'I close the door of my bedroom and go to the window.'

c. *Caminé hacia el dormitorio sin mirar el rostro* ***de mi padre*** *y sin esperar su respuesta.*
'I walked towards the bedroom without looking at the face of my father and without waiting for his response.'

(3) Dative possessor:

a. *Se iba en vacaciones a su casa y se dedicaba a arreglar****le*** *el cercado* ***a su mamá****, ...*
'He went on vacation to his home and was engaged to fix the fence of her mom, …'

b. *El coñac* ***le*** *quemó la garganta* ***a Oliveira****, ...*
'The brandy burned the throat of Oliveira, …'

c. *Y cuando decía estas palabras* ***le*** *miraba el traje* ***a José Gabino.***
'And when he said these words he looked at the suit of José Gabino.'

The sample consists of 105 tokens with a dative possessor and 307 tokens with a genitive possessor, as shown in the table below:

Table 5.1: Case of possessor in the sample

Case	Count	Percent
dative	105	25.49
genitive	307	74.51
N = 412		

As in the study of ditransitive sentences, the sentences were gathered from *Corpus del Español* (Davies, 2002-, *www.corpusdelespanol.org*). All of them belong to the 20th Century. We included in our study data from the four genres in the corpus (oral, fiction, news, and academic). The distribution of the examples according to genre is the following: oral (N = 23), fiction (N = 340), news (N = 33), academic (N = 16). We included in our study data from the Americas (N = 175) and Spain (N = 33). We were not able to determine the geographic origin of 204 examples.

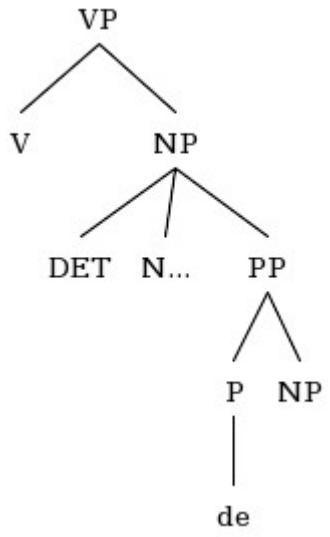
The first step in order to build the sample was to select representative verbs which occur in possessive constructions from Masullo (1992) and Delbecque and Lamiroy (1996). The verbs used in the study are the following:

Table 5.2: Verbs in the sample of possessive constructions

VERB	Count	Percent
abrir ('open')	26	6.31
acariciar ('caress')	71	17.23
admirar ('admire')	10	2.43
apretar ('press')	23	5.58
arreglar ('fix')	13	3.16
cerrar ('close')	23	5.58
examinar ('examine')	17	4.13
mirar ('watch')	61	14.81
quemar ('burn')	41	9.95
romper ('break')	24	5.83
sentir ('feel')	20	4.85
tocar ('touch')	48	11.65
tomar ('take')	35	8.5
TOTAL	412	100

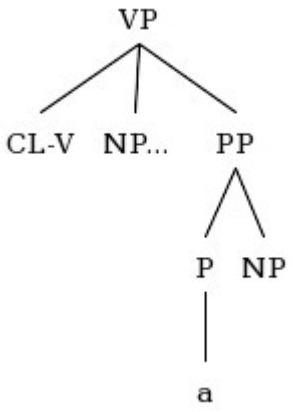
The search in *Corpus del Español* was conducted using two different queries. The query to retrieve genitive possessive constructions would retrieve any inflected form of each verb, followed by any definite or indefinite determiner, followed by the preposition *de* within a ten-word window. For instance, the query for the verb *arreglar* ('to fix') was "[arreglar] el/la/los/las/un/una/unos/unas....de". The syntactic structure that would be retrieved by this regular expression is shown in the tree-diagram below:

(4)



The query to retrieve dative possessive constructions would retrieve any inflected form of each verb, preceded by *le* or *les*, followed by the preposition *a* within a ten-word window. For instance, the query for the verb *arreglar* ('to fix') was "le/les [arreglar] ....a". The syntactic structure that would be retrieved by this regular expression is shown in the tree-diagram below:

(5)



This study of possessive constructions is not about the occurrence or not of DCLD, but a study of the choice of genitive vs. dative expression of the possessor. Therefore, there were no reasons to exclude pronominal possessors. Pronominal possessors occur as either genitive or dative. A genitive pronominal possessor is a possessive pronoun (Sp. *mi*, *mis* ('my'), *tu*, *tus* ('your'), *su*, *sus* ('his, her, their'), *nuestro*, *nuestra* ('our'), *vuestro*, *vuestra* ('your')[37]. A Dative pronominal possessor is a dative clitic. See examples from the sample below:

(6) a. Genitive pronominal possessor:

Partieron de mañana, con el sol de diciembre alumbrando y quemando tan fuerte como el dolor quemaba ***sus*** *corazones.*

'They left in the moring, the sun of December shining and burning as strong as the pain was burning their hearts'

b. Dative pronominal possessor:

Bebió con los ojos cerrados, sintiendo que el líquido le quemaba la garganta.

'He drank with his eyes closed, feeling the liquid burning his throat'

37 These examples were retrieved using a different query syntax, the verbal lemma followed by any of the possessive determiners ([VERB] mi/mis/tu/tus/su/sus/nuestro/nuestra/vuestro/vuestra).

5.3 THE VARIABLES

The examples were annotated for a set of independent variables and compiled in a Calc spreadsheet.[38] The independent variables were the same as used in the quantitative study of ditransitive sentences, with the addition of the variable "relation", which refers to alienable or inalienable possession. Grammatical category and length were not included because this study is not concerned with word order. The definitions of animacy, definiteness, givenness, region, and medium are the same used in the quantitative study of ditransitive sentences. In this section we will discuss the definition of the variable "relation" and the values used for the variable "lexical semantics", which are different from those in the study of ditransitive sentences.

5.3.1 Relation

The variable "relation" refers to the nature of the possession relation and has two values: "alienable possession" and "inalienable possession". Inalienable possession is a "possession relation in which the thing possessed is an inherent part of the possessor" (Matthews, 2007). Otherwise, a possession relation is considered alienable possession. Typically, body parts are considered to be an inherent part of the possessor. But there are also part/whole possession relations that involve inanimate objects. See examples from the sample below:

(7) a. Inalienable possession: body part:

*Me quedé mirando **la cara de Sandra** .*
'I kept looking at Sandra's face.'

38 http://www.openoffice.org/

b. Inalienable possession: part/whole relation:

*Creo que mis pies tocaron **el suelo de la piscina**, ...*

'I think my feet touched the floor of the pool, ...'

c. Alienable possession:

*Y volvió a su rutina de ayudar a lavar cubiertos, arreglar **su cama**, releer **sus pocos libros** y esperar cartas.*

'He returned to his routine to help wash utensils, fix his bed, reread his few books and wait for letters.'

The dative possessive construction has been associated in the literature with the interpretation that the possessor has been affected, especially if there is a relation of inalienable possession (Delbecque and Lamiroy, 1996: 96).

5.3.1.1 Descriptive statistics The following table shows the distribution of possession relation types in the sample:

Table 5.3: Type of possession relation in the sample

Possession relation	Count	Percent
Alienable possession	73	17.72
Inalienable possession	339	82.28
N = 412		

5.3.2 Lexical semantics

The verbs in the sample can be classified in two groups: those that denote a change of state (Levin and Rappaport Hovav, 1995) and those that do not . The notion of change of state is

relevant because if a change of state has taken place its undergoer has been causally affected. The classification of the verbs in the sample according to this criterion is shown below:

(8) a. Change of state: *abrir* ('to open'), *arreglar* ('to fix'), *cerrar* ('to close'), *quemar* ('to burn'), *romper* ('to break').

b. No change of state: *acariciar* ('to caress'), *admirar* ('to admire'), *apretar* ('to press'), *examinar* ('to examine'), *mirar* ('to watch'), *sentir* ('to feel'), *tocar* ('to touch'), *tomar* ('to take').

In order to determine which verbs denote a change of state, we use the availability of adjectival passives as a test (Levin and Rappaport, 1986; Bosque, 1999). A change of state produces a resulting state, which is denoted by an adjectival passive (a passive participle used predicatively). The following is an example of an adjectival passive in English:

(9) a. Passive:
The door was opened.

b. Adjectival passive:
The opened door.

In the adjectival passive example, the door is opened as a result of the change of state denoted by the verb *to open*. Verbs that denote a change of state allow for adjectival passive, and verbs that do not denote a change of state do not allow for it, as shown in the following examples:

(10) Adjectival passive (change of state):

a. *Juan rompió la ventana.*
Juan broke the window
'Juan broke the window.'

b. *La ventana está rota.*
the window is broken
'The window is broken.'

(11) Adjectival passive (no change of state):

a. *Juan acarició al gato.*
Juan caressed the cat
'Juan caressed the cat.'

b. *#El gato está acariciado.*
the cat is caressed
#'The cat is caressed.'

The verb *acariciar* ('to caress') does not denote a change of state and there is no resulting state to which the adjectival passive can make reference. As a result, the adjectival passive is semantically anomalous.

The following are examples from the sample of verbs that denote and do not denote a change of state:

(12) a. Change of state:
Suponte tú romper las chapas de una… una puerta...
'Suppose if you break the sheet iron of a door...'

b. No change of state:

Le vi el filo a la espada y de inmediato pensé...

'I saw the blade of the sword and immediately I thought...'

5.3.2.1 Descriptive statistics The following table shows the distribution of the lexical semantics of the verbs in the sample:

Table 5.4: Lexical semantics in the sample

Lexical semantics	Count	Percent
Change of state	127	30.83
No change of state	285	69.17
N = 412		

5.3.3 Descriptive statistics of animacy, definiteness, and givenness

The following table shows the distribution of animacy, definiteness, and givenness of possessor in the sample. The annotation of these independent variables was performed using the same guidelines as in the quantitative study of ditransitive sentences (see Chapter 4).

Table 5.5: Animacy of possessor in the sample

Animacy of possessor	Count	Percent
animate	12	2.91
human	278	67.48
inanimate	118	28.64
organization	4	0.97
N = 412		

Table 5.6: Definiteness of possessor in the sample

Definiteness of possessor	Count	Percent
definite	395	95.87
indefinite	15	3.64
specific	2	0.49
N = 412		

Table 5.7: Givenness of possessor in the sample

Givenness of possessor	Count	Percent
activated	89	21.6
in focus	112	27.18
referential	2	0.49
type identifiable	20	4.85
uniquely identifiable	189	45.87
N = 412		

5.3.4 Summary of independent variables

a. Animacy of possessor: human, organization, animate, inanimate.

b. Definiteness of possessor: definite, specific indefinite, indefinite.

c. givenness of possessor: in focus, active, uniquely identifiable, referential, type identifiable.

d. Relation: alienable possession, inalienable possession.

e. Lexical semantics of the verb: change of state, perception, contact

f. Region: Americas, Spain, unknown.

g. Medium: spoken, written.

5.4 THE METHOD

As in the quantitative study of ditransitive sentences, we will resort to multivariate analysis (logistic regression), a statistical method that "assesses the individual relative contribution of each factor to the observed variation when all factors are considered simultaneously (Walker, 2010: 38). In order to run the logistic regression we will use GoldVarbX (http://individual.utoronto.ca/tagliamonte/goldvarb.htm).

5.4.1 Simplification of variables

Before running the logistic regression, we made a some adjustments in the independent variables and their values. The independent variables which constitute hierarchies or scales (animacy, and givenness of possessor) were reduced to binary values. The values of animacy were simplified to "human" vs. "non-human" and the values of givenness to "activated" (in focus and activated) vs. "non-activated" (uniquely identifiable, referential, and type identifiable). This simplification was done because of practical reasons: there were not enough tokens in some cells to perform statistics with GoldVarbX. The independent variable definiteness of possessor was excluded from the logistic regression altogether because its distribution, nearly categorically definite, was not suitable to perform statistics.

5.5 THE RESULS

The logistic regression analysis performed using GoldVarbX found four factor groups (independent variables) as statistically significant predictors of the case of the possessor (genitive

or dative). These variables were, in order of rank: animacy of possessor, type of possession relation, givenness of possessor, and lexical semantics of the verb. The ranking of the variables as predictors is based on the range of the weight of the factors within the factor group. The weight of a factor indicates the probability of application (occurrence) of the dependent variable. The application value in this case is the occurrence of a dative possessor. A weight higher than 0.5 indicates that the individual factor favors the application, while a weight lower than 0.5 disfavors it. The range of a factor group is obtained by subtracting the largest factor weight from the smallest factor weight in each factor group (Walker, 2010:41). The third column in the table indicates the percentage of applications (dative possessors) in the factor (for instance, 36.30% of the [+human] possessors are dative). The fourth column indicates the number of tokens in the factor (for instance, 278 possessors are [+human]). The table also indicates the total N (= 412) and the input value (= 0.133), which is the value of the corrected mean (overall probability that the rule will apply, which in this case is the expression of the possessor as dative).

Table 5.8: Case of possessor (logistic regression)

Total N: 412, Input: 0.133, Application value: dative possessor

Factor Group	Factor Weight	% of rule application within factor	N within factor
Animacy			
human	0.73	36.30%	278
non-human	0.11	3.00%	134
Range:	*62*		
Relation			
alienable	0.16	15.10%	73
inalienable	0.59	27.70%	339
Range:	*43*		
Givenness			
activated	0.71	43.30%	201
non-activated	0.30	8.50%	211
Range:	*41*		
Lexical semantics			
change of state	0.73	33.90%	127
no change of state	0.39	21.80%	285
Range:	*34*		

Factor groups not selected: Region, medium.

The results show that the predictors of the expression of the possessor as dative belong to two classes: features of the possessor that contribute to its pragmatic (animacy, givenness) and semantic salience (type of possession relation, lexical semantics). The first group of factors overlaps with the predictors of DCLD in ditransitive constructions.

5.5.1 Animacy of possessor

The highest ranked predictor of dative possessors is animacy (range = 62). Table 5.9 shows that the expression of the possessor as dative is more likely when the possessor is human. 36.33% of the human possessors are dative. The percentage of datives among non-human possessors goes down to 2.99%. This distribution is significant according to the Chi-square test (Pearson Chi-Square = 52.942, DF = 1, $p \leq 0.001$).

Table 5.9: Animacy and case of possessor (cross-tabulation)

Rows: Animacy of possessor Columns: Case of possessor

	dat	gen	All
human	101	177	278
	36.33	63.67	100.00
	70.8	207.2	278.0
non-human	4	130	134
	2.99	97.01	100.00
	34.2	99.8	134.0
All	105	307	412
	25.49	74.51	100.00
	105.0	307.0	412.0

Cell Contents: Count
% of Row
Expected count

Pearson Chi-Square = 52.942, DF = 1, $p \leq 0.001$

5.5.2 Relation

The type of possessor relation was also found significant by the logistic regression (range = 43). Table 5.10 shows that the expression of the possessor as dative is more likely when there is a relation of inalienable possession. 27.73% of the instances of inalienable possession have a dative possessor. The percentage of datives in instances of alienable possession goes down to 15.07%. This distribution is significant according to the Chi-square test (Pearson Chi-Square = 5.070, DF = 1, p = 0.024).

Table 5.10: Type of possession relation and case of possessor (cross-tabulation)

Rows: Relation Columns: Case of possessor

	dat	gen	All
alienable	11	62	73
	15.07	84.93	100.00
	18.6	54.4	73.0
inalienable	94	245	339
	27.73	72.27	100.00
	86.4	252.6	339.0
All	105	307	412
	25.49	74.51	100.00
	105.0	307.0	412.0

Cell Contents: Count
% of Row
Expected count

Pearson Chi-Square = 5.070, DF = 1, p = 0.024

5.5.3 Givenness of possessor

The type of possessor relation was also found significant by the logistic regression (range = 41). Table 5.11 shows that the expression of the possessor as dative is more likely when the possessor is activated. 43.28% of the [+activated] possessors are dative. The percentage of datives among [-activated] possessors goes down to 8.53%. This distribution is significant according to the Chi-square test (Pearson Chi-Square = 65.468, DF = 1, $p \leq 0.001$).

Table 5.11: Givenness and case of possessor (cross-tabulation)

Rows: Givenness of possessor Columns: Case of possessor

	dat	gen	All
[-activated]	18	193	211
	8.53	91.47	100.00
	53.8	157.2	211.0
[+activated]	87	114	201
	43.28	56.72	100.00
	51.2	149.8	201.0
All	105	307	412
	25.49	74.51	100.00
	105.0	307.0	412.0

Cell Contents: Count
% of Row
Expected count

Pearson Chi-Square = 65.468, DF = 1, $p \leq 0.001$

5.5.4 Lexical semantics

The lexical semantics of the verb was also found significant by the logistic regression (range = 34). Table 5.12 shows that the expression of the possessor as dative is more likely when the verb

denotes a change of state. 33.86% of the examples with a change of state verb have a dative possessor. The percentage of dative possessors in examples with no change of state verbs goes down to 21.75%. This distribution is significant according to the Chi-square test (Pearson Chi-Square = 6.777, DF = 1, p = 0.009).

Table 5.12: Lexical semantics and case of possessor (cross-tabulation)

Rows: Lexical semantics	Columns: Case of possessor		
	dat	gen	All
change of state	43	84	127
	33.86	66.14	100.00
	32.4	94.6	127.0
no change of state	62	223	285
	21.75	78.25	100.00
	72.6	212.4	285.0
All	105	307	412
	25.49	74.51	100.00
	105.0	307.0	412.0

Cell Contents: Count
% of Row
Expected count

Pearson Chi-Square = 6.777, DF = 1, p = 0.009

5.6 CONCLUSIONS

The results of the study on the distribution of dative possessors show that their occurrence is more likely when the possessor is pragmatically salient (high in animacy and givennness), a fact that confirms our secondary hypothesis #3. This fact also suggests that, in spite of formal and functional differences, a unified analysis of dative case in Spanish is possible: both the overt

marking of dative case in ditransitive constructions and the choice of a dative construction in possessive constructions is favored by pragmatic salience.

In addition, the occurrence of dative possessors is favored by two factors that favor the interpretation of the possessor as a proto-patient (Dowty, 1991): change of state predicate and inalienable possession. Undergoing a change of state is the first proto-patient property listed by Dowty (1991) (see Chapter 1). Moreover, if there is a relation of inalienable possession, the possessor is perceived as affected by the change of state that has taken place (Delbecque and Lamiroy, 1996: 96). The possessive alternation consists of the promotion of an argument of the head of the DO (genitive) to core argument of the verb (dative) resulting in an increase of the verb valence from a binary to a ternary predicate. This alternation, that can be described as an instance of possessor raising (Masullo, 1992), is more likely to take place when the possessor has reached a threshold of proto-patient properties.

6.0 CONCLUSIONS

Throughout the dissertation we have argued that dative clitics are agreement markers with different functional loads in ditransitive and non-ditransitive constructions. In ditransitive constructions, they optionally overtly marked the recipient as IO. Pragmatically salient IOs are more likely to be overtly marked as dative. In non-ditransitive constructions, the dative clitic signals that an argument alternation has taken place. The argument alternation that takes place in non-ditransitive constructions consists of the promotion of an argument to IO, which is a core grammatical function in Spanish. This conclusion applies to those contexts in which DCLD is not required, such as when the IO is an stressed pronoun or it is left or right-dislocated. Therefore, DCLD is not optional across the board, but in a set of contexts.

The quantitative studies of ditransitive and possessive constructions have confirmed our main and secondary hypotheses, stated below:

a. Main hypothesis: in the contexts under study DCLD is an optional phenomenon constrained by multiple factors.

b. Secondary hypothesis #1: the occurrence of DCLD and the order of theme and recipient are dissociated phenomena.

c. Secondary hypothesis #2: the occurrence of DCLD is constrained by properties of the recipient and is dissociated from properties of the theme.

d. Secondary hypothesis #3: pragmatic salience favors the expression of an argument as dative when there is an alternation between dative and a less prominent grammatical function.

The confirmation of our main hypothesis (in the contexts under study DCLD is an optional phenomenon constrained by multiple factors) implies that DCLD in ditransitive constructions cannot be reduced to syntactic or lexical considerations. The occurrence of DCLD in a ditransitive construction is an option provided by the morphosyntax of Spanish, but the grammar underspecifies its use. The results of our study show that the use of DCLD is subject to an important degree of dialectal and stylistic variation. Regarding linguistic factors, speakers are more likely to use DCLD in a ditransitive construction when the recipient is pragmatically salient (high on animacy/givenness).

The confirmation of our secondary hypothesis #1 (the occurrence of DCLD and the order of theme and recipient are dissociated phenomena) underscores the differences between English dative-shift and Spanish DCLD that we have discussed throughout the dissertation. Overt coding of grammatical functions in Spanish is independent of word order. The Spanish data are useful in order to tell apart the effects of independent variables that constrain the coding of grammatical functions (animacy, givenness) and the determination of word order (grammatical complexity), factors which are intertwined in English. The result of our study suggest that the placement of arguments within Spanish VPs can be explained as the result of functional constraints operating

over a shallow syntactic structure, without resorting to additional syntactic structure, such as VP-shells. This option should be preferred as the more parsimonious one. Nevertheless, our results do not preclude a VP-shells analysis.

The confirmation of our secondary hypothesis #2 (the occurrence of DCLD is constrained by properties of the recipient and is dissociated from properties of the theme) also underscores the differences between English dative-shift and Spanish DCLD. It is compatible with the claim that English is a Primary/Secondary Object language and Spanish is a Direct/Indirect Object language (Dryer, 1986; Raúl Aranovich, 2007). In Spanish ditransitive constructions there is no competition for the DO position, which is always occupied by the theme. Therefore the features of the theme do not influence the occurrence of DCLD, which is only concerned with properties of the recipient.

The confirmation of our secondary hypothesis #3 (pragmatic salience favors the expression of an argument as dative when there is an alternation between dative and a less prominent grammatical function) implies that the phenomenon of dative case in Spanish is partially unified. In spite of the different functional load fulfilled by dative clitics in ditransitive and non-ditransitive constructions, dative case is favored by pragmatic prominence across different construction types.

The results of our quantitative study of ditransitive and possessive constructions shows an important difference regarding the significance of lexical semantics. Lexical semantics was not found to be a significant predictor of dative marking of recipients by the logistic regression, but it was found to be a significant predictor of the choice of a dative instead of a genitive possessor. The occurrence of a dative possessor is favored both by inalienable possession relations and

change of state predicates, both factors that allow for an affected reading of the possessor. This result is expected under the assumption that dative marking of recipients and possessors does not perform the same function. Dative marking of recipients is an instance of agreement, not the indication that an argument alternation has taken place. If there is no argument alternation, it should be expected for dative marking of recipients not to be subject to lexical constraints. On the other hand, the possessive alternation is a clear example of an argument alternation, which is expected to be subject to lexical constraints, such as the Argument Selection Principle (Dowty, 1991), and the proto-properties of the arguments involved.

The conclusions of this study help advance our understanding of dative constructions in Spanish. However, there are empirical and theoretical problems that have not been addressed in this dissertation and that would constitute possible topics of future research. Bresnan et al. (2007) also considered as an independent variable the concreteness of the theme. It would be beneficial to incorporate that independent variable in further research. Regarding the distribution of dative case in Spanish, the conclusions of this study could be enhanced by the implementation of further quantitative studies of the alternations not covered in this dissertation (benefactive alternation, experiencer alternation). It would be also beneficial to compare the results of this study regarding the order of theme and recipient with the results of further quantitative studies on the order of constituents within the VP, such as the order of complements and modifiers, objects and secondary predicates, etc. Our results show that dialectal and stilistic variation play a role in the occurrence of DCLD in ditransitive constructions, but our discussion has been focused on the role of linguistic factors. Further research on the role of dialectal and stylistic variation in DCLD would be necessary in order to provide a more complete picture of the phenomenon.

APPENDIX A

SETS OF SPANISH PRONOUNS

Table A.1: Set of stressed Spanish pronouns

	Nominative	Objective	Reflexive
1sg	yo	a mí	a mí mismo/misma
2sg	tú, vos (informal)[39] usted (formal)	a ti, a vos a usted	a ti mismo/misma a usted mismo/misma
3sg	él (masc), ella (fem), ello (neuter)	a él (masc) a ella (fem) a ello (neuter)	a sí mismo/misma
1pl	nosotros (masc), nosotras (fem)	a nosotros (masc) a nosotras (fem)	a nosotros mismos(masc) a nosotras mismas(fem)
2pl	vosotros (masc), vosotras (fem), ustedes[40]	a vosotros (masc), a vosotras (fem), a ustedes	a vosotros mismos (masc) a vosotras mismas (fem) a ustedes mismos/mismas
3pl	ellos (masc), ellas (fem)	a ellos (masc) a ellas (fem)	a ellos mismos (masc) a ellas mismas (fem)

39 *Vos* is used instead of *tú* in some varieties of Latin American Spanish.

40 *Vosotros* and *vosotras* are not used in Latin American Spanish. Instead, *ustedes* is used as both formal and informal 2pl nominative pronoun. In Spain, *vosotros* and *vosotras* are informal and *ustedes* is formal.

Table A.2: Set of unstressed Spanish pronouns

	Accusative	Dative	Reflexive
1sg		me	
2sg		te	
3sg	lo (masc), la (fem)	le/se	se
1pl		nos	
2pl (Spain)		os	
2pl (Latin America)	los (masc), las (fem)	les/se	se
3pl	los (masc), las (fem)	les/se[41]	se

41 The pronoun *se*, among its many functions in Spanish grammar, is an allomorph of *le* and *les* that occurs when a dative unstressed pronoun co-occurs with an accusative unstressed pronoun.
Yo le di el libro a María
**Yo le lo di*
Yo se lo di

BIBLIOGRAPHY

Aarts, Bas (2007) *Syntactic Gradience: The Nature of Grammatical indeterminacy*. Oxford: Oxford University Press.

Abbot, Barbara (2006) "Definiteness and Indefiniteness". In Lawrence Horn and Gregory Ward (eds.) *The Handbook of Pragmatics.* Oxford: Blackwell.

Abbot, Barbara (2010) *Reference.* Oxford: Oxford University Press.

Ackerman, Farrell and John Moore (2001) *Proto-Properties and Grammatical Encoding.* Stanford: CSLI.

Aissen, Judith (2003) "Differential Object Marking: Iconicity vs. Economy". *Natural Language and Linguistic Theory* 21: 435-483.

Alsina, Alex (1996) *The Role of Argument Structure in Grammar. Evidence from Romance*. Stanford: CSLI.

Anagnostopoulou, Elena (2006) "Clitic Doubling". In Martin Everaert and Henk van Riemsdijk (eds.) *The Blackwell Companion to Syntax.* Oxford: Blackwell.

Andrews, Avery (1990) "Unification and morphological blocking". *Natural Language and Linguistic Theory* 8: 507-557.

Aoun, Joseph (1985) *A grammar of anaphora*. Cambridge, Mass.: The MIT Press.

Aranovich, Raúl (2007) "The Spanish Dative Alternation in LMT". Workshop on Case, Word Order, and Prominence in Argument Structure. Radboud University, Nijmegen.

Aranovich, Roberto (2007) "Against verb raising in Spanish". Hispanic Linguistics Symposium 2007, University of Texas, San Antonio.

Ariel, Mira (1988) "Referring and Accessibility". *Journal of Linguistics* 24, 65-87.

Ariel, Mira (1990) *Accessing Noun-Phrase Antecedents.* London: Routledge.

Arnold, J., T. Wasow, A. Losongco, and R. Ginstrom (2000) 'Heaviness vs Newness: The Effects of Complexity and Information Structure on Constituent Ordering. *Language* 76: 28-55.

Baker, Mark C. (1988) *Incorporation: A Theory of Grammatical Function Changing.* Chicago: The University of Chicago Press.

Barss, Andrew and Howard Lasnik (1986) "A Note on Anaphora and Double Objects". *Linguistic Inquiry* 17: 347-354.

Becerra Bascuñan, Silvia (2006) *Estudio diacrónico y sincrónico del objeto indirecto en el español peninsular y de América*. Copenhaguen: Museum Tusculanum Press.

Behagel, O. (1909) "Beziehungen zwischen Umfang und Reihenfolge von Satzgliedern" *Indogermanische Forschungen* 25: 110-142.

Bleam, Tonia (2003) "Properties of the Double Object Construction in Spanish". In R. Núñez Cedeño, L. López and R. Cameron (eds.) *A Romance Perspective on Language Knowledge and Use.* Amsterdam: John Benjamins.

Bod, Rens, Jennifer Hay and Stefanie Jannedy (eds.) (2003) *Probabilistic Linguistics.* Cambridge: The MIT Press.

Bosque, Ignacio (1999) "El sintagma adjetival. Modificadores y complementos del adjetivo. Adjetivo y participio". In Ignacio Bosque and Violeta Demonte (eds.) *Gramática descriptiva de la lengua española.* Madrid: Espasa Calpe.

Bossong, Georg (1991) "Differential Object Marking in Romance and Beyond". In Dieter Wanner and Douglas A. Kibbee (eds.) *New Analysis in Romance Linguistics*. Amsterdam: John Benjamins.

Bresnan, Joan & Sam Mchombo (1987) "Topic, pronoun and agreement in Chichewa". *Language* 63: 741-782.

Bresnan, Joan (2001) *Lexical-Functional Syntax.* Oxford: Blackwell.

Bresnan, Joan, Anna Cueni, Tatiana Nikitina, and R. Harald Baayen (2007) "Predicting the Dative Alternation". KNAW Academy Colloquium. Amsterdam.

Butt, Miriam (2006) *Theories of Case.* Cambridge: Cambridge University Press.

Chomsky, Noam (1965) *Aspects of the Theory of Syntax.* Cambridge: The MIT Press.

Chomsky, Noam (1981) *Lectures on Government and Binding. The Pisa Lectures.* Berlin: Mouton de Gruyter.

Chomsky, Noam (1986) *Knowledge of Language. Its Nature, Origin, and Use*. New York: Praeger Publishers.

Chomsky, Noam and Howard Lasnik (1993) "The Theory of Principles and Parameters". In J. Jacobs, A. von Stechow, W. Sternefeld y T. Vennemann (eds.) *Syntax: An International Handbook of Contemporary Research*. Berlin & New York: Walter de Gruyter.

Company Company, Concepción (2006) "El objeto indirecto". In Concepción Company Company (ed.) *Sintaxis histórica de la lengua española*. *Primera parte: la frase verbal.* México: Fondo de Cultura Económica.

Comrie, Bernard (1989) *Language Universals and Linguistic Typology.* Chicago: The University of Chicago Press.

Croft, William (1988) "Agreement vs Case Marking and Direct Objects". In M. Barlow and C. Ferguson (eds.) *Agreement in Natural Language*: *Approaches, Theories, Descriptions.* Chicago: The University of Chicago Press.

Croft, William (2001) *Radical Construction Grammar.* Oxford: Oxford University Press.

Cuervo, Cristina (2003) "Structural asymmetries but same word order. The dative alternation in Spanish". In Anna Maria Di Sciullo (ed.) *Asymmetry in Grammar. Volume I: Syntax and semantics.* Amsterdam: John Benjamins.

Culicover, Peter (2009) *Natural Language Syntax.* Oxford: Oxford University Press.

Culicover, Peter W. and Ray Jackendoff (2005) *Simpler Syntax*. Oxford: Oxford University Press.

Davies, Mark. (2002-) Corpus del Español (100 million words, 1200s-1900s). Available online at http://www.corpusdelespanol.org.

Delbecque, Nicole and Béatrice Lamiroy (1996) 'Towards a typology of the Spanish dative'. In William van Belle and Willy van Langendonck (eds.) *The dative*. Amsterdam: John Benjamins.

Demonte, Violeta (1995) "Dative Alternation in Spanish". *Probus* 7: 5-30.

Dobrovie-Sorin, Carmen (1994) *The Syntax of Romanian. Comparative Studies in Romance.* Berlin: Mouton de Gruyter.

Dowty, David (1991) "Thematic Proto-Roles and Argument Selection". *Language* 67: 547-619.

Dryer, Matthew S. (1986) "Primary Objects, Secondary Objects, and Antidative". *Language* 62: 808-845.

Evans, Vyvyan (2007) *A Glossary of Cognitive Linguistics*. Salt Lake City: The University of Utah Press.

Fanselow, Gisbert, Caroline Féry, Ralf Vogel, and Matthias Schlesewsky (eds.) (2006) *Gradience in Grammar. Generative Perspectives.* Oxford: Oxford University Press.

Fernández Soriano, Olga (1999) "El pronombre personal. Formas y distribuciones. Pronombres átonos y tónicos. In Ignacio Bosque and Violeta Demonte (eds.) *Gramática descriptiva de la lengua española.* Madrid: Espasa Calpe.

Firbas, Jan (1964) *"*On defining the theme in functional sentence analysis". *Travaux Linguistique de Prague* 1: 267-280.

Firbas, Jan (1966) "Non-Thematic Subjecs in Contemporary English". *Travaux Linguistique de Prague* 2: 239-254.

Fontana, Josep M. (1993) *Phrase Structure and the Syntax of Clitics in the History of Spanish.* PhD Dissertation, University of Pennsylvania.

Frawley, William (1992) *Linguistic Semantics.* Hillsdale, NJ: Lawrence Erlbaum Associates Publishers.

Fulmer, S. Lee (1990) "Dual-position affixes in Afar: an argument for phonologically driven morphology. In Aaron L. Halpern (ed.) *Proceedings of the Ninth West Coast Conference on Formal Linguistics.* Stanford: CSLI Publications.

Garretson, Gregory (2004) "Coding practices used in the project Optimal Typology of Determiner Phrases". Unpublished manuscript, Boston University, Boston, MA. http://npcorpus.bu.edu/html/documentation/index.html.

Givón, Talmy (1976) "Topic, Pronoun, and Grammatical Agreement". In Charles N. Li (ed.) *Subject and Topic.* New York: Academic Press.

Givón, Talmy (ed.) (1983) Topic Continuity in Discourse: A Quantitative Cross Language Study. Amsterdam: John Benjamins.

Givón, Talmy (1984) *Syntax: A Functional-Typological Introduction. Volume I*. Amsterdam: John Benjamins.

Givón, Talmy (1988) "The pragmatics of word order: Predictability, importance and attention". In Michael Hammond, Edith A. Moravcsik and Jessica Wirth (eds.) *Studies in Syntactic Typology*. Amsterdam: John Benjamins.

Givón, Talmy (2001) *Syntax: An Introduction. Volume I.* Amsterdam: John Benjamins.

Green, G. 1974. *Semantics and Syntactic Regularity*. Bloomington: Indiana University Press.

Gries, Stefan Th. (2003) "Grammatical variation in English: a question of 'structure vs. function'?". In Günter Rohdenburg and Britta Mondorf (eds.) *Determinants of grammatical variation in English*. Berlin & New York: Mouton de Gruyter.

Gundel, Jeanette K., Nancy Hedberg, Ron Zacharski (1993) "Cognitive status and the form of referring expressions in discourse". *Language* 69: 274-307.

Halpern, Aaron L. (2001) "Clitics". In Andrew Spencer and Arnold M. Zwicky (eds.) *The Handbook of Morphology.* Oxford: Blackwell.

Hawkins, John (1978) *Definiteness and indefiniteness: A study in reference and grammaticality prediction.* London: Croom Helm.

Hawkins, John (1994) *A performance theory of order and constituency.* Cambridge: Cambridge University Press.

Hawkins, John (2004) *Efficiency and complexity in grammars*. Oxford: Oxford University Press.

Jackendoff, Ray (1987) "The status of thematic relations in linguistictheory". *Linguistic Inquiry* 18: 368-411.

Jackendoff, Ray (1990a) *Semantic Structure.*Cambridge: The MIT Press.

Jackendoff, Ray (1990b) "On Larson's treatment of the double object construction". *Linguistic Inquiry* 21: 427-456.

Jaeggli, Osvaldo (1982) *Topics in Romance syntax*. Dordrecht: Foris.

Jelinek, Eloise (1984) "Empty categories, case, and configurationality". *Natural Language and Linguistic Theory* 2: 39-76.

Keenan, Edward (1976) "Towards a universal definition of subject". In Charles Li (ed.) *Topic and Subject*. New York: Academic Press.

Kuno, Susumu (1976) "Subject, Theme, and the Speaker's Empathy -A Reexamination of Relativization Phenomena". In Charles N. Li (ed.) *Subject and Topic.* New York: Academic Press.

Krifka, Manfred, Francis Jeffry Pelletier, Gregory N. Carlson, Alice ter Meulen, Gennaro Chierchia, and Godehard Link (1995) "Genericity: an introduction". In Gregory N. Carlson

and Francis Jeffry Pelletier (eds.) *The Generic Book.* Chicago: University of Chicago Press.

Labov, William (1969) "Contraction, deletion, and the inherent variability of the English copula". *Language* 45: 715-762.

Legendre, Geraldine, William Raymond, and Paul Smolensky (1993) "An Optimality-Theoretic typology of case and grammatical voice systems". In *Proceedings of the Nineteenth Annual Meeting of the Berkeley Linguistic Society* 464-478.

Larson, Richard K. (1988) "On the Double Object Construction". *Linguistic Inquiry* 19: 335-391.

Leonetti, Manuel (1999) "El artículo". In Ignacio Bosque and Violeta Demonte (eds.) *Gramática descriptive de la lengua española. Madrid: Espasa Calpe.*

Levin, Lorraine S. (1985) *Operations on Lexical Forms: Unaccusative Rules in Germanic Languages*. PhD Dissertation, MIT.

Levin, Beth and Malka Rappaport (1986) "The Formation of Adjectival Passives". *Linguistic Inquiry* 17: 623-661.

Lyons, John (1968) *Introduction to Theoretical Linguistics*. London: Cambridge University Press.

Lyons, Cristopher (1999) "Definiteness". In Keith Brown and Jim Miller (eds.) *Concise Encyclopedia of Grammatical Categories*. Oxford: Elsevier.

Maling, Joan (2001) "Dative: The Heterogeneity of the Mapping among Morphological Case, Grammatical Functions, and Thematic Roles". *Lingua* 111: 419-464. Special Issue *On the Effects of Morphological Case.*

Masullo, Pascual José (1992) *Incorporation and case theory in Spanish: A crosslinguistic perspective.* PhD Dissertation, University of Washington.

Matthews, P. H. (2007) *The Concise Oxford Dictionary of Linguistics.* Oxford: Oxford University Press.

Mayer, Elisabeth (2006) "Optional direct object clitic doubling in Limeño Spanish". *Proceedings of LFG 2006* http://csli-publications.stanford.edu/LFG/11/lfg06mayer.pdf.

Mchombo, Sam (2001) "Chichewa". In Andrew Spencer and Arnold M. Zwicky (eds.) *The Handbook of Morphology.* Oxford: Blackwell.

Mendikoetxea, Amaya (1999) “Construcciones inacusativas y pasivas”. In Ignacio Bosque and Violeta Demonte (eds.) *Gramática descriptiva de la lengua española.* Madrid: Espasa Calpe.

Milsark, Gary (19 74) *Existencial sentences in English.* PhD Dissertation, MIT.

Monachesi, Paola (1998) “Decomposing Italian clitics”. In Sergio Balari and Luca Dini (eds.) *Romance in HPSG.* Stanford: CSLI.

Monachesi, Paola (2005) *The Verbal Complex in Romance: A Case Study in Grammatical Interfaces*. Oxford: Oxford University Press.

Nishida, Chiyo (2010) “What can a corpus study tell us about the Spanish dative alternation?”. Hispanic Linguistics Symposium 2010. Indiana University, Bloomington.

Noyer, Rolf (1994) “Mobile affixes in Huave: optimality and morphological well-formedness. In *Proceedings of the Twelfth West Coast Conference on Formal Linguistics.* Stanford: CSLI Publications.

Oehrle, Richard T. (1976) *The grammar of the English dative alternation*. PhD Dissertation, MIT.

Pensado, Carmen (1995) “La creación del complemento directo prepositional y la flexión de los pronombres personales en las lenguas románicas”. In Carmen Pensado (ed.) *El complemento directo prepositional*. Madrid: Visor Libros.

Perlmutter, David (1971) *Deep and Surface Structure Constraints in Syntax.* New York: Holt, Rinehart, and Winston.

Pinker, Steven (1984) *Language Learnability and Language Development.* Cambridge: Harvard University Press.

Pinker, Steven (1989) *Learnability and Cognition.* Cambridge: The MIT Press.

Prince, Ellen (1980) “Toward a taxonomy of given-new information”. In P. Cole (ed.) *Radical Pragmatics.* New York: Academic Press.

Rivero, María Luisa (1975) “Referential Properties of Spanish Noun Phrases”. *Language* 51: 32-48.

Romero Morales, Juan (2008) *Los dativos en español.* Madrid: Arco Libros.

Russell, Bertrand (1905) “On denoting”. *Mind* 14: 479-493.

Strozer, Judith Reina (1976) *Clitics in Spanish*. PhD Dissertation, UCLA.

Suñer, Margarita (1988) "The role of agreement in clitic doubled constructions". *Natural Language and Linguistic Theory* 6: 391-434.

Tagliamonte, Sali A. (2011) "Probabilistic syntax from a sociolinguistic perspective: The dative alternation". LSA Annual Meeting. Pittsburgh, Pennsylvania, January 6-9, 2011.

Toman, Jindrich (1995) *The magic of a common language: Jakobson, Mathesius, Trubetzkoy, and the Prague Linguistic Circle*. Cambridge: The MIT Press.

Walker, James A. (2010) *Variation in Linguistic Systems.* London: Routledge.

Wasow, Thomas (2002) *Postverbal Behavior.* Stanford: CSLI.

Woods, Anthony, Paul Fletcher, and Arthur Hughes (1986) *Statistics in language studies*. Cambridge: Cambridge University Press.

Woolford, Ellen (1999) "Animacy hierarchy effects on object agreement". In Paul Kotey (ed.) *New Dimensions in African Linguistics and Languages*. Trenton, NJ: Africa World Press.

Yamamoto, Mutsumi (1999) *Animacy and Reference: a Cognitive Approach to Corpus Linguistics.* Amsterdam: John Benjamins.

Zaenen, Annie, Jean Carletta, Gregory Garretson, Joan Bresnan, Andrew Koontz-Garboden, Tatiana Nikitina, Mary C. O'Connor, and Tom Wasow (2004) "Animacy encoding in English: Why and how". In Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics (ACL'04), Workshop on Discourse Annotation, pages 118–125. Barcelona, Spain.

Zagona, Karen (2002) *The Syntax of Spanish.* Cambridge: Cambridge University Press.

Zwicky, Arnold M. and Geoffrey K. Pullum (1983) "Cliticization vs. Inflection: English N'T". Language 59: 502-513.