

ABSTRACT

Title: PRE-VERBAL STRUCTURE BUILDING IN
ROMANCE LANGUAGES AND BASQUE

Leticia Pablos, PhD in Linguistics, 2006

Directed By: Professor Colin Phillips
Professor Juan Uriagereka
Department of Linguistics.

The main goal of the work in this dissertation is to investigate pre-verbal structure building effects in languages with different configurations such as Spanish, Galician and Basque, by means of using different pre-verbal cues in order to show that syntactic structure can be both interpreted and anticipated before the verbal head.

I examine the syntax of Clitic-Left Dislocations (CLLDs) and other kinds of left-dislocations in Spanish and then analyze their processing. I concentrate on the use of clitic pronouns in Spanish and Galician in CLLD constructions that require the presence of the clitic pronoun to interpret the left-dislocated phrase and where I examine if this left-dislocation is interpreted at the clitic pronoun. Experimental results from three self-paced reading experiments provide evidence that the clitic in these constructions is required and used to interpret the thematic features of the topicalized NP before the verb. Thus, I demonstrate that clitic pronouns are used as pre-verbal cues in parsing and that the active search mechanism is also triggered in long-distance dependencies involving clitic pronouns. I conclude that the active search mechanism is a more general architectural mechanism of the parser that is

triggered in all kinds of long-distance dependencies, regardless of whether the search is triggered by gaps or pronouns.

In Basque, verbal auxiliaries overtly encode agreement information that reflects the number of arguments of the verbal head. In negatives, auxiliaries are obligatorily fronted and split from the verbal head with which they otherwise form a cluster. Thus, verbal auxiliaries in Basque are a pre-verbal morphological cue that can assist the parser in predicting structure. Specifically, I examine how predictions for the upcoming structure of the sentence are determined by agreement information on the number of arguments specified in the auxiliary and by the mismatch of this auxiliary with the case features of the NP that follows it. I provide results from a self-paced reading experiment to argue that the parser uses the information encoded in the auxiliaries and demonstrate that the mismatch of the auxiliary with the following NP can prevent the reader from following a garden-path analysis of the sentence.

PRE-VERBAL STRUCTURE BUILDING IN ROMANCE LANGUAGES AND
BASQUE

By

Leticia Pablos

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2006

Advisory Committee:
Professor Colin Phillips, Chair
Professor Juan Uriagereka, Co-chair
Professor Tonia Blears
Professor Mary Ellen Scullen
Professor Amy Weinberg

© Copyright by
Leticia Pablos
2006

Dedication

A mis padres, Julia Robles y Ángel Pablos, por arroparme en la distancia cada día.

A Jose Lorenzo, por permanecer siempre a mi lado en este largo viaje.

Acknowledgements

The time has come when I have to thank all the people who have contributed one way or another to the work in this dissertation. I would like to thank the following people:

Colin Phillips, my main advisor and to whom I owe most of what I know about psycholinguistics. Thank you for being so patient and for believing in that it is possible to learn everything that one needs to know to run an experiment from scratch. Even stats! Thank you also for the fruitful discussions we have had over the years and for teaching me how to be a critical thinker and how to become a psycholinguist. Thank you for always being so helpful in any situation where your help as an advisor was required.

Juan Uriagereka, thank you for being such a good advisor for linguistics and for life in general, for being always ready to set me into the next challenge and for passing on the feeling that nothing is impossible in this life. Thank you for being so stubborn and for getting me thinking in the Galician experiment. Graciñas Juan!

Tonia Bleam, who I got to meet late in my graduate years (but whom I wish I would have met before), I thank her for being so generous with her time and for the challenging and fruitful discussions we had about Spanish syntax. Gracias Tonia!

Amy Weinberg, thank you for awaking my curiosity in psycholinguistics during my first year and for always posing the most challenging questions (that I never managed to answer back). Mary Ellen Scullen, thank you for being always available and so positive about everything and for the first tips to be a TA.

In the Basque Country, I would like to especially thank Jon Franco and Jon Ortiz de Urbina for always being there available for advice and for helping me recruiting people for the experiments. To them I'm also deeply in debt for having sent me through the linguistics' path...If it wasn't for them this dissertation would have never existed! Thanks also to Alazne Landa, Myriam Uribe-Etxebarria, Ricardo Etxepare and Itziar Laka who were always very helpful. Outside the University, thanks to Jasone Aldekoa and Irene Gonzalez for their help in recruiting people in the Berritzegune of Galdakao. Mila esker bioi!

In Galiza, I would like to thank Rosario Álvarez, Mónica Ares, Yolanda Pena, Xose-Luis Regueira and Flory Xouvanova for their help in recruiting participants in Santiago. Special thanks go to Manuel Lamas, who was my patient guide during the days I run the experiment and to the Lorenzo family for being my hosts during those days. I would also like to thank Carlos Acuña-Fariña and Manolo Carreiras for valuable comments on different presentations of the Galician experiment at SLE and Penn State.

Among my colleagues, I would like to thank first Nina Kazanina and Masaya Yoshida for being excellent colleagues and friends. Thanks Nina for all your advice and support during all those years and especially for your irreplaceable friendship. Spasivo Ninotchka! Thanks to Masaya with whom I shared the experience of writing a thesis in parallel and who always provided me with a new enlightening linguistics reading for the day. In the early days, Aitziber Atutxa and Itziar San Martin for helping me survive through my first year and for providing me with the tips to make it through. Thanks especially to Aitzi who was always so supportive! Later on, I

thank Cilene Rodrigues and Max Guimaraes for their friendship and for always being so combative. Ana Gouvea and Sachiko Aoshima for being the pioneers in running sentence processing experiments in the department and for opening the path to future generations. Elixabete Murguia for always giving good advice about the future and for being such an incomparable colleague and friend. Rob Fiorentino, Tomohiro Fujii and Utako Minai for being great listeners and for sharing the experiences of graduate school day by day with me. To the visiting students Angel Gallego and Maia Duguine for making my last year of grad school a little brighter and for making me recover the enthusiasm of the first years. Thanks also to all the students and research assistants at the UMD Linguistics Department, particularly: Diogo Almeida, Clara Cabezas, Ilhan Cagri, Brian Dillon, Soo-Min Hong, Hirohisa Kiguchi, Moti Liberman, Mitsue Motomura, Phil Monahan, Chizuru Nakao, Ivan Ortega-Santos, Usama Soltan, Eri Takahashi, Heather Taylor, Akira Omaki, Hajime Ono, Kaori Ozawa, Acrisio Pires, Matt Wagers and Ming Xiang. Outside UMD, I would like to thank Lara Reglero for always being ready for a good laugh and for always keeping in touch.

I immensely thank my family and friends here and in the other side of the sometimes-too-wide-Atlantic Ocean. I would like to deeply thank my parents for being my constant support over these years and those friends here and overseas who always gave me words of encouragement: Edna Durand, Rossella and Loredana DiStravolo, Nicole Spiridakis and Michael Ross, Ilya Arakelyan, Arseni Goussev, the Sicard Family, Virginia Manzanares, Maitane Castro and Ion González, Irene González, Alicia Tejeda, Josefi Subirats, Jose Román and Manu Suarez. Thank you all from the top of my heart!

Last but not least, Jose Lorenzo, without whom this trip would have not been so smooth. Thank you for your love and support over all these years and for always being there for me whenever I needed you. Thank you for always listening patiently to all the crazy linguistic ideas I had to brainstorm about. I am sure you know as much about linguistics as I do by now!

Finally this thesis would have not been written without the music of great jazz musicians such as Bill Evans, John Coltrane, Sony Rollins, Miles Davis and Thelonious Monk. To them, I owe the strength to continue writing in the moments where I was more desperate and tired.

Part of the work in this dissertation has been financially supported by a grant from the Department of Education, Universities and Research from the Government of the Basque Country.

Table of Contents

Dedication.....	ii
Acknowledgements.....	iii
CHAPTER 1: PRE-VERBAL STRUCTURE BUILDING.....	1
1.1. Evidence for pre-verbal structure building.....	5
1.1.1 Head-driven models vs. fully incremental parsing models.....	6
1.1.2. Evidence in favor of incremental structure building.....	8
1.1.3. Evidence in favor of pre-verbal construction of long-distance dependencies.....	12
1.2. How does the parser create structure pre-verbally?.....	13
1.2.1. Characteristics of the parser.....	14
CHAPTER 2: THE SYNTAX OF LEFT-PERIPHERAL ELEMENTS IN SPANISH.....	20
2.1 Properties of Spanish movement constructions.....	22
2.1.1 Gap Site.....	22
2.1.2 Island sensitivity.....	24
2.1.3 Connectivity.....	26
2.1.4 Parasitic Gaps.....	29
2.1.5 Weak Crossover.....	30
2.2 Properties and classification of constructions involving left-dislocated phrases in Spanish.....	31
2.2.1 CLLDs vs. LDs.....	32
2.2.2 Clitic-less Dislocations: PPs, Bare Nouns and Subjects.....	57
2.2.3 Parsing CLLD constructions.....	72
2.3 Conclusion.....	74
CHAPTER 3: ACTIVE-SEARCH FOR CLITIC PRONOUNS IN LEFT-DISLOCATED CONSTRUCTIONS IN SPANISH.....	77
3.1 Processing long-distance dependencies.....	77
3.1.1. Gap as-a-first-resort vs. gap as-a-last-resort.....	79
3.1.2. Classic Filled-Gap Effects.....	80
3.1.3. Active Filler Hypothesis.....	82
3.2. Are the parsing mechanisms involved in forming CLLD dependencies active? 86	
3.2.1. Long-distance dependencies involving pronouns.....	86
3.2.2. Previous work on long-distance dependencies in Romance languages	88
3.2.3. Experiment 1a: off-line grammaticality-rating questionnaire.....	91
3.2.4. Experiment 1b: Online Sentence Reading Experiment.....	93
3.2.5. How active are dependency construction mechanisms involving clitic pronouns?.....	102
3.2.6. Subject Filled-gap Effects.....	104
3.2.7. Pre-verbal dependency completion effects in head-final languages.	107
3.2.8. Experiment 2.....	109
3.3. General Discussion.....	119

Appendix 3-A: Off-line Sentence Completion Experiment 1.....	124
Appendix 3-B: Off-line Sentence Completion Experiment 2.....	129
Appendix 3-C: Data items for online sentence reading experiment 1b.....	135
Appendix 3-D: Data items for online sentence reading experiment 2.....	139
CHAPTER 4: THE ACTIVE-SEARCH FOR CLITIC PRONOUNS IN GALICIAN	
.....	146
4.1. Active search for clitic pronouns in CLLD constructions vs. active search for gaps.....	146
4.1.1. Active search for pronouns: CLLD dependencies.....	147
4.2. Galician as a case study for a stronger clitic prediction than Spanish.....	148
4.3. Spanish vs. Galician clitic placement.....	150
4.3.1. Spanish clitic placement.....	150
4.3.2. Galician clitic placement.....	151
4.4. Experiment 3a: Off-line Grammaticality-rating Experiment.....	152
4.5. Experiment 3b: On-line Sentence Reading Experiment.....	154
4.6. General Discussion.....	163
4.6.1. Comparison with results obtained for Spanish CLLD Dependencies	163
4.7. Conclusion.....	165
Appendix 4-A: Data items for online reading experiment in Galician.....	166
CHAPTER 5: BASQUE AUXILIARY FRONTING STUDY.....	172
5.1. Basque: a rich agreement language.....	173
5.2. Direct vs. indirect cues for predicting structure ahead.....	179
5.2.1. Case markers and word order as inducers of clause boundaries.....	179
5.2.2. Direct cues = (English) verbal argument structure and subcategorization information.....	182
5.2.3. Indirect cues= Japanese classifier mismatch.....	183
5.3. Expectation-driven parser.....	187
5.4. Predictions for this experiment.....	190
5.4.1. Experiment 4: On-line Reading Experiment on Basque Auxiliary Fronting	192
5.5. General Discussion.....	200
5.5.1. Reanalysis in Basque.....	205
5.5.2. How to build an embedded clause from a mismatch.....	216
5.6. Conclusions.....	229
Appendix 5-A: Data items for online reading experiment on basque auxiliary fronting.....	231
CHAPTER 6: CONCLUSIONS.....	237
Bibliography.....	240

CHAPTER 1: PRE-VERBAL STRUCTURE BUILDING

This dissertation focuses on the use of pre-verbal cues to interpret structure in advance of syntactic heads in human sentence processing. Specifically, it explores the processing of long-distance dependencies involving a left-dislocation and a clitic in Spanish and Galician and the processing of fronted auxiliaries in Basque negative sentences.

In this chapter, I will show commonalities in the way different languages with a varying internal configuration are parsed in terms of interpreting structure in advance of the verbal head. I will demonstrate how speakers of Spanish and Galician on the one hand, and Basque on the other, can interpret elements before the verb is encountered bottom-up, by means of using morphological or agreement cues specific to each language. These two cues have a different status in that in Spanish and Galician the clitic is the element that is anticipated and where the fronted phrase can be interpreted, whereas in Basque the cue is the combination of the auxiliary and the following NP.

Many of the studies on the processing of long-distance dependencies in head-initial-languages (i.e. languages in which heads appear in the beginning of each clause or where the head is clause initial) have focused on how structures are projected in parsing from the information stored in lexical heads (mainly verbal heads) (Trueswell et al. 1993, MacDonald et al. 1994, Garnsey et al 1997, Boland et al. 1997, Tanenhaus et al. 1997). These studies focused on how the argument

structure of the verb can guide parsing by providing direct cues about the number of complements that the structure will have and their likely probability. Considering this previous evidence, I would like to explore whether there are additional cues that the parser can use to project structure in an anticipatory fashion, before the bottom-up information of the verbal head becomes available. In particular, I aim to investigate whether agreement information can assist parsing structure before heads in both head-initial and head-final languages. Since agreement information is represented differently from language to language, I will examine different phenomena in two Romance languages (Spanish and Galician) and Basque.

In Romance languages, clitic pronouns have been argued to be overt realizations of agreement with the verbal head (Kayne 1991, 1994, Uriagereka 1995). Clitics are words that syntactically function as a free morpheme and that are part of the argument structure of the verb. More importantly, they can appear in pre-verbal position (*proclisis*). Therefore, clitics are pre-verbal morphosyntactic cues that can assist the parser in interpreting structure in advance of the verb. In this thesis, I will concentrate on the use of clitic pronouns in Spanish and Galician¹ in topicalization constructions or more specifically Clitic-Left Dislocations (henceforth CLLDs) constructions such as those illustrated in (1a) and (1b).

- (1) a. A Jose, el editor le ha dicho [que lo va a contratar].
PA Jose, the editor cl(IO) has said that cl(DO) going to hire.
'To Jose, the editor has told him that he will hire him'
- b. Ó Xose, o editor sempre lle di [que o vai contratar].
PA Jose, the editor always cl(IO) said that cl(DO) going to hire.

¹ Even if clitics in both Spanish and Galician can be found post-verbally (*enclisis*), I will only concentrate here on pre-verbal clitics. The reason for this will be that I want to focus on dependency completion effects prior to the verbal head.

‘To Jose, the editor always tells him that he will hire him’

CLLD constructions in these two languages require the presence of the clitic pronoun to interpret the left-dislocated phrase. Hence, they provide a good testing ground to examine if the left-dislocated phrase can only be interpreted at the verb or if this left-dislocation is interpreted earlier at the clitic pronoun. As I will discuss in chapters 3 and 4 of this dissertation, the experimental results provide evidence to show that the clitic in these constructions is required and used to interpret the thematic features of the topicalized NP before the verb. Thus, I will demonstrate that clitic pronouns are used as pre-verbal cues in parsing and that, more generally speaking, there are pre-verbal cues in head-initial languages that can assist the parser in projecting structure in an anticipatory fashion.

In Basque, verbal auxiliaries overtly encode agreement information that reflects the number of arguments of the verbal head. Interestingly, in negative and emphatic sentences, auxiliaries are obligatorily fronted and split from the verbal head with which they otherwise form a cluster. Thus, verbal auxiliaries in Basque are a pre-verbal morphological cue that can assist the parser in predicting structure. Specifically, I will examine how predictions for the upcoming structure of the sentence are determined by agreement information on the number of arguments specified in the auxiliary and by the mismatch of this auxiliary with the case features of the noun phrase that follows it. For this purpose, I will use two types of auxiliaries, monotransitive and ditransitive, that contrast in that they morphologically encode whether the verb takes two or three arguments. This contrast in the morphology of the auxiliary is illustrated in (2) and (3).

- (2) Ninak; ez dio Masayari [pro; liburua erosi duela esan].
N-Erg neg aux-ditrans M-dat book-abs buy aux-trans-Comp say
 ‘Nina has not told Masaya that she has bought the book’
- (3) Tomok ez du [Utakori liburua eman diola] esan.
T-Erg neg aux-trans U-dat book-abs give aux-trans-Comp say
 ‘Tomo has not said that he has given the book to Utako’

In the examples in (2) and (3) the auxiliary following negation determines the number of arguments that the parser will expect the verbal head to have. We will show that this expectation is temporarily maintained or adjusted depending on whether the NP following the auxiliary matches in case with the expectations previously created by the auxiliary. In chapter 5 of the dissertation, I will provide results from a self-paced reading experiment to argue that the parser uses the information that is encoded in the auxiliaries. I will also demonstrate that the mismatch of the auxiliary with the following NP as in (3) can prevent the reader from following a garden-path analysis of the sentence.

Although the phenomena under investigation are quite different, they all speak to the same point regarding how structure can be anticipated based on a partial input. In all cases, I will show how experimental results can be accounted for by assuming an incremental parser with a strong predictive mechanism that uses either morphosyntactic (auxiliary information) or syntactic cues (uninterpreted left-dislocation). In this kind of parser, a morphological mismatch or an uninterpreted dislocated phrase is the trigger to predict structure and help anticipating structure before the verbal head is reached. Based on experimental evidence from typologically distinct constructions that argues for the necessity of a powerful projection

mechanism in the parser, I will assume a “left-corner parser” with a strong predictive mechanism (as discussed in section 1.2.1. of this chapter).

In what follows, in section 1.1, I will first examine the crosslinguistic psycholinguistic evidence that has shown that there is pre-verbal structure building and that has provided evidence against head-driven parsing models. Then, I discuss other studies on pre-verbal structure building that speak to the interpretation of long-distance dependencies before the verb. Finally, in section 1.2., I will focus on the type of parsing mechanism that will be able to generate the structures of the experiments.

1.1. Evidence for pre-verbal structure building

Head-final languages have raised a challenge for theories of parsing. Since the head of the phrase comes clause-finally in these languages, these languages posit an interesting question for human parsing. There are two possible alternatives for how the parsing of these languages could proceed. The first alternative is the head-driven parsing which assumes that the parser needs to wait to see the verbal head to interpret all the relevant information. The second hypothesis is the fully incremental parsing which assumes that the arguments are interpreted incrementally and connected to the structure of the sentence being processed by using the available cues provided by each language (e.g. word order, case marking, etc.). The second hypothesis presupposes that the parser can still interpret the structure and connect the arguments with each other even before the verbal head becomes available. In the literature, there have been different accounts of parsing models that relate to this particular issue in

these languages. In the following section, I present the differences between the parsing theories that have been proposed for these languages (incremental vs. head-driven) and I discuss the evidence that has been used to show that the parser builds structure incrementally and immediately.

1.1.1 Head-driven models vs. fully incremental parsing models

Conceivably, sentences can be parsed in various ways. On one hand, there are head-driven models of parsing (Abney 1989, Prittchet 1992, Mulders 2002) that advocate strictly bottom-up parsing of the sentence. In this kind of model, only the information found in the input is considered and no prediction about the upcoming material is made. A commitment to a syntactic analysis is made only when a licensing head is found in the input. Before that, words and phrases from the input are left unattached and stored until the head is reached. Specific challenges for this framework are presented by languages like Basque, in which the head appears at the end of the sentence. Hence, under a head-driven parsing model, all the arguments that precede the verb should be left unattached until the verb is encountered.

This type of parsing model has been proposed by researchers before because of the advantage of reducing the need for massive syntactic reanalysis. Since no early commitments are made before reaching the heads, there is no need to reanalyze structure if the wrong choice of structure has been made. The disadvantage of this approach is that it also places a high memory cost on the parser because it implies that

the parser has to keep all the unconnected arguments in memory until the head is encountered.

The implications of a head-driven parser in head-final languages could be reflected in the following way. If the parser finds an input like that in (4) in a language like Japanese, the three initial NPs will be buffered in memory and they will not be combined with each other until the first verbal head (in this case the verb *to introduce*) appears and assigns a theta role to them. Following Pritchett (1992), attachments in this model of parsing cannot be made until they are licensed by the head.

- (4) Tom-ga Guy-ni Frank-o shookai sita
T-Nom G-Dat F-Acc introduced PAST
'Tom introduced Frank to Guy'

On the other hand, fully-incremental models of parsing (Crocker 1994, Stabler 1994, Gorrell 1995, Sturt & Crocker 1996, Lombardo & Sturt 2002) use the available bottom-up information and they make predictions about the upcoming structure even if the head is not present. In these models, each word is incorporated within the syntactic structure as it is encountered. Therefore, the disadvantage of this kind of model is that reanalysis may be more frequent due to the prior commitments and expectations that are made about the structure and that might have to be revised later on in the sentence. Nevertheless, sentences will not always have to be reanalyzed since in many cases there is only one possible analysis of the sentence. The advantage of these models is that the memory cost is reduced to a certain extent and the structure that has already been seen does not have to be kept unconnected within memory,

since as much bottom-up information as possible is tried to be incorporated into the projected structure.

The implications of a fully incremental parser in head-final languages could be summarized as follows. There is no delay in parsing, and phrases are immediately associated with each other and connected to the current syntactic structure when local information (e.g. case marking) is available. As opposed to head-driven models, parsing decisions are never delayed and they can be made in advance of the verbal head. In example (4), that would imply that all the NPs are introduced in the structure and related with each other before the verb *to introduce* is reached.

1.1.2. Evidence in favor of incremental structure building

A previous set of studies has tried to demonstrate that there is an incremental attachment of phrases in the structure before the verbal head is accessed in languages with a head-final configuration. The evidence that has been used to advocate an incremental parsing model for head-final languages comes mainly from Japanese and German.

Mazuka & Itoh (1995) showed in an eye-tracking study that the three initial NPs that had identical nominative case marking were read slower in sentence (5a) than the first three NPs that had different cases in (5b). They considered this as evidence to show that the information before the verb is used.

- (5) a. Yoko-ga [Hiromi-ga [Asako-ga Ø kaita] genkoo-o kakinaosita]
Y-Nom [H-Nom [A-Nom wrote] draft-Acc re-wrote]
 syorui-o yonda.
papers-acc read.
 ‘Yoko read the papers that Hiromi re-wrote based on the draft Asako wrote’
- b. Yoko-ga Hiromi-o Asako-ni shookai sita
Y-Nom H-Acc A-Dat introduced
 ‘Yoko introduced Hiromi to Asako’

Nevertheless, in this study it is hard to tease apart the reasons for the slow reading time in the first three NPs in (5a). This slowdown could be related to the fact that the NPs are being associated with each other before the verbal head is processed. But it could also be due to the fact that it is not frequent to find three nominative NPs in Japanese, and that an uncommon case array slows the reader down. Moreover, there could also be potential interference in memory encoding because of case similarity.

A second piece of evidence against the use of a head-driven parser in Japanese comes from a sentence completion study by Inoue and Fodor (1995) using fragments like (6a). They expected Japanese speakers to give more continuations like that in (6b) than that in (6c), indicating that participants would have a preference to expect the three NPs to be co-arguments and would normally expect a verb that takes these three arguments. Otherwise, if speakers initially considered the three NPs not to be co-arguments as in (6c), one of the NPs would have to be left unassigned and it will not be included in the projected structure. Japanese speakers consistently completed the fragment in (6a) as if the three NPs were the arguments of the same verb (the equivalent will be (6b)).

- (6) a. Bob-ga Mary-ni ringo-o....
B-Nom M-Dat apple-Acc
 b. Bob-ga Mary-ni ringo-o ageta
B-Nom M-Dat apple-Acc gave
 ‘Bob gave Mary the apple’
 c. Bob-ga Mary-ni [_sringo-o tabeta] inu-o ageta
B-Nom M-Dat apple-Acc ate dog-Acc gave
 ‘Bob gave Mary the dog that ate the apple’

Finally, another piece of evidence that has been used to argue that association of the NPs is done before reaching the verbal head is Bader & Laser’s (1994) self-paced reading study in German. In German embedded clauses introduced by a complementizer, verbs appear in clause-final position and therefore verbal elements occur clause-finally. The sentences in (7) are the experimental sentences they used.

- (7) Maria dachte.....
Maria thought
- a. [dass sie_{NOM} [nach dem Ergebnis zu fragen] tatsächlich erlaubt HAT]
that she for the result to ask indeed permitted has
 ‘that she indeed has given permission to ask for the result’
- b. [dass [sie_{ACC} nach dem Ergebnis zu fragen] tatsächlich erlaubt
that her for the result to ask indeed permitted
 WORDEN IST]
been is.
 ‘that permission indeed has been given to ask her for the result’
- d. [dass er_{NOM} [nach dem Ergebnis] zu fragen tatsächlich erlaubt HAT]
that he for the result to ask indeed permitted has
 ‘that he indeed has given permission to ask for the result’
- e. [dass [ihn_{ACC} nach dem Ergebnis] zu fragen tatsächlich erlaubt
that him for the result to ask indeed permitted
 WORDEN IST]
been is.
 ‘that permission indeed has been given to ask him for the result’

In the sentences in (7a-b) the pronoun *sie*, whose case is ambiguous between a nominative and an accusative pronoun, can either act as the subject of the main clause

(7a) or as the object of the embedded infinitival clause (because the verb *zu fragen* is optionally transitive). The most relevant part of the construction is the finite clause because that is where the sentence is disambiguated at the clause-final auxiliary. Moreover, these two conditions are contrasted with conditions that contain the unambiguous masculine pronouns *er* (7c) or *ihn* (7d), which do not lead the parser to any ambiguity.

The contrast exemplified in (7) makes a good testing case to examine the different implications of the two theories of parsing that we have discussed. Under the incremental approach, a reading time difference at the auxiliaries in (7a) and (7b) would be predicted if the parser were trying to attach incrementally the pronoun *sie* to the structure as soon as it is encountered. On the other hand, under a head-driven approach, there should not be any processing difference at the auxiliaries if the parser is waiting for the verb at the end of the clause before attaching all of the arguments.

Results showed evidence in favor of the incremental type of parsing models. Reading times were significantly longer for the passive auxiliary in (7b) than for the active auxiliary in the ambiguous feminine case in (7a). This implies that subjects took longer to process those ambiguous sentences where the auxiliary reveals that the ambiguous pronoun *sie* must be the object of the embedded verb and not the subject of the main clause. There are two conclusions that these results yield: (i) the parser strives to incorporate every new argument into the structure, and (ii) there is a preference to attach the pronoun *sie* as the subject of the matrix clause. Note that under head-driven models (Abney 1989 and Pritchett 1992), the ambiguous pronoun *sie* should have been interpreted as the object of the infinitival clause in (7a) as soon

as the verb *zu fragen* was encountered. Under such a structural representation, a garden path should have occurred when the active auxiliary was processed, and we should have obtained the opposite results.

Other arguments in favor of incremental structuring of the sentence come from the processing of scrambled wh-phrases in Japanese (Miyamoto & Takahasi 2002) and from the processing of backward anaphora in Japanese (Aoshima et al. 2006).

As I will discuss in chapter 5, Basque speakers also use case information and the morphological mismatch to attach every word incrementally and without delay until the verbal head is found.

1.1.3. Evidence in favor of pre-verbal construction of long-distance dependencies

Another type of evidence for pre-verbal structure building in head-final languages comes from experiments with pre-verbal filler-gap dependencies. Numerous studies have focused on preverbal dependency completion effects in long distance dependencies involving wh-gaps and structure building effects that occur ahead of disconfirming input (e.g. Nakano et al. 2002, Aoshima et al. 2004).

As we will discuss in section 3.2.7 in chapter 3, these studies have shown that the parser does not need the information from the verbal head to predict the position of a gap and that the case-features of the NPs before the verb and the word order can predict the structure ahead.

In chapter 3, I will show how Spanish speakers can also construct topicalization dependencies before the verb and interpret them using clitic pronouns. Even if Spanish has a head-initial configuration, Spanish speakers can anticipate the structure in left-dislocation constructions and predict the occurrence of the clitic pronoun guided by the grammatical knowledge of Spanish before the verbal head is processed. Effects from Spanish and Galician related to the prediction of the clitic before it is encountered in the input, will be reported in chapters 3 and 4 respectively. The fact that the clitic might be able to be projected before it is encountered in the input also emphasizes the projection power of the algorithm that generates these structures, since the clitic position is predicted before the overt realization of the clitic.

1.2. How does the parser create structure pre-verbally?

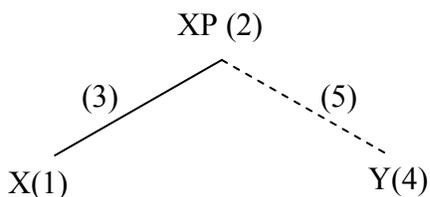
In this section, I concentrate on the mechanism that can explain how pre-verbal structure building takes place and I state the assumptions that need to be made so that structure is built ahead of time.

1.2.1. Characteristics of the parser

The parsing model I will be adopting for this study has similar features to the basic characteristics of a Left-corner parser (Johnson-Laird 1983, Abney & Johnson 1991, Resnik 1992, Stabler 1994, Schneider 1999, Aoshima 2003).

In a left-corner parser, structure is built bottom-up using the lexical information encoded in the input words. When a constituent is completed (the left corner²) the parent node of this constituent is built and its potential siblings are predicted³. The steps followed by the parser are characterized in the diagram in (8).

(8) Input word : X



This parser builds a maximal, partial syntactic analysis of the sentence with each new word in the input using the syntactic knowledge specific to the grammar of the language being processed (Crocker 1992, 1994). Moreover, as mentioned earlier, we need a parser with a powerful predictive mechanism that goes beyond predictions based on the lexical information provided by the input and that projects structure in advance of the heads of phrases (Lombardo & Sturt 2002).

² The left-corner of any constituent is its leftmost immediate constituent.

³ Steps in the derivation are marked by (n)

1.2.1.1. Strong prediction based parser

In this parser, building projections in advance of encountering the head of the phrase is permitted if it allows connecting each new input word to the structural representation.

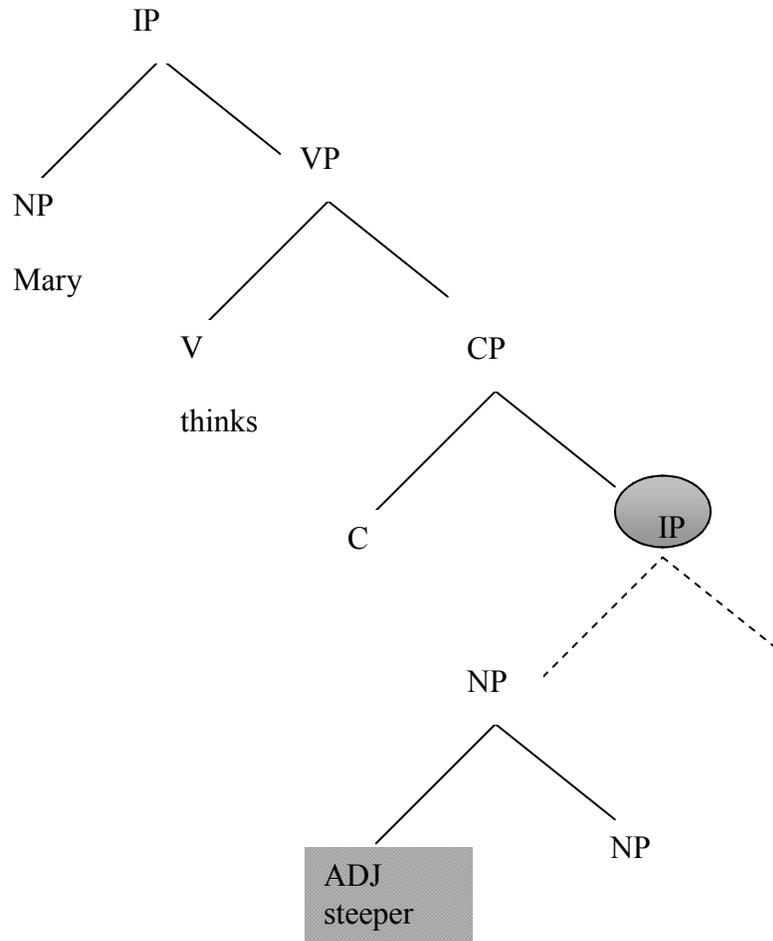
Let us illustrate this point by looking at the English example in (9) from (Lombardo & Sturt 2002) and its representation in (10).

(9) He thinks [_{CP}[_{IP}[_{NP}[_{ADJ} [*steeper*] prices] have come about because...]]

If the steps of a left-corner parser are followed, at the time the adjective *steeper* has to be incorporated into the structure that has been built so far, the parser has previously projected a CP based on the information of the verb *to think*, which normally takes clausal complements. At this point, the parser has to connect the adjective *steeper* to the current structure. Therefore, the parser has to find a way to connect the adjective to the CP that is already projected. The adjective can project an NP because only a noun head would license it. In order for the NP to be part of the CP, the NP has to be the embedded subject of the clausal complement of *think*. Therefore, IP is the projection required to connect these two maximal projections (NP & CP), but IP cannot be projected based on bottom-up information.

It is at this point that the predictions of a left-corner parser would not be able to handle completely the derivation of this structure. Based only on the bottom-up information (an adjective) and the previous prediction (a CP), it is not possible to project an IP.

(10)



Therefore, in order for the parser to incorporate the adjective into the current structure, there is a need for other inferences besides the projection established by the verb to have a CP and the prediction established by the adjective to have an NP. For that, we need a parser that projects beyond lexical structure and that can perform inferences about the input.

As I will argue in chapter 5 for Basque, the parser will face a similar type of problem as described above. In this case, the parser will have to derive the existence of an embedded sentence from a morphological mismatch. Moreover, in the case of

Basque, the parser will not even have the information provided by the matrix verbal head, so the number of inferences that are needed in order to process the sentence incrementally is greater than for the English example in (9).

The organization of this thesis is as follows.

Chapter 2 centers on the syntax of different left-peripheral elements in Spanish and on its implications for language processing. I focus first on the properties of Spanish movement constructions to establish the basic assumptions about movement. I then compare the properties of two different kinds of left-peripheral phrases: Left-dislocations (LD) and Clitic Left-dislocations (CLLD). Then I discuss the divergent arguments that have been used in the literature for their syntactic derivation (particularly for CLLDs) and I discuss whether they involve movement or not showing that CLLD involves movement but LD does not. Finally, I compare the syntactic behavior of a third kind of left-dislocation (referred to as Clitic-less Left? Dislocation), with CLLDs. I argue that clitic-less dislocations behave like CLLDs. I conclude that the clitic pronoun is not the reason that CLLD has special properties and I compare the results of the syntactic analysis to those of parsing.

Chapter 3 discusses the existing literature on the processing of long-distance wh-dependencies and discusses the active search mechanism that is triggered when dependencies involving gaps are processed. I present a parallel between the active search mechanism in dependencies that involve gaps with those that involve clitic pronouns in Spanish and I investigate whether dependencies involving clitic pronouns have the same active-search mechanism in constructions with a left-dislocation. I present two self-paced reading experiments in Spanish. The first experiment examines whether an active search is also triggered for left-dislocated constructions

with clitic pronouns. The second experiment is aimed at examining how active the search for a clitic pronoun can be in these dependencies if they behave like wh-dependencies. I argue that an active search mechanism applies when sentences with left-dislocation that require clitic pronouns are processed and that clitics are used to interpret those long distance dependencies. I conclude that the active search mechanism is a general parsing mechanism intrinsic to the architecture of the human parser that is used for processing of various types of long-distance dependencies, regardless of whether there is a pronoun or a gap involved in the dependency.

Chapter 4 covers the processing of long-distance topicalization dependencies involving clitic pronouns in Galician. I discuss how the clitic requirement is stronger in Galician than in Spanish. Then, I present a self-paced reading experiment in Galician in which I test the active search mechanism in topicalized constructions. I argue for the fact that the active search mechanism is also triggered in the processing of topicalization dependencies in this language. I conclude that Galician provides additional evidence for claiming that the active search mechanism is a more general mechanism used in parsing of long-distance dependencies.

Chapter 5 focuses on the real-time processing of structures containing morphologically rich auxiliaries in Basque. I present the basic configurational properties of Basque and I discuss work in previous literature on the use of direct and indirect cues in processing. I then present a self-paced reading experiment aimed at testing how an indirect cue provided by a mismatch of the agreement information in the auxiliary and the case marking in the following NP can guide parsing. I show that the information provided by the mismatch is used to prevent the parser from a garden-path effect. I show how the results of the experiment fit into the existing literature on

reanalysis and I present some assumptions about the parser that are necessary to account for our experimental results. I conclude that the combination of the agreement information encoded in the auxiliaries and case marking is used in incremental parsing in head-final languages.

Chapter 6 presents the conclusions for the whole thesis.

CHAPTER 2: THE SYNTAX OF LEFT-PERIPHERAL ELEMENTS IN SPANISH

In this chapter, I will focus on the syntax of Clitic Left-dislocation constructions (henceforth CLLD). First, I compare CLLD with wh-movement in Spanish. The purpose of this comparison is to determine whether the parallel established in the processing of chains that involve gaps and those that involve pronouns, as will be discussed in chapter 3, holds for how these two constructions are syntactically derived. What we find is that CLLD shows reliable properties of movement in that it is sensitive to island constraints. However, it does not exhibit all of the properties of wh-movement. For example, CLLD is not sensitive to weak-crossover effects and the gap does not license a parasitic gap. I argue that CLLD does in fact involve movement, but the derivation is different from Wh-movement in that the left peripheral element is not a wh-operator. Next, I compare CLLD constructions to another kind of left-peripheral phrase,⁴ called Left-Dislocation (henceforth LD), which does not involve movement. Finally, I will examine if CLLD constructions, which have a clitic pronoun at the bottom of the chain, are derived in a similar way as left-peripheral constructions that have a null element at the bottom of the chain (what I will call clitic-less dislocation) or if they should, instead, be considered a different kind of construction. It has been claimed that the presence of the clitic is responsible

⁴ I will use the term left-dislocated phrase as a generic way to name left-peripheral elements. When referring to the specific construction of Left-Dislocation, in the sense alluded to in the text, I will be using the nomenclature LD.

for the quirky properties of CLLDs and the nature of the chain, which make these constructions sometimes resemble movement constructions and sometimes not. The goal of this section (2.2.2) is to investigate whether the role of the clitic is crucial in CLLDs, by comparing it with clitic-less dislocations.

The chapter is organized as follows. In the first section I will discuss the basic properties of *wh*-constructions in Spanish. This is done to establish the basic tests for *wh*-movement that will be used throughout the chapter.

In the second section I will compare different kinds of left-dislocation constructions in Spanish to CLLD constructions. In section 2.2.1, I will discuss the basic syntactic properties of CLLD and LD constructions separately and relate that to their movement/ non-movement nature. Then, in the same section, I will review the arguments that have been used to support the non-movement nature of LD constructions and also the controversial arguments that have been used to discuss the movement/non-movement nature of CLLDs. Then, I will give arguments for why one hypothesis is to be preferred over the other.

In section 2.2.2 I will try to support the hypothesis arguing that these clitic-less constructions can be taken as CLLDs due to the resemblance of their syntactic properties with the CLLDs. Specifically, I will argue that dislocated bare nouns and PPs are cases of CLLD in that they are derived in the same way that CLLD constructions are derived since they display the same properties as CLLD with respect to the movement tests that we establish in section 2.1. Moreover, the fact that in the clitic-less cases there is no overt clitic could be relevant for finding out whether the overt clitic is a central, defining, property of CLLD constructions in general. In

what follows, I will argue for a movement analysis for CLLDs and clitic-less constructions.

Finally, I will relate the conclusions for the syntactic analysis of these constructions to the results found for parsing CLLD dependencies that will be introduced in chapter 3 and discuss the predictions for the other kinds of constructions based on our experimental results.

2.1 Properties of Spanish movement constructions

In this section I will discuss the properties that have been taken as diagnostics for A-bar movement and show how these properties are manifested in Spanish. The properties of A-bar movement in Spanish will be shown through wh-movement, taken as a canonical instance. This will serve as a basis for further discussion on the CLLD, LD and clitic-less dislocation phenomena I will be covering in the rest of the sections.

2.1.1 Gap Site

Wh-movement constructions in Spanish leave an empty position (which I will represent as “*e*” in my examples) behind, as exemplified in (1a-b), and cannot be

generally doubled by a pronoun or clitic, as shown in (1b-d)⁵. Basically the same generalization holds whether relevant wh-phrases are D-linked or not.⁶

- (1) a. ¿Quiénes [dices [que no han venido e]?
 Who(pl.) (you) say that (they) not had come
 ‘Who did you say that they did not come?’
- b. *¿Quiénes [dices [que ellos no han venido e]?
 Who(pl.) (you) say that they not have come
 ‘Who did you say that they did not come?’
- c. ¿Qué [ha comprado e María para Pedro]?
 What has bought Maria to Pedro
 ‘What did Maria buy for Pedro?’
- d. *¿Qué [lo ha comprado e María para Pedro]?
 What CL(DO) has bought Maria for Pedro
 ‘What did Maria buy for Pedro?’

⁵ There is a distinction in the acceptance of wh-phrase extraction and clitic presence depending on whether the extracted element related to clitic doubling is a direct object or an indirect object. This is illustrated in the examples in (i) and (ii) by Suñer (1988): in Porteño Spanish the clitic presence in indirect object extraction is allowed, whereas comparable direct object extraction is not.

- (i). ¿A quién leⁱ regalaron un auto eⁱ?
 PA who CL(IO) gave a car
 ‘Whom did they give a car to?’
- (ii). *¿A quién loⁱ condecoraron eⁱ?
 PA who CL(DO) decorated
 ‘Whom did they decorate?’

Suñer (1988) also relates the ungrammaticality of (ii) to the kind of wh-word that is used.

- (iii). *¿A qué candidato loⁱ condecoraron eⁱ?
 PA which candidate CL(DO) decorated
 ‘Which candidate did they decorate?’
- (iv). ¿A cual de los dos candidatos loⁱ condecoraron eⁱ?
 PA which of the two candidates CL(DO) decorated
 ‘To which of the two candidates did they decorate?’

⁶D-linked wh-phrases are also incompatible with a clitic as illustrated in (i) and (ii).

- (i). ¿[Qué libro] *(lo) has comprado e?
 Which book (you) (CL-DO) have bought
 ‘Which book did you buy it?’
- (ii). ¿[A qué hombre] *(lo) has conocido e?
 Which man (you) (CL-DO) have met
 ‘Which man did you meet?’

2.1.2 Island sensitivity

Wh-movement in Spanish is sensitive to island constraints (Ross 1967), but wh-constructions in Spanish are subject to slightly different restrictions from those in English. In particular the bounding nodes⁷ are NP and CP in Spanish, instead of the customary NP and IP (Torrego 1984). Subjacency conditions in Spanish are therefore different from those of English.

Subjacency blocks extraction from indirect questions in English because IP is a bounding node in this language as shown by (2a)⁸. In contrast, in Spanish, as in Italian (Rizzi 1978), it is possible to wh-move out of indirect questions (2b).

- (2) a. * What will you wonder [where he will put *e*]?
b. ¿Qué no sabes [dónde él pondrá *e*]?
What (you) not know where he will put
'What do you wonder where he will put?'

However, Spanish wh-movement is sensitive to strong islands, and thus does not allow extraction out of any type of wh-phrase that is within a complex NP island, relative clause island, coordinate island, sentential subject island or adjunct island.

⁷ A bounding node is a node that plays a role in determining whether a movement is local enough. Traditionally, NP and S (in English) or S' (in Italian) are considered bounding nodes. More recently, bounding nodes have been defined in terms of barriers. (Chomsky 1981, 1986b, Rizzi 1982)

⁸ Notice that in Spanish the verb 'to wonder' *preguntarse* does not permit wh-extraction, therefore, there can't be a direct correspondent in Spanish of the sentence in (2).

- (i). * ¿Quién se pregunta Juan si habrá salido?
'Who wonders Juan if will have left?'
(ii). * ¿A quién te preguntas si ellas entregaron el libro?
'Who do you wonder if they have given the book to?'

- (3) a. *¿[A quién] [[el hecho de [que Pedro conozca *e*]] no tiene importancia]?
 (Complex NP island)
 PA whom the fact of that Pedro knows not have relevance
 ‘Who does the fact that Pedro met have no relevance?’
- b. *¿[A quién] Juan dice [que visitó [al hombre [que odia *e*]]]?
 (RC island)
 PA who Juan says that (he) visited PA+the man that (he) hates
 ‘Who does Juan say that he visited the man who hates?’
- c. *¿[[[A quién] [ves *e* un par de veces]] y [te emocionas]]]?
 (Coordinate island)
 PA Who (you) see a couple of times and (you) CL(DO) are moved
 ‘Who do you see a couple of times and you are moved?’
- d. *¿[A quién] [[que María haya conocido *e*] le preocupa a Pedro]?
 (Sentential Subject island)
 PA whom that Maria has met CL(IO) worries PA Pedro
 ‘Who does that María has met worry Pedro’
- e. *¿[A quién] [Juan se desmayó después de [que conoció *e*]]]?
 (Adjunct island)
 PA whom Juan refl. fainted after of that (he) met
 ‘Who did Juan fainted after he meet?’

In summary, wh-phrases cannot be extracted out of so-called “strong islands” in Spanish (Torrego 1984), but movement out of certain weak islands (indirect questions) is possible (as seen in (2b) above).

The same island sensitivity restriction holds for those wh-phrases that are D-linked⁹.

⁹ This sets a contrast with Greek wh-cases (McMahon 2002), where non-D linked wh-phrases are sensitive to weak islands but D-linked wh-phrases are not. McMahon (2002) relates the fact that D-linked wh-phrases are not sensitive to weak islands with the fact that CLLDs in Greek are not sensitive to weak islands either. Since D-linked wh-phrases behave in the same way as non-D-linked phrases in Spanish, this comparison is not available for Spanish.

2.1.3 Connectivity

Wh-movement in Spanish exhibits connectivity effects. This can be seen through various diagnostics.

Before proceeding in the discussion, it has to be specified that there are two classes of connectivity that we consider following Cinque (1977). In the CLLD literature the term “connectivity” has been used with a broader sense than is normal in connection with the binding theory.

Cinque (1977) gives an extensive list of the different connectivity effects found in CLLD constructions in Italian. On the one hand, Cinque (1977) shows connectivity effects related to subcategorization and case information, where connectivity can be understood as the identity of case and subcategorization between the dislocated element and the gap site.

On the other hand, connectivity can be understood as coreference between two syntactic constituents in a configuration where one element c-commands the other even if at surface structure the bound element is outside the c-command domain of the antecedent. Examples of this can be found in section 2.1.3.3.

2.1.3.1 Pied-piping

In cases of movement, if a preposition is required when the argument is in-situ, then this preposition must be expressed on the moved element. There is no preposition stranding in Spanish and the preposition cannot be dropped at the gap site. It needs to be pied-piped together with the wh-phrase. Contrast the examples in (4a)

and (4c) with the corresponding non-pied-piped counterparts in (4b) and (4d) respectively.

- (4) a. ¿**De quién** estabas acordándote al ver las fotos?
Of who were (you) thinking about when see the pictures
'Who where you thinking about while seeing the pictures?'
- b. *¿**Quién** estabas acordándote **de e** al ver las fotos?
Of who were (you) thinking about when see the pictures
'Of whom where you thinking about while seeing the pictures?'
- c. ¿**De que hijos** está Pedro muy orgulloso *e*?
Of which sons is Pedro very proud
'(Of) His sons, Pedro is very proud of'
- d. *¿**Que hijos** está Pedro muy orgulloso **de e**?
Which sons is Pedro very proud of
'Which sons is Pedro very proud of?'

2.1.3.2 Case: A-marking

Indirect objects and animate direct objects in Spanish are marked for case through the preposition *a* (this is also referred to as a-marking). Therefore, extracted indirect objects and animate direct objects always need to be a-marked as shown in the examples below:

- (5) a. ¿A quién dice Juan que Pedro ha conocido *e*?
PA who says Juan that Pedro has met
'Whom does Juan say that Pedro has met?'
- b. ¿A quién dice Juan que Pedro le ha leído un poema *e*?
To who says Juan that Pedro CL(IO) has read a poem
'To whom has Juan said that Pedro has read a poem?'

If there were no connectivity in Spanish between the base position and the surface position of the wh-phrase, sentences like (6) should be possible sentences in Spanish, but they are not.¹⁰

- (6) a. ??¿Quién dice Juan que Pedro ha conocido *e*?
 PA who says Juan that Pedro has met
 ‘Whom does Juan say that Pedro has met?’
- b. *¿Quién dice Juan que Pedro le ha leído un poema *e*?
 To who says Juan that Pedro CL(IO) has read a poem
 ‘To whom has Juan said that Pedro has read a poem?’

2.1.3.3 Binding facts

Anaphor binding facts are respected in Spanish wh-extractions. In wh-extractions the wh-phrase must be interpreted in its base position so that the anaphor within the wh-phrase can be c-commanded by its antecedent.

- (7) a. ¿Qué foto de si mismo_i dió Juan_i *e* a María?¹¹
 Which picture of himself gave Juan to Mary
 ‘Which picture of himself gave Juan to Mary?’

We can also use Principle C effects as a diagnostic for connectivity, where the R-expression contained in the fronted wh-phrase is not bound in its base position by the pronoun *he* and is locally free.

¹⁰ Direct object extraction cases like (6a) are taken to be grammatical by some Spanish speakers. So it seems the need for a-marking is lost somehow when there is a long-distance wh-extraction. Contrast (6a) with (i) where there is a shorter extraction.

(i). *¿Quién Pedro ha conocido *e*?
 who Pedro has met
 ‘Who has Pedro met?’

¹¹ It has to be noted that some speakers find these anaphor cases ungrammatical.

- (8) ¿Qué foto de Juan_i cree él^{*i/j} que María vió *e*?
 Which picture of Juan *e* thinks Mary saw
 ‘Which picture of Juan Mary thinks he saw?’

Connectivity effects also arise in quantifier variable binding. In the example in (9), only if the *wh*-phrase is interpreted in its base position where the quantifier NP can c-command it, can the proper binding relationship be established for this reading.

- (9) ¿A cuál de sus_i profesores admira *e* cada alumno_i?
 PA which of his professors admires each student
 ‘Which of his professors does each student admire?’

The tests shown with respect to preposition stranding, case and binding show that Spanish *wh*-questions do give rise to connectivity effects as expected. That is, the *wh*-expression in these examples behaves as if it is interpreted in the gap site. Connectivity effects are taken to be canonical of movement¹².

2.1.4 Parasitic Gaps

Parasitic gaps, which are normally taken to be a diagnostic of A-bar movement, are also apparent in Spanish *wh*-questions. A parasitic gap is licensed by an A-bar trace that does not c-command it. This is the case in the Spanish *wh*-constructions in (10a-b).

- (10) a. ¿[A quién_i] [han contratado t_i] [sin entrevistar *e*]?
 PA who (they) have hired without interviewing
 ‘Who have they hired without interviewing?’

¹² But see work on pseudoclefts by Heycock & Kroch (1999), which argues that connectivity effects are not restricted to movement constructions, and therefore dispute this type of argument.

- b. ¿[A qué hombre_i] [han contratado t_i] [sin entrevistar e]?
 PA which man (they) have hired without interviewing
 ‘Which man have they hired without interviewing?’

The gaps (*e*) in (10a-b) are parasitic on the trace of the wh-phrase (*t*) and the sentences are grammatical because the wh-trace does not c-command the parasitic gap.

2.1.5 Weak Crossover

Weak cross over effects are also found in Spanish wh-constructions. In the case of (11c-f), a WCO violation occurs when the wh-phrase crosses over the NP containing the pronoun *su*, whereas in (11a-b) that does not happen because the wh-phrase does not have to cross the NP containing the pronoun *su*. The examples in (11c-e) show that wh-constructions in Spanish exhibit WCO effects.

- (11) a. ¿[Qué niño_i] [t_i quiere a su_i madre]?
 which child loves PA his mother
 ‘Which child loves his mother?’
- b. ¿[Quién_i] [t_i quiere a su_i madre]?
 who loves PA his mother
 ‘Who loves his mother?’
- c. ?¿[A qué niño_i] [quiere su_i madre t_i]?
 PA which child loves his mother
 ‘Which child does his mother love?’
- d. ?¿[A quién_i] [quiere su_i madre t_i]?
 PA whom loves his mother
 ‘To who does his mother love?’
- e. ?¿[A quién_i] [Pedro dice que su_i madre quiere t_i]?
 PA whom Pedro says that his mother loves
 ‘Who does Pedro say that his mother loves?’

In summary, I have shown that Spanish *wh*-interrogatives show the prototypical properties of A-bar movement. These properties found in Spanish *wh*-constructions will be used as diagnostics when examining the properties of both LD and CLLD constructions, and of clitic-less dislocation cases.

2.2 Properties and classification of constructions involving left-dislocated phrases in Spanish

In the remaining two sections, I will analyze the syntactic properties of CLLDs with respect to other constructions that involve a left-dislocated phrase. One of the remaining goals of these sections is to see if the clitic pronoun that is specific to these CLLD constructions is playing a crucial role in their syntactic behavior or whether it is more related to the kind of phrase being topicalized. In the literature it has been argued that CLLD constructions involve binding and not movement (Cinque 1990) and that the non-movement properties of this construction are a direct consequence of the clitic pronoun. In section 2.2.2, I challenge the arguments that center the syntactic behavior of CLLDs in the clitic pronoun by showing that clitic-less left-dislocation cases share syntactic properties with CLLDs constructions.

Through the different syntactic tests used for testing the movement or non-movement properties of constructions (illustrated in section 2.1), I argue that CLLDs are derived by movement and that the clitic does not itself play a defining role in determining the type of chain that is involved in CLLDs. This being the case, I will conclude that CLLDs should be related to the other left-dislocated constructions that

do not have a clitic at the bottom of the chain, and also that the difference in whether there is a remnant clitic pronoun or not is related to the discourse properties of the dislocated phrase- not the kind of derivation at stake.

In the next section, I examine the properties of CLLD and LD, using *wh*-movement as a backdrop. I conclude that CLLD left-peripheral elements are derived via movement, whereas LD left-peripheral elements are base-generated in this position as hanging topics. Clitic-less dislocations will be examined in section 2.2.2. I show that these expressions are derived via movement just like CLLD expressions. Interestingly, the base-generated hanging topic analysis seems to be unavailable in the clitic-less cases.

2.2.1 CLLDs vs. LDs

Spanish shows two main strategies for fronting phrases: Left-dislocation (henceforth LD) and Clitic Left-Dislocation (henceforth CLLD)¹³. These are illustrated in (12a) and (12b) respectively.

- (12) a. **A María**, Pedro **la** quiere *e* con locura. (CLLD)
 PA María, Pedro CL(DO) loves with craziness
 'Maria, Pedro madly loves her'
 b. **María**, Pedro **a ella** la quiere con locura.¹⁴ (LD)

¹³ There is a third kind of strategy called Contrastive Focus, but we will not discuss it here because it involves Focus and a special intonation. For a discussion of this construction the reader is referred to (Casielles-Suarez 2004, Zagana 2002)

(i) Esos ZAPATOS no se pone en verano.
 Those shoes (he/she) not refl. wear in summer
 'Those shoes he/she doesn't wear in the summer'

¹⁴ It has to be noted that in the cases of LD, when the strong pronoun is pronounced it always has an emphatic effect. This characteristic holds for all the LD cases discussed in this chapter.

Maria, Pedro PA she CL(DO) loves with craziness
'Maria, Pedro madly loves her'

The goal of this section is to examine the properties of these two constructions and to determine the nature of the relationship between the fronted element and the gap site. As I will show, the fronted phrase in an LD construction is clearly base-generated, whereas CLLD constructions are arguably derived by movement but do not show all of the characteristic properties of wh-movement.

The discussion will be organized in the following way. First, I introduce the properties specific to each of the constructions. These properties will allow us simply to identify whether a given expression instantiates CLLD or LD. I then focus on the syntactic derivation of each construction. First, I discuss LDs, which are taken to be base-generated, and then, I discuss CLLDs, which have been shown to have both movement and non-movement properties.

The following distinctions between CLLD and LD are based on the accounts that have been given for these constructions both in Italian (Cinque 1977, 1990) and in Spanish (Olarrea 1996, Zagona 2000)¹⁵. I first give a list of the properties exhibited by each construction and then elaborate on several of the properties. The properties of both LDs and CLLDs in Spanish are taken to be the following¹⁶.

¹⁵ The first account given for these constructions in Spanish is Rivero (1980) but she takes the Spanish constructions to behave in a similar way to the English left-dislocation and topicalization cases discussed in Ross (1967) and therefore draws a different kind of distinction from the one we will be using.

¹⁶ Note that these properties can be taken as descriptive, pre-theoretical properties, but that they clearly define two separate constructions in that the properties within each grouping can be shown to interact and pattern together. Thus, there are implicational relationships between the properties as discussed below.

LD Properties:

1. The dislocated constituent must be a non-case marked NP (DP)
2. There can be a coreferential element in the “gap site” which may be an overt phrase, a tonic pronoun or an epithet
3. The dislocated element appears in root clauses only
4. The dislocated constituent might be preceded by *topicalizing expressions*
5. The dislocated constituent doesn’t display grammatical and selectional connectivity with the coreferential element in the “gap site”
6. There can only be one LD dislocated per sentence
7. LDs are not sensitive to strong or weak islands
8. There is necessarily a pause separating the dislocated phrase from the rest of the sentence

CLLD properties:

1. The dislocated constituent is not restricted to NP (DP), it can also be a PP (an a-marked DP)¹⁷
2. There must be a gap and no overt category can appear in the gap site, although typically there is a coreferential clitic¹⁸.
3. The dislocated constituent may appear both in root and embedded clauses
4. The dislocated constituent cannot be preceded by a *topicalizing expression*

¹⁷ This property is related to property (5).

¹⁸ This is the clitic that lends its name to the construction and hence is generally taken to be an essential marker of the construction. I argue below that there are cases of CLLD with no clitic, i.e., left-peripheral subjects, bare nouns, and PPs, none of which has a corresponding clitic in Spanish.

5. The dislocated constituent displays grammatical and selectional connectivity with the coreferential element
6. There can be more than one CLLD dislocated per sentence
7. CLLDs are only sensitive to strong islands
8. No pause is necessary between the dislocated phrase and the rest of the sentence

Below I provide further explanation and examples of these basic descriptive properties for differentiating between LDs and CLLDs.

The form of the dislocated phrase is different depending on whether we are looking at LD or CLLD cases. In the case of LDs, the dislocated constituent must be an NP (DP) whereas in the case of CLLDs, the dislocated constituent is not restricted to DPs and can be an a-marked DP or a PP (as will be argued for further below).

In the LD cases in (13) and (14), only a bare DP can appear in the left-peripheral position. The examples in (13b) and (14b) show that including a preposition or morphological case marking on the DP is incompatible with having an overt tonic pronoun in the gap site (interaction of property 1 with property 2).

- (13) a. **Juan**, no me acuerdo **de él**.
 Juan, not CL remember of him
 ‘Juan, I don’t remember him’
- b. ***De Juan**, no me acuerdo **de él**.
 Of Juan, not CL remember of him
 ‘Of Juan, I don’t remember him’
- (14) a. **Juan, lo** vimos **a él** en la fiesta.
 Juan, CL(DO) saw P him at the party
 ‘Juan, we saw him at the party’
- b. ***A Juan, lo** vimos **a él** en la fiesta.

To Juan, CL(DO) saw P him at the party
'To Juan, we saw him at the party'

In the CLLD cases in (15), however, the dislocated constituent can be an a-marked DP as in (15a).

- (15) a. **A Juan, lo** vimos en la fiesta.
PA Juan, CL(DO) saw at the party
'Juan, we saw him at the party' (Olarrea 1996)
- b. ***A Juan**, vimos en la fiesta.
PA Juan, saw at the party
'Juan, we saw at the party'

Thus, the possibility of having a “dislocated” PP or an a-marked DP correlates with whether the “gap site” contains a full tonic pronoun or whether it contains no overt material. In the LD cases (13-14), the dislocated phrase is associated with an overt tonic pronoun in the gap site (e.g.: de él, a él), while in (15), the coreferential phrase is not overt.

The only overt constituent in the CLLD cases is a clitic. Clitics are possible (and in fact required) for some types of constituents, but not others, as seen in the contrast of (15a) and (15b) above. In (15a) the dislocated constituent corresponds to the direct object of the verb, and the corresponding clitic is present and indeed required. If the clitic is not present in this case the sentence is ungrammatical as in (15b).

Left-dislocation constructions, however, may also have a clitic as shown in (14a). Thus, the presence of a clitic is not a reliable cue to distinguish the two constructions, and there are many examples of sentences where no cues are available to allow us to determine which of the two constructions might be involved. For

example, (16) is ambiguous between an LD and a CLLD case since the fronted DP is inanimate and hence does not exhibit case-marking. In addition, all other properties of this sentence are compatible with either construction. I assume that both analyses are available for this sentence.

- (16) **El libro, lo** compré ayer. (CLLD/LD)
 The book,CL(DO) bought yesterday
 ‘The book, I bought it yesterday’

Because sentences such as (16) lack surface cues that would indicate which construction is involved, and hence are ambiguous, they provide a poor testing ground for identifying other syntactic properties of the two constructions. Consequently, in the remainder of this chapter, I will avoid using examples like (16) and instead use examples which contain one of the surface properties listed so that we have an independent indicator of which construction we are dealing with. For example, I will use animate object DPs such as those in (17a) that exhibit a-marking in CLLDs but not in LDs, and for LD constructions such as that in (17b-c) I include an overt tonic pronoun in the gap site.

- (17) a. **Al bibliotecario, lo** conocí ayer. (CLLD)
 To+the librarian, CL(DO) met yesterday
 ‘The librarian, I met him yesterday’
- b. **El bibliotecario, lo** conocí **a él**¹⁹ ayer. (LD)
 the librarian, CL(DO) met PA he yesterday
 ‘The librarian, I met him yesterday’
- c. **El bibliotecario, él** sí que sabe cómo encontrar cualquier libro. (LD)
 The librarian, he yes that knows how to find any book
 ‘The librarian, he does know how to find any book’

¹⁹ As we have discussed in a previous footnote, the tonic pronouns in these constructions are accompanied by an emphatic intonation.

Finally, another way of identifying LDs with respect to CLLDs is that they can be introduced by “topicalizing expressions”, such as “*hablando de*” (speaking of) or “*con respecto a*” (with respect to) as illustrated in (18a-b).

- (18) a. **Hablando de** Pedro, él no va a poder irse de vacaciones (LD)
Talking about Pedro, he not will be able to go on holidays
‘Talking about Pedro, he won’t be able to go on holidays’
- b. * Te he dicho que **con respecto a** Pedro, no le han dado vacaciones
(CLLD)
I told you that with regard to Pedro, not CL(IO) have given holidays
‘I told you that with regard to Pedro, they haven’t given him holidays’

Here an interaction between property 3 and property 4 is shown.

2.2.1.1 Syntactic derivation of LDs and CLLDs

In what follows I will concentrate on the syntactic analyses of CLLD and LD. It is standardly accepted that LDs involve base-generation rather than movement of the left-peripheral constituent. However, the analysis of CLLD has been much more controversial. I will address the question of whether CLLDs should be considered to be derived by movement or not, and discuss the arguments made in the literature both for and against a movement analysis. The reason that an analysis for CLLD has been so elusive is that according to some diagnostics, such as sensitivity to islands, CLLD looks like a typical case of A-bar movement, but CLLD does not pass all of the diagnostics for A-bar movement, and instead has some properties of A-movement or non-movement (Cinque 1990, McMahon 2002). Taking previous arguments as a starting point, I will argue that the CLLD construction is derived by movement despite the fact that some of its properties (such as not licensing parasitic gaps or not

having weak crossover effects) do not correspond with A-bar movement constructions. I will suggest that the discourse properties of this construction are the source of its hybrid A/A' movement properties and that the clitic is merely a by-product of this hybrid nature of the movement chain.

2.2.1.1.1 Left-dislocation (LD)

I take LDs to be base-generated in a Topic phrase following Hernanz & Bruccart (1987) and Rivero (1990)²⁰. The following facts support a non-movement analysis of this construction: the dependency is not sensitive to islands, the dislocated phrase does not exhibit connectivity effects with the “gap site” and an overt element can appear at the bottom of the chain (besides the clitic).

The fact that LDs are not sensitive to islands is the first piece of evidence for concluding that this construction does not involve movement. In each of the examples in (19) there is a left-dislocated (LD) DP, as evidenced by the lack of a-marking on the DP and by the overt tonic pronoun in the base position (shown in bold). In each case, the dependent pronoun is inside an island, but the examples are grammatical. I illustrate with various types of islands, as labeled.

- (19) a. **Esa mujer**, [el hecho de que Pedro la conozca a ella no tiene
 That woman, the fact of that Pedro CL(DO) knows PA her not have
 importancia]. (Complex NP island)
 importance
 ‘That woman, the fact that Pedro knows her has no relevance’

²⁰ A construction that behaves in a similar way to Spanish LD is what has been called Hanging Topic Left Dislocation (HTLD). For a discussion of this construction in German and references therein see (Grohmann 2000).

- b. **Holanda**, [un amigo [que vive **allí/en ese país**] quiere que lo
The Netherlands, a friend that lives there/in that country wants that CL(DO)
visite. (RC island)
visit.
'The Netherlands, a friend who lives there/in that country wants me to visit
him'
- c. **Esa actriz**, ves un par de películas y te enamoras de **ella**.
(Coordinate Island)
That actress, (you) see a couple of movies and CL(DO) fall-in-love of she
'That actress, you see a couple of movies and you fall in love with her'
- d. **Juan**, que María lo haya conocido **a él** le preocupa a Pedro.
(Sentential Subject island)
Juan, that Maria CL(DO) has met PA he CL(IO) worries PA Pedro
'Juan, that María has met him worries Pedro'
- e. **María**, Juan se desmayó tras conocerla **a ella**. (Adjunct island)
Maria, Juan refl. fainted after meet-CL(DO) PA her
'Maria, Juan fainted after meeting her'

If the dependency between the left-peripheral DP and the pronoun in the argument position (direct object of *conocer* "to know") were the result of movement, the movement would be expected to incur a subjacency violation and thus be ungrammatical. Compare these examples with the parallel but ungrammatical wh-movement examples given in (3).

For completeness, I also give an example of a weak island in LDs. However, since wh-movement is not sensitive to weak islands in Spanish anyway, we wouldn't expect this example to be ungrammatical regardless of whether LD were derived by movement or not. So we cannot conclude anything about island sensitivity in LDs from the example in (20).

- (20) **Ese director**, María no se cree que **él** haya hecho esa película.
that director, Maria not refl. think that he has made that movie
'That director, Maria does not believe that he has made that movie'

The second argument for the non-movement nature of these constructions comes from connectivity facts. That LDs do not show connectivity effects between the dislocated phrase and “gap site” provides evidence for the non-movement analysis of these constructions.

- (21) a. **Pedro**, estaba acordándome **de él** en este preciso instante.
 Pedro, I was thinking about him in this concrete moment
 ‘Pedro, I was thinking of him right now’
- b. **Los hijos**, Pedro está muy orgulloso **de ellos**.
 The sons, Pedro is very proud of them
 ‘The sons, Pedro is very proud of them’
- c. **Pedro**, María sabe que no quiere volver a ver **a semejante caradura**²¹.
 Pedro, Maria knows that not want again see such a cheeky devil
 ‘Pedro, Maria knows that she does not want to see such a cheeky devil again’
- d. **Pedro**, tú sabes que **a él** nadie le regala nada.
 Pedro, you know that PA he nobody CL(IO) gives anything
 ‘Pedro, you know that to him nobody gives anything’

In the examples in (21), the dislocated phrase does not exhibit the preposition (a-b) or case-marker (c-d) that would be required on an in-situ phrase, and which would be expected if the phrase originated within the clause and then moved to the left-peripheral position. Compare these examples to the wh-movement examples in (4), (5) and (6). Recall from section 2.2.1 that adding the preposition or case-marking

²¹ Contrast this with the examples of CLLD in (i-ii) where the fronted phrase needs to match in case with the clitic as shown below:

- (i) A **Pedro**, María sabe que no lo quiere volver a ver.
 PA Pedro, Maria knows that not CL(DO) want again see
 ‘To Pedro, Maria knows that she does not want to see him again’
- (ii) ***Pedro**, María sabe que no lo quiere volver a ver.
 Pedro, Maria knows that not CL(DO) want again see
 ‘Pedro, Maria knows that she does not want to see him again’

results in ungrammaticality when there is a tonic pronoun in the “gap site” (as in examples in (13) and (14)).

LD constructions also do not show connectivity with respect to binding. The examples in (21e-f) show a pronoun in the dislocated phrase, which is interpreted as bound by the subject DP. As the examples show, this binding is impossible.

- (21) e. ??**Sus**_i **hijos**_j, Pedro_i está muy orgulloso **de ellos**_j.
 His sons, Pedro is very proud of them
 ‘His sons, Pedro is very proud of them’
- f. ***Su**_i **novia**, [**cada chico**]_i le regalará flores **a ella**.
 His girlfriend, each boy will give her flowers
 ‘His girlfriend, each boy will give her flowers’

If the dislocated phrase were interpreted in the gap site, the subject should be able to bind the pronoun *su* at that point. However, the examples in (21e-f) are ungrammatical, suggesting that there is no point at which the dislocated element is ever in the gap site. Compare the examples in (21) with the grammatical wh-movement versions in (7), (8) and (9).

Finally, the fact that the referential element at the gap site can be a strong pronoun or an epithet (22a-c), also suggests that there is no movement in these cases.

- (22) a. **Pedro, ese hombre** vuelve loco a todo el mundo. (overt phrase)
 Pedro, that man drives crazy PA all the world
 ‘Pedro, that man drives everyone crazy’
- b. **Diana Krall, ella** sí que me gusta. (tonic pronoun)
 D. K., she yes that CL(IO) (I) like
 ‘Diana Krall, she I really like’
- c. **La actriz**, Aristarain venera a **esa belleza**. (epithet)
 The actress, Aristarain worships PA that beauty
 ‘The actress, Aristarain worships that beauty’

Generally movement is taken to require an empty gap. And the fact that wh-movement in Spanish requires an empty gap (see section 2.1.1.) supports this assumption.

2.2.1.1.2 Other properties of LDs

A property of LDs that I assume to follow from their discourse function is the fact that they can be preceded by “topicalizing expressions”²². These expressions are used to mark elements that are presupposed, background information from a previous discourse and that then are introduced as a new discourse topic. The element does not have to be discourse linked in the sense that it does not have to have been mentioned in the immediately preceding discourse. Moreover, as pointed out by Cinque (1977), the LD sentences with and without the topicalizing expression are equivalent and, since the topicalizing expression is optional, the meaning should not be changed if it is deleted.

- (23) a. **Con respecto a** Pedro, él no va a poder irse de vacaciones.
With respect to Pedro, he not will be able to go on holidays
'With respect to Pedro, he won't be able to go on holidays'
- b. **Hablando del** libro, al final no he podido encontrar ese volumen.
Talking about+the book, at the end not have able find that volume
'Talking about the book, at the end I haven't been able to find that volume'

In addition, there cannot be more than one left-dislocated phrase of an LD type. This is illustrated in (24).

- (24) a. * **Pedro, el alumno**, él no lo ha conocido.
Pedro, the student, he not CL(DO) has met

²² This term is discussed by Cinque (1977) who refers to Rodman's (1974) work. Zagana (2002) and Olarrea (1996) also use the same term.

‘Pedro, the student, he has not met him’

- b. * **Holanda, esa actriz**, allí/ en ese país nunca ha estado ella.
The Netherlands, that actress, there/in that country never has been she
‘The Netherlands, that actress, she has never being there/in that country’

Moreover, as illustrated in (24 a-b), LDs occur only in root clauses and cannot be embedded²³. Compare the examples in (25a-b) to the examples in (25c-d) where the LDs are in root clauses.

- (25) a. *Estoy segura de [que **la actriz**, [Medem quiere contratar **a esa belleza**]]
I am sure that the actress, Medem wants to hire PA that beauty
‘I’m sure that the actress, Medem wants to hire that beauty’
- b. *Me pregunto si [**el libro**, [Pablo se quiere leer **ese culebrón**]]
I wonder if the book, Pablo refl. wants to read that soap-opera
‘I wonder if the book, Pablo wants to read that soap-opera’
- c. **La actriz**, Medem quiere contratar **a esa belleza**
the actress, Medem wants to hire PA that beauty
‘The actress, Medem wants to hire that beauty’
- d. **El libro**, Pablo se quiere leer **ese culebrón**
the book, Pablo refl. wants to read that soap-opera
‘The book, Pablo wants to read that soap-opera’

One last fact that I want to mention about LDs is about the position of the LD constituent. It can be shown that the position for the LD is outside of CP and not inside CP because the LD constituent must appear outside of a fronted wh-expression as reflected in examples (26a-b).

- (26) a. ? Me pregunto, **esa actriz**, por qué Medem quiere contratar **a esa belleza**²⁴.

²³ Suñer (1994) suggests that LDs can be adjoined to subordinate clauses too, but the example that is given could be ambiguous between a LD and CLLD case since the embedded phrase (*the textbooks*) is an inanimate NP and that could also be considered a CLLD.

(i) Briana me dijo que **los libros de texto**, Drea ya los había vendido.
Briana told me that the textbooks, Drea had already sold.
(Hernanz & Brucart 1987)

(I) CL(DO) wonder the actress why Medem wants to hire PA that beauty
'I wonder why the actress, Medem wants to hire that beauty'

- b. *Me pregunto por qué **esa actriz** Medem quiere contratar **a esa belleza**.
(I) CL(DO) wonder why the actress Medem wants to hire PA that beauty
'I wonder why the actress, Medem wants to hire that beauty'

In this section, I have shown that LDs are not sensitive to islands, that they do not show connectivity effects and that there can be an overt element in the "gap site".

Together with other properties of LDs indicating that they cannot be embedded, that there cannot be multiple LDs fronted and that they can be preceded by a "topicalizing expression", the previous evidence suggests that LDs are base-generated and that, most likely, they are generated above CP.

Moreover, we have seen that there are some cases of LDs that are ambiguous between LD and CLLD constructions. These are non-animate DPs, such as that in (16) repeated here as (27).

- (27) **El libro, lo** compré ayer. (CLLD/LD)
The book, CL(DO) bought yesterday
'The book, I bought it yesterday'

2.2.1.2 **Clitic Left-Dislocation (CLLDs): Movement vs. base generated approach**

In the following section, I first discuss the various arguments that have been used to show the movement properties of CLLD (such as being sensitive to island constraints). Then, I discuss the properties in this construction that do not exhibit the properties of wh-movement. These properties are that CLLDs are not sensitive to

²⁴ For this example to sound natural parenthetical intonation is needed.

weak-crossover effects and that the gap in these constructions does not license a parasitic gap. Finally, I argue that CLLD involves movement, but the derivation is different from Wh-movement in that the left peripheral element is not a wh-operator.

Different analyses have been proposed for the derivation of CLLDs with regard to whether this construction involves movement or not. On the one hand, there are analyses that have considered CLLDs in different Romance languages to involve A-bar movement, mainly due to the fact that these constructions are sensitive to strong islands and to the fact they show connectivity effects (Cinque 1977 for Italian, Dobrovie-Sorin 1990 for Romanian, Sportiche 1993 and Kayne 1994 for French, Grohmann 2000 for Greek), and in some cases due to “partial reconstruction” effects found in CLLDs (Cecchetto 2000 for Italian)²⁵.

On the other hand, there are analyses that consider the dislocated phrase to be based generated in its surface position. Under this view, the CLLD phrase is often taken to be a clausal adjunct linked to a clause internal covert pronominal (if there is one) via an A-bar chain (Cinque 1990 for Italian, Contreras 1991, Zagana 2002 and Casielles-Suarez 2004 for Spanish, Iatridou 1995 and Anagnostopoulou 1997 for Greek). Even though these constructions show island and connectivity effects and they have an empty category at the foot of the chain, that they lack Weak Crossover (WCO) effects and that they fail to license parasitic gaps is taken to be a reflection of their non-movement nature.

²⁵ Cecchetto (2000) only focuses on object CLLD cases (DPs) in Italian because those are the CLLD cases where the clitic is required. He only considers the object CLLD cases to be derived by movement.

For CLLDs that involve fronted PPs, where the clitic is non-obligatory in Italian, Cecchetto (2000) suggests that the analysis of these constructions might not involve movement (Cecchetto & Chierchia 1999).

Finally, the third kind of approach is the 2-move analyses (which combine A-movement and A-bar movement within the same analysis) in which the quirky movement properties of CLLD constructions are taken to be due to a combination of both an A-movement at a first stage in the derivation (when it moves within the vP) and A-bar movement at a second stage (when the constituent adjoins to TP). Under this analysis, CLLD and clitic doubling constructions have the same underlying properties, and they only differ in the number of movement operations the fronted NP undergoes in CLLDs (Agouraki 1992 and McMahon 2002 for Greek, Bleam 1996 for Spanish, Cecchetto 2000 for Italian).

2.2.1.2.1 Arguments that CLLD involves Movement

The main properties that make the CLLD resemble wh-constructions are their sensitivity to strong islands and connectivity effects. As illustrated in examples (28a-d) and (29), CLLDs are not sensitive to weak islands (32) but they are sensitive to strong islands (28a-d).²⁶

- (28) a. ***A esa mujer**, el hecho de [que Pedro la conozca *e*] carece de
 PA that woman, the fact of that Pedro CL(DO) knows lacks of
 importancia. (Complex NP island)

²⁶ As discussed in Cinque (1977), the copies left behind by CLLDs that are PPs cannot be inside an island, whereas the copies of NPs in LD constructions are not sensitive to island constraints. It is interesting to note at this point that strong CLLD islands are not as strong violations in Italian as they are in Spanish. As Cecchetto (2000) points out, this is due to the ambiguity of a fronted NP that can either be interpreted as a hanging topic or as a CLLD. This is not the case in Spanish where for the CLLD interpretation, the fronted NP would have to be a-marked by a preposition. Contrast (i) in Italian and (ii) in Spanish.

- (i) ?? Gianni, temo la possibilità che lo arrestino
 Gianni, (I) fear the possibility that (they) CL-(DO) arrest
 ‘Gianni, I fear the possibility that they will arrest him’ (Cecchetto 2000)
- (ii) * A Juan, temo la posibilidad de que lo arresten
 PA Juan, (I) fear the possibility of that (they) CL-(DO) arrest
 ‘To Juan, I fear the possibility that they will arrest him’

importance

‘To that woman, the fact that Pedro knows her is not important’

- b.* A **María**, [alguien [que la alagó *e*]] no se presentó a la fiesta.
(RC island)
PA Mary, somebody that CL(DO) praised not CL(IO) go to the party
‘María, someone who praised her did not go to the party’
- c. ??A **ese actor**, [ves un par de películas] y [te lo
PA that actor, (you) see a couple of movies and (you) CL(IO) CL(DO)
quieres llevar *e* a casa]. (Coordinate island)
want bring home
‘To that actor, you see a couple of movies and you want to take him
home with you’
- d. ??A **Juan**, [que María lo haya conocido *e*] le preocupa a Pedro.
(Sentential Subject island)
PA Juan, that Maria CL(DO) has met CL(IO) worries PA Pedro
‘To Juan, that María has met him worries Pedro’

Sensitivity to strong islands is expected if CLLD constructions involve A-bar movement. Moreover, if these constructions are directly compared with wh-movement constructions, they behave exactly the same with respect to weak island sensitivity –in that they are not sensitive to weak islands either. As I have previously discussed, in comparison to languages like Greek (Grohmann 2000, McMahon 2002), Spanish does not show a real contrast in the acceptability of the extraction out of weak islands, depending on whether the wh-phrases are D-linked constituents.

As illustrated in (29), CLLDs are not sensitive to weak islands:

- (29) A **ese director**, María no cree [que *e* lo hayan premiado por esa película].
(Negative island)
PA that director, Maria not think that CL(DO) have awarded for that movie
‘To that director, Maria does not think that they have given him an award for that movie’

Therefore, whatever accounts for the grammaticality of the wh-movement in (3) (section 2.1.) may also explain the grammaticality of the example in (29).

The second argument for the movement nature of CLLDs comes from connectivity facts. CLLDs show connectivity effects between the dislocated phrase and the clitic pronoun, and in that the fronted element obligatorily manifests the morphological case that would be required of an overt object DP if it appeared in the base (normal object) position as in (30a). Examples (30b-d) show that this is the case in CLLD constructions, where there needs to be case identity between the fronted phrase and the clitic or remnant element.

- (30) a. *María sabe que no quiere volver a ver a Pedro.*
 Maria knows that not (she) want again to see PA Pedro
 ‘Maria knows that she does not want to see Pedro again’
- b. *A Pedro, María sabe que no lo quiere volver a ver.*
 PA Pedro, Maria knows that not (she) CL(DO) want again see
 ‘To Pedro, Maria knows that she does not want to see him again’
- c. *María sabe que a Pedro no lo quiere volver a ver.*
 Maria knows that PA Pedro not (she) CL(DO) want again to see
 ‘Maria knows that to Pedro, she does not want to see him again’
- d. **María sabe que Pedro, no lo quiere volver a ver.*
 Maria knows that Pedro not (she) CL(DO) want again to see
 ‘Maria knows that Pedro, she does not want to see him again’

In contrast, as shown in (31), a-marking is not possible on dislocated elements that are clearly left-dislocations, when there is an overt pronoun in the gap site.

- (31) **A Pedro, María sabe que no lo quiere volver a ver a él.*
 PA Pedro, Maria knows that not CL(DO) want again see PA he
 ‘Pedro, Maria knows that she does not want to see him again’

Connectivity related to subcategorization is also relevant. This kind of connectivity addresses how the preposition that has been fronted and is selected by a verb (or an adjective) that subcategorizes for it has to be governed by this verb or

adjective. This kind of connectivity is related to the pied-piping facts for wh-phrases that we have discussed in section 2.1.3.1.

Examples (32a-b) show that CLLDs need to display this kind of connectivity between the dislocated phrase and its coreferential element. Otherwise the sentences turn out to be ungrammatical as those in (32c-d).

- (32) a. Estaba acordándome **de Pedro** en este preciso instante.
 I was thinking about Pedro in this exact moment
 ‘I was thinking of Pedro right now’
- b. Pedro está muy orgulloso **de sus hijos**.
 Pedro is very proud (of) his sons
 ‘Pedro is very proud of his sons’
- c. ***(De) Pedro**, estaba acordándome en este preciso instante.²⁷
 (of) Pedro, I was thinking about him in this concrete moment
 ‘(of) Pedro, I was thinking of him right now’
- d. ***(De) sus hijos**, Pedro_i está muy orgulloso.
 (of) his sons, Pedro is very proud
 ‘(of) his sons, Pedro is very proud’

Another type of connectivity that CLLDs display is with respect to Binding. The dislocated phrase can be bound by its antecedent even if at surface structure it is outside the binding domain, as shown in examples (33 a-b) below.

- (33) a. **A sus hijos_j**, Pedro_i los quiere por igual.
 PA his sons, Pedro CL(DO) loves equally
 ‘To his sons, Pedro loves them equally’
- b. **A su_i novia**, **cada chico_i** le regalará flores en San Valentín.
 PA his girlfriend, each boy CL(IO) give-FUT flowers in St. V.
 ‘To his girlfriend, each boy will give her flowers in Valentine’s Day’
- c. ***A su_i novia**, [**cada chico_i** que se acuerde] le regalará
 PA his girlfriend, each boy who refl. remember CL(IO) give-FUT

²⁷ Remember that this CLLD is one of the clitic-less cases we are testing.

flores en San Valentín.
 flowers in St. Valentine.
 ‘To his girlfriend, each boy who will remember will give her flowers
 in Valentine’s Day’

If there were no movement, the phrase *a su novia* in (33b) would never be in a position where it could be c-commanded (and thus bound) by its binder *cada chico*. We would expect a weak crossover effect in (33b). But in fact the sentence is grammatical. Contrast the sentence in (33b) with the sentence in (33c), where the binder *cada chico* is not in a position where it can c-command the fronted phrase *a su novia* and the sentence is ungrammatical.

Moreover, as illustrated in (34a), CLLDs can be embedded. Compare the examples in (34a-b) where the sentence is grammatical when the CLLD is embedded in (34a) and ungrammatical when the CLLD is above CP (34b).

- (34) a. Creo que **a la actriz**, [Medem **la** quiere contratar]]
 (I) think that PA the actress, Medem CL(DO) wants to hire
 ‘I think that the actress, Medem wants to hire her’
- b. *Creo **a la actriz**, que Medem **la** quiere contratar
 (I) think PA the actress, that Medem CL(DO) wants to hire
 ‘I think that the actress, Medem wants to hire her’

In this section I have shown that CLLDs display strong island sensitivity and connectivity effects. This strongly suggests that CLLDs involve movement. Moreover, I have also shown that CLLDs can appear in embedded clauses.

2.2.1.2.2 Diffusing Arguments against CLLD being Base-generated

In this section, I review arguments that have been used to reject a movement analysis of CLLDs. We will see that these arguments are not against movement, but

are instead arguments that merely show CLLD movement cases are not completely parallel with Wh-movement cases, due to the fact that they don't involve wh-traces, but a different kind of empty element.

Cinque (1990) has argued that CLLD constructions cannot be derived by movement because the trace left behind is not a variable but a *pro*. Among the reasons he discusses for not considering CLLDs to be derived by movement is the fact that the traces left by CLLDs do not behave like wh-traces.

One of the arguments that Cinque (1990) uses for demonstrating that the clitic in CLLDs is not the overt spell-out of a wh-trace, but is a *pro* instead, is related to the fact that the CLLD construction could be regarded as another instance of clitic doubling. The objection to this view is that CLLDs are witnessed even in languages where clitic doubling is not possible, such as Italian or Greek²⁸²⁹.

The second argument that Cinque (1990) offers is the inability of the clitic to license parasitic gaps. If these constructions were derived via A-bar movement, the trace of the relevant elements should be a wh-trace, and it should be able to license parasitic-gaps. However parasitic-gaps are not licensed by CLLD constructions in Italian or Romanian (see Cinque (1990) for Italian and Dobrovie-Sorin (1990) for Romanian). Moreover, both Cinque (1990) and Chomsky (1982) discuss the inability of clitics to license parasitic-gaps in Spanish. This is illustrated in (35), where the example in (35a) is a parasitic gap and the one in (35b) is not a parasitic gap.

²⁸ Although under a different perspective, it could also be seen as if in these languages overt clitics can't be doubled.

²⁹ For a different perspective see Cecchetto (2000) and his discussion of PG and WCO cases within Clitic Doubling from Rio-Platense Spanish and Romanian.

- (35) a. *Al hombre_i, lo han contratado *t_i* sin entrevistar *e*³⁰
 PA man, (they) CL(DO) have hired without interviewing
 ‘The man, they have hired without interviewing’
- b. Al hombre_i, lo han contratado *t_i* sin entrevistarlo
 PA man, (they) CL(DO) have hired without interviewing-CL(DO)
 ‘The man, they have hired without interviewing him’

Interestingly, inanimate CLLDs do seem to license parasitic gaps. Compare the contrast in (35) with the ambiguous inanimate left-dislocations in (36). As I previously discussed in section 2.2.1.1, these examples are ambiguous between a CLLD and a LD due to the lack of a-marking. The parasitic gap case in this instance seems to be licensed as shown in (36a).

- (36) a. ? El libro_i, lo ha vendido *t_i* sin leer *e*
 The book, CL(DO) has sold without reading
 ‘The book, he has sold it without reading’
- b. El libro_i, lo ha vendido *t_i* sin leerlo
 The book, CL-(DO) has sold without reading-CL(DO)
 ‘The book, he has sold it without reading it’

Related to this piece of evidence, there have also been arguments that have shown that clitics can license parasitic-gaps in Spanish (Bordelois 1986, Suñer & Yépez 1988, Campos 1991, García-Mayo & Kempchinsky 1994, Culicover 2001). Specifically, Campos (1991) argues that parasitic gaps can be licensed by clitics and he provides the following contrast where the inanimate parasitic gap in (37a) is grammatical and the animate one in (37b) is ungrammatical. Interestingly, this correlates with our paradigm in (35) and (36).

³⁰ Campos (1991) argues that the difference with respect to whether a *pro* or an operator is expected in cases like (35a-b) has to do with whether the argument position is bound from an A’ or A-position. For Campos (1991), if the argument position is not A’-bound then, the empty category associated with the clitic must be *pro* and this would account for the impossibility of parasitic gaps. This cannot be the reason for the ungrammaticality of the CLLD case in (35a) if we take the left-dislocation to be in an A’ position.

- (37) a. Lo archivaron t_i sin leer e_i
 CL(DO) file without reading
 ‘They filed it without reading it’
- b. *Lo visitaron t_i sin llamar e_i
 CL(DO) visit without calling’
 ‘They visited him without calling’

That clitics can license parasitic-gaps in Spanish is also attested by García-Mayo (1992) and Rizzi (1986), who also note that only [-animate] clitics may license parasitic-gaps in Italian and Catalan as exemplified in (38a) for Catalan and (38b) for Italian respectively. This stands in sharp contrast with French as exemplified in (39) from Tellier (1991).

- (38) a. El_i vaig cuinar t_i sense posar e_i al forn. (Catalan)
 It (I)have cooked without to-put in+the oven
 ‘I have cooked it without putting it in the oven’
- b. L’ho cucinato t_i senza mettere e_i al forno³¹. (Italian)
 It’(I) have cooked without to-put in+the oven
 ‘I have cooked it without putting it in the oven’
- (39) * Vous l’avez rangé t_i sans avoir lu e_i . (French)
 You it’have put’away without to ‘have read
 ‘You put it away without having read’

The question then is why [-animate] clitics can license parasitic gaps in Spanish, Catalan and Italian³² and why these cases are not possible in French³³. These questions are beyond the scope of this chapter and I leave them for future research.

³¹ For contrasting data on inanimate Parasitic-gaps in Italian see Cecchetto (2000).

³² The reason why there exists a contrast between (37a-b) and why the animate clitic in (37b) does not license parasitic gaps could be related to the fact that clitic doubling only comes with animate NPs. It could be that in the case of (37b) there is a *pro* for the double that prevents the local A’-binder from licensing the parasitic gap.

³³ Garcia-Mayo (1992) attributes the difference between French and the other three languages to the fact the clitic is adjoined to V in French and to IP (therefore higher, in a position to c-command the parasitic gap) in Spanish, Catalan and Italian.

The various issues involved in the acceptability of parasitic-gaps in Spanish make the argument for clitic pronouns not being the overt realization of wh-traces less strong, and thus the parasitic gap tests quite questionable in Spanish. Therefore, the parasitic gaps test is simply not conclusive as an argument in favor of the non-movement nature of CLLDs.

The other diagnostic for A' movement is Weak crossover effects. Cases of WCO arise when the operator binds both a pronoun and a variable, neither of which c-command each other. In (40) there is a pronoun that is co-indexed with the CLLD dislocated phrase intervening between the surface and base positions of the constituent, no WCO effect arises in this case.

- (40) a. A Juan_i, su_i madre lo quiere
PA Juan, his mother CL(DO) loves
'Juan, his mother loves him'
- b. A todo hijo_i, su_i padre lo quiere.
PA all son his father CL(DO) loves
'To all son his father loves him'

The sentences in (40a-b) illustrate that CLLD cases do not show WCO effects. Under this test CLLDs do not resemble movement constructions.

So considering both the PG and the WCO arguments, it seems the main argument for not regarding CLLD to have undergone movement has been the fact that the clitic in CLLDs does not license PGs and that CLLD does not show WCO effects.

The discussion of both PG and WCO effects shows that the trace left by the left-dislocated element in CLLD constructions does not behave like a variable.

As discussed in Iatridou (1995), the fact that the examples in (40) do not show WCO effects means that the empty categories left after the verb in cases like (40) are

not behaving like variables as expected if the construction was derived by movement, but are instead instances of *pro*. The CLLD cases then seem to have no WCO violation because the empty category is not a variable but a *pro*.

Cecchetto (2000), who supports a movement analysis of CLLD constructions, also discusses the lack of PG and WCO effects in these constructions and relates the obligatory presence of the clitic in CLLD object constructions (what he refers to as DPs) to the fact that the traces in these constructions cannot be considered syntactic variables (or *wh*-traces) but an A' bound *pro*. For Cecchetto (2000) the *pro* in the CLLD object cases is only legitimated by the relevant morphology, which is the accusative clitic³⁴.

The question then is whether, even though the trace is surely not the one left by a *wh*-movement as such, we can still consider CLLDs to involve some kind of A'-movement. The reason for the difference in behavior may well be in the displaced element not being a *wh*-operator. The XP in the case of CLLD constructions is surely more referential than a *wh*-phrase, which is presumably why it has some specific properties to it, correlating with the presence of a clitic pronoun (making the displaced element akin to a D-linked *wh*-phrase).

In this section, I have shown that CLLD constructions do not license parasitic gaps when the dislocated phrase is animate but they do license parasitic gaps when

³⁴ However, Cecchetto (2000) also emphasizes how this argument cannot apply to traces left by PPs and argues that that is the reason for a trace left by a fronted PP not to obligatorily co-occur with a clitic. Cecchetto's observation of relating the kind of trace found in these constructions with the presence of the clitic and the clitic having to co-occur obligatorily with the dislocated phrase is particularly relevant for the cases of left-dislocated PPs, Bare Nouns and subjects that we are interested in, since those cases do not have a clitic but still share the same special properties of CLLDs in that they are sensitive to strong islands and they do not license parasitic gaps or WCO effects.

the dislocated element is inanimate, suggesting that this test is not reliable for the relevant constructions. On the other hand, I have illustrated that WCO effects do not arise in these constructions, which makes it a counter-argument for the fact that they are derived by movement.

2.2.2 Clitic-less Dislocations: PPs, Bare Nouns and Subjects

It has been claimed that the presence of the clitic is responsible for the quirky properties of CLLDs and the nature of the chain, which makes these constructions sometimes resemble movement constructions in some ways but not others. The goal of this section is to investigate the role of the clitic and the nature of the chain in CLLD constructions. In order to do this I focus particularly on constructions such as those in (41) through (43), where there is a fronted element but no clitic or other resumptive element in the gap site. In (41), there are examples of dislocated Bare Noun cases (mass nouns (41a), inanimate nouns (41b) and animate nouns (41c)) where there is no resumptive partitive clitic unlike in languages such as Italian or Catalan that show a partitive clitic for those cases. (42) shows examples of Prepositional Phrases (both argument and adjunct PPs) that do not show an overt realization of the locative clitic like Italian or Catalan. Finally, (43) shows an example of a subject dislocation that does not have a resumptive clitic in Spanish either.

- (41) a. **Dinero**, dicen que no tiene *e*. (Bare Noun) (Casielles-Suarez 2004)
Money (they) say that (he) not have
'Money, they say he doesn't have'

- b. **Manzanas**, dicen que no ha cogido *e*. (Bare Noun)³⁵
 Apples, (they) say that (he) not has get
 ‘Apples, they say that he didn’t get’
- c. **Hijos**, la mujer ha dicho que no tiene *e*. (Bare Noun)
 Children, the woman has said that (she) not have
 ‘Children, the woman has said that she doesn’t have’
- (42) a. **Del examen**, nadie ha hablado *e* todavía (PP) (C-S 2004)
 Of+the exam nobody has talked yet
 ‘About the exam, nobody has said anything yet’
- b. **En el jardín**, los niños se divierten mucho *e*. (adjunct PP) (C-S 2004)
 In the garden the children refl. enjoy much
 ‘In the garden, the children enjoy themselves very much’
- c. **En la estantería**, Juan colocó el libro *e*. (argument PP)
 In the bookshelf, Juan placed the book
 ‘In the bookshelf, Juan placed the book’
- (43) **El niño**, parece que *e* está protestando siempre (subject) (Rivero 1980)
 The child it seems that (he) is always complaining
 ‘It seems the child is always complaining’

These constructions resemble CLLD and LD constructions in that there is a left-dislocated phrase and a gap site that is connected to the dislocated phrase, either by movement or otherwise, in both constructions. However, these constructions pose an interesting challenge since, unlike LDs or object CLLD cases, they do not have any kind of resumptive element in the gap site. Hence, these clitic-less cases could

³⁵ We are not going to cover the discussion of singular bare nouns as those in (i) and (ii) respectively in the discussion of Bare Nouns because their distribution in general is much more restricted (Dobrovie-Sorin et al. (in press)).

- (i) ?**Manzana**, siempre como una *e* a media tarde. (Bare Noun)
 Apple, (I) always eat one at mid afternoon
 ‘Apple, I always eat one in the middle of the afternoon’
- (ii) **Sombrero**, siempre lleva *e*. (Bare Noun)
 Hat, (he) always wears.
 ‘Hat, he always wears’

provide more information about the movement/non-movement nature of these two types of left-dislocations (LDs and CLLDs).

If, on the one hand, they are considered to be CLLD cases, they present an interesting issue, which is whether the clitic is a necessary element to CLLDs, or rather whether it is a by-product of the derivation of these constructions. If, on the other hand, they are considered to be LD cases, the question is if there is an empty position at all, since these constructions are assumed to be base generated.

If they do not undergo movement, they should be closer in behavior to LD constructions, whereas if they undergo movement, they would be closer to CLLD constructions, as will be discussed.

In what follows, I compare two hypotheses: (i) (41)-(43) have the same type of structure and discourse functions as CLLD constructions; (ii) these sentences have the same structure and discourse functions as LD constructions; (iii) they don't resemble either CLLD constructions or LDs, meaning that they have their own syntactic and discourse function characteristics.

In the following section I show that these cliticless dislocations most closely resemble CLLD in that they are derived by movement. More specifically, I will conclude that subject dislocations are compatible with both CLLD and LD (an overt element can appear in the gap site showing that it can be an LD, and the dislocation can appear in an embedded clause and with more than one dislocation showing that it can be CLLD). Thus, sentences such as (43) are ambiguous: the same string can be CLLD or an LD, depending on say, the discourse context. PP clitic-less dislocations such as (42) are compatible with CLLDs in that they are sensitive to islands. Finally,

the Bare Noun cases such as (41) are not ambiguous. I show that they are only compatible with a CLLD analysis. Thus, all three types of clitic-less dislocations have an analysis involving movement and will be considered cases of CLLLD.

To support my analysis, I demonstrate that these cases show some of the movement properties found in CLLD cases and that they are in the same position where CLLDs are in the syntactic structure.

2.2.2.1 Island sensitivity effects in BN, PP and subject cases

The island test applied to BN, PP³⁶ and subject cases seems to support the fact that clitic-less left-peripheral phrases behave like CLLDs in that they are sensitive to strong islands but not to weak islands, as shown in examples (44) through (48). The (a) examples illustrate weak islands and the (b) examples illustrate strong islands.

- (44) a. Dinero, no sé cómo se puede saber quién gasta *e*.
 Money, (I) not know how refl. can know who spends
 ‘Money, I don’t know how it can be determined who spends it’
- b. *Dinero, alguien que gasta *e* tendría que ser más precavido.
 Money, somebody who spends should that be more cautious
 ‘Money, somebody who spends should be more cautious’
- (45) a. Manzanas, no sé quién dicen que no ha cogido *e*.
 Apples, (I) not know who (they) say that (he) not has get
 ‘Apples, they say that he didn’t get’
- b. *Manzanas, alguien que come *e* tendría que ser más sano.
 Apples, somebody who eats (he) should that be more healthy
 ‘Apples, somebody who eats should be more healthy’
- (46) a. Niños, no sé cuántos se puede saber quién cuida *e*.

³⁶ Here it should be noted that the argument PPs in (48a-b) work in the same way as the adjunct PPs in (47a-b). In previous literature this distinction has not been made. It is not trivial, particularly if we take accounts such as Cecchetto & Chierchia (1999) into consideration, which do not distinguish between a-marked dative PP and other PPs.

Children, (I) not know how many refl. can know who takes-care-of
'Children, I don't know how many it can be determined who takes care of'

b. *Niños, alguien que cuida *e* tendría que ser más paciente.
Children, somebody who takes-care-of (he) should that be more patient
'Children, somebody who takes care of them should be more patient'

(47) a. En el jardín, no sé cómo se puede saber que los niños
In the garden, (I) not know how refl. can know that the kids
han jugado *e*.
have played
'In the garden, I don't know how it can be known that the kids had played'

b. *En el jardín, los niños que han jugado *e* tendrían que haber recogido
In the garden, the kids that have played should that have gathered
los juguetes.
the toys.
'In the garden, the kids that had been playing should had gathered the toys'

(48) a. En la estantería, Juan se pregunta por qué el bibliotecario colocó el libro *e*.
In the bookshelf, Juan wonders why the librarian placed the book
'In the bookshelf, Juan wonders why the librarian placed the book'

b. *En la estantería, el librero que colocó el libro *e* tendría que
In the bookshelf, the librarian who placed the book should that
haberla limpiado.
have-CL(DO) cleaned
'In the bookshelf, the librarian who placed the book should have cleaned it'

(49) a. El niño, no sé cómo se puede saber cuándo ha llegado *e*.
The child, (I) not know how refl. can know when has arrived
'The child, I don't know how can it be known that he has arrived'

b. ?El niño, conocí a las niñas con las que ha jugado *e*.
The child, (I) met PA the girls with CL(DO) that has played
'The child, I met the girls with whom he played'

Except for the subject case in (49), it seems clear that the examples in (44) through (48) behave exactly like CLLD cases with respect to island sensitivity. Note that if an LD analysis were available for these cases, we should see insensitivity to islands and the (b) examples above should be grammatical. This is, in fact, what we

find in the case of subjects in (49). The example in (49b) is grammatical because there is an analysis for dislocated subjects in which the DP is base-generated in the left-peripheral position in the matrix clause. This possibility is made more explicit in (50), where the base position is filled with a tonic pronoun, thereby forcing the LD analysis. Here, the island violation is obviated.

- (50) El niño, conocí a las niñas con las que él ha jugado
 The child, (I) met PA the girls with CL(DO) that he has played
 ‘The child, I met the girls with whom he played’

On the other hand, we can force a CLLD analysis for the subject-dislocation case by putting the dislocated DP in an embedded clause as in (51a-b).

- (51) a. *Creo que el niño, Juan conoció a las niñas con las que
 (I) think that the child, Juan met PA the girls with CL(DO) that
 ha jugado.
 (he) has played
 ‘I think that the child, Juan met the girls with whom he played’
- b. Creo que el niño, Juan dice que parece cansado.
 (I) think that the child, Juan says that (he) looks tired
 ‘I think that the child, Juan says that he looks tired’

When we force the CLLD analysis, the subject dislocation is sensitive to islands, as expected (compare this with (51b) which shows that embedding the dislocated element is fine if there is no island). This option of analyzing the dislocation as LD is not available for Bare Nouns or PPs, as shown by the strong ungrammaticality of the relative clause islands in (44-48b). Thus, we have shown evidence that clitic-less dislocations involving bare nouns, PPs, and subjects are derived via movement. For subjects there is also a non-movement LD derivation. For Bare nouns and PPs there is no non-movement option.

In the next sections, we establish that, in addition to being derived by movement, bare nouns and PPs appear in the same position in the left-periphery as CLLD elements and not in the position of LD elements.

2.2.2.2 Position of BNs and PPs with respect to the complementizer

As discussed in section 2.2.1, one difference between LD and CLLD is their positions within the left-periphery. LD elements are high in the left periphery (above CP), whereas CLLD elements are low (below CP). As exemplified in (52b) and (53b), CLLDs can only appear below the complementizer and within the CP, but never outside. Whereas (52c) shows that LDs can be above the complementizer and outside CP, but never inside the CP³⁷ as in (52d).

- (52) a. Me pregunto por qué a María Juan la quiere. (CLLD)
 (I) CL(DO) wonder why PA Mary Juan CL(DO) loves
 ‘I wonder why Mary Juan loves’
- b. *Me pregunto a María, por qué Juan la quiere. (CLLD)
 (I) CL(DO) wonder PA Mary why Juan CL(DO) loves
 ‘I wonder Mary why Juan loves’
- c. María, por qué Juan la quiere a ella. (LD)
 Mary, why Juan CL(DO) loves PA her
 ‘Mary, why does Juan loves her’

³⁷ Suñer (1994) shows that LDs are adjoined to CP in matrix but not in embedded sentences.

- (i) *Mara me dijo a ese hijo de puta, quién lo iba a invitar.
 Mara to me said that bastard, who he was going to invite
 ‘Mara asked me that bastard, who would invite him’
 (Hernanz and Brucart 1987)
- (ii) Las acelgas, ¿quién las detesta?
 The spinach, who CL(DO) detests
 ‘Spinach, who detests them?’
 (Hernanz and Brucart 1987)

d. *Por qué, María, Juan la quiere a ella. (LD)
 Why, Mary, Juan CL(DO) loves PA she
 'Why Mary Juan loves her'

(53) a. Creo que a María, Juan la quiere. (CLLD)
 (I) think that PA Mary, Juan CL(DO) loves
 'I think that Mary, Juan loves'

b. *Creo a María, que Juan la quiere. (CLLD)
 (I) think PA Mary, that Juan CL(DO) loves
 'I think Mary, that Juan loves'

When BNs and PPs are in the left-periphery, they can only appear below the complementizer as illustrated in (54a) and (55a) and not above the complementizer as illustrated in (56) and (57).

(54) a. Creo que dinero, Juan gasta *e*. (BN)
 (I) think that money, Juan spends
 'I think that money, Juan spends'

b. *Creo, dinero, que Juan gasta *e*.
 (I) think, money, that Juan spends
 'I think that money, Juan spends'

(55) a. Me pregunto, por qué dinero Juan gasta *e*.
 (I) CL(DO) wonder why money Juan spends
 'I wonder why Juan money spends'

b. *Me pregunto dinero, por qué Juan gasta *e*.
 (I) CL(DO) wonder money why Juan spends
 'I wonder why money Juan spends'

c. *Dinero, por qué Juan gasta *e*.
 Money, why Juan spends
 'Money, why Juan spends'

d. Por qué, dinero Juan gasta *e*.
 Why money Juan spends
 'Money, why Juan spends'

(56) a. Creo que en la estantería, Juan coloca un libro *e*.
 (I) think that in the shelf, Juan puts a book
 'I think that in the shelf, Juan puts a book'

- b. *Creo, en la estantería, que Juan coloca un libro *e*.
 (I) think in the shelf, that Juan puts a book
 ‘I think that in the shelf, Juan puts a book’
- (57) a. Me pregunto, por qué en la estantería Juan coloca un libro *e*.
 (I) CL(DO) wonder why in the shelf Juan puts a book
 ‘I wonder why Juan in the shelf puts a book’
- b. *Me pregunto en la estantería, por qué Juan coloca un libro *e*.
 (I) CL(DO) wonder why in the shelf Juan puts a book
 ‘I wonder why in the shelf Juan puts a book’
- c. *En la estantería, por qué Juan coloca un libro *e*.
 In the shelf, why Juan puts a book
 ‘In the shelf, why Juan puts a book’
- d. Por qué, en la estantería, Juan coloca un libro *e*.
 why in the shelf Juan puts a book
 ‘In the shelf, why Juan puts a book’

(54) and (57) provide evidence that dislocated BNs and PPs are derived in the same way that CLLDs are in that they appear in the same position as CLLD elements.

2.2.2.3 Multiple CLLD occurrences

An additional argument that the cliticless BN dislocations are a case of CLLD and not LD comes from examining combinations of multiple dislocations. CLLDs can be recursive, which means that in Spanish we can have several CLLDs in any order in the same sentence as reflected in (58).

- (58) a. [A mi novio]_{CLLD} [a mi madre]_{CLLD} no se lo
 PA my boyfriend PA my mother not CL(IO) CL(DO)
 he presentado todavía.
 have introduced yet
 ‘I have not introduced my boyfriend to my mother yet’
- b. [A mi madre]_{CLLD} [a mi novio]_{CLLD} no se lo

PA my mother PA my boyfriend not CL(IO) CL(DO)
 he presentado todavía.
 have introduced yet
 ‘I have not introduced my boyfriend to my mother yet’

Contrast this with multiple LD cases (and cf. (24)):

(59) ***[El organizador]**_{LD}, **[el conferenciante]**_{LD}, él no lo ha presentado
 The organizer, the presenter, he not CL(DO) has introduced
 ‘The organizer, the presenter, he has not introduced him’

The question now is how the combination of fronted LDs and CLLDs behaves and whether there are any order restrictions associated with them. Consider the data in (60) and (61) below. In (60) we see that the combination of a CLLD with a LD is ungrammatical when the CLLD precedes the LD.

(60) a. ***[A mí]**_{CLLD} **[Diana Krall]**_{LD}, ella sí que me gusta
 PA me Diana Krall , she yes that CL(IO) pleases
 ‘Diana Krall, I like her very much’
 b. ***[A los estudiantes]**_{CLLD}, **[Pablo]**_{LD}, él sí que los asesora.
 PA the students Pablo, he yes that CL(DO) advices
 ‘Pablo really advices the students’

Contrast these examples with those in (61) that have the opposite order (LD+CLLD):

(61) a. **[Diana Krall]**_{LD}, **[a mí]**_{CLLD}, ella sí que me gusta.
 Diana Krall PA me, she yes that CL(IO) pleases
 ‘Diana Krall, I like her very much’
 b. **[Pablo]**_{LD}, **[a los estudiantes]**_{CLLD}, él sí que los asesora.
 Pablos PA the students, he yes that CL(DO) advices
 ‘Pablo really advices the students’

As seen in the previous discussion, multiple CLLDs are possible and only one LD per sentence is possible; if a CLLD and an LD appear in the left-periphery of the

same clause, the LD must precede the CLLD³⁸. (This makes sense because, as seen above, the LD constituent is higher in the structure than the CLLD constituent).

We can now use these facts to test the clitic-less BN/PP cases. If it is true that the dislocation of a PP or a bare NP is necessarily an instance of CLLD, then we expect to be able to combine it with another instance of CLLD in any order in the same sentence.

If, on the other hand, dislocations of PPs and bare NPs were cases of LD, we would not expect CLLDs to precede the bare NPs or PPs in the same sentence. That is, the following order would be impossible: *CLLD+BN and *CLLD+PP. As exemplified in (62) and (63), it is possible to have a dislocated element with an overt clitic and a dislocated BN with no overt clitic in the same sentence.

- (62) a. [**A mí**]_{CLLD} [**dinero**]_{BN} Juan nunca **me** deja.
 PA me money Juan never CL(IO) lends
 ‘Juan never lends me money’
- b. [**Dinero**]_{BN} [**a mí**]_{CLLD}, Juan nunca **me** deja.
 money PA me Juan never CL(IO) lends
 ‘Juan never lends me money’
- (63) a. [**A mí**]_{CLLD} [**en la estantería**]_{PP} Juan nunca **me** ordena los libros.
 PA me in the shelf Juan never CL(IO) organizes the books
 ‘Juan never organizes the books in the shelf for me’
- b. [**En la estantería**]_{PP} [**a mí**]_{CLLD}, Juan nunca **me** ordena los libros.
 In the shelf PA me Juan never CL(IO) organizes the books
 ‘Juan never organizes the books in the shelf for me’

Crucially, as shown in (56a) and (57a), the cliticless dislocations (BN and PP) can occur to the left of a CLLD within the same clause. Because only CLLD constituents, and not LD constituents, can occur in this position. I take this fact to be

³⁸ See Cinque (1977: fn 13) and Olarrea (1996: 69) who also point out that the combination *CLLD LD is not grammatical. Whenever both types of dislocation are possible the LD constituent must precede the CLLD.

another argument that they are cases of CLLD and not of LD³⁹.

The fact that these dislocated bare nouns can appear in any order intermingled with other cases of clitic left-dislocated elements constitutes proof that we are dealing with the same structure, a CLLD.

Given that the results of the tests I have applied to BN and PP clitic-less cases all point to the fact that BNs and PPs resemble CLLDs in behavior, I conclude the following about the BN and PP cases. The fact that they are sensitive to islands constitutes evidence for considering the clitic-less BN and PP cases to be derived by movement. Moreover, the evidence that they can be embedded and that they can be combined with CLLDs leads us to think that they are in the same structural position as CLLDs. The question, as in the CLLD cases, is whether the position where they sit is an A' or A-position, and what the precise derivation is that brought them to this position.

2.2.2.4 The clitic system of other languages has BN and PP clitics

We now address the question about why is it that clitic-left-dislocation can occur without a clitic. One thing to note is that Spanish doesn't have clitics for dislocated Bare Nouns, PPs and subjects, but other Romance languages such as Italian and Catalan do have partitive and locative clitics for these dislocations.

Spanish lost the partitive clitic that a dislocated bare noun triggers in other Romance languages such as Catalan, as the example in (64) shows.

(64) **(De) diners** diuen que no **en** té. (Catalan)
(of) money (she) says that not cl(partitive) have

³⁹ The other order is possible and is consistent with CLLD but also with LD.

‘She says she doesn’t have any money’ (Vallduví 1988)

The same loss in the Spanish clitic system applied to the dislocation of PPs. In languages like Catalan, which retained this kind of clitic an overt clitic appears, whereas in languages that have lost it, the CLLD of a PP shows no clitic. Compare the Catalan examples in (65) and (66) with the Spanish counterparts in (67) and (68).

(65) **De l’examen**, ningú no n’ha parlat encara. (PP - partitive)
Of the’exam nobody not **CL-(partitive)**’have talked yet
‘About the exam, nobody has said anything yet’
(Hernanz and Bruccart 1987)

(66) **Al jardí**, els nens s’hi diverteixen molt. (PP - locative)
In+the garden the children refl.’**CL-(locative)** enjoy much
‘In the garden, the children enjoy themselves a lot’
(Hernanz and Bruccart 1987)

(67) **Del examen**, nadie ha hablado *e* todavía. (argument PP)
Of+the exam nobody have talked yet
‘About the exam, nobody has said anything yet’

(68) **En el jardín**, los niños se divierten mucho *e*. (adjunct PP)
In the garden the children refl. enjoy much
‘In the garden, the children enjoy themselves a lot’

Thus, the absence of a clitic in the examples in (44)-(49), (67) and (68) could be due to the incomplete clitic system of modern Spanish, which has not preserved the type of clitics that are triggered in other languages when a BN or a PP are dislocated.

The fact that this clitic is still apparent in other languages suggests that in these languages there is no distinction between the constructions we are exploring in this section and CLLDs. Even if left-dislocated phrases involving PPs and BNs do not show a clitic overtly in languages like Spanish, the fact that in other languages there is no distinction between CLLD and clitic-less cases suggests that these constructions

are derived in the same manner. Whether there is a clitic or not at the bottom of the chain is secondary to the derivation of these dependencies and it will depend on whether the language still preserves a clitic realization for the fronted phrase.

2.2.2.5 Parasitic Gaps

We have established that bare noun and PP dislocations are instances of CLLD and not LD, and also that subject dislocations have a possible CLLD analysis (and the LD analysis is also available). We are now in a position to ask whether the special properties (the quirky A-bar movement properties) of CLLD are due to the presence of the clitic itself. Cinque (1990) and others have posited that the presence of the clitic directly determines the properties of the chain, forcing the bottom of the chain to be a *pro* rather than a *wh*-trace. As discussed above, I subscribe to the idea that the bottom of the CLLD chain is not a *wh*-trace; however, it is not as clear that it is the clitic itself that is responsible for the properties of CLLD. Recall that CLLD resembles *wh*-movement in all ways except that, unlike *wh*-movement, CLLD does not license a parasitic gap and it does not exhibit weak-crossover effects. If these properties (what I am calling the “special properties” of CLLD) are due to the clitic, then we should find that the clitic-less cases we are considering here would behave more like *wh*-movement and not exhibit the special properties. In fact, what we find is that (at least with respect to parasitic gaps) the bare noun clitic-less cases behave exactly like CLLD. The following example shows that dislocated bare nouns do not license parasitic gaps⁴⁰.

⁴⁰ Unfortunately, we cannot test PPs since PPs are known not to license parasitic gaps in any case. As Pesetsky points out in Chomsky (1982), parasitic gaps are impossible if a PP is moved to COMP and the parasitic gap is an NP.

- (69). *Dinero, han buscado sin encontrar *e*
 Money, (they) have searched without finding
 ‘Money, they have looked for without finding it’
- (70). *Niños, han cuidado sin conocer *e*.
 Children, (they) have looked after without knowing.
 ‘Children, they have looked after them without knowing them’
- (71). *Manzanas, han traído sin comer *e*.
 Apples, (they) have brought without eating
 ‘Apples, they have brought without eating them’

The BN and PP cases behave similarly to CLLDs in that they do not license parasitic gaps.⁴¹ Given that bare noun and PP dislocations can only receive a CLLD analysis, which involves derivation by movement, this example provides evidence that the special properties of CLLD (such as not licensing parasitic gaps) are not due to the presence of an overt clitic. Here we see a case of movement that does not license a parasitic gap, but there is no overt clitic. Thus, we conclude that the clitic itself is not responsible for the special properties of CLLD.

-
- (i) * A book from which I copied *t* without buying *e*
 It has to be noted, though, that the cases we would like to consider are all cases of PP parasitic gaps as those in (ii) through (iv).
- (ii). ??En la estantería, han colocado un libro sin hacer sitio *e*.
 in the shelf, (they) have placed a book without making some space.
 ‘In the shelf, they have placed a book without making some space’
- (iii). *En el jardín, han jugado los niños sin pisar *e*.
 In the garden, (they) have played the children without step on
 ‘In the garden, the children have played without stepping on’
- (iv). *Del examen, nadie ha hablado sin olvidarse *e*.
 Of the exam, nobody has talked without forgetting
 ‘Of the exam, nobody has talked without forgetting’

⁴¹ It should be noted that LDs do not license a parasitic gaps.

- (i). *El hombre_i, lo han contratado **a él** sin ver *e_i*.
 The man, CL-DO have hired **PA he** without seeing
 ‘The man, they have hired him without seeing’

Thus, this example by itself does not constitute an argument that bare nouns should be analyzed as CLLD. However, I have shown independently in the previous sections that bare noun dislocations only have a CLLD analysis. Therefore, the fact that the BN does not license a parasitic gap does show that the special properties of CLLD carry over to the bare noun cases even in the absence of a clitic.

2.2.2.6 Future Research

The next question to ask is regarding the differences of these clitic-less cases with respect to CLLD cases. If they are derived in a similar way and if they sit in the same structural position as CLLDs, then we should wonder why these cases do not show a clitic pronoun in the empty position and whether that is dependent on the referentiality of the fronted phrase. Is the status of the bare noun that of a wh-operator or should it rather behave like the XP in a CLLD case? Is the trace of the moved BN or a moved PP a *pro* or does it behave like a wh-trace? If it does not show the same referentiality as the CLLD cases through the appearance of an overt element at the bottom of the chain, then it might be the case that this kind of movement resembles regular wh-movement. These are questions I would like to address in future research.

2.2.2.7 Implications for Parsing

In the following section, I discuss the implications of what I have concluded for the syntax of CLLD, LD and clitic-less dislocations with respect to parsing.

2.2.3 Parsing CLLD constructions

In chapter 3, I present evidence for how CLLD constructions are processed and I show that long-distance dependencies with a clitic at the foot of the chain⁴² (or

⁴² In fact, the clitic is part of the chain but it is not strictly at the bottom of the chain. For simplicity of the argument, I will be referring to the position of the clitic throughout the thesis as at the bottom of the chain.

CLLDs) and dependencies that end in an unpronounced gap are treated the same in terms of the active search strategy⁴³. We will conclude that the parsing mechanism to interpret these two kinds of dependencies is the same and that the form of the bottom of the chain (e.g. gap vs. pronoun) does not affect the activation of this mechanism.

As I discuss in chapter 3, the need to interpret the left-dislocated phrase in CLLD constructions will trigger the activation of a search for a clitic pronoun.

2.2.3.1 Parsing ambiguous CLLD/LD constructions

Since LDs do not involve movement and some of the [-animate] dislocations such as (75) are ambiguous with CLLD cases, these sentences raise interesting questions with respect to parsing predictions for these constructions.

- (72) **El libro, lo** compré ayer. (CLLD/LD)
The book,CL(DO) bought yesterday
'The book, I bought it yesterday'

These are the cases that might be more complicated for the parser in making on-line predictions since the parser would find the left-dislocated phrase to be ambiguous between a CLLD and a LD. This entails that the parser's expectation after processing the fronted phrase would not only be restricted to a clitic pronoun but would also include any of the other constituents that can appear at the gap site in LDs (e.g. epithets, tonic pronouns or overt phrases).

⁴³ A term that will be introduced in chapter 3.

2.2.3.2 Parsing clitic-less cases

The fact that these constructions that do not show any overt element at the bottom of the dependency behave like CLLDs makes stronger the hypothesis of having an active search mechanism for constructions that either involve gaps or pronouns. If both kinds of dependencies were derived in the same way (e.g. by movement), then we would expect they would be processed similarly. This means that there should be an active search for an empty category triggered as soon as the parser recognizes the dislocated phrase as having the specific properties of a BN or a PP (and maybe a subject).

2.3 Conclusion

In this chapter I have given an account for those left-dislocations involving PPs, Bare Nouns and subjects which do not show a clitic pronoun and I have tried to relate them to CLLD cases. This was done by showing that both constructions might be sharing the same kind of properties that are specific to movement constructions. Moreover, this evidence suggests that the clitic is not the reason for why CLLDs have both movement and non-movement like properties, but rather the referentiality of the dislocated XP. I have first given the basic properties of Wh-questions in Spanish both for D-linked and non D-linked questions. Then I have given a clear distinction of LD and CLLD constructions and their properties. Then I have covered the different reasons why these constructions have been said to undergo movement or not and why CLLDs might have such a hybrid nature. Finally, I have tried to compare those hybrid

properties of CLLDs with the clitic-less cases. I have concluded that BN cases behave like CLLD cases and that they both should be considered left-dislocations with the same properties. The difference with respect to whether a clitic is found or not is secondary and it is only due to the nature of the fronted phrase. Finally, LD cases are clearly considered hanging-topics, which do not share any property with either CLLDs nor BNs.

We have seen that the arguments used to argue for the non-movement nature of CLLDs are not intrinsically about the movement/non-movement nature of these constructions, but about the kind of trace found in these constructions. One of the main things I would like to emphasize is this: the fact that the clitic is not the overt realization of a *wh*-trace in CLLD cases does necessarily imply that these constructions do not involve movement. All the tests seem to indicate that these constructions involve some kind of movement, even if it cannot be directly identified as *wh*-movement. The clitic, on the other hand, is taken to be a by-product of the nature of the phrase that has been dislocated and does not seem to be an intrinsic characteristic related to movement (the clitic is compatible with both movement and non-movement).

The implications of the properties of these constructions for their processing and the active search mechanism have been discussed in section 2.2.2.7. I have shown that LD and clitic-less cases are conflictive when trying to predict how the parser predicts the upcoming element that interprets the dislocated phrase in these constructions. In the case of LDs, the challenge for the parser is to know what element to expect (e.g. a clitic or a strong pronoun) in the inanimate cases where CLLDs and LDs overlap, whereas in the case of BNs or PPs, the challenge is in

determining that there is a null element just from the form of the dislocated phrase. This, together with some of the syntactic questions stated above, are open questions that will be covered in future research.

CHAPTER 3: ACTIVE-SEARCH FOR CLITIC PRONOUNS IN LEFT-DISLOCATED CONSTRUCTIONS IN SPANISH

3.1 Processing long-distance dependencies

The processing of structural dependencies is an important aspect of sentence comprehension that requires relating two elements in the sentence that might not be in a local relation with each other. Processing of elements that are in a long-distance dependency requires a number of resources from the parser. First, it requires the parser to maintain the un-interpreted information in memory and second, it requires it to determine the specific point in the sentence at which the element that started the dependency can be interpreted in the course of comprehension.

The central interest of this chapter is to examine the processing of Spanish structural dependencies that involve a left-dislocated noun phrase and a clitic pronoun that co-refers with it. The goal is to explore possible parallels between the processing of these constructions and that of wh-dependencies. Specifically, I test whether the active search mechanism found in the processing of wh-dependencies is also applied during the processing of long-distance dependencies between a left-dislocated NP and a clitic pronoun. If this active mechanism is also evident in long distance dependencies that have a clitic pronoun at the foot of the chain, then this mechanism could be taken to be a more general property of the parser, and not a mechanism that is only apparent in wh-constructions.

A number of previous studies have focused on the processing of long-distance dependencies involving wh-phrases and have tested how a wh-phrase that was distant from the gap was interpreted on-line (e.g., Crain & Fodor 1985, Stowe 1986, Frazier 1987, Aoshima et al. 2004, Lee 2004). These studies focused on the gap search mechanism that is triggered after encountering a fronted wh-phrase, and examined where in the processing of the sentence the gap could be posited. The general consensus after several studies is that there is an active gap search mechanism involved in the interpretation of long-distance wh-dependencies. This entails that the parser posits a gap at every position in the bottom-up input where it is grammatically licit to do so and that the parser uses the bottom-up information available in the sentence to predict the gap at every available grammatical position.

The focus of this chapter is to examine the processing of structural dependencies that involve a left-dislocated phrase and a clitic pronoun that co-refers with it and to test whether the parser uses an active search mechanism to process long-distance dependencies that involve overt clitic pronouns similar to the one found in the processing of wh-dependencies. I report results from an off-line grammaticality questionnaire and two self-paced reading experiments in Spanish that involve clitic left-dislocation constructions. I show that an active dependency building mechanism is implicated in the processing of these constructions. Upon encountering a left-dislocated NP the parser triggers an active search for an associated clitic in the upcoming input. This implies that the active search mechanism is a more general property of the parser and not a mechanism that is only specific to dependencies that involve gaps. It could be thought that since gaps are more difficult to identify in a bottom-up manner this mechanism might only be apparent in constructions that have

a gap at the foot of the chain. In this chapter, I show that this is not the case and that this active search mechanism is also triggered in long distance dependencies involving clitic pronouns. This suggests that the active dependency formation mechanism is a general mechanism in the human processor.

3.1.1. Gap as-a-first-resort vs. gap as-a-last-resort

In an influential paper, Fodor (1978) presented a pair of alternative hypotheses about how the search for a gap might proceed. The proposals that were suggested initially by Fodor (1978) differed in whether the search for a gap was a last or a first resort strategy in the parser, where the gap-as-last-resort hypothesis predicted that the gap would not be posited until clear confirmation is found in the input about the presence of an empty argument slot. The gap-as-first-resort hypothesis predicted that the gap would be hypothesized at every possible grammatical position in the sentence without waiting for confirming evidence. These two hypotheses were also named the gap-driven versus filler-driven hypotheses.

Fodor (1978) discussed the two hypotheses for gap search and argued that the gap-as-last-resort strategy did not reflect the parser's gap-finding routines because it would not predict temporary ambiguities for gap location within the language. If the parser only considered positing a gap when there is clear confirmation of there being an empty argument slot, (which in a sentence like (1) will be after the verb *write*), then doubtful gaps (gaps that are temporarily possible) such as the gap found after the verb *say* in (1) would not occur because the parser would not find clear evidence to posit a gap there and therefore there would not be any temporary ambiguities in

language. The gap-as-last resort hypothesis would therefore disregard the fact that in a sentence like (1) people temporarily assign the PP *to whom* as an argument of the verb *to say*.

(1) To whom did the father say ____that he was planning to write____?

The gap as first-resort strategy could have many variants according to whether the gap would be posited based on the lexical expectations created by the verb, the strictness in the application of grammatical constraints on parsing and on how much of the bottom-up information is used to confirm the postulation of this gap. The variant that has received most attention is the Active Filler Hypothesis.

Fodor raised these two hypotheses and the experimental records subsequently confirmed the gap-as-first-resort option as the most plausible one (Frazier 1987, Frazier & Flores D'Arcais 1989). More specifically, the gap as first-resort strategy predicted that if it turns out that there is a lexical NP in the first position where the gap is posited, there should be some processing difficulty. This effect had actually been previously demonstrated in experiments with long-distance dependencies involving *wh*-phrases with what was called the filled-gap effect (Crain & Fodor 1985, Stowe 1986).

3.1.2. Classic Filled-Gap Effects

The first studies to show that the parser posits a gap at the first available position after processing a *wh*-phrase were conducted by Crain and Fodor (1985) and by Stowe (1986).

Crain & Fodor (1985) showed in a self-paced reading experiment that the reading times at the wh-question sentence in (2a) for the noun phrase *us* occupying the potential gap position were longer than the reading times for the same noun phrase in its declarative counterpart (2b) where no gap is predicted. This could only be possible if the parser is constructing a gap as soon as it encounters an appropriate verb and it is surprised at encountering an overt NP in the position where it was expecting a gap.

- (2) a. Who had the little girl expected **us** to sing those stupid French songs for __ at Christmas?
- b. The little girl had expected us to sing those stupid French songs for Cheryl at Christmas.

Stowe (1986) showed that the same ‘filled-gap effect’ is found in an embedded wh-question like (3a) at the noun phrase *us* with respect to its baseline condition in the non-extraction (3b).

- (3) a. My brother wanted to know who Ruth will bring **us** home to __ at Christmas.
- b. My brother wanted to know if Ruth will bring us home to Mom at Christmas.

Both studies suggested that the slowdown at the noun phrase *us* in (2a) and (3a) reflected reader’s expectation for a gap rather than for an overt noun phrase, when there is an uninterpreted wh-filler in the sentence. There, finding the overt noun phrase instead results in a processing difficulty. They showed that the parser constructs object gap sites at least by the time it encounters an appropriate verb, and that it does not wait for confirming evidence about the gap location. Subsequently, evidence for gap postulation has been found in different studies (see for example

Frazier & Clifton 1989, Traxler & Pickering 1996, Aoshima et al. 2004, Lee 2004, Phillips et al. 2005).

3.1.3. Active Filler Hypothesis

Following work done by Frazier and others in Dutch (Frazier 1987b, Frazier & Flores D'Arcais 1989) to test if the parser expected a gap at the first available position, Clifton & Frazier (1989) and Frazier & Clifton (1989) proposed a version of the gap as the first resort hypothesis named the Active Filler Hypothesis (henceforth AFH).

Active Filler Hypothesis: (Clifton & Frazier 1989)

When a filler of category XP has been identified in a non-argument position, such as COMP, rank the option of assigning its corresponding gap to the sentence over the option of identifying a lexical phrase of category XP.

The AFH predicted that readers actively construct dependencies on-line and do not wait to see them confirmed at the gap position. The identification of a filler or a wh-phrase that is case-marked induces a special mechanism in the parser, which automatically starts searching for a gap to complete its interpretation at the thematic position. AFH emphasizes the active gap creation of the parser once a displaced wh-phrase has been processed and assumes that gaps will be posited even if there is no direct evidence of an unfilled argument slot. The active nature of this parsing mechanism entails that it will try to posit a gap at every possible grammatical gap

position, and, only when encountering evidence disconfirming the appearance of the gap will it consider the next available gap position in the sentence. This active search for a gap will go on until the definite gap position is encountered and the uninterpreted filler is finally interpreted. By interpreting it we mean that the phrase that was displaced can finally get its semantic role assigned.

One of the remaining questions is whether this active search mechanism applies for every kind of long-distance dependency. Would the active search be active in a long-distance dependency where the identification of a filler requires a pronoun instead of a gap? If active filling is triggered by the appearance of an item in an A position (i.e., a position to which a theta-role can be assigned), from which it cannot be interpreted, we would expect the AFH to be triggered in the same way whether or not the displaced constituent is linked to a clitic pronoun or to a gap.

But, what would it mean to actively construct a phonetically overt element and how does that differ from actively constructing a phonetically null element? If a displaced constituent is linked to a gap, when the gap that has been actively looked for is encountered in the input, this leaves very few things for the parser to do. If the displaced constituent is linked to a clitic pronoun, it is hard to determine whether the process of interpretation at the position where the clitic is confirmed is the same as in the case of gaps, since an overt element might require additional steps for its interpretation.

If this were the case, we would expect processing facilitation at the clitic pronoun position when it is finally encountered and a slower reading time in the cases where a clitic pronoun could have been predicted but where the prediction was not verified by encountering an actual pronoun. This is the focus of experiment 1, where

whether the search for clitic pronouns in topicalization dependencies in Spanish is as active as the search for gaps in wh-dependencies is tested.

I show that the AFH applies in these cases. Given this, I then ask what information is required to license a pronoun in these dependencies. Since the clitic pronoun comes before the verb in Spanish left-dislocation constructions, we cannot use verb argument structure information to constrain the choice of clitic positions between the topicalized, or left-dislocated position and the clitic site. Spanish is an interesting case because argument structure information is not available, and case marking is not as determining as in the South East Asian languages that have previously been studied in this context (Yamashita 1997, Miyamoto 2002, Aoshima et al 2004). The goal of experiment 2 is to test the specific timing of these dependencies and to examine when pronouns can be posited in these dependencies.

Since the bottom-up information in the case of clitic pronouns is richer than in the case of gaps, it would be interesting to know whether the parser still requires an active mechanism to process topicalization dependencies that involve the search for a pronoun or whether this mechanism is no longer active when there is enough information. In the case of gaps, the search is active due to the scarce information available about gaps in the input that can be processed in a bottom-up fashion but it is not clear whether that should be the case for pronouns. Therefore, if dependencies involving clitic pronouns result in an active search, this would make the active search a more general property of the parser.

The overt appearance of the clitic is required to complete the dependency in the Spanish CLLD cases I will discuss. Then, the difference between predicting a gap vs. a pronoun lies in how much is necessary for the completion to be fulfilled. The

cues needed to infer the gap vs. the pronoun position might be different just because of the intrinsic characteristics of each of them. Moreover, if we think about the filler-driven versus gap-driven processing, pronouns and gaps make different predictions with respect to these two hypotheses. In the case of gaps, there is always the need to check back on the displaced element so that the gap is completely interpreted, whereas in the case of pronouns that is not completely necessary. Pronouns can be interpreted partially without the requirement to check back on the displaced element that co-refers with them. Therefore, in the case of pronouns it seems a filler-driven processing seems to be active whereas in the case of gaps that is not completely evident.

3.2. Are the parsing mechanisms involved in forming CLLD dependencies active?

3.2.1. Long-distance dependencies involving pronouns

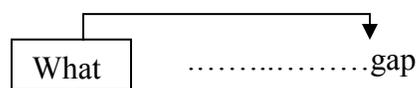
This study focuses on the processing of long-distance dependencies that involve overt clitic pronouns. In particular it examines the processing of topic-clitic dependencies (also referred to as Clitic Left Dislocations, Cinque 1990) where there is a fronted phrase that requires a clitic pronoun for its interpretation.

This study examines CLLD sentences like those in (4) where a clitic must obligatorily complete the dependency started at the dislocated phrase and where the fronted NP is marked with a preposition that provides it with case.

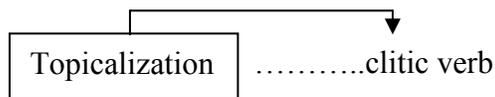
- (4) A María, Pedro la quiere con locura.
PA Maria, Pedro CL (DO) loves with craziness
'Maria, Pedro loves her very much'

Topicalization has much in common with other kinds of filler-gap dependencies. In classic English filled-gap studies a gap is required after a wh-phrase filler has been processed whereas in Spanish topic-clitic dependencies a clitic pronoun is required once the topic phrase has been processed.

English: What do children like ____?



Spanish: A estas escritoras, mis amigos las veneran.
 ‘These writers, my friends them worship’



The requirement of a clitic following the dislocated phrase is the starting point for this study. Based on speaker’s intuitions and previous syntactic literature (Rivero 1980, Contreras 1991, Zagana 2002) that has been extensively discussed in chapter 2, it is predicted that in a linear sequence like the one in (5) Spanish participants should require a clitic after having processed the fronted noun phrase, as shown in (5a). Moreover, since this topicalized prepositional phrase is case ambiguous due to the preposition *a*, which acts both as a dative and accusative marker in Spanish, both dative and accusative object clitics could be predicted in the upcoming sequence. However the parser could never take the dislocated phrase to be a subject. The parser’s expectations are measured by having a clitic at the first available potential clitic position like in (5b) or by delaying the completion of the dependency until the position of the clitic in the embedded clause like in (5c).

- | | | | |
|-----|---|-----|--|
| (5) | A estas escritoras, mi madre | (a) | ...las adora.
...them (acc) adores. |
| | <i>(To) these writers, my mother...</i> | (b) | ...les dijo que las adora.
...them(dat) told them(acc)
adores. |
| | | (c) | dijo que las adora
said that them (acc) adores |

3.2.2. Previous work on long-distance dependencies in Romance languages

I would like to turn now to studies on the processing of long-distance dependencies to examine whether results from languages with a similar configuration to Spanish (head-initial) can provide additional information on the processing of these dependencies.

Specifically, I focus on the work that has been done on the processing of long-distance dependencies in Italian and Brazilian Portuguese (BP henceforth) and I examine what the evidence in these languages tells us about the processing of long-distance dependencies involving gaps and pronouns. Previous studies on long-distance dependencies in Romance have focused on the processing of wh-dependencies in Italian (De Vincenzi, 1991) and on the processing of object anaphora bound by topicalization in Brazilian Portuguese (Maia, 1997).

De Vincenzi (1991) on the one hand, tested d-linked (inanimate) and non-d-linked wh-questions (animate) where the wh-question was ambiguous and could be considered either the subject or the object of the verb. She examined whether the AFS mechanism could also be active in Italian wh-questions and whether this active search for a gap would favor the initial consideration of the wh-word as a subject, since that would be the first possible position in the structure that could be associated with an empty element. Results showed a subject interpretation preference for non-d-linked wh-words (animate) but not for d-linked wh-words (inanimate) and she attributed this difference both to the fact that the disambiguating lexical NP comes post-verbally in Italian and to the fact that bare wh-words and d-linked wh-words don't have the same kind of syntactic/semantic properties.

Maia (1997), on the other hand, compared the processing of overt pronouns and empty categories in object position both in structures where these anaphors were bound by a subject and when they were bound by an antecedent in topic position. More specifically, he investigated whether pronouns and empty categories in object position in BP facilitated the access to their antecedents and whether this facilitation effect would vary with respect to the nature (topic/ subject) of the antecedent.

Maia used a cross-modal priming technique where subjects were given a probe recognition task in which target sentences were orally presented and reaction times were measured for visual probes corresponding to the antecedent of the overt pronoun and the gap. Results showed participants preferred to co-index antecedents in subject positions with overt pronouns and antecedents in topic position with empty categories.

Maia's data looks at pure left-dislocation constructions in BP where the requirements with respect to what kind of element is necessary to complete the chain are different from the Spanish clitic left-dislocation cases that this study analyzes. BP left-dislocations require a gap to come later in the input whereas Spanish clitic left-dislocations require a clitic.

The difference between predicting the appearance of a clitic pronoun with respect to predicting the appearance of a gap is that the predicted category in this case is an overt category instead of a category without phonological realization. In the case of gaps, there is no need to find extra phonological material to confirm their presence because they are null elements, whereas in the case of clitic pronouns, the overt realization of the pronoun is required to complete the dependency. One could argue that the AFS is only active when the parser needs to find phonologically null elements

and that in the case of pronouns there is no requirement to be active because they are easily identifiable in a bottom-up manner. Although clitic pronouns can be identified more easily because they are overt categories in the input, I take the AFS to be as active in these constructions as when there are gaps involved because the requirement to have a clitic pronoun as the foot of the dependency would trigger a search for a clitic. The urgency to interpret the fronted uninterpreted phrase would trigger the active search for the element that can complete the interpretation regardless of the nature of the element that can do it (e.g.: null vs. overt).

The first experiment in this study addresses this question and examines why is the active search mechanism active to begin with in a topic-clitic dependency.

In sum, the questions that this study will try to answer are first, whether there is an active search mechanism in CLLD dependencies in Spanish and if they have an active pronoun search mechanism just as *wh*-gap dependencies in English do. Second, how active is the search in these dependencies and how early can the parser construct the position of the clitic pronoun in these type of dependencies. And third, what kind of information does the parser use to predict the appearance of the clitic pronoun in these constructions where the verbal information comes after the clitic. The experiments discussed in this study speak directly to these questions

3.2.3. Experiment 1a: off-line grammaticality-rating questionnaire.

To test whether speakers require the clitic after processing a fronted phrase in Spanish and whether this clitic continuation was widely accepted an off-line grammaticality-rating task was conducted.

Thirty-nine⁴⁴ native speakers of Castilian Spanish⁴⁵ completed the off-line questionnaire. They all participated in the grammaticality-rating study after completing the on-line reading study because we did not want to bias them for a particular reading before the on-line reading was completed.

The sentences included in the questionnaire were simplified and shortened versions of the target items prepared for the on-line self-paced reading experiment presented in section 3.2.4 with no embedding and where the verb was obligatorily transitive. The clitics used were direct object clitics in accusative case (either feminine or masculine in gender).

Sixteen items with two conditions each were created for this off-line questionnaire. The two conditions created for each item differed in the presence or absence of a direct object clitic. Each subject saw the two conditions of each sentence together so that they would consider both the clitic and the clitic-less condition at the same time.

⁴⁴ The subjects on this experiments participated also on the online self-paced reading experiment presented in section 3.2.4. The subjects excluded in the on-line experiment data analysis were automatically excluded from the grammaticality rating experiment analysis.

⁴⁵ All the speakers were from the Basque Country and neighboring regions (Cantabria) except one subject who was from the south of Spain (Huelva). Moreover, eight of these thirty-nine speakers were bilingual speakers of Basque and Spanish, but had Spanish as their predominant language.

The critical sentence was preceded by a context sentence. A complete set of experimental conditions is shown in (6).

- (6) *Context*
 Caminando por la playa, mi madre y yo nos encontramos a dos señoras mayores.
 ‘(When we were) walking on the beach, my mother and I met two old ladies’
- a. *Clitic-less condition*
 *A esas señoras, mi madre conoce.
 (To) these old ladies, my mother knows
 ‘(To) These old ladies, my mother knows’
- b. *Clitic condition*
 A esas señoras, mi madre las conoce.
 (To) these old ladies, my mother them-acc knows
 ‘These old ladies, my mother knows them’

Subjects were asked to rate each of the conditions in the sentences on a scale from one (unacceptable) to five (acceptable), depending on how natural the sentences sounded to them, where the expectation was that the cliticless condition in (6a) would be considered unacceptable.

Table 1 shows the average of ratings of clitic continuation preference in the two conditions.

Condition	Mean	StDev	StdErr
Clitic	4.43	0.28	0.07
Clitic-less	1.9	0.22	0.05

Table 1: Grammaticality-rating Questionnaire, Rating of the off-line clitic continuation preference test.

These results clearly show that subjects have a preference for a clitic continuation and that they require a clitic after a topicalized phrase has been introduced.

3.2.4. Experiment 1b: Online Sentence Reading Experiment

Experiment 1 examines whether there is active search for a clitic pronoun in a construction involving a left-dislocated phrase that needs to be bound by a clitic. The prediction is that this is the case and so we expect processing facilitation of the clitic pronoun that co-refers with the left-dislocated phrase. If the requirement to predict a clitic in the upcoming sequence of words activates a clitic-search mechanism in the left-dislocation cases, then this active search mechanism should be reflected in the processing speed of the clitic and it would provide evidence for the active search mechanism being a more general property of the parser's architecture.

Method

Self-paced Reading experiment

Participants

Fifty-nine native speakers of Spanish participated in the experiment (nineteen Latin-American⁴⁶ and forty Castilian). All were students either at the University of Maryland, College Park, or at the University of Deusto, Bilbao. They were paid \$10 or its equivalent in Euros for their participation in the experiment, which lasted approximately forty-five minutes.

Materials and Design

⁴⁶ The Latin-American dialects tested included Argentinean (n=3), Bolivian (n=2), Chilean (n=2), Colombian (n=3), Ecuadorian (n=2), Nicaraguan (n=1), Peruvian (n=2) and Venezuelan (n=1).

The two conditions that were generated for each item differed in the absence/presence of a clitic preceding the main clause verb. This meant that the completion of the dependency took place either early, at the clitic preceding the matrix verb as in (7a), or late, at the clitic in the most embedded clause as in (7b).

(7) a. *Early Completion condition*

A estas chicas, mi hermana mayor más tarde **les** dijo que
To these girls, my sister old later on them-dat told that
 ya lo cree que las conoce desde hace tiempo.
indeed it-acc(masc) thinks that them-acc(fem) knows since long time.
 ‘To these girls, my old sister later told them that she indeed thinks she knows them for a long time’

b. *Late Completion condition*

A estas chicas, mi hermana mayor más tarde dijo que ya
To these girls, my sister old later on said that indeed
 lo cree que **las** conoce desde hace tiempo.
it-acc(masc) thinks that them-acc(fem) knows since long time.
 ‘These girls, my old sister later said that she indeed thinks she knows them for a long time’

The verbs chosen for the matrix verb position were optionally ditransitive verbs (e.g. *decir* ‘to say’) that could take a dative argument or not (what Levin (1993) calls verbs of communication that take dative alternation). The reason for choosing optionally ditransitive verbs was to allow an equal bias for the presence or absence of a preceding clitic. The logic was that if the parser does not find a clitic at the matrix verb to complete the dependency earlier in the sentence it would try to do it at the clitic in the most embedded sentence. An intermediate embedded sentence to delay the completion of the dependency was included. A clitic was placed in this intermediate sentence⁴⁷, which did not match in features with the dislocated phrase⁴⁸.

⁴⁷ The clitic in the intermediate sentence co-refers with the entire most deeply embedded clause.

⁴⁸ The reason for including this intermediate sentence with a clitic was to test if there could be a slowdown at the clitic *lo* due to a mismatch in features of the clitic (masculine, singular)

referent for the topicalization at the beginning of the sentence. One complete set of experimental conditions is shown in (8).

- (8) *Context:*
Yendo a la escuela, mi hermana mayor y yo vimos a mis amigas Ana e Irene.
'Going to school, my sister and I saw my friends Ana and Irene'
- a. *Late Completion condition*
A /estas/ chicas,/ mi /hermana/ mayor/ más tarde/ dijo/ que/ ya/
To these girls, my sister old later on said that indeed
lo / cree/ que / las / conoce /desde/ hace/ tiempo.
it-acc(masc) thinks that them-acc(fem) knows since long time

'These girls, my old sister later said that she indeed thinks she knows them for a long time'
- b. *Early Completion condition*
A /estas /chicas,/ mi /hermana/ mayor/ más tarde/ les/ dijo/ que/
To these girls, my sister old later on them-dat told that
ya / lo / cree / que / las / conoce /desde/ hace/ tiempo.
Indeed it-acc(masc) thinks that them-acc(fem) knows since long time.

'To these girls, my old sister later told them that she indeed thinks she knows them for a long time'.

Procedure

The experiment was conducted on Macintosh G3 desktop and laptop computers running the *mw_run* software developed at MIT. Participants were timed in a word-by-word self-paced non-cumulative moving-window reading task (Just, Carpenter, & Woolley, 1982). All critical sentences were presented on a single line. The context sentence appeared all at once and was sometimes more than a single line long. The segmentation indicated with slashes in (8) was the actual segmentation used in the experiment, where every word was a region except the adverbial preceding the main clause clitic, which was a two-word region. Words initially appeared as a row of dashes, and participants pressed the spacebar of the keyboard to reveal each

subsequent region of the sentence. When subjects pressed the spacebar for the first time, the context sentence was displayed at once. A subsequent key press replaced the first two lines with dashes and revealed the first word of the target sentence.

In order to ensure that participants attended to the stimuli, a yes/no comprehension question was included for each passage. The questions either made reference to the context sentence or to different parts of the experimental items.

Data Analysis

Analyses were conducted on comprehension task accuracy, item accuracy and residual reading times. The residual reading time for each segment corresponded to the difference between the observed reading time and the reading time predicted based upon a regression equation that was calculated for each participant predicting reading time from segment length. The regression equations were based upon data from all fillers and experimental items.

Only trials on which the comprehension question was answered correctly were included in the analysis. Four participants who showed overall accuracy below 80% or accuracy below 80% on target sentences were removed from further analysis. Residual reading times longer than 425 ms were discarded, removing 10% of trials. Separate analyses were conducted on Castilian and Latin American participants and on monolingual and bilingual speakers (Basque and Spanish, n=8) to see whether dialect or bilingualism had any effect on the results. These yielded no significant differences. The means and analyses presented below are based on the remaining trials.

Results

Comprehension accuracy and residual reading times at each region were entered into a repeated-measures ANOVA with *dependency late/early* completion as the within subject factor.

Comprehension Task Accuracy

Among the 55 subjects included in the analysis, average comprehension accuracy was 93%. The accuracy did not differ significantly across the two conditions and was 94% in the early completion condition and 93% in the late completion condition. In addition, there was no significant difference in the comprehension task accuracy between the subjects who live in Bilbao (93%) and those who live in Maryland, USA (92%).

Self-paced Reading

The reading time analysis yielded the following results. Residual reading times for early and late completion conditions in all regions are shown in Figure 3.

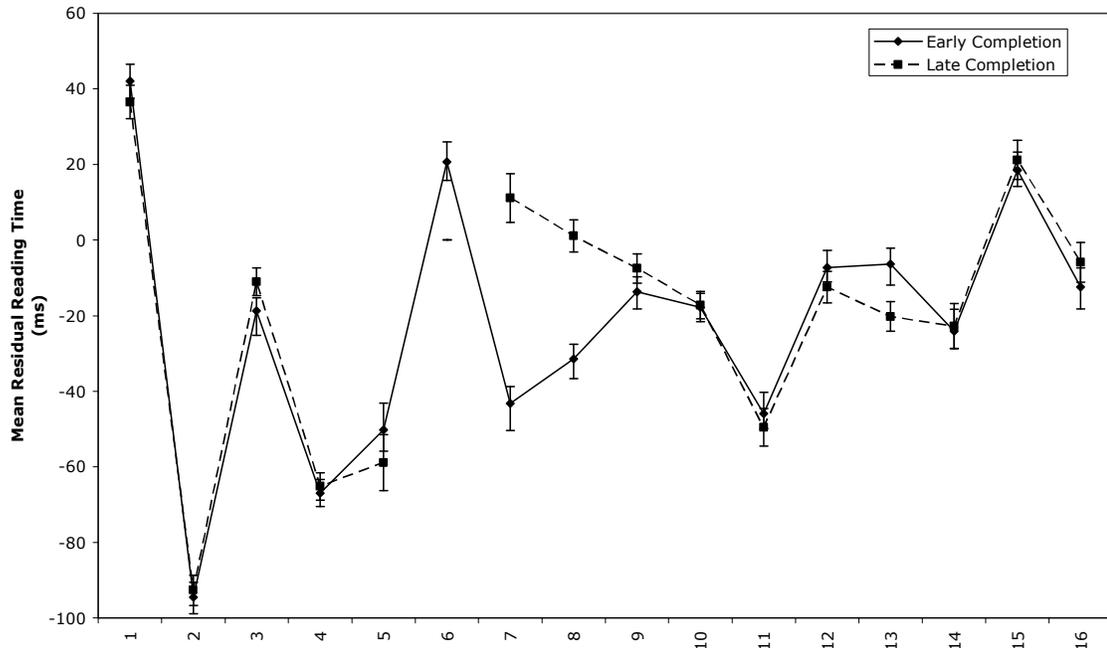


Figure 3: Reading times per region for the early and late completion conditions.
 A₁ estas_chicas₂ mi₃ hermana_mayor₄ mas_tarde₅ LES₆ dijo₇ que₈ ya₉ lo₁₀ cree₁₁ que₁₂ LAS₁₃
 conoce₁₄ desde₁₅ hace₁₆ (tiempo)

There were no significant differences between reading times in the early and the late completion conditions in regions 1-6 (all $F_s < 1$).

In region 7, the matrix verb *dijo*, there was a significant main effect of clitic presence at the matrix verb ($F_1(1, 54)=42.69$, $MSe=849052$, $p<.001$; $F_2(1, 23)=40.57$, $MSe=865717$, $p<.001$). Matrix verbs were read on average 52 milliseconds slower in the late completion condition than in the early completion condition. In region 8, the first complementizer *que*, there was a significant main effect of condition ($F_1(1,54)=36.27$, $MSe=310794$, $p<.001$; $F_2(1,23)=27.67$, $MSe=307850$, $p<.001$). The complementizers in the late completion condition were read 32 milliseconds slower than complementizers in the early completion condition.

At regions 9-12, there were no significant differences between reading times in the two conditions (all $F_s < 1$).

In region 13, the accusative clitic *las*, the effect was significant both in the items and participants' analysis ($F_1(1,54)=4.08$, $MSe=53241.6$, $p<.05$; $F_2(1, 23)=5.44$, $MSe=62267.9$, $p<.05$). The accusative clitic *las* was read 14 milliseconds slower in the early completion condition than in the late completion condition. It is unlikely that this is an artifact of preceding regions, since there were no significant differences between conditions at regions 9-12, and the differences at regions 7-8 were in the opposite direction. Moreover, there were no differences at any subsequent regions.

It is also worth noting that the effect at the embedded clitic went in the opposite direction from the effect at the main verb. The late completion condition was read slower at the main verb region whereas it was read faster at the embedded clitic region.

Discussion

We hypothesized that the processing of the left-dislocated phrase (filler) would result in an active search for a clitic pronoun in the sentence and that the delay in the appearance of the clitic that could complete the dependency would result in a reading time difference in the clitic pronoun that comes the latest. This predicts the effect at region 13, (the embedded clitic *las*) and then the one at region 7 (the main verb region).

Hence, the result of central interest in this experiment is a shorter reading time observed in the late completion condition at the accusative clitic *las* (Region 13),

relative to the reading time for the same region in the early completion condition. In the late completion condition the topicalized phrase was not previously bound in the preferred way therefore there is an existing prediction for an upcoming clitic in order to interpret the dependency started at the left-dislocated element. Since a clitic matching in features with the left-dislocated phrase does not appear until the most embedded sentence in the late completion condition, the faster reading time of the clitic in this condition is taken as an indication of its high prediction. I interpret this effect as reflecting the use of clitic pronouns as underlying arguments that help to complete the unbounded dependency before encountering the main verb in Region 14. This result is interpreted as a facilitation effect.

An explanation based on predictability also explains the existence of the second robust effect - the slowdown in reading time in the late completion condition at the matrix verb (Region 7). This effect arises due to the reader's expectation for a clitic after processing the topicalization in Regions 1-2 in the late completion condition. Since there is no clitic preceding the matrix verb that would allow interpretation of the topicalization and confirm the interpretation of this category, readers slow down. However, since the early completion condition does have a clitic at Region 6 preceding the matrix verb, the reading time of the matrix verb region in this condition might reflect that the verb in this case is highly predicted once the dative clitic has appeared with respect to the condition where there is no clitic present and any verb can be expected. Therefore, it is not clear what could be causing the effect at the main verb and there could be two different explanations for this effect: the lack of the confirmation of the expected clitic or the high predictability of a

ditransitive verb once a clitic appears in the input. The effect at the first complementizer (Region 8) is assumed to be a residue of the effect at the previous region.

The previous effects (Region 7 and Region 13) are taken to be evidence of the expectation for a clitic by the parser once the left-dislocated phrase has been processed. In the region corresponding to the main verb (region 7), the failure to have a clitic preceding the verb causes a slow down in the late completion condition, whereas in the clitic region (region 13) this condition is read faster due to the anticipated prediction of this clitic. Nevertheless, this explanation can only be possible if the left-dislocated phrase is creating an expectation for a clitic in the parser.

3.2.5. How active are dependency construction mechanisms involving clitic pronouns?

Experiment 2, addresses the question of how active this AFH mechanism is in the parser in constructions that involve overt pronouns and at what kind of information does the parser use to predict the position of the clitic in the sentence. The first experiment examined if dependency construction involving pronouns could be considered as active as in dependencies involving gaps and found a facilitation effect at the clitic that suggested that both wh-gap dependencies and topic-clitic dependencies could be considered equally active. Building from experiment 1, the second experiment looks at how active the AFH mechanism is in these constructions

and it tries to look for the specific place in the sentence where the clitic position is projected.

The clitic pronoun that completes the dependency in this case appears in a preverbal position, which means that the subcategorization information provided by the verb cannot be used to guide the active search for the clitic pronoun as in the case of English *wh*-dependencies where the object gap is immediately adjacent to the verb. It is less clear then what kind of information is used in a head-initial language to interpret these active dependencies where the foot of the dependency appears pre-verbally.

Most evidence on the processing of gaps in English involves gaps associated with arguments that go after a verb. In these cases, it is clear how the grammar indicates the availability of a potential gap position. For gaps that are in preverbal position, though, it is less clear whether active dependency formation occurs before the subcategorizer and how these gap positions are identified if the subcategorizer information is not available to the parser. The second experiment in this study addresses this question and examines what kind of information the parser uses to predict the foot of the dependency in a head-initial language if there is no verb available. More specifically, it examines if clitic pronouns that precede the verbal head in Spanish topicalization constructions are used to interpret the left-dislocated phrase or whether these dependencies can only be interpreted at the verb. This is particularly relevant for disentangling the discussion on whether the verb is completely necessary to project the position of the gap in filled-gap studies where the gap comes post-verbally (Pickering & Barry 1991, Gibson & Hickok 1993, Sag & Fodor 1995, Pablos 2003). If there is a way to interpret long-distance dependencies in

head-initial languages without the use of the verb, then it means it is not the only information used to project a gap position.

3.2.6. Subject Filled-gap Effects

A context that has been examined in English where the gap might come before the verb that subcategorizes for it is that of subject gaps. Subject-gaps are the only circumstance in English under which the gap precedes the subcategorizer, therefore, they are a good test case to examine if active dependency formation occurs before the verb by looking at the presence of a slowdown at a potential subject gap position.

Stowe (1986) was the first one to test if a filled-gap effect could be obtained for an overt NP in subject position. In Stowe's experiment 1, subject and object positions were compared separately with respect to their control in (9a) where there was no extraction to see if the presence of a lexical noun phrase in these two positions in the sentence caused reanalysis.

- (9)
- a. My brother wanted to know if Ruth will bring us home to Mom at Christmas. (IF-CLAUSE)
 - b. My brother wanted to know WHO ___ will bring us home to Mom at Christmas. (WH-SUBJ)
 - c. My brother wanted to know WHO Ruth will bring ___ home to Mom at Christmas. (WH-OBJ)
 - d. My brother wanted to know who Ruth will bring us home to ___ at Christmas. (WH-POBJ)

Whereas a 'filled-gap effect' was visible in the embedded wh-question in (9b) at the noun phrase in object position as with respect to its baseline condition in the

non-extraction condition in (9a), the difference between the if-clause in (9a) and the wh-complement (wh-obj/pobj) clauses in (9c) or (9d) was not significant at the NP in subject position *Ruth*, providing no evidence that people expect a gap in subject position. Stowe (1986) suggested the lack of processing difficulty in the subject noun phrase could be because people do not expect a gap in subject position or because they do not have any trouble recovering from the expectation at subject position, unlike object position.

Considering the different hypotheses that previous studies (Clifton & Frazier 1989, Clifton & De Vincenzi 1990, De Vincenzi 1991, Stevenson 1993, Gibson & Hickok 1993, Gibson et al. 1994, Gorrell 1998) had given for the lack of a subject filled-gap effect in Stowe's (1986) study, Lee (2004) constructed an experiment where he examined the processing of relative clauses in English. He tested whether a filled-gap effect would surface at the subject NP position filled by *Irene* in the preposition stranding case in (10a) in contrast to the construction in (10b). In addition, he included a length manipulation to increase reanalysis difficulty in his items (10c) and (10d), and added a prepositional phrase that intervened between the wh-phrase and the NP where the filled-gap effect was expected. Due to this length effect, reanalysis at *Irene* should be harder in (10c) than (10a).

- (10) a. That is the laboratory [which_i] **Irene** used a courier to deliver the samples to t_i.
b. That is the laboratory [to which_i] Irene used a courier to deliver the samples t_i.
c. That is the laboratory [which_i], on two different occasions, **Irene** used a courier to deliver the samples to t_i.
d. That is the laboratory [to which_i], on two different occasions, Irene used a courier to deliver the samples t_i.

The subject gap manipulation resulted in the NP *Irene* being read significantly more slowly in those conditions which had a subject gap, concretely (10a) and (10c). The effect at *Irene* can be interpreted as a subject filled-gap effect since having adopted a subject gap analysis at *which*, the appearance of a lexical subject interferes with the subject gap prediction and results in a slowdown.

The other studies that looked at subject gaps in Romance are Clifton & De Vincenzi (1990) and De Vincenzi (1991). Clifton & De Vincenzi (1990) and De Vincenzi (1991) examined whether the mechanism implied by the AFH could also be operative in Italian ambiguous subject and object wh-questions and whether this active search for a gap would favor the initial consideration of the wh-word as a subject, since that would be the first possible position in the structure that could be associated with an empty element. More specifically, they tested inanimate discourse-linked (*which N*) wh-questions and animate wh-questions (*who*) where in both cases the wh-question was ambiguous and could be initially considered either the subject or the object of the verb until the post-verbal noun phrase would disambiguate it.

They used a phrase-by-phrase self-paced reading technique. Results showed a subject interpretation preference for animate wh-words but not for inanimate discourse-linked wh-words. The animate wh-words were read faster when they were subject than when they were object. This difference did not appear for the inanimate discourse-linked wh-words.

This result showed a clear contrast with Stowe's (1986) results, which showed an advantage of object extraction over subject extraction. They attributed this difference to the fact that the disambiguating lexical NP comes post-verbally in Italian as opposed to English where it comes pre-verbally. Their conclusion is that the

occurrence of the verb between the filler and the disambiguating noun phrase in Italian gives the reader more time to commit to an initial analysis and provide a possible semantic interpretation of the wh-word.

3.2.7. Pre-verbal dependency completion effects in head-final languages

A different context where gaps occur before the main subcategorizer is that of object gaps in head-final languages. Previous studies have focused on preverbal dependency completion effects in long distance dependencies involving wh-gaps (e.g. V2 constructions in German, Clahsen and Featherston 1999; long-distance scrambled wh-phrases in Japanese: Nakano et al. 2002, Aoshima et al. 2004) and have found completion effects before the verb at the foot of the chain.

Most of the research undertaken to look for pre-verbal effects in head-final languages has been conducted in Japanese. In Japanese scrambled structures where the object has been displaced within the sentence, the gap is posited in the position prior to the verb. Attempts to show if the parser posits a trace in preverbal position in long distance wh-gap scrambled structures have been made by Nakano et al. (2002) and Aoshima et al. (2004). Nakano et al. (2002) used a cross-modal lexical priming technique to look for pre-verbal antecedent priming effects and Aoshima et al. (2004) used the self-paced reading technique to look for a slowdown in pre-verbal position.

Both Nakano et al.'s (2002) study and Aoshima et al.'s (2004) results suggest that in the case of Japanese the case marking is playing an important role in predicting the position where the trace of the scrambled element is going to be. If

there is a fronted noun phrase that cannot be interpreted in the main clause the only information that the parser has available is each noun phrase's case marking. If the scrambled object cannot be interpreted in the matrix sentence and there is a nominative marked noun-phrase marking the beginning of the embedded clause, the only place where the fronted phrase could be interpreted is in preverbal position.

So it seems that the evidence being used by the parser to predict the position of the gap before encountering the subcategorizer both in the scrambled sentences used by Nakano et al (2002) and Aoshima et al. (2004) is the case marking on the NPs and the word order. The evidence given in these studies is important to consider what kind of information is the parser using in these cases to project the position of the gap. If there is an active search for a gap the moment the scrambled phrase is processed, it seems the parser can be guided by the case marking in the noun phrases and the word order information provided by these.

3.2.8 Experiment 2

Experiment 2 investigates whether the initial topicalization affects the processing of the most embedded clitic pronoun in the sentence. If the incomplete dependency triggered the search for the embedded pronoun in experiment 1b, we would predict no reading time difference between cases with or without the embedded clitic pronoun if there was no topicalization.

Method

Participants

Forty-eight native speakers of Spanish participated in the experiment (eleven Latin-American⁵⁰ and thirty seven Castilian⁵¹). All were students either at the University of Maryland, College Park, or at the University of Deusto, Bilbao. They were paid \$10 or its equivalent in Euros for their participation in the experiment, which lasted approximately forty-five minutes.

Materials and Design

Twenty-four items in four conditions each were used in the experiment⁵². The four conditions for each item differed in the absence/presence of the clitic preceding the matrix verb and in the absence/presence of topicalization. Materials differed from those in experiment 1b in that they had only one embedding (the intermediate

⁵⁰ The Latin-American dialects tested included Argentinean (n=1), Chilean (n=1), Ecuadorian (n=1), Mexican (n= 2), Paraguayan (n=1), Puerto Rican (n=4) and Uruguayan (n=2).

⁵¹ All the speakers were from the Basque Country except six participants that were from other regions of Spain: Andalucía (n=3), Aragón (n=1) and Castilla y León (n=2).

⁵² Experimental items are listed in Appendix 3-D.

embedding was eliminated⁵³) and the subject of the embedded sentence was overt. The reason to have the embedded subject overt in this experiment was to avoid the confusion of the clitic pronoun *las* with the determiner, which has the same form in Spanish. By adding the embedded subject, it was assured that the clitic pronoun was unambiguously considered a clitic and not a determiner. Experiment 2 compared topic-clitic constructions with other Spanish constructions in which topicalization is absent but the linear order of NPs is identical. The new conditions took advantage of the fact that Spanish allows post-verbal subjects, by placing an NP just like the topicalized NP in final position of the context sentence (11) for all the conditions. Thus, the linear distance between the pronouns and their antecedents is matched across topicalization (11ab) and non-topicalization conditions (11cd) and the clitic pronouns in (11cd) have an antecedent in the nearby discourse. Regardless of the fact that the referent is in an inter-sentential or intra-sentential position, the number of words from the antecedent to the clitic pronoun is the same in the four conditions.

As in experiment 1b, the verbs in the main clause were optionally ditransitive and were equally biased for having a clitic pronoun preceding them or not. An auxiliary was added to the main verb to see whether the robust effect seen in experiment 1b at the main verb region will disappear when an auxiliary would be added. The verbs in the embedded clause on the other hand were obligatorily transitive and had an auxiliary.

Four lists were created by distributing the twenty-four paired items in a Latin Square design. Each subject saw exactly one of the lists intermixed with seventy-two

⁵³ Since there was no slowdown due to mismatch in features at the embedded clitic pronoun *lo*, it was decided to exclude the embedded sentence containing the mismatching clitic *lo* from experiment 2.

filler items in a random order. A context sentence preceded each of the items to make the topicalization more natural and to provide an antecedent for the clitic in the non-topicalized cases. The context sentence was identical in the four conditions. A complete set of experimental conditions is shown in (11).

(11) Context sentence:

Cuando / abrieron/ las/ puertas/ del/ teatro/ en/ el/ que/ se/ celebraba/ el/
casting, /entró/ precipitadamente/ un/ grupo /de/ chicas. /
*When the doors of the theatre where the casting was given opened, a group of
girls entered precipitately.*

(11a) *Topicalization/ Late Completion:*

A /estas /chicas, / el /organizador /del /casting/ha / explicado/
To these girls, the organizer of the casting explained
con /todo /tipo /de /detalles /que /el /manager /las / iba /a /llamar /por/apellido.
with all sort of details that the manager them-acc(fem) will call by name.

‘These girls, the organizer of the casting has explained with all sorts of details
that the manager would call them by surname’

(11b) *Topicalization/ Early Completion:*

A /estas /chicas, / el/ organizador /del /casting /les / ha / explicado/
To these girls, the organizer of the casting them-dat explained
con/ todo/ tipo/ de/ detalles/ que/ el/ manager/ las/ iba/ a/ llamar/por /apellido.
with all sort of details that the manager them-acc(fem) will call by name.

‘These girls, the organizer of the casting has explained them with all sorts of
details that the manager would call them by surname’

(11c) *Non-topicalization/ Late Completion:*

El /organizador /del /casting /ha /explicado /con /todo /tipo /de /detalles/
The organizer of the casting explained with all sort of details
que/ el /manager / las / iba /a /llamar /por /apellido.
that the manager them-acc(fem) will call by name.

‘To these girls, the organizer of the casting has explained with all sorts of
details that the manager would call them by surname’

(11d) *Non-topicalization/ Early Completion:*

El /organizador /del /casting /les /ha / explicado /con /todo /tipo/ de /detalles/
The organizer of the casting them-dat explained with all sort of details
que / el / manager/ las/ iba /a /llamar /por /apellido.
that the manager them-acc(fem) will call by name.

‘To these girls, the organizer of the casting has explained them with all sorts of details that the manager would call them by surname’

Procedure

The experiment was conducted on a Macintosh G3 iBook laptop computer running the *Linger* software developed at MIT. Participants were timed in a word-by-word self-paced non-cumulative moving-window reading task (Just, Carpenter, & Woolley, 1982). All sentences were presented on a single line, except the context sentence, which was sometimes more than a single line long. The segmentation indicated with slashes in (11) was the actual segmentation used in the experiment. Words initially appeared as a row of dashes, and participants pressed the spacebar of the keyboard to reveal each subsequent region of the sentences. When subjects pressed the spacebar for the first time,⁵⁴ the first word of the context sentence was revealed and the rest of the context sentence and the target sentences appeared in dotted lines.

In order to ensure that participants attended to the stimuli, a yes/no comprehension question was included for each passage. The questions either made reference to the context or to different parts of the experimental items.

Data Analysis

Only sentences for which the corresponding comprehension question was answered correctly were included in the analysis. Residual reading times longer than 475 ms were discarded, removing 2.7 % of trials. Separate analyses were conducted

⁵⁴ Since part of our segmentation included the clause final subject in the context, which was actually a region of interest for the analysis, the words in the context sentence were displayed one by one (in contrast to the presentation of the context in experiment 1b).

on Castilian and Latin American participants and on monolingual and bilingual speakers (Basque and Spanish, n=8) to see whether dialect or bilingualism had any effect on the results and no significant differences were found. The means and analyses presented below are based on the remaining trials.

Reading times from topicalized conditions and non-topicalized conditions were compared pairwise in a one-way ANOVA. ANOVAs were computed on the participant mean residual reading times collapsing over items (F1), and on item means collapsing over participants (F2). All significant effects with $p < .05$ are reported.

Results

Comprehension Task Accuracy

Two participants out of 50 whose task accuracy was below 80% in target sentences and below 85 % in total were removed from further analysis. Among the 48 subjects included in the analysis, average comprehension accuracy was 91%. The average correct response percentage did not differ significantly across the four conditions. The average comprehension accuracy in the topicalized early completion condition was 90% and 92% in the late completion condition. The average comprehension accuracy in the non-topicalized early completion condition was 90% and 90% in the late completion condition. In addition, there was no significant difference in the comprehension task accuracy between the subjects who live in Bilbao, Basque Country (91%) and those who live in Maryland, USA (91%).

Self-paced Reading

The reading time analysis yielded the following results. Residual reading times for non-topicalized early and late completion conditions in all regions are shown in Figure 4 and the residual reading times for topicalized early and late completion conditions in all regions are shown in Figure 5.

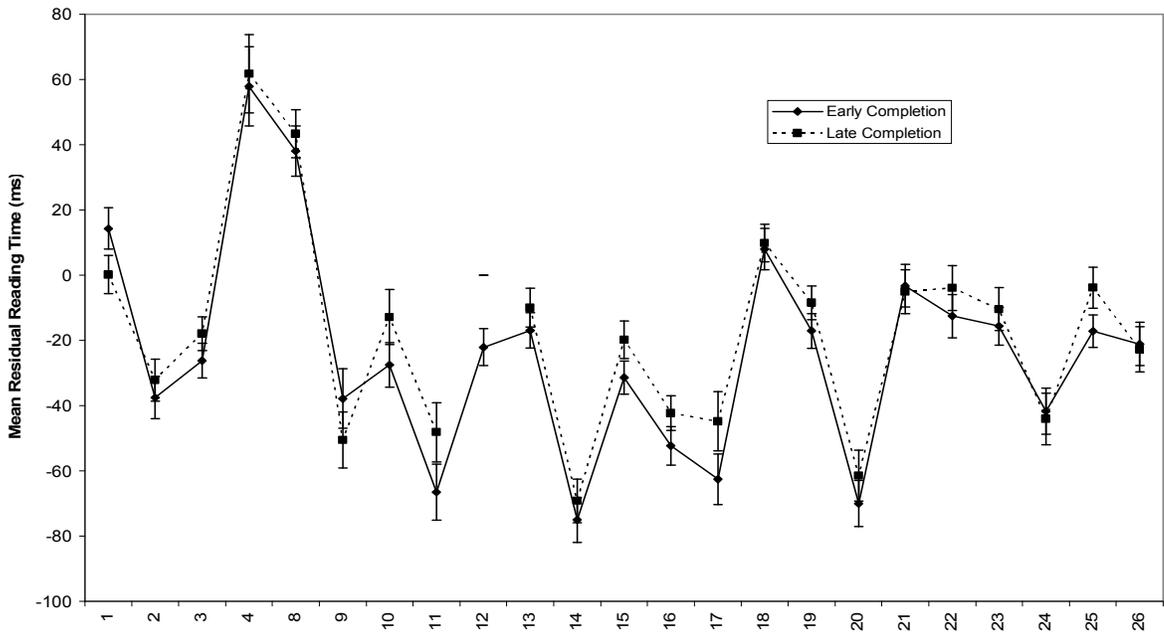


Figure 4: Mean residual reading times in milliseconds per region for the non-topicalized early and late completion conditions.

[Context...]...n₁ grupo₂ de₃ chicas₄. El₈ organizador₉ del₁₀ casting_ teatral₁₁ LES₁₂ **ha**₁₃ explicado₁₄ con₁₅ **mucho**₁₆ detalle₁₇ que₁₈ el₁₉ **director**₂₀ LAS₂₁ iba₂₂ a₂₃ llamar₂₄ por₂₅ apellido₂₆

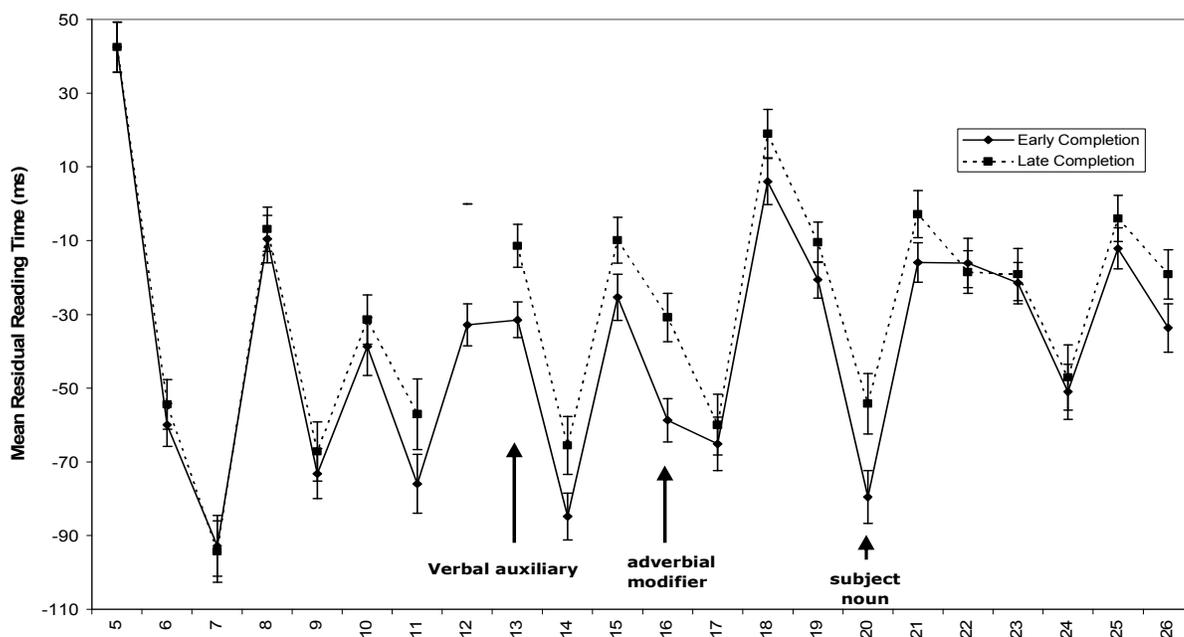


Figure 5: Mean residual reading times per region for the topicalized early and late completion conditions.

A₅ estas₆ chicas₇, el₈ organizador₉ del₁₀ casting_teatral₁₁ LES₁₂ **ha**₁₃ explicado₁₄ con₁₅ **mucho**₁₆ detalle₁₇ que₁₈ el₁₉ **director**₂₀ LAS₂₁ iba₂₂ a₂₃ llamar₂₄ por₂₅ apellido₂₆

In the *non-topicalized conditions* in figure 4 ((11c) vs. (11d)), there were no reliable main effects or interactions in any region.

The pairwise comparison for the *topicalized conditions* showed the following results.

At all regions prior to region 13 there were no significant differences between the reading times in the early and late completion conditions (all $F_s < 1$).

In region 13, the verbal auxiliary *ha*, there was a significant main effect of *dependency* or clitic presence ($F_1(1,47)=6.15$, $MSe=49378$, $p<.05$; $F_2(1, 23)=4.28$, $MSe=48591.3$, $p<.05$). The effect was due to longer reading times for the late completion condition than for the early completion condition. The verbal auxiliaries

in the former condition were read on average 20 milliseconds slower than in the latter. However, the interaction of the *dependency* factor with *topicalization* did not reach significance (both $F_s \ll 1$).

In regions 14 and 15 (verbal participle and adverbial preposition) the effect was not significant (all $F_s < 1$).

In region 16, the adverbial modifier *mucho*, there was a significant main effect of the dependency factor (clitic presence) ($F_1(1,47)=9.03$, $MSe=90999.4$, $p<.005$; $F_2(1, 23)=10.82$, $MSe=104011$, $p<.005$). The effect was due to longer reading times for the late completion condition than for the early completion condition, where the adverbial modifier was read 28 milliseconds slower in the late completion condition. However, the interaction of the *dependency* factor with *topicalization* did not reach significance (both $F_s \ll 1$).

At regions 17, 18 and 19 there were no significant differences between the reading times in the two topicalized conditions ($F_s < 1$).

In region 20, the noun of the embedded subject *director*, there was a significant main effect ($F_1(1,47)=4.68$, $MSe=65902.7$, $p<.05$; $F_2(1, 23)=4.84$, $MSe=77400.1$, $p<.05$). The effect was due to longer reading times for the late completion condition than for the early completion condition, where the embedded subject was read 25 milliseconds slower than in the late completion condition. However, the interaction of the *dependency* factor with *topicalization* did not reach significance (both $F_s \ll 1$).

In region 21 (accusative clitic *las*) there were no significant differences between the reading times in the two topicalized conditions ($F_s < 1$).

At regions 22, 23, 24, 25 and 26 there were no significant differences between the reading times in the two topicalized conditions ($F_s < 1$).

Discussion

The result of central interest in this experiment is the slowdown in reading time observed in the late completion condition at the noun of the embedded subject *director* (Region 20), relative to the reading time for the same region in the early completion condition. This effect is interpreted as a filled-subject effect and as a reflection that there is an expectation for a clitic in the linear sequence of words right after the complementizer is processed. The appearance of an overt noun phrase at this position causes a slowdown and delays the presence of the clitic that was predicted after a clitic was not found preceding the main verb. This effect can be interpreted as previous filled-gap effects in that there is an element interfering in the appearance of the clitic in the immediate context and delaying the interpretation of the fronted phrase. The preferred continuation is that that presupposes a null subject that co-refers with the main subject of the sentence (*organizador* in this case). Hence, when the continuation introduces a new entity, there is a cost associated with that in that it interferes with the active clitic search process started at the filler. There was not significant slowdown observed at the accusative clitic pronoun *las* (Region 21). This could be due to the bigger effect found in the preceding region.

There could be several reasons why the subject of the embedded sentence is read slower in the late completion condition. It could be that a null subject is expected in order for the clitic to come next in the input and the parser is therefore surprised to find the null subject position filled with an overt subject. The other option could be

that the null subject is expected to avoid a subject-shift or a change in the antecedent (information structure) and when there is an overt subject introducing a new discourse referent, the parser slows down. A final option, assuming no incremental parsing of the material is that all the noun phrases do not have a theta-role yet and have to be left unconnected. Then, the load of processing would be higher at the time the embedded subject is read. Considering the evidence found and the expectations that I think the parser has at the complementizer, I take the first option to be the most plausible one.

The second effect of interest was the slowdown observed in the late completion condition at the verbal auxiliary relative to the reading time in the early completion condition. This result stems from readers' expectations of a clitic. After processing the fronted phrase longer reading times follow when the subjects do not find a clitic preceding the matrix verb that would allow them to interpret the topicalization. In experiment 1b, I also suggested that this pattern could result from the verb being highly predicted once the dative clitic pronoun has appeared versus the late completion condition where any verb could be expected. However, this explanation is now ruled out since the verbal auxiliary and the adverbial modifier were also read slower in the late completion condition relative to the early completion condition.

The third effect of interest was the slowdown observed in the late completion condition at the adverbial modifier relative to the reading time in the early completion condition. This effect could be a result of the expectation of a resumptive pronoun. It could be that there is a continuation of the active gap filling started at the left-dislocated phrase and that the effect at the main verb has extended in this case to the following regions, which are those of the adverbial.

One could also question why an interference effect is not apparent at the adverbial position of the intermediate embedded clause of experiment 1b. If the clitic is projected at the complementizer, then why is there no effect at the adverbial preceding the clitic pronoun? If we assume that the null subject continuation is the parser's preference both to avoid the change of referent and to have the clitic come next, the expectation to have a null subject is fulfilled and there is no effect because what could be the intervening element is an adverb. The next question then is why the adverbial does not show a slowdown effect if it is interfering with the clitic being the next word after the complementizer. One of the reasons for this could be that there is no referent shift with the intervening element as in the case of the embedded subject in experiment 2. The other could be that since the adverbial chosen to precede the clitic could be considered in a cluster with it, it is not considered an interference and its parsing comes for free. The third reason could be that non-arguments do not cause so robust interference effects as arguments or complements do.

3.3. General Discussion

The goal of this study was to investigate the processing of clitic pronouns in topic-clitic dependencies in Spanish. In particular, I was interested in testing whether the requirement for a clitic in the upcoming sentence after processing a topicalization causes an active search for the clitic in the sentence. The hypotheses in the two experiments stemmed from the claim that the presence of an uninterpreted 'filler' would trigger a search for a clitic, the only way to interpret this displaced element. By interpretation, we mean association with a case or a theta role.

Most existing evidence in the active search literature is related to the search for gaps. Therefore, it could be hypothesized that the parser resorts to a more active approach when processing long-distance dependencies that involve gaps because gaps are difficult to identify in a bottom-up fashion. Since they are phonetically null, the search mechanism might need to be more active to look for them than if they were overt elements in the sentence. This would then predict a less active approach in dependencies that involve overt elements such as pronouns since they can be identified more easily in a bottom-up manner. If after identifying the filler as being moved the category that has been predicted is an overt pronoun as in the case of topic-clitic dependencies, then the search might not need to be so active. We might not see robust effects of dependency completion for overt pronouns as we saw for gaps. Both the facilitation effect in experiment 1 and the filled-subject effect in experiment 2 suggest that there seems to be an active search mechanism in the processing of topic-clitic dependencies that involve clitic pronouns that is as active in these cases as it is in wh-dependencies that involve gaps. This means that having richer bottom-up information does not stop the parser from doing incremental processing of the input where the urge to interpret the information encountered as soon as possible triggers an active search of material necessary to complete the dependency. Based on the results, this active search mechanism seems to apply in any kind of long distance dependency that requires the interpretation of some uninterpreted filler, regardless of whether the element that completes the dependency (the foot of the dependency) is an overt or null category. This suggests that the filler's status as an uninterpreted or incomplete element is what drives the Active Filler.

Spanish is a good test case in that even if not every NP is marked for case as in a language like Japanese, it still has a relatively rich agreement (reflected in the possibility to drop subjects) that allows the parser to interpret things before the verb. Clitic pronouns are an example of the moderately rich agreement system in Spanish. Moreover, through clitic pronouns we can examine if interpretation of material can happen before the verbal head is processed. If Spanish is contrasted to English, where it is very unlikely to find agreement marks that allow interpretation prior to the verbal head, the advantages of testing pre-verbal effects in a head-initial language like Spanish become apparent.

In the case of English, results on processing long-distance filled-gap dependencies are consistent with head-driven parsing theories (Abney 1989, Pritchett 1992, Mulders 2002) that presuppose that the parser waits for the head to interpret all the uninterpreted information because the only place where dependencies can be interpreted is at the verb. There is no possible previous place to interpret them unless subject gaps are considered (Lee 2004). In Japanese, results are not consistent with head-driven parsing theories because case marking provides a way to interpret things before the verbal head. So the question is whether this pre-verbal interpretation of long-distance dependencies can also happen in a language less rich in case than Japanese but with a head-initial configuration like English.

The literature on filled-gap effects both in head-initial and head-final languages has shown that different kinds of information are used to predict the gap position and that the gap is projected at different positions depending on the configuration of the language. In the case of English filled-gap effects (Crain & Fodor 1985, Stowe 1986) where the gap position immediately follows the verb, the gap can

be predicted from the moment the verb is processed. So the filled-gap effect is right at the noun phrase that occupies the gap position because the gap has already been constructed at the previous word (the verb) and the parser is surprised to find that gap position filled. The gap can be constructed at the verb because it is possible to determine through its subcategorization that there is a missing argument.

In the case of Japanese filled-gap effects (Aoshima et al. 2004) where the gap position precedes the verb, the information used to construct the gap position is different from English. The gap is predicted in these cases through the information provided by the case marking of the noun phrases and the word order. If there is a scrambled fronted noun phrase with dative case, the gap for this noun phrase will be constructed as soon as the nominative marked noun phrase preceding the gap position is processed, since the case array in Japanese would demand a dative marked phrase to come after the nominative. The filled-gap in this case is found at the dative-marked noun phrase that follows the nominative-marked phrase because the gap for a dative phrase has already been constructed at the preceding nominative phrase and the overt phrase occupying the gap position has the same case-marking as the predicted gap.

In the case of the Spanish data that concerns us it seems the position of the clitic is projected earlier than the gap position in pre-verbal filled-gap studies (Aoshima et al. 2004). In the later, the gap is projected at the noun phrase immediately preceding the gap position whereas in Spanish topic-clitic constructions, the clitic is projected earlier at the complementizer preceding the noun phrase. This is suggested in that the filled-subject effect happens to be at the embedded subject instead of at the clitic, indicating that by the time the embedded subject is processed, the clitic pronoun has already been constructed. The next question then, is what kind

of information the parser uses to predict the foot of the dependency at the complementizer. When the expectation to have a clitic has failed at the region preceding the main verb, the next possible clitic position is at the complementizer position, since there is the expectation to have a null subject continuation, which will favor the appearance of the clitic as the next word in the input. Therefore, it seems the information used by the parser in Spanish to predict the foot of the dependency is closely related to the expectation for a null subject continuation and the preference to keep the referents already present within the piece of sentence that has been parsed.

It is also worth to point out another difference of the Spanish topic-clitic dependencies with respect to other filled-gap dependencies. If the dependency is constructed at the complementizer in Spanish topic-clitic dependencies, this means that more structure rather than just the pronoun needs to be projected ahead and that it could be the system is more active in this case than in dependencies that involve a gap. In Spanish topic-clitic dependencies, the parser seems to be projecting both the clitic and the subject that precedes the clitic (which most likely will be preferred to be null) at the complementizer, whereas in the case of either English filled-gap effects or Japanese pre-verbal filled-gap effects, the only thing that the parser projects either at the verb or at the nominative marked phrase is the gap position.

Appendix 3-A: Off-line Sentence Completion Experiment 1

Building upon the results of the off-line grammaticality questionnaire, a sentence completion test was prepared in order to check how the presence or absence of a previous clitic in the sentence could affect the clitic requirement and to show that, when asked to complete a sentence containing a topicalization, subjects would spontaneously generate a clitic.

Participants

Fifteen native speakers of Castilian Spanish participated in this experiment.

Materials and Design

Eighteen items of three conditions each were used. Three lists were created by distributing the eighteen items in a Latin Square design. Each subject saw exactly one of the lists intermixed with thirty-six filler items in a random order. Items were derived from the experimental items for the on-line study (experiment 1b, chapter 3). No context sentence was included in this test.

There were three conditions that included initial sentence fragments such as those in (1). Participants were instructed to provide a completion that sounded natural to them.

- (1) a. *Single NPs condition*
A estas chicas, mi hermana ...
(To)these girls, my sister...

b. *Non-clitic condition*

A estas chicas, mi hermana dijo que...
(To)these girls, my sister said that...

c. *Clitic condition*

A estas chicas, mi hermana les dijo que...
To these girls, my sister them told that...

The following results were expected considering that speakers would require a clitic after a topicalized phrase. Condition (1a) should show a clitic continuation. Condition (1c) should show a clitic continuation in the embedded clause. Condition (1b) should have no clitic continuation predicted for the embedded clause because the clitic requirement is already fulfilled in the main clause.

Results

The completions were classified according to the use of clitics in the completion of the sentence fragment. Completions were classified as clitic presence responses if a clitic was provided and as clitic absence responses if a clitic was not provided. Results are shown in Table 2.

Clitic Continuations				
n=15 (90-6each condition)	Clitic Absence		Clitic Presence	
Conditions	N	%	N	%
a. Single NPs	9	10	81	90
b. Non-clitic	58	64.4	32	35.5
c. Clitic	81	90	9	10

Table 2: Completion Test, classification of sentence completions according to clitic continuations.

Participants provided completions for the sentence fragments on 100 % of the trials (270 trials). Of these, 45% had a clitic present (122 trials). The proportion of clitic continuations was 90% in the single NP condition (81 trials), 35.5 % in the non-

clitic condition (32 trials) and 10% in the clitic condition (9 trials). Fisher Exact tests were used to determine whether differences among conditions were reliable. Pairwise comparisons of (1a) and (1c) conditions showed a significant effect of clitic continuation ($p < .001$, 2-tailed), due to many fewer clitic continuations in condition (1c) than in (1a). Fisher Exact tests were used in condition (1b) because the proportion of clitic absent continuations in this condition was higher than expected (58%). Although clitic continuations were not so common in (1b) and (1c) conditions, pairwise comparisons showed a highly reliable effect of clitic continuation ($p < .0001$, 2-tailed) in condition (1b).

While the contrast observed in the (1a)/(1c) condition pair was expected from the grammaticality questionnaire results presented in section 3.2.3, the continuations produced in (1b) were not. Completions within condition (1b), the non-clitic condition, were classified in order to see what interfered in the production of clitic continuations leading to the high percentage values observed in Table 2 for Clitic Absence. Results are shown in Table 3.

Non-clitic Continuation						
	Subject pro-drop		Overt subject		Impersonal	
Condition	N	%	N	%	N	%
Non-clitic	43	74.1	11	18.9	4	6.8

Table 3: Completion Test, Classification of sentence completions according to non-clitic continuations.

Participants provided non-clitic continuations for the sentence fragments in the non-clitic condition on 64.4% of the trials (58/90 trials). Of these, 74.1 % were subject pro-drop continuations (43/58 trials) where the antecedents were: the subject (2a), the topicalized phrase (2b) or both (2c). The most common continuation was

that in (2b) where the subject pro-drop was co-referent with the topicalized phrase (36/43 trials)⁵⁵.

- (2) a. A aquellas famosas_i, la irónica dibujante_j comentó que...
To those famous women, the ironic cartoonist said that ...
pro_i pronto tendría una caricatura de ellas_i. (sample completion)
(she) soon have a caricature of they
'Those famous women, the ironic cartoonist said that...
she would have a caricature of them pretty soon'
- b. A estas ladronas_i, la comprensiva psicóloga mencionó que...
To these thieves, the understanding psychologist mentioned that....
pro_i padecían cleptomanía. (sample completion)
(they_i) suffered kleptomania
'These thieves, the understanding psychologist mentioned that...
they suffered from kleptomania'
- c. A esas montañeras_i, la enérgica guía_j aseguró que...
To those climbers, the energetic guide assured that....
pro_{i+j} no abandonarían la marcha. (sample completion)
(they_{i+j}) not abandon the walk
'Those climbers, the energetic guide assured that...
they would not stop hiking'

It should be pointed out that all the sentences in (3) can only be grammatical if speakers interpret a covert dative clitic at the matrix verb. This condition can only be grammatical if interpreted as if it had this underlying clitic. Without this covert clitic interpretation the sentence is ungrammatical because the topicalization cannot be interpreted anywhere.

⁵⁵ Notice that this result confirms the presupposition that we have done for the filled-subject effect in the Spanish experiment 2 where the most frequent continuation is that of a null subject that co-refers with the topic phrase.

- (3) a. A aquellas famosas_i, la irónica dibujante_j *(les_i) comentó que...
 To those famous, the ironic cartoonist *(CL-IO) said that ...
pro_i pronto tendría una caricatura de ellas_i. (sample completion)
 (she)_i soon have a caricature of they
 ‘*(To) those famous women, the ironic cartoonist told (them) that...
she would have a caricature of them pretty soon’

Discussion

Results of this completion experiment suggest that Spanish speakers use clitics to continue sentence fragments that contain a topicalization. In almost 45% of all cases the completed sentence had a clitic presence. Results for the off-line grammaticality questionnaire showing a high clitic requirement were replicated in the Single NP condition. The Clitic condition also confirmed that whenever there is a previous clitic in the sentence, subjects do not need an upcoming clitic. Finally, participants did not predominantly require an upcoming clitic in the embedded clause in the Non-clitic condition due to various interfering effects.

Taken together, the results of the completion experiment add evidence in favor of a clitic requirement when a topicalized phrase has been processed. Nevertheless, further tests are required since the Non-clitic condition had fewer clitic continuations than were initially expected. With this in mind, a second completion experiment was conducted to increase the number of clitic continuations in the conditions that did not present any clitic.

Appendix 3-B: Off-line Sentence Completion Experiment 2

The items used for the second experiment were constructed taking as a starting point the items from the first sentence completion experiment (appendix 3-A). Items were modified in the following way: First, the verb form was changed to past participle in an effort to prevent the covert clitic interpretation in the clitic-less condition. Then, an overt subject was included at the end of the sentence to prevent participants from completing the sentences with a subject pro-drop. Finally, a fourth condition was included where the verb could not take a dative argument and where the interpretation of the fronted phrase in the main clause should be banned.

Participants

Fifty-two native speakers of Castilian Spanish participated in this experiment.

Materials and Design

Eighteen items of four conditions each were used. Four lists were created by distributing the eighteen items in a Latin Square design. Each subject saw exactly one of the lists intermixed with thirty-six filler items in a random order. Items were derived from the experimental items for the completion experiment 1 (appendix 3-A). No context sentence was included in this test.

There were four conditions that included an initial sentence fragment such as those in (1). Participants were instructed to provide a completion that sounded natural to them. The new condition with respect to the completion experiment 1 was (1d) where the main verb could not take a dative argument.

- (1) a. *Single NPs condition:*
A estas chicas, mi hermana...
(To)these girls, my sister...
- b. *Non-clitic condition (say_type verb):*
A estas chicas, mi hermana ha dicho que el doctor...
(To)these girls, my sister has said that the doctor...
- c. *Clitic condition:*
A estas chicas, mi hermana les ha dicho que el doctor...
(To)these girls, my sister them has told that the doctor...
- d. *Non-clitic condition (think_type verb):*
A estas chicas, mi hermana ha pensado que el doctor...
(To)these girls, my sister has thought that the doctor...

The following results were expected considering that speakers would require a clitic after a topicalized phrase. Condition (1a) should show a clitic continuation and replicate the results for the same condition in experiment 1. Condition (1b) should show a clitic continuation in the embedded clause. Condition (1c) should have no clitic continuation predicted for the embedded clause because the clitic requirement is already fulfilled. Condition (1d) should have a clitic continuation expected since, as mentioned before, the type of verb used cannot take a dative argument and, therefore, the interpretation of the topicalization is not possible there.

Results

The completions were classified according to the use of clitics to finish the sentence fragment. Completions were classified as clitic presence responses if a clitic was provided and as clitic absence responses if a clitic was not provided. Results are shown in Table 4.

Clitic Continuations				
n=52 (260-5each condition)	Clitic Absence		Clitic Presence	
Conditions	n	%	n	%
a. Single NPs	18	6.9	242	93
b. Non-clitic (say_type verb)	157	60.4	103	39.6
c. Clitic (say_type verb)	166	63.8	94	36.2
d. Non-clitic (think_type verb)	97	37.3	163	62.7

Table 4: Completion Test, classification of sentence completions according to clitic continuations.

Participants provided completions for the sentence fragments on 100 % of the trials (1, 040 trials). Of these, 57.8% had a clitic present (602 trials). The proportion of clitic continuations was 93% in the single NP condition (242 trials), 39.6 % in the non-clitic (say_type verb) condition (103 trials), 62.7 % in the non-clitic (think_type verb) condition (163 trials) and 36.2% in the clitic condition (94 trials). Fisher Exact tests were used to determine whether differences among conditions were reliable. Pairwise comparisons of all conditions showed a reliable effect of clitic continuation ($p < .001$, 2-tailed) with the exception of the condition pair (1b)/(1c).

For the clitic-condition, having an embedded subject increased the number of completions with a clitic, which was 36.2 % (94/260 trials), as opposed to the 10% (9/90trials) of the completion experiment 1. In conditions (1b) and (1d) the proportion of clitic absent continuations was higher than expected, 60% and 37% respectively. Completions within these two conditions were classified in order to see what interfered in the production of clitic continuations leading to the high percentage values observed in Table 4 for Clitic Absence. Particularly in condition (1b), which was also tested in the completion experiment 1⁵⁶. Results are shown in Table 5.

⁵⁶ With the difference that in experiment 1 it had no overt embedded subject.

Clitic Absence Continuations				
n=52	Grammatical		Ungrammatical	
Conditions	N	%	n	%
b. NC_say_type (n=157)	157	100		
d. NC_think_type (n=97)	30	31	67	69(25.8)

Table 5: Completion Test, Classification of sentence completions according to non-clitic continuations.

Participants provided non-clitic continuations for the sentence fragments in the non-clitic condition (say_type) on 60.4% of the trials (157/260 trials) and the non-clitic condition (think_type) on 37.3% of the trials (97/260 trials).

In 51.1% cases of clitic absence in the non-clitic (say_type) condition, the topic could not be interpreted at the most embedded clause. These continuations are ungrammatical unless a (null) gap is interpreted in the main clause as exemplified in the examples in (2).

- (2) a. A esas estudiantes, la secretaria del decano *(les) ha contado que el decano...
*To these students, the secretary of the dean *(cl-IO) has said that the dean...
 estará de baja un mes. (sample completion)*
will be on a sick leave for a month.
 ‘These students, the secretary of the dean has said that the dean will be on a sick leave for a month’
- b. A aquellos pintores, la presentadora *(les) ha contado que la prensa...
*To those painters, the presenter *(cl-IO) has said that the press...
 engaña. (sample completion)*
deceives.
 ‘Those painters, the presenter of the show has said that the press deceives’

25.8% cases of clitic absence continuations in the non-clitic (think_type) condition consisted of ungrammatical sentences where subjects left the topicalized phrase uninterpreted as exemplified in the examples in (3a-b). Since the think-type verbs in this condition do not take dative arguments, the only way to interpret the left-

dislocation would be through a clitic in the embedded clause. As seen in the completion samples in (3a-b), there is no clitic provided in the embedded clause to interpret the topicalized phrase. This leaves the topicalized phrase uninterpreted.

- (3) a. *A aquellos delincuentes, la jefe de policía ha calculado que la abogada...
To those delinquents, the police chief has guessed that the lawyer...
no iba a tener esperanza. (sample completion)
not would have any hope.
'To those delinquents the police chief has guessed that the lawyer would not have any hope'
- b. *A aquellos delincuentes, la jefe de policía ha calculado que la abogada...
To those delinquents, the police chief has guessed that the lawyer...
llegará en diez minutos. (sample completion)
would arrive in ten minutes.
'To those delinquents the police chief has guesses that the lawyer would arrive in ten minutes'

Discussion

When subjects have to produce sentence continuations results showed that the clitic requirement was not so strong in certain conditions, particularly the two non-clitic conditions in (1b) and (1d). Moreover, the fact that in this completion experiment the embedded subject was added, disfavored clitic continuation completions and favored completions where the topicalization could not be interpreted in the most embedded clause unless subjects interpreted a (null) gap in the upstairs clause. This was particularly the case in the non-clitic (*say_type*) condition.

The non-clitic (*think_type*) condition allowed a direct comparison with the non-clitic (*say_type*) condition, useful for checking the hypothesis that speakers put gaps at the foot of the topicalization dependency and to test whether some of the clitic-less continuations were interpreted with a null gap upstairs.

Nevertheless, participants gave more ungrammatical continuations in this completion experiment than in completion experiment 1. The condition where this was more prominent was the non-clitic (think_type) condition where subjects could not interpret a null gap at the main verb and still gave clitic absence continuations on 37.3% of the trials. These continuations suggest that in this particular task, sometimes participants do not keep the topicalization in memory and they can give a completion to the sentence where the topicalization is not interpreted. This stands in contrast with results from the off-line grammaticality questionnaire where results showed a stronger clitic requirement when subjects were given the continuation to the topicalization.

It seems then that the clitic requirement has strength among Spanish speakers but this strength varies depending on the experimental task, where the sentence completion task does not show this clitic requirement as consistently as the off-line grammaticality-rating experiment 1a.

Appendix 3-C: Data items for online sentence reading experiment 1b

The experiment contained 24 sets of two conditions as presented below.

item 1.

Yendo a la escuela, mi hermana mayor y yo vimos a mis amigas Ana e Irene.

1a. A estas chicas, mi hermana mayor más tarde dijo que ya lo cree que las conoce desde hace tiempo.

1b. A estas chicas, mi hermana mayor más tarde les dijo que ya lo cree que las conoce desde hace tiempo.

¿Conoce la hermana a las chicas desde hace tiempo? Y

item 2

En la residencia de ancianos, la enfermera observaba a un grupo de ancianas con cariño.

2a. A esas ancianas, la joven enfermera recientemente dijo que ya lo cree que las puede ayudar en sus tareas.

2b. A esas ancianas, la joven enfermera recientemente les dijo que ya lo cree que las puede ayudar en sus tareas.

¿Fueron las ancianas las que dijeron que podían ayudar a la joven enfermera? N

item 3

En los pasillos de la facultad, un grupo de alumnas miraba el tablón de anuncios en busca de sus notas.

3a. A aquellas alumnas, la amable profesora vagamente comentó que ya lo deduce que las aprobarán en el examen final.

3b. A aquellas alumnas, la amable profesora vagamente les comentó que ya lo deduce que las aprobarán en el examen final.

¿Era amable la profesora? Y

item 4

Entre las bailarinas que hicieron la prueba, algunas destacaron por su profesionalidad.

4a. A esas bailarinas, la imaginativa coreógrafa reservadamente mencionó que ya lo supone que las quieren contratar para la próxima actuación.

4b. A esas bailarinas, la imaginativa coreógrafa reservadamente les mencionó que ya lo supone que las quieren contratar para la próxima actuación.

¿Fueron todas las bailarinas las que destacaron por su profesionalidad? N

item 5

Estando en zona conflictiva, algunas activistas arriesgaron la vida por su país.

5a. A esas activistas, la presidenta del partido satisfechamente aseguró que ya lo aprueba que las condecoren por su valor.

5b. A esas activistas, la presidenta del partido satisfechamente les aseguró que ya lo aprueba que las condecoren por su valor.

¿Aprueba la presidenta del partido que condecoren a las activistas? Y

item 6

Varias estudiantes visitaron el departamento al que habían mandado su inscripción.

6a. A aquellas estudiantes, la secretaria del decano poco después contó que ya lo espera que las admitan en el programa.

6b. A aquellas estudiantes, la secretaria del decano poco después les contó que ya lo espera

que las admitan en el programa.

¿Fueron unos profesores los que visitaron el departamento? N

item 7.

Durante los últimos minutos, varias atletas no pudieron llegar al final del maratón.

7a. A esas atletas, la exigente entrenadora rudamente comentó que ya lo supone que las excluirán de la selección de los juegos olímpicos.

7b. A esas atletas, la exigente entrenadora rudamente les comentó que ya lo supone que las excluirán de la selección de los juegos olímpicos.

¿Supone la entrenadora que excluirán a las atletas de la selección? Y

item 8.

Algunas chicas jóvenes del internado querían asistir al baile de fin de curso acompañadas.

8a. A estas adolescentes, la madre superiora sensatamente dijo que ya lo aprueba que las acompañen al baile.

8b. A estas adolescentes, la madre superiora sensatamente les dijo que ya lo aprueba que las acompañen al baile.

¿Era la madre superiora la que quería asistir al baile? N

item 9.

Durante el campeonato, varias gimnastas fueron maleducadas con sus contrincantes.

9a. A esas gimnastas, la juez del campeonato seguidamente mencionó que ya lo intuye que las vayan a descalificar por su poca deportividad.

9b. A esas gimnastas, la juez del campeonato seguidamente les mencionó que ya lo intuye que las vayan a descalificar por su poca deportividad.

¿Fueron algunas de las gimnastas maleducadas con sus contrincantes? Y

item 10.

En la comisaría, varias detenidas esperaban a ser interrogadas.

10a. A aquellas delincuentes, la jefe de policía bruscamente dijo que ya lo sospecha que las condenarán a algunos días de cárcel.

10b. A aquellas delincuentes, la jefe de policía bruscamente les dijo que ya lo sospecha que las condenarán a algunos días de cárcel.

¿Era la jefe de prisiones la que sospechaba que condenarían a las delincuentes a varios días de cárcel? N

item 11.

Unas cuantas mujeres embarazadas se desmayaron debido al calor que afectó a la región.

11a. A esas mujeres, la joven comadrona expertamente comentó que ya lo desea que las ayuden a la hora del parto.

11b. A esas mujeres, la joven comadrona expertamente les comentó que ya lo desea que las ayuden a la hora del parto.

¿Es a la hora del parto cuando la comadrona desea que ayuden a las mujeres embarazadas? Y

item 12.

En la agencia de traductores, algunas de las traductoras tenían que traducir hasta a cinco idiomas.

12a. A estas traductoras, la audaz editora confidentemente dijo que ya lo confiesa que las exigen mucho en la agencia.

12b. A estas traductoras, la audaz editora confidentemente les dijo que ya lo confiesa que las exigen mucho en la agencia.

¿Era hasta diez idiomas a lo que las traductoras tenían que traducir? N

item 13.

En las clases de canto, algunas de las niñas se quedaban afónicas de tanto ensayar para las funciones.

13a. A esas niñas, la maestra de canto fríamente comentó que ya lo reconoce que las pide demasiado.

13b. A esas niñas, la maestra de canto fríamente les comentó que ya lo reconoce que las pide demasiado.

¿Era a esas niñas a las que la maestra reconoce que exige demasiado? Y

item 14.

Entre las empleadas del supermercado, algunas sólo trabajaban los fines de semana.

14a. A estas empleadas, la jefa de personal maliciosamente comentó que ya lo duda que las vayan a necesitar en otras ocasiones.

14b. A estas empleadas, la jefa de personal maliciosamente les comentó que ya lo duda que las vayan a necesitar en otras ocasiones.

¿Eran las empleadas de un video-club a las que la jefa de personal dudaba que fuesen a necesitar? N

item 15.

Las chicas que robaron el banco fueron sometidas a un examen psicológico en la comisaría.

15a. A aquellas ladronas, la comprensiva psicóloga cuidadosamente mencionó que ya lo teme que las vayan a encarcelar por unos cuantos años.

15b. A aquellas ladronas, la comprensiva psicóloga cuidadosamente les mencionó que ya lo teme que las vayan a encarcelar por unos cuantos años.

¿Mencionó la psicóloga que teme que vayan a encarcelar a las ladronas? Y

item 16.

De toda la expedición, el grupo de montañeras era el más capacitado para llegar a la cumbre.

16a. A esas montañeras, la enérgica guía complacientemente aseguró que ya lo garantiza que las acompañará hasta la cima.

16b. A esas montañeras, la enérgica guía complacientemente les aseguró que ya lo garantiza que las acompañará hasta la cima.

¿Era poco enérgica la guía? N

item 17.

Varias actrices televisivas protestaron por ser ridiculizadas en el último cómic de la dibujante.

17a. A aquellas famosas, la irónica dibujante divertidamente comentó que ya lo lamenta que las caricaturice en su libro.

17b. A aquellas famosas, la irónica dibujante divertidamente les comentó que ya lo lamenta que las caricaturice en su libro.

¿Era irónica la dibujante que caricaturizó a las famosas en su libro? Y

item 18.

Algunas inmigrantes no tenían permiso legal para entrar en el país, pero quisieron intentar entrar de todas formas.

18a. A estas inmigrantes, la agente policial inoportunamente mencionó que ya lo sospecha que las retendrán en la frontera.

18b. A estas inmigrantes, la agente policial inoportunamente les mencionó que ya lo sospecha que las retendrán en la frontera.

¿Tenían todas las inmigrantes permiso legal para entrar en el país? N

item 19.

El programa de radio dedicó un especial a un grupo de científicas premiadas por su labor investigadora.

19a. A esas científicas, la locutora de radio oportunamente aseguró que ya lo destaca que las admira por su trabajo.

19b. A esas científicas, la locutora de radio oportunamente les aseguró que ya lo destaca que las admira por su trabajo.

¿Es por su trabajo por lo que la locutora de radio admira a las científicas? Y

item 20.

Varias cantantes no acudieron al programa de televisión porque no las avisaron a tiempo.

20a. A estas cantantes, la organizadora del programa ásperamente dijo que ya lo reconoce que las llamaron demasiado tarde.

20b. A estas cantantes, la organizadora del programa ásperamente les dijo que ya lo reconoce que las llamaron demasiado tarde.

¿Llamaron a las cantantes a tiempo para acudir al programa de televisión? N

item 21.

Tres niñas saltaban a la cuerda en la entrada de la casa mientras su abuela y una amiga las contemplaban.

21a. A aquellas nietas, su cariñosa abuela tiernamente dijo que ya lo confiesa que las quiere con locura.

21b. A aquellas nietas, su cariñosa abuela tiernamente les dijo que ya lo confiesa que las quiere con locura.

¿Confiesa la abuela que quiere a sus nietas con locura? Y

item 22.

Varias de las camareras del bar estuvieron enfermas en casa durante una larga temporada.

22a. A esas camareras, la encargada del bar torpemente contó que ya lo presiente que las reemplazarán en pocos días.

22b. A esas camareras, la encargada del bar torpemente les contó que ya lo presiente que las reemplazarán en pocos días.

¿Fue la encargada la que estuvo durante un largo tiempo en casa enferma? N

item 23.

Entre los nuevos empleados de la empresa, varias chicas destacaban por su trabajo.

23a. A estas empleadas, la secretaria jefe complacidamente dijo que ya lo supone que las vayan a ascender en menos de un mes.

23b. A estas empleadas, la secretaria jefe complacidamente les dijo que ya lo supone que las vayan a ascender en menos de un mes.

¿Supone la secretaria jefe que se ascenderá a las empleadas en menos de un mes? Y

item 24.

Algunas alumnas de medicina eran también madres que no encontraban tiempo libre para sus hijos.

24a. A aquellas madres, la dulce niñera tristemente mencionó que ya lo piensa que las extrañan sus hijos.

24b. A aquellas madres, la dulce niñera tristemente les mencionó que ya lo piensa que las extrañan sus hijos.

¿Tenían las madres tiempo libre para sus hijos? N

Appendix 3-D: Data items for online sentence reading experiment 2

The experiment contained 24 different items with 4 test conditions, that are presented below, preceded by the context used for each of them.

item 1

Cuando abrieron las puertas del teatro en que se celebraba el casting, entró precipitadamente un grupo de chicas.

1a. A estas chicas, el organizador del casting ha explicado con mucho detalle que el manager las iría llamando por apellido.

1b. A estas chicas, el organizador del casting les ha explicado con mucho detalle que el manager las iría llamando por apellido.

1c. El organizador del casting ha explicado con mucho detalle que el manager las iría llamando por apellido.

1d. El organizador del casting les ha explicado con mucho detalle que el manager las iría llamando por apellido.

¿Era en un teatro donde se celebraba el casting? Y

item 2

Mientras el tenor cantaba la última aria del repertorio durante el ensayo, se deslizó silenciosamente por la sala un grupo de periodistas.

2a. A estas periodistas, el director del teatro ha dicho con mucha calma que el tenor las iría recibiendo al finalizar el ensayo.

2b. A estas periodistas, el director del teatro les ha dicho con mucha calma que el tenor las iría recibiendo al finalizar el ensayo.

2c. El director del teatro ha dicho con mucha calma que el tenor las iría recibiendo al finalizar el ensayo.

2d. El director del teatro les ha dicho con mucha calma que el tenor las iría recibiendo al finalizar el ensayo.

¿Es antes del ensayo cuando el director del teatro ha dicho que el tenor recibirá a las periodistas? N

item 3 a

En el momento en que apareció el famoso cantante, se concentró repentinamente frente al escenario un grupo de fans.

3a. A estos fans, el agente de seguridad ha gritado con tono frustrado que la policía los iba a desalojar si no se mantenían a la distancia indicada.

3b. A estos fans, el agente de seguridad les ha gritado con tono frustrado que la policía los iba a desalojar si no se mantenían a la distancia indicada.

3c. El agente de seguridad ha gritado con tono frustrado que la policía los iba a desalojar si no se mantenían a la distancia indicada.

3d. El agente de seguridad les ha gritado con tono frustrado que la policía los iba a desalojar si no se mantenían a la distancia indicada.

¿Es el agente de seguridad quien ha gritado? Y

item 4

Mientras la maestra se ausentaba para hacer fotocopias, salió sigilosamente al patio un grupo de niños.

4a. A esos niños, el conserje del colegio ha asegurado con aire descontento que el director los iría a buscar si no volvían inmediatamente a clase.

4b. A esos niños, el conserje del colegio les ha asegurado con aire descontento que el director los iría a buscar si no volvían inmediatamente a clase.

4c. El conserje del colegio ha asegurado con aire descontento que el director los iría a buscar si no volvían inmediatamente a clase.

4d. El conserje del colegio les ha asegurado con aire descontento que el director los iría a buscar si no volvían inmediatamente a clase.

¿Es a la maestra a quien el director iría a buscar si no volvía inmediatamente a clase? N

item 5

Cuando el equipo de baloncesto acabó el entrenamiento, entró alegremente a la cancha un grupo de animadoras.

5a. A aquellas animadoras, el entrenador del equipo ha contado con poca discreción que el coreógrafo las iba a suplantar en la próxima temporada.

5b. A aquellas animadoras, el entrenador del equipo les ha contado con poca discreción que el equipo las iba a suplantar en la próxima temporada.

5c. El entrenador del equipo ha contado con poca discreción que el coreógrafo las iba a suplantar en la próxima temporada.

5d. El entrenador del equipo les ha contado con poca discreción que el coreógrafo las iba a suplantar en la próxima temporada.

¿Ha sido poco discreto el entrenador? Y

item 6

En el momento en que se acabó la película en el estreno del festival de cine, apareció inmediatamente en el escenario un grupo de actores.

6a. A esos actores, el director de la película ha mencionado en tono relajado que el presentador los iría nombrando por orden de aparición.

6b. A esos actores, el director de la película les ha mencionado en tono relajado que el presentador los iría nombrando por orden de aparición.

6c. El director de la película ha mencionado en tono relajado que el presentador los iría nombrando por orden de aparición.

6d. El director de la película les ha mencionado en tono relajado que el presentador los iría nombrando por orden de aparición.

? ¿Iba a nombrar el presentador a los actores de manera aleatoria? N

item 7

Cuando salieron los niños de clase, se concentró paulatinamente frente a la puerta del aula un grupo de madres.

7a. A esas madres, la secretaria del colegio ha dicho de forma casual que la profesora las iría llamando una a una para recoger a sus hijos.

7b. A esas madres, la secretaria del colegio les ha dicho de forma casual que la profesora las iría llamando una a una para recoger a sus hijos.

7c. La secretaria del colegio ha dicho de forma casual que la profesora las iría llamando una a una para recoger a sus hijos.

7d. La secretaria del colegio les ha dicho de forma casual que la profesora las iría llamando una a una para recoger a sus hijos.

? ¿Será la profesora la que llamará a las madres para recoger a sus hijos? Y

item 8

Estando a punto de cerrar el taller, llegó a última hora en una furgoneta estropeada un grupo de jugadores de fútbol.

8a. A estos jugadores, el dueño del taller ha comentado sin mucho interés que el mecánico los iba a llamar para decirles cuánto costaría el arreglo.

8b. A estos jugadores, el dueño del taller les ha comentado sin mucho interés que el mecánico los iba a llamar para decirles cuánto costaría el arreglo.

8c. El dueño del taller ha comentado sin mucho interés que el mecánico los iba a llamar para decirles cuánto costaría el arreglo.

8d. El dueño del taller les ha comentado sin mucho interés que el mecánico los iba a llamar para decirles cuánto costaría el arreglo.

? ¿Estaba la furgoneta en buenas condiciones? N

item 9

Tras salir todos los pasajeros del avión, se reunió poco a poco en la cafetería del aeropuerto un grupo de azafatas.

9a. A esas azafatas, el piloto del vuelo ha asegurado de forma convincente que el autobús las iba a llevar a cada una a su hotel.

9b. A esas azafatas, el piloto del vuelo les ha asegurado de forma convincente que el autobús las iba a llevar a cada una a su hotel.

9c. El piloto del vuelo ha asegurado de forma convincente que el autobús las iba a llevar a cada una a su hotel.

9d. El piloto del vuelo les ha asegurado de forma convincente que el autobús las iba a llevar a cada una a su hotel.

? ¿Es el piloto quien ha asegurado que las azafatas serían llevadas a su hotel en autobús? Y

item 10

Durante el entrenamiento de las bailarinas profesionales, irrumpió escandalosamente en la academia de baile un grupo de niñas.

10a. A aquellas niñas, el profesor ha contado de forma impaciente que el instructor las iba a echar si no se comportaban.

10b. A aquellas niñas, el profesor les ha contado de forma impaciente que el instructor las iba a echar si no se comportaban.

10c. El profesor ha contado de forma impaciente que el instructor las iba a echar si no se comportaban.

10d. El profesor les ha contado de forma impaciente que el instructor las iba a echar si no se comportaban.

? ¿Irrumpieron las niñas de manera silenciosa en la clase? N

item 11

Antes de acabar la clase de natación infantil, se zambulló ágilmente en la piscina un grupo de nadadoras.

11a. A estas nadadoras, el socorrista ha asegurado de mal humor que el entrenador las iba a echar del club si no salían de la piscina inmediatamente.

11b. A estas nadadoras, el socorrista les ha asegurado de mal humor que el entrenador las iba a echar del club si no salían de la piscina inmediatamente.

11c. El socorrista ha asegurado de mal humor que el entrenador las iba a echar del club si no salían de la piscina inmediatamente.

11d. El socorrista les ha asegurado de mal humor que el entrenador las iba a echar del club si no salían de la piscina inmediatamente.

? ¿Debían salir las nadadoras de la piscina inmediatamente? Y

item 12

Cuando la bandera roja que prohibía el baño fue retirada, se remojó cautelosamente en la orilla del mar un grupo de bañistas.

12a. A esos bañistas, el guardacostas ha dicho de forma preocupada que el socorrista los estaría vigilando desde su puesto.

12b. A esos bañistas, el guardacostas les ha dicho de forma preocupada que el socorrista los estaría vigilando desde su puesto.

12c. El guardacostas ha dicho de forma preocupada que el socorrista los estaría vigilando desde su puesto.

12d. El guardacostas les ha dicho de forma preocupada que el socorrista los estaría vigilando desde su puesto.

? ¿Ha sido el socorrista quien se ha remojado en el mar? N

item 13

En el momento en que abrieron las puertas de la cafetería, se apresuró a la barra del local a pedir un grupo de estudiantes.

13a. A estos estudiantes, el encargado del local ha explicado con buen humor que el camarero los podía atender en la mesa.

13b. A estos estudiantes, el encargado del local les ha explicado con buen humor que el camarero los podía atender en la mesa.

13c. El encargado del local ha explicado con buen humor que el camarero los podía atender en la mesa.

13d. El encargado del local les ha explicado con buen humor que el camarero los podía atender en la mesa.

? ¿Es a la barra donde un grupo de estudiantes se ha apresurado a pedir? Y

item 14

Mientras el guía del tour revisaba el itinerario, se puso a sacar fotos al parlamento un grupo de turistas.

14a. A estos turistas, el guía ha comentado con mucha discrección que el guarda los podía detener si no dejaban de hacer fotos.

14b. A estos turistas, el guía les ha comentado con mucha discrección que el guarda los podía detener si no dejaban de hacer fotos.

14c. El guía ha comentado con mucha discrección que el guarda los podía detener si no dejaban de hacer fotos.

14d. El guía les ha comentado con mucha discrección que el guarda los podía detener si no dejaban de hacer fotos.

? ¿Era el guía quien se había puesto a hacer fotos al parlamento? N

item 15

Al comenzar el segundo tiempo del partido de fútbol, salió rápidamente al campo un grupo de suplentes.

15a. A estas suplentes, el entrenador ha asegurado de forma nerviosa que el árbitro las podría descalificar a la mínima falta.

15b. A estas suplentes, el entrenador les ha asegurado de forma nerviosa que el árbitro las podría descalificar a la mínima falta.

15c. El entrenador ha asegurado de forma nerviosa que el árbitro las podría descalificar a la mínima falta.

15d. El entrenador les ha asegurado de forma nerviosa que el árbitro las podría descalificar a la mínima falta.

? ¿Son las suplentes las que podrían ser descalificadas a la mínima falta? Y

item 16

Cuando la entrega de premios literarios hubo acabado, se concentró en la entrada de la sala un grupo de redactores.

16a. A aquellos redactores, el organizador de la entrega ha comentado con mucha seriedad que el ganador los podía recibir en persona.

16b. A aquellos redactores, el organizador de la entrega les ha comentado con mucha seriedad que el ganador los podía recibir en persona.

16c. El organizador de la entrega ha comentado con mucha seriedad que el ganador los podía recibir en persona.

16d. El organizador de la entrega les ha comentado con mucha seriedad que el ganador los podía recibir en persona.

? ¿Era una entrega de premios cinematográficos para lo que los redactores se habían concentrado? N

item 17

En la expedición que subía a la escarpada montaña, se quedó incomunicado a mitad de camino hacia la cima un grupo de montañeras.

17a. A esas montañeras, el equipo de salvamento ha dicho de forma esperanzadora que el helicóptero las iba a buscar hasta que oscureciese.

17b. A esas montañeras, el equipo de salvamento les ha dicho de forma esperanzadora que el helicóptero las iba a buscar hasta que oscureciese.

17c. El equipo de salvamento ha dicho de forma esperanzadora que el helicóptero las iba a buscar hasta que oscureciese.

17d. El equipo de salvamento les ha dicho de forma esperanzadora que el helicóptero las iba a buscar hasta que oscureciese.

? ¿Ha sido el equipo de salvamento el que ha dicho que buscarían a las montañeras hasta que oscureciese? Y

item 18

Cuando las luces del campamento se hubieron apagado, se reunió sigilosamente en la sala común un grupo de monitores.

18a. A esos monitores, el director del campamento ha contado de forma secreta que el inspector los iba a supervisar a la mañana siguiente.

18b. A esos monitores, el director del campamento les ha contado de forma secreta que el inspector los iba a supervisar a la mañana siguiente.

18c. El director del campamento ha contado de forma confidencial que el inspector los iba a supervisar a la mañana siguiente.

18d. El director del campamento les ha contado de forma confidencial que el inspector los iba a supervisar a la mañana siguiente.

? ¿Ha sido de forma pública que el director ha mencionado la supervisión del inspector? N

item 19

En el momento en que se abrieron las puertas del avión, salió pausadamente un grupo de voluntarias.

19a. A estas voluntarias, el médico ha comentado con gran satisfacción que el embajador las iba a invitar a su mansión como agradecimiento.

19b. A estas voluntarias, el médico les ha comentado con gran satisfacción que el embajador las iba a invitar a su mansión como agradecimiento.

19c. El médico ha comentado con gran satisfacción que el embajador las iba a invitar a su mansión como agradecimiento.

19d. El médico les ha comentado con gran satisfacción que el embajador las iba a invitar a su mansión como agradecimiento.

? ¿Iban a ser las voluntarias invitadas a la mansión del embajador? Y

item 20

Después de que acabaron las obras de remodelación, se instaló en el nuevo edificio una cooperativa de artistas.

20a. A estos artistas, el agente inmobiliario ha contado de forma breve que el arquitecto los iba a visitar en pocos días.

20b. A estos artistas, el agente inmobiliario les ha contado de forma breve que el arquitecto los iba a visitar en pocos días.

20c. El agente inmobiliario ha contado de forma breve que el arquitecto los iba a visitar en pocos días.

20d. El agente inmobiliario les ha contado de forma breve que el arquitecto los iba a visitar en pocos días.

? ¿Es el agente inmobiliario quien se ha instalado en el nuevo edificio? N

item 21

Mientras el médico visitaba a los pacientes, pasó riendose por el pasillo adyacente un grupo de enfermeras.

21a. A esas enfermeras, el médico ha dicho de forma confidencial que el supervisor las iba a entrevistar para evaluar su rendimiento.

21b. A esas enfermeras, el médico les ha dicho de forma confidencial que el supervisor las iba a entrevistar para evaluar su rendimiento.

21c. El médico ha dicho de forma confidencial que el supervisor las iba a entrevistar para evaluar su rendimiento.

21d. El médico les ha dicho de forma confidencial que el supervisor las iba a entrevistar para evaluar su rendimiento.

? ¿Iba a evaluar el supervisor a las enfermeras? Y

item 22

Al pasar la excursión de arte de la escuela secundaria por la zona del museo, se detuvo de repente ante el edificio un grupo de alumnos.

22a. A aquellas alumnos, el profesor ha asegurado con aire cansado que el guía los podía llevar allí en cualquier otro momento.

22b. A aquellas alumnos, el profesor les ha asegurado con aire cansado que el guía los podía llevar allí en cualquier otro momento.

22c. El profesor ha asegurado con aire cansado que el guía los podía llevar allí en cualquier otro momento.

22d. El profesor les ha asegurado con aire cansado que el guía los podía llevar allí en cualquier otro momento.

? ¿Quería el profesor ir al museo? N

item 23

Tras llegar a la frontera con el país vecino, esperó pacientemente en la aduana un grupo de refugiados.

23a. A esos refugiados, el agente de aduanas ha mencionado sin ninguna emoción que el gobierno los iba a deportar al entrar en el país vecino.

23b. A esos refugiados, el agente de aduanas les ha mencionado sin ninguna emoción que el gobierno los iba a deportar al entrar en el país vecino.

23c. El agente de aduanas ha mencionado sin ninguna emoción que el gobierno los iba a deportar al entrar en el país vecino.

23d. El agente de aduanas les ha mencionado sin ninguna emoción que el gobierno los iba a deportar al entrar en el país vecino.

? ¿Era en la frontera con el país vecino donde el grupo de refugiados esperaba? Y

item 24

Antes de que se cerrasen las puertas del centro comercial, entró rápidamente un grupo de clientes.

24a. A aquellos clientes, el agente de seguridad ha dicho sin mucho interés que el encargado los iba a desalojar en pocos minutos.

24b. A aquellos clientes, el agente de seguridad les ha dicho sin mucho interés que el encargado los iba a desalojar en pocos minutos.

24c. El agente de seguridad ha dicho sin mucho interés que el encargado los iba a desalojar en pocos minutos.

24d. El agente de seguridad les ha dicho sin mucho interés que el encargado los iba a desalojar en pocos minutos.

? ¿Entro el grupo de clientes al centro comercial de manera pausada? N

CHAPTER 4: THE ACTIVE-SEARCH FOR CLITIC PRONOUNS IN GALICIAN

4.1. Active search for clitic pronouns in CLLD constructions vs. active search for gaps

As I have previously discussed in chapter 3 for the processing of Spanish CLLD dependencies, it is an agreed fact that there is an active search for a gap in the case of *wh*-dependencies (Crain & Fodor 1985, Stowe 1986). Since the position where the dislocated phrase is interpreted in these *wh*-dependencies is null, it could be that the active search for a gap in these cases is a specific response to poor input. This would then predict a less active approach in dependencies that involve overt elements since they can be identified more easily in a bottom-up manner. Nevertheless, in the discussion of the experiments on Spanish CLLD dependencies, I concluded that active dependency formation can occur regardless of whether the tail of the dependency is phonetically null or overt. This means the active search mechanism is also triggered in dependencies involving clitic pronouns.

Moreover, there is evidence from other kinds of long-distance dependencies involving pronouns that suggest that the active search mechanism is a more general mechanism in the parser. This evidence comes from backward anaphora cases (van Gompel & Liversedge 2003, Sturt et al. 2004, Kazanina 2005, Kazanina et al. 2006,

Aoshima et al. 2006) and shows that there is also an active search for the antecedent of a pronoun.

4.1.1. Active search for pronouns: CLLD dependencies

Results from the previously discussed self-paced reading experiments using Spanish CLLD constructions have demonstrated that topicalization does affect the processing of the clitics in the sentence and that there is an active search for clitics in these constructions. But what does it exactly imply to say that there is an active search for clitics? Having an active search for clitics would mean that the position of the clitic is anticipated before its overt realization in the linear order. As in the case of wh-dependencies, the clitic in topic-clitic dependencies would be expected at every available position so that the interpretation of the dislocated phrase is made as soon as possible. When the interpretation of the dislocated phrase is not possible due to a delay in the appearance of the clitic, we expect a surprise on the part of the parser with effects similar to those observed in wh-dependencies when the gap position is filled with an extra argument (Crain & Fodor 1985, Stowe 1986, Aoshima et al. 2004, Lee 2004).

4.2. Galician as a case study for a stronger clitic prediction than Spanish

This study uses Galician to test if the same active search mechanism for clitic pronouns that was used in Spanish is used in a language where the requirement to have clitics in topic-clitic constructions and other constructions involving clitic pronouns is more robust and consistent across speakers. The off-line grammaticality experiment that I report in section 4.5 of this chapter shows that Galician has a strong clitic-requirement.

As seen in chapter 3 on the processing of Spanish CLLD dependencies, results from the Completion experiments on CLLD dependencies showed fewer clitic continuations than expected suggesting that some participants may have been able to accommodate a null clitic reading even if that reading is normally considered ungrammatical. This is the case in example (2) in Appendix 3-B repeated here as (1)

- (1) A esas estudiantes, la secretaria del decano *(les) ha contado que el decano...
*To these students, the secretary of the dean *(cl-IO) has said that the dean...
estará de baja un mes.
will be on a sick leave for a month.
'These students, the secretary of the dean has said that the dean will be on a sick leave for a month'*

Galician seems to be less flexible with regard to the use of null clitic interpretations as exemplified in the contrast in (2a-b). The null clitic interpretation is impossible for Galician speakers and the only way to get that interpretation is by having an overt realization of the clitic (2b).

- (2) a. * Ás estudantes, a secretaria do decano (**Ø clitic**) contou que o decano... estará de baixa un mes.
To these students, the secretary of the dean said that the dean...will be on a sick leave for a month.
- b. Ás estudantes, a secretaria do decano contou-lles que o decano... estará de baixa un mes.
To these students, the secretary of the dean told them that the dean...will be on a sick leave for a month.
 ‘These students, the secretary of the dean has said that the dean will be on a sick leave for a month’

Similarly, clitic doubling, a construction somewhat related to Clitic Left-Dislocation as discussed in chapter 2 (Cinque 1990, Iatridou 1995, Cecchetto 2000), seems to be more robust in Galician than Spanish, as exemplified in the contrast in (3a-b) and (4a-b), where the sentence in (4b) is possible in Spanish but ungrammatical in (3b) in Galician⁵⁷.

- (3) a. Quitamos-lle una foto ós nenos.
 (we) Took-CL-(IO) a picture to+the children
- b. *Quitamos una foto ós nenos.
 (we) Took a picture to+the children
 ‘We took a picture of the kids’ (Juan Uriagereka, p.c.)
- (4) a. Les sacamos una foto a los niños.
 CL-(IO) (we)Took a picture PA the kids
- b. ?Sacamos una foto a los niños.
 (We) Took a picture of the children
 ‘We took a picture of the children’

⁵⁷ European Portuguese, on the other hand is very restrictive with respect to the availability of clitic doubling and only the sentence without the clitic is allowed. (Ana María Martins, p.c.)

- (i) *Tirámos-lhe(s) uma fotografia ao(s) menino(s)
 (we) took-CL-OI-PL/SG a picture to the kid(s)
- (ii) Tirámos uma fotografia ao(s) menino(s)
 (we) took a picture to the kid(s)

This raises an interesting contrast for clitic doubling within these three languages where clitic doubling is impossible in European Portuguese, optional in Spanish and obligatory in Galician.

Galician > Spanish > European Portuguese

4.3 Spanish vs. Galician clitic placement

In the following section, we discuss differences in clitic placement between Spanish and Galician. The data on Galician clitic placement has been adapted from the data on European Portuguese discussed in (Raposo & Uriagereka 2003).

4.3.1. Spanish clitic placement

The pattern of clitic placement in Spanish is sensitive to the finiteness of the clause that hosts the clitic, irrespective of the main/ subordinate clause distinction.

Proclisis is obligatory in finite clauses in Spanish, as exemplified by (5) through (7).

- (5) Te vi ayer (*vite)
*CL(DO) see yesterday (*see-CL(DO))*
'I saw you yesterday'
- (6) ¿Quién te vió ayer?
Who CL(DO) saw yesterday
'Who saw you yesterday?'
- (7) Pienso que te vi ayer (*vite)
*(I) think that (I) CL(DO) see yesterday (*see-CL(DO))*
'I think I saw you yesterday'

Enclisis, on the other hand, is obligatory in non-finite clauses in Spanish as exemplified in (8).

- (8) Después de verte, me fui (*tever)
*After of to see-CL(DO), CL(REFL) leave (*CL(DO)-to see)*
 ‘After seeing you, I left’

4.3.2. Galician clitic placement⁵⁸

Enclisis and *proclisis* are allowed in both finite and infinitival clauses in Galician.

In root clauses, *enclisis* and *proclisis* are in complementary distribution. *Proclisis* is used when there is an affective verb or adverb to the left of the verb (e.g. a *wh*-phrase), as shown in (9). *Enclisis* is elsewhere, in simple declarative root clauses, as shown in (10).

- (9) Quen che viu onte? (*viuche)
*Who CL(DO) see yesterday (*see-CL(DO))*
 ‘Who saw you yesterday?’

- (10) (Eu) vinche onte (*che vin)
*(I) saw-CL(DO) yesterday (*CL(DO) saw)*
 ‘I saw you yesterday’

On the other hand, in subordinate clauses, *proclisis* is virtually obligatory.

- (11) Ela di que che chamou onte (*chamouche)
*She say that CL(DO) call yesterday (*call-CL(DO))*
 ‘She said that she called you yesterday’

⁵⁸ The placement of Galician clitics will have some relevance when presenting the stimuli for the self-paced reading experiment in section 4.5 since, to test pre-verbal structure building effects, all the clitics have to appear in pre-verbal position.

4.4 Experiment 3a: Off-line Grammaticality-rating Experiment

The requirement to have a clitic pronoun in topic-clitic constructions is considered to be quite robust across Galician speakers. The strength of this clitic requirement in topic-clitic constructions in Galician was tested through a grammaticality-rating experiment.

Forty-two native speakers of Galician completed the off-line questionnaire. Participants were asked to rate each of the conditions in the sentences on a scale from one (unacceptable) to five (acceptable), depending on how natural the sentences sounded to them, where the expectation was that the cliticless condition in (12b) would be considered ungrammatical.

The sentences included in the questionnaire were simplified and shortened versions of the target items in the on-line study with no embedding and where the verb was obligatorily transitive. The clitics used were direct object clitics in accusative case (either feminine or masculine in gender) and since the sentence had no embedding, the clitic appeared in post-verbal position (see section 4.3.2).

Sixteen items with two conditions⁵⁹ each were created for this off-line questionnaire. The two conditions created for each item differed in the presence or

⁵⁹ Two more conditions were originally included in this experiment to test a syntactic phenomenon specific to Galician called “interpolation”. Unlike other Romance languages, Galician allows a constituent (e.g.: a subject, an adverb or negation) to intervene between the clitic pronoun and the verb, so long as all the constituents are within the same clause (Alvarez et al. 1986).

(i). Fai o que che eu digo.

Do it that you I tell

‘Do what I tell you’

(ii). Quen me alí pillara.

absence of a direct object clitic. Each subject saw the two conditions of each sentence together.

The critical sentence was preceded by a context sentence. A complete set of experimental conditions is shown in (12).

(12) *Context:*
De tódolos obreiros da construción, algúns teñen moitos anos de experiencia.
From all+the workers of+the construction, some have many years of experience.

a. *Clitic condition:*

Ós⁶⁰ obreiros, o capataz necesitaos.
To+the workers, the foreman needs-CL(DO).
'All the workers, the foreman needs'

b. *Clitic-less condition:*

*Ós obreiros, o capataz necesita.
To+the workers, the foreman needs.
'All the workers, the foreman needs'

Table 1 shows the average of ratings of clitic continuation preference in the two conditions.

Who me there catch

'If I could be there!'

(iii). *Mire que eu lle non podo.*

Look that I it-dat not can

'Look, I can't do it'

This so-called "interpolation" effect would test if effects apparent at the clitic pronoun in existing results of the two self-paced reading experiments on topic-clitic dependencies in Spanish (chapter 3) were completely independent of the verb. Unfortunately, the grammaticality-ratings had very low scores (signaling that the construction is not accepted by Galician speakers nowadays) and it was decided to exclude these two conditions from the analysis.

⁶⁰ There is one thing that should be noted about the configuration of Galician and it is that the preposition and the article come together forming a cluster.

Condition	Mean	StDev	StdErr
Clitic	4.32	0.06	0.02
Clitic-less	2.37	0.14	0.03

Table 1: Grammaticality-rating Questionnaire, rating of the off-line clitic continuation preference experiment.

These results clearly show that subjects have a preference for a clitic continuation and that they require a clitic after a topicalized phrase has been introduced. The clitic-less condition had a higher rate than expected due to two other conditions that were included in the test⁶¹.

4.5 Experiment 3b: On-line Sentence Reading Experiment

Experiment 3b examines whether there is active search for a clitic pronoun in CLLD constructions in Galician. The active search mechanism should be effective for clitic pronouns in Galician CLLD constructions if it is true that an uninterpreted dislocated phrase triggers the search for a clitic pronoun. Additionally, if the presence of the clitic pronoun is postponed to the embedded clause and if there is an

⁶¹ The other two conditions (c/d) were bi-clausal sentences with a topicalized phrase, that differed in the clitic pronoun at the embedded sentence being interpolated or not (having the subject intervening between the clitic and the verb). This condition tested how acceptable the interpolation was for Galician speakers.

c. Ós obreiros, o capataz di que a empresa constructora **os** necesíta.

To+the workers, the foreman says that the construction company CL(DO) needs.

d. Ós obreiros, o capataz di que os a empresa constructora necesíta.

To+the workers, the foreman says that CL(DO) the construction company needs.

The contrast between conditions (c) and (d) shows that Galician speakers do not find the clitic pronoun interpolation acceptable, at least not if the intervening subject is a full NP. The cases reported in Alvarez et al. 1986 use first person singular personal pronouns. The reason to avoid using these pronouns in the experiment was to avoid an effect due to the discourse already presupposed in this type of pronouns. Nevertheless, there are also cases with full NPs with a third person referent reported in Alvarez et al. (1986):

(i). Trigo que lle *a palla* doura.....
wheat that CL-dat the straw goldens
‘Wheat that the straw it goldens....’

intervening embedded subject that interferes with the appearance of this clitic, we would expect a reading slowdown at the embedded subject the same way as in the Spanish experiment 2 reported in chapter 3 there is a reading slowdown at the embedded subject. Moreover, since the clitic presence requirement is stronger in Galician than in Spanish, it could be that in Galician this effect can be seen earlier than in the Spanish experiment, where the effect appears at the noun of the embedded subject.

Method

Self-paced Reading experiment

Participants

Forty-two native speakers of Galician⁶² participated in the experiment. All were students at the University of Santiago de Compostela, Santiago. They were paid 10 Euros for their participation in the experiment, which lasted approximately forty-five minutes.

Materials and Design

This experiment resembles the Spanish experiment reported in chapter 3.

In the stimuli of this experiment clitics appear always in pre-verbal position. The reason for only focusing on pre-verbal clitics is that, as in the case of the Spanish experiments presented in chapter 3, this experiment focuses on pre-verbal dependency completion effects.

⁶² Subjects were from different regions of Galiza: A Coruña (n=25), Lugo (n=5), Ourense (n=2) and Pontevedra (n=10).

Since, as discussed in section 4.3.2, clitics can appear post-verbally in finite root clauses in Galician, the placement of the clitic was restricted to a pre-verbal position. Clitics appear pre-verbally in Galician finite root clauses when there is an adverbial to the left of the verb. Therefore, the sentences in the experiment had an adverbial to the left of the verb to do the appearance of the preverbal clitic obligatory.

On the other hand, since in finite embedded clauses pre-verbal clitics are obligatory, the presence of the clitic pre-verbally was guaranteed. The sentence in (13) shows the pre-verbal position of the clitics in the early completion condition.

- (13) *Ás marisqueiras, o xefe da cooperativa nunca lles asegura*
To+the fisherwomen, the boss of+the cooperative never CL(IO) assures
que a fabrica as ha contratar.
that the factory CL(DO) will hire
 ‘To the fisherwomen, the boss of the cooperative never assures that the factory will hire them’

Twenty-four items of two conditions⁶³ each were used in the experiment. The two conditions that were generated for each item differed in the absence/presence of a clitic preceding the main clause verb. This meant that the completion of the dependency took place either early, at the clitic preceding the matrix verb as in (14a), or late, at the clitic in the most embedded clause as in (14b).

⁶³ As in the off-line experiment, this experiment also had the following two conditions with interpolation. Conditions (c) and (d) differed in the absence/presence of a clitic at the matrix verb and in whether the clitic at the embedded clause was interpolated or not. These conditions tested how Galician speakers interpreted sentences where the clitic in the embedded clause was interpolated and distance apart from the verb.

(c). *Ás marisqueiras, o xefe da cooperativa nunca lles asegura que as a fabrica ha contratar.*
 To+the fisherwomen, the boss of the cooperative never CL(IO) assured that CL(DO) the factory will hire

(d). *Ás marisqueiras, o xefe da cooperativa nunca asegura que as a fabrica ha contratar.*
 To+the fisherwomen, the boss of the cooperative never assured that CL(DO) the factory will hire

Since the results of the off-line had very low acceptability ratings for the constructions with interpolation, the discussion of the interpolation conditions is not included.

(14) *Context:*

A situación laboral das marisqueiras da rexión sempre foi moi instable e incerta.

The work circumstances of the fisherwomen of the region always was very unstable and uncertain.

a. *Early Completion*

Ás/ marisqueiras,/ o /xefe/ da / cooperativa /poucas/ veces /**lles**

To+the fisherwomen, the boss of+the cooperative few times them

/asegura/ que /a/ fabrica/ **as**/ ha/ contratar.

assures that the factory them will hire

‘The boss of the cooperative assured the fisherwomen a few times that the factory will hire them’

b. *Late completion*

Ás/ marisqueiras,/ o/ xefe/ da / cooperativa/poucas/ veces/

To+the fisherwomen, the boss of+the cooperative few times

asegura /que /a / fabrica /**as** / ha / contratar.

*assures that the factor **them** will hire*

‘The boss of the cooperative assured a few times that the factory will hire them’

Two lists were created by distributing the twenty-four paired items in a Latin Square design together with the interpolated conditions. Each subject saw exactly one of the lists intermixed with seventy-two filler items in a random order. The slashes indicate the region segmentation, where every word was a region.

Procedure

The experiment was conducted on a Macintosh G3 iBook laptop computer running the *Linger* software developed at MIT. Participants were timed in a word-by-word self-paced non-cumulative moving-window reading task (Just, Carpenter, & Woolley, 1982). All sentences were presented on a single line, except the context sentence, which was sometimes more than a single line in length. The segmentation indicated with slashes in (14) was the actual segmentation used in the experiment.

Words initially appeared as a row of dashes, and participants pressed the spacebar of the keyboard to reveal each subsequent region of the sentences. When subjects pressed the spacebar for the first time, the first word of the context sentence was revealed and the rest of the context sentence and the target sentences appeared in dotted lines.

In order to ensure that participants attended to the stimuli, a yes/no comprehension question was included for each passage. The questions either made reference to the context or to different parts of the experimental items.

Data Analysis

Only sentences for which the corresponding comprehension question was answered correctly were included in the analysis. 2 participants whose task accuracy was below 80% in target sentences and below 85 % in total were removed from further analysis. 4 participants, whose residual reading times were longer than the average, were excluded of the analyses. Raw reading times longer than 2000 ms were discarded, removing 3 % of trials. The means and analyses presented below are based on the remaining trials.

Results

ANOVAs were computed on the participant mean residual reading times collapsing over items (F1), and on item means collapsing over participants (F2). All significant main effects and interactions with $p < .05$ are reported.

Comprehension Task Accuracy

Among the 36 subjects included in the analysis, average comprehension accuracy was 84%. The average correct response percentage did differ significantly across the two conditions. The average comprehension accuracy in the early completion condition was 81% and 79% in the late completion condition. In general, both conditions had low accuracy ratings. This could be due to the fact that the slowdown during the sentence reading interfered in the comprehension question.

Self-paced Reading

The reading time analysis yielded the following results. Raw reading times for early and late completion conditions in all regions are shown in Figure 6.

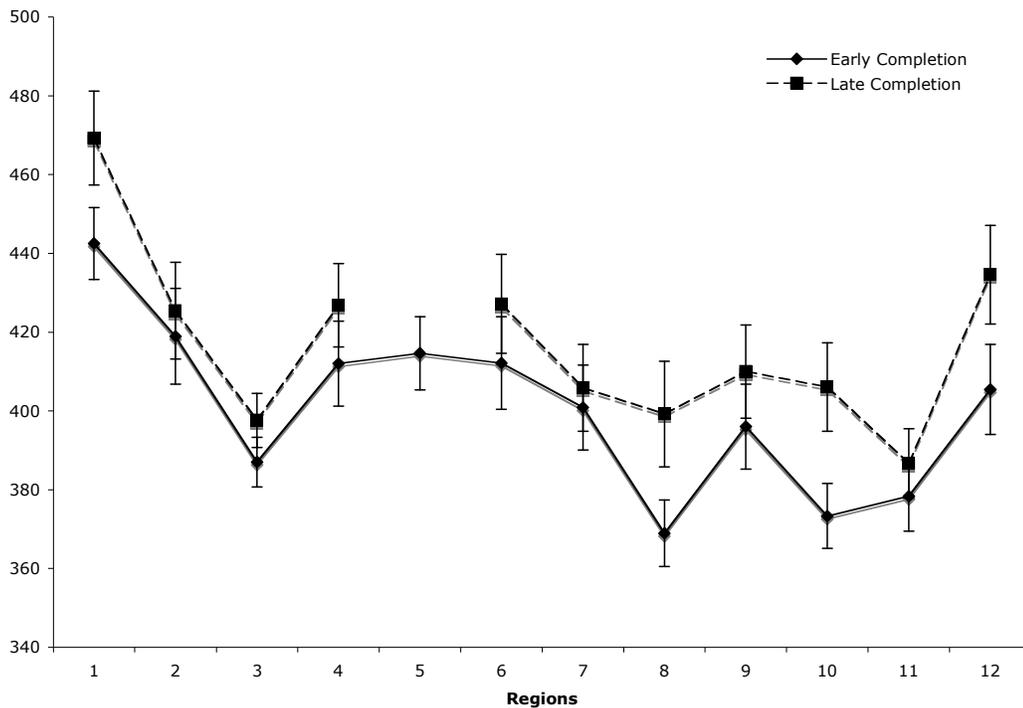


Figure 6: Mean raw reading times in milliseconds per region for the early and late completion conditions.

Ás_marisqueiras,₁ o_xefe_da_cooperativa,₂ poucas_veces,₃ lles,₄ asegura,₅ que,₆ a_fabrica,₇ as_ha,₈ contratar,₉ ha,₁₀ contratar,₁₁ contratar,₁₂

Region 1(1&2 combined) ($F_1(1, 20)=4.48$, $MSe=120263$, $p=.11$; $F_2(1,35)=2.94$, $MSe=135395$, $p=.10$) showed no significant effects and no significant differences between the reading times in the early and the late completion conditions. (all $F_s < 1$).

At all regions prior to region 8 there were no significant differences between the reading times in the early and late completion conditions (all $F_s < 1$).

In region 8, the determiner *a*, there was a significant main effect of *dependency* or clitic presence ($F_1(1,20)=4.62$, $MSe=89535.5$, $p=.04$; $F_2(1, 35)=4.20$, $MSe=62479.5$, $p=.04$). The effect was due to longer reading times for the late completion condition than for the early completion condition. The determiner in the late completion condition was read on average 31 milliseconds slower than in the early completion condition.

In region 9, the noun of the embedded subject *factory*, the effect was not significant (all $F_s < 1$).

In region 10, the embedded accusative clitic *as*, there was a significant main effect ($F_1(1, 20)=4.74$, $MSe=95437.1$, $p<.05$; $F_2(1,35)=6.36$, $MSe=75962.1$, $p<.05$). The effect was due to longer reading times for the late completion condition than for the early completion condition, where the embedded accusative clitic was read 33 milliseconds slower than in the late completion condition.

In region 11, the verbal auxiliary *ha*, the effect was not significant (all $F_s < 1$).

In region 12, the embedded verb *contratar* ‘to hire’, there was a significant effect ($F_1(1, 20)=5.10$, $MSe=94967.4$, $p<.05$; $F_2(1,35)=4.05$, $MSe=65788$, $p<.05$). The effect was due to longer reading times for the late completion condition than for

the early completion condition, where the embedded verb was read 29 milliseconds slower than in the late completion condition.

Discussion

The result of central interest in this experiment is the slowdown in reading times observed in the late completion condition at the determiner of the embedded subject, *a* (Region 8), relative to the reading time for the same region in the early completion condition. After the expectation for a clitic fails in the matrix clause when no overt clitic precedes the matrix verb, then the next position where the clitic is predicted is following the complementizer *que*. The appearance of an overt noun phrase following the complementizer causes a slowdown. More specifically, the slowdown is evident as soon as the determiner of the noun phrase is processed. This slowdown is caused by the fact that the presence of an overt noun phrase delays the presence of the predicted clitic. The preferred continuation is one that presupposes a null subject that co-refers with the main subject of the sentence (*xefe da cooperativa* 'owner of the coop'). Hence, when the continuation introduces a new entity, there is a cost associated with that in that it interferes with the active clitic search process started at the filler. Therefore, this effect at the embedded determiner is interpreted as evidence of an expectation for a clitic.

Since determiners and clitics are homophonous in Galician, this effect could also be due to a mismatch in agreement features between the predicted clitic (plural) and the encountered determiner (singular). If the parser is actively searching for a

clitic and the clitic that it finds mismatches in number features with the fronted phrase, then, the parser is left with the other alternative, which is the determiner.

(15) *Late completion condition*

Ás marisqueiras, o xefe da cooperativa nunca asegura
To+the(PL) fisherwomen, the boss of+the cooperative never assures
 que a...
that the(SG)

‘To the fisherwomen, the boss of the cooperative never assures that the...’

Nevertheless, if this were the only reason to have a slowdown, reading time differences at subsequent regions would not be predicted⁶⁴.

The second effect of interest was the slowdown observed in the late completion condition at the embedded clitic, *as* (Region 10), relative to the reading time in the early completion condition. This result stems from readers’ expectations of a clitic and maybe the fact that once there has been a slowdown in the determiner and an interference effect, when the clitic finally comes it is not expected any more. Alternatively, this result might reflect a spillover effect from the effect at the embedded subject.

Finally, the effect at the embedded verb *contratar* ‘to hire’ (Region 12) where the late completion condition is read more slowly than the early completion condition

⁶⁴ Clitics and determiners are also homophonous in Spanish. In the Spanish experiment 1 in chapter 3 (section 3.2.4), there is also an embedded clitic appearing immediately after the complementizer which could be mistaken with a determiner, but this clitic matches in number and gender features with the topicalized phrase and this clitic is read more slowly in the early completion condition than in the late completion condition.

a. *Early Completion condition*

A estas chicas, mi hermana mayor más tarde les dijo que
 [To these girls](PL), my sister old later on them-dat told that
 ya lo cree que LAS conoce desde hace tiempo.
 indeed it-acc(masc) thinks that them-acc(fem-PL) knows since long time.
 ‘To these girls, my old sister later told them that she indeed thinks she
 knows them for a long time’

could be due to a spillover effect from the accusative clitic region or alternatively it could be an integration effect once all the required elements have been found

4.6 General Discussion

4.6.1 Comparison with results obtained for Spanish CLLD Dependencies

Since the experimental design for this experiment was based on the experimental design of the Spanish experiments discussed in chapter 3, in what follows, I compare the results of the second on-line reading experiment in Spanish with the results of the Galician on-line experiment.

Among similar results, Galician shows an effect at the embedded subject where the late completion is read slower than the early completion. This resembles results obtained for Spanish experiment 2 where there was also a slowdown in the embedded subject. These two slowdowns seem to reflect the active search for a clitic pronoun in CLLD dependencies. In the case of Galician, the effect occurs one word earlier than in Spanish. This could be due to two different reasons: either because the Galician determiner can be mistaken with a clitic or because the requirement for the clitic in Galician is stronger and the clitic is expected earlier.

With respect to results where Spanish and Galician differ, the reading time slowdown at the matrix verb in the late completion relative to the early completion condition in the Spanish experiment is absent in the Galician experiment. This lack of

reading slowdown effect at the main verb could be due to the fact that Galician clitics can appear post-verbally and there could always be an expectation for the clitic to appear post-verbally. But, as reflected in the sentence in (16) taken from the off-line experiment in section 4.4, the clitics that appear post-verbally are always suffixes on the verb and it would not have been natural for the reader to have the verb and the affixed clitic be presented in separate words. Therefore, this could not be the reason for the absence of effect at the matrix verb.

- (16) Ós obreiros, o capataz necesitaos.
To+the workers, the foreman needs-CL(DO).
'All the workers, the foreman needs'

The second result where Spanish and Galician differ is in the effect at the embedded clitic. The results of the Galician experiment show that the embedded clitic is read more slowly in the late completion condition than in the early completion and the results of the Spanish experiment 2 (reported in section 3.2.8) show that the effect at the clitic is not significant even if it is in the same direction (late completion read more slowly than early completion). Moreover, if the expectations for the clitic were still maintained because there has not been a clitic in the main verb and the fronted phrase has not been interpreted yet at any position, following the results of the Spanish on-line experiment 1, there should be a facilitation effect at the embedded clitic region (where the late completion condition should be read faster relative to the early completion condition). The effect in the embedded clitic goes in the opposite direction (the late completion is read more slowly), which makes this effect look like a spillover effect.

4.7 Conclusion

This study shows that there is an active search for clitic pronouns in topic-clitic constructions in Galician. This is shown in that the delay in the appearance of the expected clitic caused by intervening material (such as an overt subject) results in a reading slowdown. The filled-subject effect at the embedded subject in the Galician experiment provides more evidence of there being an active search for clitic pronouns in CLLD constructions in a language with a strong requirement for clitic pronouns, which does not allow null interpretations of clitics in the matrix clause so easily as Spanish seems to do (see chapter 3, appendix 3-B). Moreover, clitics have a broader distribution in Galician and therefore can be expected in more positions (i.e. pre-verbally and post-verbally) along the sentence. Which means that if there is an active search activated for topicalizations in Galician, this active search would have to be more fine-grained in this language where clitics can be expected both in pre-verbal and post-verbal positions. The parser will have to know with precision where to expect a clitic in this language and that will require a parser with a very accurate knowledge of the constraints of grammar in this specific language.

All in all, the results suggest that active search mechanisms are not restricted to dependencies that terminate in a gap and that they apply to a wider range of long-distance dependencies (such as topic-clitic dependencies). Therefore, this suggests that active search might reflect a more general architectural property of parsing and not a specific response to the poor input encountered in wh-dependencies

Appendix 4-A: Data items for online reading experiment in Galician

item 1

Na cantina víase sempre un vello bébedo que lle falaba a tódalas rapazas que pasaban pola rúa.

1a. Ás rapazas, o vello da cantina sempre lles di que algún rapaz as ha agarimar cando chegue o inverno.

1b. Ás rapazas, o vello da cantina sempre di que algún rapaz as ha agarimar cando chegue o inverno.

1c. Ás rapazas, o vello da cantina sempre lles di que as algún rapaz ha agarimar cando chegue o inverno.

1d. Ás rapazas, o vello da cantina sempre di que as algún rapaz ha agarimar cando chegue o inverno.

? ¿Era na cantina onde se via sempre ó vello bébedo? Y

item 2

Na casa, á noite, ós cativos xogaban ó arredor dos avós sen preocuparse por nada.

2a. Ós cativos, o avó sempre lles explica que a vida os ha ensinar cando lles chegue o tempo.

2b. Ós cativos, o avó sempre explica que a vida os ha ensinar cando lles chegue o tempo.

2c. Ós cativos, o avó sempre lles explica que os a vida ha ensinar cando lles chegue o tempo.

2d. Ós cativos, o avó sempre explica que os a vida ha ensinar cando lles chegue o tempo.

? ¿É o avó quen ha ensinar ós cativos? N

item 3

A situación laboral das marisqueiras da rexión sempre foi moi inestable e incerta.

3a. Ás marisqueiras, o xefe da cooperativa nunca lles asegura que a fabrica as ha contratar aínda que leven anos traballando nela.

3b. Ás marisqueiras, o xefe da cooperativa nunca asegura que a fabrica as ha contratar aínda que leven anos traballando nela.

3c. Ás marisqueiras, o xefe da cooperativa nunca lles asegura que as a fabrica ha contratar aínda que leven anos traballando nela.

3d. Ás marisqueiras, o xefe da cooperativa nunca asegura que as a fabrica ha contratar aínda que leven anos traballando nela.

? ¿Son ás marisqueiras a quen a fábrica pode que non contrate? Y

item 4

A Xunta entrega cada ano dez premios en distintas categorías dirixidos ós medios de comunicación.

4a. Ós xornalistas, a editora soamente lles comenta que a xunta os pode premiar na modalidade de actualidade informativa.

4b. Ós xornalistas, a editora soamente comenta que a xunta os pode premiar na modalidade de actualidade informativa.

4c. Ós xornalistas, a editora soamente lles comenta que os a xunta pode premiar na modalidade de actualidade informativa.

4d. Ós xornalistas, a editora soamente comenta que os a xunta pode premiar na modalidade de actualidade informativa..

? ¿É a editora quen pode premiar ós xornalistas? N

item 5

As ONG traballan na mellora da situación dos refuxiados buscando unha participación mais activa dos gobernos.

5a. Ás refuxiadas, o dirixente da ONG continuamente lles di que o goberno as ha axudar en canto a súa situación se estabilice.

5b. Ás refuxiadas, os dirixentes da ONG continuamente di que o goberno as ha axudar en canto a súa situación se estabilice.

5c. Ás refuxiadas, os dirixentes da ONG continuamente lles di que as o goberno ha axudar en canto a súa situación se estabilice.

5d. Ás refuxiadas, os dirixentes da ONG continuamente di que as o goberno ha axudar en canto a súa situación se estabilice.

? ¿É poucas veces que o dirixente da ONG di que o goberno ha axudar ás refuxiadas? Y

item 6

Os actores xoves sempre teñen mais dificultades para conseguir premios nos primeiros anos da súa carreira profesional.

6a. Ós actores, a conselleira de cultura repetidamente lles explica que a academia os ha nomear para os premios cinematográficos anuais.

6b. Ós actores, a conselleira de cultura repetidamente explica que a academia os ha nomear para os premios cinematográficos anuais.

6c. Ós actores, a conselleira de cultura repetidamente lles explica que os a academia ha nomear para os premios cinematográficos anuais.

6d. Ós actores, a conselleira de cultura repetidamente explica que os a academia ha nomear para os premios cinematográficos anuais.

? ¿Son anuais os premios para os que a academia ha nomear ós actores? N

item 7

Sempre que chegan as festas ó cura busca xente que axude para arranxa-la igrexa a tempo.

7a. Ás mulleres, o cura da parroquia rapidamente lles comenta que o sancristán as ha axudar a prepara-la igrexa.

7b. Ás mulleres, o cura da parroquia rapidamente comenta que o sancristán as ha axudar a prepara-la igrexa.

7c. Ás mulleres, o cura da parroquia rapidamente lles comenta que as o sancristán ha axudar a prepara-la igrexa.

7d. Ás mulleres, o cura da parroquia rapidamente comenta que as o sancristán ha axudar a prepara-la igrexa.

? ¿É o sancristán quen ha axudar a prepara-la igrexa? Y

item 8

En Bruxelas, a nova reforma agraria europea foi admitida e estase poñendo en funcionamento.

8a. Ós gandeiros, a comunidade europea frecuentemente lles promete que a lexislación os ha beneficiar de agora en diante.

8b. Ós gandeiros, a comunidade europea frecuentemente promete que a lexislación os ha beneficiar de agora en diante.

8c. Ós gandeiros, a comunidade europea frecuentemente lles promete que os a lexislación ha beneficiar de agora en diante.

8d. Ós gandeiros, a comunidade europea frecuentemente promete que os a lexislación ha beneficiar de agora en diante.

? ¿É desde sempre que a lexislación beneficiou ós gandeiros? N

item 9

En zonas rurais, as escolas soamente teñen unha mestra ou mestre para tódolos nenos.

9a. Ás mestras, a asociación de pais sempre lles contan que a aldea as ha lembrar moitos anos.

9b. Ás mestras, a asociación de pais sempre contan que a aldea as ha lembrar moitos anos.

9c. Ás mestras, a asociación de pais lles sempre contan que as a aldea ha lembrar moitos anos.

9d. Ás mestras, a asociación de pais lles sempre contan que as a aldea ha lembrar moitos anos.

? ¿Teñen as zonas rurais soamente unha mestra ou mestre para tódolos nenos? Y

item 10

Tódolos domingos pola mañá a banda xuntase na praza para tocar diante da xente que sae de misa.

10a. Ós músicos, o director da banda a miúdo lles explica que o público os ha admirar aínda que non toquen o repertorio habitual.

10b. Ós músicos, o director da banda a miúdo explica que o público os ha admirar aínda que non toquen o repertorio habitual.

10c. Ós músicos, o director da banda a miúdo lles explica que os o público ha admirar aínda que non toquen o repertorio habitual.

10d. Ós músicos, o director da banda a miúdo explica que os o público ha admirar aínda que non toquen o repertorio habitual.

? ¿É os venres que os músicos se xuntan para tocar na praza? N

item 11

No xulgado da barriada mais perigosa da cidade hai soamente un garda en todo o edificio.

11a. Ás avogadas, o xuíz de oficio raramente lles asegura que o garda as ha protexer durante os días do xuízo.

11b. Ás avogadas, o xuíz de oficio raramente asegura que o garda as ha protexer durante os días do xuízo.

11c. Ás avogadas, o xuíz de oficio raramente lles asegura que as o garda as ha protexer durante os días do xuízo.

11d. Ás avogadas, o xuíz de oficio raramente asegura que as o garda as ha protexer durante os días do xuízo.

? ¿É ás avogadas que o garda ha protexer? Y

item 12

O goberno dedica un moi escaso porcentaxe anual do presuposto ó desenvolvemento tecnolóxico.

12a. Ós estudantes, o director da investigación poucas veces lles explica que o goberno os ha financiar soamente un ano mais.

12b. Ós estudantes, o director da investigación poucas veces explica que o goberno os ha financiar soamente un ano mais .

12c. Ós estudantes, o director da investigación poucas veces lles explica que o goberno os ha financiar soamente un ano mais.

12d. Ós estudantes, o director da investigación poucas veces explica que o goberno os ha financiar soamente un ano mais.

? ¿É o director da investigación quen financia ós estudantes? N

item 13

O partido celebra as súas eleccións internas cada tres anos para elixir as candidaturas a deputados na cámara.

13a. Ás deputadas, o presidente do partido raramente lles asegura que o consello as ha elixir dúas veces consecutivas.

13b. Ás deputadas, o presidente do partido raramente asegura que o consello as ha elixir dúas veces consecutivas.

13c. Ás deputadas, o presidente do partido raramente lles asegura que as o consello ha elixir dúas veces consecutivas.

13d. Ás deputadas, o presidente do partido raramente asegura que as o consello ha elixir dúas veces consecutivas.

? ¿É o consello quen ha elixir ás deputadas? Y

item 14

Moitos escritores recorren hoxe por hoxe a axuda de axentes para contactar coas diferentes editoriais.

14a. Ós escritores, o axente a miúdo lles di que a editorial os ha publicar soamente se gañan un concurso literario.

14b. Ós escritores, o axente a miúdo di que a editorial os ha publicar soamente se gañan un concurso literario.

14c. Ós escritores, o axente a miúdo lles di que os a editorial ha publicar soamente se gañan un concurso literario.

14d. Ós escritores, o axente a miúdo di que os a editorial ha publicar soamente se gañan un concurso literario.

? ¿Ha publicar a editorial ós escritores aínda que non gañen un concurso literario? N

item 15

Os membros de grupos musicais a miúdo aproveítanse da admiración de seguidoras novas.

15a. Ás seguidoras, o batería do grupo por veces lles confesa que o cantante as ha esquecer pasados dous días.

15b. Ás seguidoras, o batería do grupo por veces confesa que o cantante as ha esquecer pasados dous días.

15c. Ás seguidoras, o batería do grupo por veces lles confesa que as o cantante ha esquecer pasados dous días.

15d. Ás seguidoras, o batería do grupo por veces confesa que as o cantante ha esquecer pasados dous días.

? ¿É ás seguidoras que o cantante a esquecer? Y

item 16

A líder do sindicato mantíña unha opinión moi crítica do comportamento dos representantes empresariais.

16a. Ós traballadores, a líder sindical decote lles explica que a patronal os ha enganar nas negociacións.

16b. Ós traballadores, a líder sindical decote explica que a patronal os ha enganar nas negociacións.

16c. Ós traballadores, a líder sindical decote lles explica que os a patronal ha enganar nas negociacións.

16d. Ós traballadores, a líder sindical decote explica que os a patronal ha enganar nas negociacións.

? ¿É o líder sindical quen ha enganar ós traballadores nas negociacións? N

item 17

O grupo de investigadoras do laboratorio biolóxico acadaron os seus obxectivos con fondos propios.

17a. Ás investigadoras, o comité de revisión a miúdo lles di que a deputación as ha subvencionar o vindeiro ano.

17b. Ás investigadoras, o comité de revisión a miúdo di que a deputación as ha subvencionar o vindeiro ano.

17c. Ás investigadoras, o comité de revisión a miúdo lles di que as a deputación ha subvencionar o vindeiro ano.

17d. Ás investigadoras, o comité de revisión a miúdo di que as a deputación ha subvencionar o vindeiro ano.

? ¿Foi con fondos propios que as investigadoras acadaron os seus obxectivos? Y

item 18

Naquela empresa, os programadores pasan o tempo xogando no ordenador e falando entre eles.

18a. Ós programadores, a xestora do proxecto poucas veces lles explica que o cliente os pode visitar sen avisar.

18b. Ós programadores, a xestora do proxecto poucas veces explica que o cliente os pode visitar sen avisar.

18c. Ós programadores, a xestora do proxecto poucas veces lles explica que os o cliente pode visitar sen avisar. 18d. Ós programadores, a xestora do proxecto poucas veces explica que os o cliente pode visitar sen avisar.

? ¿Son os programadores quen poucas veces explican que o cliente pode visitar sen avisar?
N

item 19

Varias das corredoras da maratón cruzáronse no camiño das outras para frea-lo seu avance.

19a. Ás participantes, o organizador da carreira caladiñamente lles di que o xurado as ha descualificar polo seu comportamento.

19b. Ás participantes, o organizador da carreira caladiñamente di que o xurado as ha descualificar polo seu comportamento.

19c. Ás participantes, o organizador da carreira caladiñamente lles di que as o xurado ha descualificar polo seu comportamento.

19d. Ás participantes, o organizador da carreira caladiñamente di que as o xurado ha descualificar polo seu comportamento.

? ¿É caladiñamente que o organizador da carreira di que o xurado ha descualificar ás participantes? Y

item 20

Hai poetas que seguen mantendo un estilo rigoroso e tradicional na súa obra malia que desa maneira cheguen a menos lectores.

20a. Ós poetas, a crítica literaria sempre lles asegura que a xuventude os ha rexeitar senón que modernicen o seu estilo.

20b. Ós poetas, a crítica literaria sempre asegura que a xuventude os ha rexeitar senón que modernicen o seu estilo.

20c. Ós poetas, a crítica literaria sempre lles asegura que os a xuventude ha rexeitar senón que modernicen o seu estilo.

20d. Ós poetas, a crítica literaria sempre asegura que os a xuventude ha rexeitar senón que modernicen o seu estilo.

? ¿Chegan os poetas a todo tipo de lectores? N

item 21

Durante o conflito político no país, pouca xente puido facer fotos do acontecido nas cárceres.

21a. Ás fotografas, a redactora xefe abertamente lles comenta que o xornal as ha ascender se conseguen a exclusiva.

21b. Ás fotografas, a redactora xefe abertamente comenta que o xornal as ha ascender se conseguen a exclusiva.

21c. Ás fotografas, a redactora xefe abertamente lles comenta que as o xornal ha ascender se conseguen a exclusiva.

21d. Ás fotografas, a redactora xefe abertamente comenta que as o xornal ha ascender se conseguen a exclusiva.

? ¿Teñen as fotografas que conseguir unha exclusiva para ascender? Y

item 22

Tódolos domingos, a xente da aldea xúntase a xoga-la partida na cantina a beira da praia.

22a. Ós mariñeiros, o vello da cantina paseniño lles conta que o mar os ha desafiar toda a vida.

22b. Ós mariñeiros, o vello da cantina paseniño conta que o mar os ha desafiar toda a vida.

22c. Ós mariñeiros, o vello da cantina paseniño lles conta que os o mar ha desafiar toda a vida.

22d. Ós mariñeiros, o vello da cantina paseniño conta que os o mar ha desafiar toda a vida.
? ¿É ó vello a quen o mar desafia? N

item 23

Na televisión retransmíten o primeiro concurso de música tradicional feita por mulleres.

23a. Ás gaiterías, o entrevistador de televisión decote lles asegura que a audiencia as ha acoller con tanto entusiasmo coma ós seus compañeiros.

23b. Ás gaiterías, o entrevistador de televisión decote asegura que a audiencia as ha acoller con tanto entusiasmo coma ós seus compañeiros.

23c. Ás gaiterías, o entrevistador de televisión decote lles asegura que as a audiencia ha acoller con tanto entusiasmo coma ós seus compañeiros.

23d. Ás gaiterías, o entrevistador de televisión decote lles asegura que as a audiencia ha acoller con tanto entusiasmo coma ós seus compañeiros.

? ¿Asegura o entrevistador que a audiencia ha acoller con entusiasmo ás gaiterías? Y

item 24

Durante o verán, a facultade de medicina ofrece cursos nos que se intenta coñecer mellor os métodos tradicionais.

24a. Ós menciñeiros, o voceiro da facultade repetidamente lles comenta que a universidade os ha invitar para falar das súas tradicións ancestrais.

24b. Ós menciñeiros, o voceiro da facultade repetidamente comenta que a universidade os ha invitar para falar das súas tradicións ancestrais.

24c. Ós menciñeiros, o voceiro da facultade repetidamente lles comenta que os a universidade ha invitar para falar das súas tradicións ancestrais.

24d. Ós menciñeiros, o voceiro da facultade repetidamente comenta que os a universidade ha invitar para falar das súas tradicións ancestrais.

? ¿É poucas veces que o voceiro comenta que a universidade ha invitar ós menciñeiros? N

CHAPTER 5: BASQUE AUXILIARY FRONTING STUDY

Basque provides a good testing ground for incremental sentence processing theories because of its rich nominal and verbal morphology system. Importantly for our purposes, as discussed in the Introduction in chapter 1, this rich agreement information might be used to interpret structure before the head is encountered. As in the case of other head-final languages like Japanese, it seems unlikely that the parser waits for the verb to interpret all the previous arguments. Instead, there is pre-verbal structure building with nominal arguments associated together before the verbal head is encountered bottom-up.

The aim of the study described in this chapter is to look for anticipatory structure building in Basque and to see which is the scope of this anticipatory structure building (i.e. whether it goes beyond heads and complements). Such structure building is possible due to the case-marking information available in the noun phrases and morphology about the number of arguments in the verbal auxiliaries. The experiment on Basque auxiliary fronting aims at providing more evidence for building structure ahead of time. It extends the existing evidence for incremental attachment of arguments reviewed in chapter 1 for German (Bader & Lasser 1994) and Japanese (Mazuka and Itoh 1995, Inoue and Fodor 1995) by exploring the role of rich morphological agreement as a cue for pre-verbal structure building. The main objective is to examine how the combination of agreement and case marking is employed by the parser to assist processing given that the information

on the verb becomes available late. Previous studies on head-final languages revealed the important role of case marking for building structure pre-verbally. In this chapter, in addition to the case marking system, I take advantage of another property of the language, its rich agreement. More specifically, I will use a morphological mismatch between the auxiliary and the NP that follows it to show that this kind of cue can be used to predict structure. I will argue that the parser uses every piece of information that it encounters not only to interpret things before the head, but also to predict the upcoming structure and to make inferences about the upcoming input. I will show that this is only possible if the parser uses more than bottom-up information and resorts to its knowledge of grammar and grammatical constraints to make inferences about the upcoming input.

In sum, the two questions that this study aims to answer are: (i) what is the information used in advance of the verb in Basque to interpret things, and (ii) whether a morphological mismatch can be used to predict the presence of an embedded clause.

5.1. Basque: a rich agreement language

Basque has rich agreement. This is reflected especially well in its verbal auxiliary system. Verbal auxiliaries in Basque agree with all arguments in a clause providing morphological information about agreement, number, tense and mood as illustrated in the auxiliaries in (1).

(1) a. dut = transitive auxiliary

d - u - t
present tense marker - obj - subj

b. diot = ditransitive auxiliary

d - i - o - t
present tense marker - obj - dat - subj

Thus, auxiliaries in Basque provide the reader with exact information about the number of internal arguments of the verb. Therefore, it is possible to determine the exact number of arguments even in sentences where an (optional) argument is dropped. For example, in (2) and (3) the auxiliary reveals clearly whether the verb is used monotonically or ditransitively and whether the elided absolutive argument is plural. Note, importantly, that (2) vs. (3) are indistinguishable until the end of the sentence when the auxiliary is encountered.

(2) Nik erosi ditut.

I-ERG buy tense marker - number marker- ABS⁶⁵ - SUBJ

⁶⁵ It has to be noted that Absolutive case plays several distinct syntactic functions in Basque. The different syntactic functions that the Absolutive case can have in Basque are the following:

i) The case of the direct object:

Txakurra ikusi du.
Dog-ABS see aux
'S/he has seen the dog'

ii) The case of the subject of most intransitive predicates:

Txakurra etorri da.
Dog-ABS come aux
'The dog has come'

iii) The case for the subject and object of progressive constructions with *ari* (independent morpheme correspondent with *-ing* form):

Nire laguna eskutitza idazten ari da.
My friend-ABS letter-ABS write ari aux
'My friend is writing the letter'

iv) The case of the subject and the nominal predicate of copular sentences:

Nire laguna gure herri-ko alkatea da.
My friend-ABS our town-REL mayor is
'My friend is the mayor of our town'

v) The case of vocatives:

Mutil!
Boy-ABS
'Boy!'

‘I have bought them’

- (3) Nik erosi dizkiot.
I-ERG buy tense marker – number marker- ABS - DAT - SUBJ
‘I have bought them for him/her’

The verbal cluster in Basque generally consists of the participle, which is a non-finite form, and the auxiliary, which is the inflected form.

- (4) bidali dut
send (participle) present-object-subject
‘I have sent’

Moreover, negation is one of the two contexts⁶⁶ where the verbal cluster is split and where the auxiliary precedes the verbal head in the sentence as shown in (6).⁶⁷

- (5) Nik liburua erosi dut.
I-ERG⁶⁸ book-ABS buy have

(Hualde & Ortiz de Urbina 2003: 179f.).

It is important to note that the Absolutive case involved in this study will always have the syntactic function of a direct object (1 above), which is the most frequent in Basque and this syntactic function will never be ambiguous in the stimuli with the other syntactic functions listed above.

⁶⁶ The other is emphatic affirmation. Emphatic affirmation is a bound morpheme, which could be related to the free assertion morpheme *bai* ‘yes’ (Hualde & Ortiz de Urbina 2003: 537f.). It occupies, apparently, the same surface position as sentential negation (Oyharcabal 1984, Laka 1990):

- (i) Badu erosi
Ba.Aux bought
‘S/he did buy it’

⁶⁷ The inflected auxiliary adjoins to negation in Basque negative sentences. Neg has been argued to take IP as a complement in Basque and is therefore assumed to be generated above IP (Laka 1990)

⁶⁸ It has to be noted that Basque is an Ergative language, where the Ergative is the case of transitive subjects. (Hualde & Ortiz de Urbina 2003: 180f., 364)

- (i) Jonek loreak erosi ditu.
J-ERG flowers-ACC bouy aux-trans
‘Jon has bought flowers’

Some intransitive verbs, specifically unergatives, also take transitive morphology and the subject in the ergative case. These unergative predicates take the transitive auxiliary.

- (ii) Jonek dantzatu du.
Jon dance aux-trans
‘Jon has danced’

‘I have bought the book’

- (6) Nik ez dut liburua erosi
I-ERG Neg have book-ABS buy
‘I haven’t bought the book’

Hence, despite the ambiguity exemplified in examples (2) and (3), when fronted under negation as illustrated in (6), Basque auxiliaries make it possible to anticipate the argument structure of a clause final verb and eliminate the temporary ambiguity that is created when the verb can have more than one argument structure and when arguments are dropped.

Together with the rich agreement information, Basque has other properties that make it very suitable for testing theories about syntactic structure prediction. One of its properties is that it is a head-final language with SOV order in “neutral contexts”. The neutral case array order is subj-dat-obj-attributive-verb (Lafitte 1944, de Rijk 1969, Villasante 1980, Hualde & Ortiz de Urbina 2003). The neutral order subj-dat-obj features a reverse order of the markers inside the auxiliary as illustrated in (7) (Laka 1993).

- (7) Nik Maitaneri liburua erosi diot.
I-ERG M-DAT book-ABS buy tense marker - ABS - DAT - SUBJ
‘I have bought the book to Maitane’

The second relevant property is that Basque is a pro-drop language. Due to the rich verbal morphology, the object and the subject can be elided with all relevant information preserved in the morphology of the auxiliary.

- (8) (pro_j) (pro_i) Irakurri du_it_j
pro_j pro_i read have-obj_i-subj_j
‘I have read it’

- (9) (pro_j) (pro_i) (pro_k) Irakurri di_io_kt_j
pro_j pro_i pro_k read have-obj_i-dat_k-subj_j
 ‘I have read it to him/her’

These last two properties generate a lot of ambiguity for the parser since the verbal head is delayed and the number of arguments that the verb takes cannot normally be known until the auxiliary is processed. The sentences in (8) and (9) illustrate this particular property: the verb *to read* can take either one or two internal arguments depending on whether it is used monotonically or ditransitively as illustrated in (8) and (9) above and that cannot be known until the auxiliary is processed. Recall, however, that under negation, the auxiliary that contains information on the number of arguments, appears before the verbal head. This is the configuration that this study will examine.

The goal of this study is to see if speakers of Basque use the rich agreement information in the auxiliary to predict the upcoming structure and to examine their ability in using the cue given by agreement mismatching information between the auxiliary and the following NP. Previously, various researchers have shown that purely lexical information from the verb can provide direct cues about the verb argument structure (Trueswell et al. 1993, MacDonald et al 1994, Garnsey et al 1997). In turn, this study uses the agreement information provided by auxiliaries in Basque with two objectives in mind. One is to examine whether agreement creates a positive expectation for certain arguments in the input when the auxiliary is fronted and the verbal information is not available until the end of the sentence. The other is to test whether the indirect cue provided by the morphological mismatch between the auxiliary and the following NP is used by the parser to perform inferences about the presence of an additional clause and to avoid a garden-path. Specifically, agreement

mismatches between the auxiliary and a following NP may provide cues to the parser for establishing clause boundaries and predicting the clause before the specific lexical-semantics of this verb becomes known. If so, this will provide evidence that the parser can use more than bottom-up projection of lexical information on a head and that it can perform inferences about the upcoming material based on indirect cues.

Evidence from head-final languages suggests that information prior to the verb (such as case marking) is used in determining what structure is being processed and that there is a bias to analyze pre-verbal NPs as co-arguments (Inoue 1991, Bader & Lasser 1994, Yamashita 1997, Miyamoto 2002). Furthermore, this evidence has shown that the combination of word order and case marking can help to induce clause boundaries (Miyamoto 2002). The present study contributes to this literature and extends the evidence by showing that the information provided by this mismatch can be used to perform inferences about how the NPs before the verb will be associated together and avoid a garden-path.

Anticipating the results, I will show that an indirect cue provided by a morphological mismatch can help the parser build an embedded sentence and prevent it from following a garden-path. Additionally, I will also show that recovery from the misanalysis is rather fast for the structural reanalysis that this presupposes.

5.2. Direct vs. indirect cues for predicting structure ahead

5.2.1. Case markers and word order as inducers of clause boundaries

The logic of this experiment follows a long tradition that has pointed out that information about argument structure appears in many forms that do not involve the predicate or head. Previous literature has claimed that case markers and word order can be inducers of clause boundaries in Japanese. Different studies in Japanese have shown that the NPs found before the verb are associated together before the verbal head becomes available. The difference with respect to head-initial languages like English is that Japanese case markers of the NPs provide argument information. So the parser need not be head-driven (by head-driven meaning verb-driven specifically) in order to interpret these arguments.

Within the literature on the use of case markers in Japanese, Inoue (1991) was among the first to discuss the use of case-markers to induce the correct structure of a sentence. He investigated mainly the processing of the verb or the following word to test how reanalysis would apply, but he did not focus specifically on whether the NPs would be associated together before the verb was processed. Yamashita (1997), on the other hand, reported that Japanese readers are sensitive to the types of case markers that appear in NP sequences by showing disruption in reading times at the NPs before the verb was processed, when there were two Nominative marked NPs together (in the sequence NP-Nom NP-Dat NP-Nom). Miyamoto (2002, 2003)

extended Yamashita's work and showed that the slowdown found in a particular sequence of NPs was a reflection of a general case driven strategy to assign clause boundaries. More specifically, he argued for the claim that in Japanese a second nominative NP can mark the left-edge of a clause and serve as a marker. The argument he provides for this is that the nominative NP in Japanese can only be the argument of a tensed clause and would therefore require inflection (e.g. IP).

The results of these studies could be taken to be direct evidence for incremental structure building models since they show that the case marking in the NPs is used to associate the NPs together and derive what kind of structure is being processed (e.g. monoclausal vs. biclausal) before the verbal head is processed. Alternatively, these results could instead be understood as an activation of case array templates depending on the frequency of the NP-case array in Japanese, where the difference in reading time would be due to the fact that the case array of the NPs is less frequent or that it does not convey the expectations of the parser.

At this point, and considering the evidence from Japanese studies, it would be useful to differentiate between two issues regarding case marking expectations. One is the case order expected within sentences (canonical vs. non-canonical) and the other is the case order that can cause the parser to postulate a clause boundary.

On one hand, the target sentences in this study show ERG-DAT-ABS case marking combination, which for Basque is the canonical word order. As pointed out by Aoshima et al. (in press), the frequency of the case array combination might affect the reliability of the prediction of a case marking, having some case marking combinations that are more reliable than others (e.g. NOM-DAT vs. NOM-ACC in

Japanese). In this respect, there should not be side effects related to case order frequency when predicting the order of the arguments in this study because the target conditions have the reliable ERG-DAT-ABS order. The sequence ERG-DAT should strongly predict a following absolutive-marked NP in Basque. Moreover, I use this expectation to have an absolutive-marked NP following the dative-marked NP as a basic manipulation in the matching condition of this study. In this condition, the parser is deceived by the canonical case array in that it takes it to be part of a single clause.

On the other hand, the control sentences in this study show a word order combination where the second ERG case clearly marks a clause boundary. The same way as the second nominative NP is taken to be part of an embedded clause in Japanese (Miyamoto 2002, 2003), the second ergative NP in Basque is taken to be part of a second clause. The word order in the control conditions is either ERG-DAT-ERG word order or ERG-ERG-DAT⁶⁹.

In the case of the targets, I am relying on the case information combination between the agreement information in the auxiliary and the NP following this

⁶⁹ The order ERG-ABS in Basque would be less reliable to unambiguously indicate the presence of an embedded clause because, as previously discussed in a previous footnote, the absolutive case has various syntactic functions in Basque.

As illustrated in (i) and (ii), the absolutive case could be the case of the direct object or the case of the subject of an intransitive predicate. Therefore, when processing the sequence ‘Maitane-ERG Alaitz-ABS’ on-line, the function of the absolutive marked argument could be temporarily mistaken between being the direct object (i) or the subject of an embedded clause (ii) until the verb that disambiguates the structure is found. Therefore, these cases would create more ambiguity and would be less reliable. Even if they are very common occurrences in Basque, this case array will not be used to deduce a clause boundary.

- (i) Maitanek Alaitz ikusi du.
M-ERG A-ABS see aux-trans
 ‘Maitane has seen Alaitz’
- (ii) Maitanek [Alaitz etorri dela] esan du.
M-ERG A-ABS come aux-intrans-Comp say aux-trans-Comp
 ‘Maitane has said that Alaitz has come’

auxiliary, plus the expectation of the parser for a particular case-marked NP considering the case array of the previous NPs, to predict these clause boundaries. Therefore, in the Basque target examples, I am relying on a much more indirect cue than case-marking combination or word order combination alone to determine that there is a clause boundary.

5.2.2. Direct cues = (English) verbal argument structure and subcategorization information

Previous studies on English ambiguity resolution have shown that, among other factors, the specific lexical information encoded in the verb can provide direct cues about its complement structure and about the upcoming sentence structure. These studies have shown how the verb-based lexical information can guide parsing by manipulating different factors: the plausibility of the material following the verb or how likely it is for the verb to get the NP complement it gets considering its subcategorization and its frequency biases⁷⁰ (Garnsey et al. 1997), the argument structure of verbs (Trueswell et al. 1993, Boland et al. 1994), the frequency with which individual lexical items occur in opposing thematic and syntactic relationships (MacDonald, 1994, MacDonald et al. 1994, Trueswell & Tanenhaus 1994), the thematic fit or event-specific world knowledge of ambiguous arguments (Taraban &

⁷⁰ In a study that manipulated the plausibility of the NP following the verb as a direct object, (Garnsey et al. 1997) showed that verb bias information is immediately used in parsing and that this information interacts with plausibility during sentence comprehension. When the complementizer was absent, the NP following the verb in sentences like (ia) was easier to process than in (ib).

- (i) a. The senator regretted (that) the decision had ever been made public.
- b. The senator regretted (that) the reporter had ever seen the report.

MacClellan 1988, 1990, Tabossi et al. 1994, MacDonald, 1994, McRae et al. 1998) and the referential presupposition or the interaction with context (Altmann & Steedman 1988, Spivey-Knowlton et al. 1993, Spivey-Knowlton & Tanenhaus 1994, Spivey-Knowlton & Sedivy 1995). Moreover, there have been specific accounts of how the process of accessing the lexical information of the verb interacts with the contextual domain in so-called constraint-based models (Spivey-Knowlton et al 1993, MacDonald et al. 1994).

Even though the previous studies have mainly been concerned with the factors affecting ambiguity resolution, the fact that the verb-specific information is relevant and guides parsing in these studies is important for understanding the basic process of analysis generation within sentence processing that I have discussed in chapter 1 and that is important to analyze the data under consideration in this study.

Therefore, as seen in the evidence from English ambiguity resolution studies, the cues from the verb about the upcoming structure are direct. Structure prediction is guided by the lexical information encoded in the verb, plus the plausibility of the element that follows and the frequency with which that combination happens in the language among other factors.

5.2.3. Indirect cues= Japanese classifier mismatch

Apart from the direct cues given by the verb in languages like English, there have also been studies that have used indirect cues to guide the parser in carrying out distant inferences about the upcoming structure. Both the study reported here and the

study on Japanese classifier mismatch (Yoshida et al. 2004, Yoshida 2006) show that either a morphological mismatch of the auxiliary's argument information with the case linearly adjacent NP in Basque or a semantic mismatch of the morphology required by the classifier with the linearly adjacent noun in Japanese, can generate expectations about the upcoming structure and can help the parser to anticipate upcoming structure.

Although in these two studies there is no direct lexical verb information that can guide parsing because the verb comes sentence finally, the mismatches make it easier for the parser to presume that there is an embedded sentence. This means the parser can project structure in advance and make distant inferences about the kind of elements that will satisfy the grammatical constraints established at the beginning of the sentence by either the auxiliary in the case of Basque (as it will be seen in the discussion of the current experiment) or the classifier in the case of Japanese⁷¹.

Yoshida et al. (2004) and Yoshida (2006) show that a classifier mismatch in Japanese can cue the parser in predicting an upcoming relative clause. The fact that the noun immediately adjacent to the classifier mismatches semantically with the classifier triggers the prediction of a head noun and forces the parser to analyze the adjacent noun as part of an embedded clause.

Consider the sentences in (9) where the relevant contrast is exemplified. In (9a) the classifier is associated with the subject NP *students* because its features match semantically with the noun's features, whereas in (9b) the classifier has to be

⁷¹ For a similar kind of classifier mismatch manipulation in Chinese see (Hsu et al. 2006, Hsu 2006), where Chinese speakers don't seem to use the cue provided by the mismatch so effectively and where the lack of case marking makes it difficult for the parser to know which is the embedded clause boundary.

associated with the head of the RC *book* because its features mismatch semantically with those of the adjacent NP *student*.

- (9) a. [3-nin-no gakusee-ga Yonda] Hon
 3-cl(*human*)-GEN student-NOM read book
 ‘The book that three students read’
- b. [3-satu-no [gakusee-ga Yonda] Hon]
 3-cl(*printed matter*)-GEN student-NOM read book
 ‘The three books that the student read’

Yoshida et al. (2004) [experiment 2] show in a self-paced reading study how the classifier mismatch can help readers to avoid a garden-path. Their reasoning is that if the classifier mismatch in (10b) creates an expectation for a RC, then the parser should not be surprised when encountering the embedded verb that lacks the complementizer suffix *-to* (indicating that it is within a RC). In the case where there is no mismatch as in (10a), since there is no prior indication of the upcoming embedded relative clause, the parser would be led into a garden-path when processing the embedded verb and realizing that it is within an embedded clause. This is precisely the evidence that the results show. Results show an expected reading time difference at the noun immediately following the classifier that reflects the mismatch and a reading time difference across conditions at the embedded verb, where the embedded verb in the match condition is read slower relative to the one in the mismatch condition and where there is a reversal in reading times relative to the difference in the classifier.

- (10) a. Classifier Match Condition
 Tannin-wa [[[[san-nin-no tosioita] sensee-ga] atarasii
 Class-teacher-Top three-cl(*human*)-Gen aged teacher-Nom new
 koochoo-ni yorokonde okutta] hon-o] aru-seeto-ni kyoositu-de

*president-Dat gladly gave book-ABS a-student-Dat classroom-at
yomase-masita.
made-read.*

‘The teacher made a student read the book that three aged teachers
gladly gave to the new president at the classroom.’

b. Classifier Mismatch Condition

Tannin-wa [san-satu-no [tosioita sensee-ga
Class-teacher-Top three-cl(book)-Gen aged teacher-Nom
atarasii koochoo-ni yorokonde okutta] hon-o] aru-seeto-ni
new president-Dat gladly gave book-ABS a-student-Dat
kyoositu-de yomase-masita.
class-room-at made-read.

‘The teacher made a student read three books that an aged teacher
gladly gave to the new president at the classroom.’

Yoshida’s (2006) classifier mismatch study differs from the Basque study reported in this chapter in that the mismatch can predict a specific type of embedded clause, an RC, whereas in the Basque study the mismatch generates a prediction for an embedded clause that is not restricted to just one type of clause. Moreover, in the Japanese experiment the nominative marker is the indicator that there is an embedded clause. What the classifier mismatch cues in this case is the specific prediction for a relative clause. In the Basque experiment, the case markings in the target conditions do not indicate that there is a clause boundary. This information only comes from the agreement and case mismatch between the auxiliary and the following NP, but case marking does not play such a direct role as in Japanese⁷². Moreover, the expectation that the parser has in the case of Basque is connected to a specific verbal head and a specific type of VP configuration that will match the main auxiliary, whereas in

⁷² The same difference with respect to case applies to the Chinese classifier mismatch relative clause cases reported in Hsu (2006), where the case for the classifier is ambiguous (underspecified) because it depends on the head that licenses it and the search for a particular head for this classifier is not narrowed down to a single case marking, leading the parser to a parsing break down. In contrast to the Japanese case in Yoshida (2006), the mismatch in Chinese does not predict a specific type of clause, the RC in this case.

Japanese the parser has an expectation for a noun head that will match in features with the classifier and that will have to be the head of a relative clause. The two studies have in common the fact that they are using indirect but unambiguous cues (such as a case mismatch or a classifier mismatch) to predict the upcoming structure of the sentence.

5.3. Expectation-driven parser

We want to test whether the morphological information encoded in Basque auxiliaries creates expectations about the upcoming structure. Expectations are created when the agreement morphology contained in the auxiliary is processed. The specific prediction created by the auxiliary in Basque is the projection of the argument structure of the verb. This entails the projection of both the argument slots of the verb and the verbal head. That the parser uses this information accurately to make predictions about the input is reflected in the fact that the failure to fulfill these predictions will always result in a reading time difference. There are different ways in which the expectations created by the auxiliary could fail. On one hand, the auxiliary can project information about the arguments that cannot be locally verified (e.g. if one of the complements is sentential) and that would require reanalysis when the predicted arguments are not locally found. On the other hand, it could be that arguments that have not been predicted by the auxiliary appear and force to restructure the sentence.

Furthermore, it is important to note that fronted auxiliaries in Basque are not so frequent since they are only fronted in negative and emphatic contexts, thus, this is not a widespread cue for the speakers of this language and it is unlikely that under other circumstances (e.g. declarative contexts) Basque speakers can anticipate the upcoming arguments in the sentence based on the auxiliary or verb information⁷³.

We manipulated the combination of the agreement morphology of the auxiliary and the case marking of the NP linearly adjacent to it to see how the match or mismatch between these two elements guides the parsing of the sentence. More specifically, we were interested in examining what takes place when the morphology of the auxiliary indicates the upcoming appearance of a particular subcategorization for the verb and the case of the NP adjacent to this auxiliary does not fit this specific subcategorization. We tested whether this would be a cue to infer that the complement that the auxiliary takes is sentential and therefore avoid the garden-path. In contrast, the other manipulation examined what takes place when the morphology of the main auxiliary agrees with the case of the NP adjacent to it. We tested whether the sentence would be considered to be monoclausal until the embedded auxiliary is processed and an embedded sentence has to be posited based on the agreement information in this new auxiliary.

When the auxiliary is ditransitive, both the dative argument and the absolutive argument slots are projected together with the verbal ditransitive head. This means that the parser can predict at the auxiliary what the subcategorization of the verb is,

⁷³ Even if negative sentences are not as frequent as declarative sentences in the language, they are still common enough to be able to make a processing generalization. Therefore, I take the cue provided by the auxiliary in the Basque stimuli I'm presenting to be completely reliable and to indicate how rich agreement information in Basque can help anticipating upcoming structure.

even if the lexical form of the verb is not known yet. The question of interest is what the parser does when the projected structure does not match the word in the input.

The basic contrast used for this study is illustrated in (11).

(11) a. *Mismatching condition*

[Maitanek_i ez du [ikasleari txostena idatzi
[M_i-ERG neg aux-trans [pro_i student-DAT document-ABS write
diola] esan.]
aux-ditrans-C] say.]
 ‘Maitane hasn’t said that she has written the document to the student’

b. *Matching condition*

[Maitanek_i ez dio ikasleari [txostena idatzi
M_i-ERG neg aux-ditrans student-DAT [pro_i document-ABS write
duela] esan.]
aux-trans-C] say.]
 ‘Maitane hasn’t told the student that she has written the document’

We could assume two different ways in which the conditions in (11a-b) could be parsed. A parser that does not anticipate structure before disconfirming input is found would assume no contrast between the matching or mismatching conditions until the verb in the embedded clause is parsed. This kind of parser would not predict any difference in reading time before processing the embedded verb in (11a) with respect to the embedded verb in (11b). Even if the mismatch in agreement could be noticed, this parser would predict that there is no commitment made about the structure until the verb is found. Within this kind of view, the parser cannot learn that the structure is bi-clausal until the embedded verb is processed. An expectation-based parser, on the other hand, would assume that every cue prior to the verb is used to project structure in advance and to draw inferences about the upcoming input. This kind of parser would therefore exhibit a reading time difference prior to the verb if the information does not conform to the expectations generated by different cues.

Within this kind of parser, if it is the case that the indirect cues provided by the mismatch are used to guide the parsing of the sentence, the bi-clausal structure of the sentence would be predicted earlier than at the embedded verb, since it would predict that the mismatch in agreement is used to infer the right structure for the sentence.

We would like to test if the predictions of an expectation based parser hold and if the indirect cue of a mismatch is used to avoid a garden-path in this study's conditions. In this case, the parser would assume the input to be well formed and not just an agreement violation, which means that it would try to use the mismatch cue for building new structure. If, on the contrary, structure is not built in an anticipatory way, then there should not be any difference between these two conditions before the embedded verb.

5.4. Predictions for this experiment

In the mismatching condition, if it is assumed that the parser does not build structure in an anticipatory fashion, there would be no reading time difference expected in the mismatching dative case-marked NP region. Even if the agreement mismatch could be noticed in this kind of parser, it is not clear that the parser would try to make something out of it. No reading time difference is expected before the embedded verb under this kind of parser, since that is the first place where it encounters confirming evidence that allows it to interpret the previous arguments. . If on the contrary, an incremental parser is assumed, there should be a reading time difference at the morphological mismatch. Whether this mismatch should suffice to

predict the embedded clause and whether the incremental parser is able to recover a good parse on the basis of the mismatch is not guaranteed. The evidence for the prediction of the embedded sentence could only be reflected in the lack of a reading time difference at the embedded auxiliary in this condition. The reading time difference in the mismatch region would only reflect the parser's ability to detect an agreement mismatch, but it is not direct evidence for the creation of an embedded clause and it is not guaranteed that an incremental parser can do that.

On the other hand, in the matching condition, if the parser were not processing the words in the input incrementally and trying to accommodate them in the syntactic structure, then there should be no contrast in reading times between the two mismatching conditions at the auxiliary in the embedded clause. If, on the contrary, a fully incremental parser that uses the information prior to the heads such as case and morphological information is assumed, then a reading time difference would be expected at the embedded auxiliary since that would be the place where the syntactic structure parsed up to that moment would have to be reconsidered.

This study will address the following questions. First, I will investigate whether the parser uses the information prior to the verb to predict the kind of syntactic structure that the sentence will have. Second, I will examine how the parser uses this information to try to associate arguments of a sentence together incrementally and before the head of the sentence is processed (by means of deceiving the parser and forcing it to reanalyze). Finally, I will aim to determine if, from the information provided by the mismatch between the auxiliary and the

following NP, it can be inferred that there is a clause boundary and a garden-path can be avoided.

5.4.1. Experiment 4: On-line Reading Experiment on Basque Auxiliary Fronting

The following on-line experiment examines whether the parser makes use of the morphological information in the auxiliaries about argument structure, together with the case information in the following NPs, to make predictions about the upcoming sentence structure.

I predict that morphological information encoded in the auxiliary about the subcategorization of the verb helps in interpreting the structure ahead of the main subcategorizer and provides cues for the processing of upcoming material (e.g. whether the sentence contains more than one clause). Therefore, by means of a morphological match/ mismatch contrast between the auxiliary's information and the NP immediately following the auxiliary I will induce different expectations in the parser and examine if the parser is to use the available information to make inferences about upcoming material.

Method

Self-paced reading experiment

Participants

Forty-one native speakers of Basque participated in the experiment. They were students either at the University of Deusto, Bilbao or at the University of the

Basque Country, Gasteiz. They were all from Basque speaking families and small Basque speaking towns in the Basque Country. They were all bilingual speakers of Basque and Spanish, differing in their level of proficiency in Spanish. They were paid \$10 or its equivalent in Euros for their participation in the experiment, which lasted approximately forty-five minutes.

Materials and Design

A set of twenty-four items of four conditions each was used in the experiment⁷⁴. The four conditions that were generated for each set differed in the absence or presence of the embedded subject and in whether the NP matched or mismatched the agreement marking on the preceding auxiliary. The verbs chosen for the embedded sentence were optionally ditransitive verbs (e.g. *idatzi* ‘to read’) that could optionally take a dative-marked NP (e.g. *ikasleari* ‘student-dat’) as an argument. The reason to choose optionally ditransitive embedded verbs was so that the ambiguity about the argument structure in this verb would not be resolved until the embedded auxiliary was processed.

Four lists were created by distributing the twenty-four items in a Latin Square design. Each subject saw exactly one of the lists intermixed with seventy-two filler items in a random order. A complete set of experimental conditions is shown in (12).

- (12) a. *Match-null subject*
Maitanek_i ez dio ikasleari [pro_i ikastaroko txostena
M-ERG Neg aux-ditrans Student-DAT [pro_i workshop’s document-ABS
oso presaka idatzi duela] esan.
very fast write aux-trans-Comp] say.

⁷⁴ Experimental items are listed in Appendix 5-A.

‘Maitane hasn’t told the student that she/he has written the workshop’s document very fast’

b. *Match-overt subject*

Maitanek_i ez dio ikasleari [zuzendariak ikastaroko txostena
M_i-ERG neg aux-ditr student-DAT [director-ERG workshop document_{ABS}
oso presaka idatzi diola] esan.
very fast write aux-ditrans-Comp] say.

‘Maitane hasn’t told the student that the director has written the workshop’s document very fast’

c. *Mismatch-null subject*

Maitanek_i ez du [pro_i ikasleari ikastaroko txostena
M_i-ERG neg aux-trans [pro_i student-DAT workshop’s document-ABS
oso presaka idatzi diola] esan.
very fast write aux-ditrans-Comp] say.

‘Maitane hasn’t said that she has written the workshop’s document to the student very fast’

d. *Mismatch-overt subject*

Maitanek ez du [zuzendariak ikasleari ikastaroko txostena
M-ERG neg aux-trans [director-ERG Student-DAT workshop document
oso presaka idatzi diola] esan.
very fast write aux-ditrans-Comp] say.

‘Maitane hasn’t said that the director has written the workshop’s document to the student very fast’

Procedure

The experiment was conducted on Macintosh G3 iBook laptop computer running the *Linger* software developed at MIT. Participants were timed in a word-by-word self-paced non-cumulative moving-window reading task (Just, Carpenter, & Woolley, 1982). All critical sentences were presented on a single line. The segmentation indicated with spaces in (12) was the actual segmentation used in the experiment. Words initially appeared as a row of dashes, and participants pressed the spacebar of the keyboard to reveal each subsequent region of the sentence. A key

press replaced the first two lines with dashes and revealed the first word of the target sentence.

In order to ensure that participants attended to the stimuli, a yes/no comprehension question was included. The questions made reference to different parts of the experimental items and, even if the sentences were negative statements, the questions were asked such that they could have a yes/no answer.

Data Analysis

Only sentences for which the corresponding comprehension question was answered correctly were included in the analysis. Residual reading times that were longer than 1000 ms were discarded. This procedure affected 3.5% of trials. The means and analyses presented below are based on the remaining trials.

The data from the four conditions were entered into a 2x2 ANOVA with the factors *match* (match, mismatch) and *embedded-subject* (presence/absence). Reading times from matching conditions and the mismatching conditions were compared pairwise in a one-way ANOVA. ANOVAs were computed on the participant mean residual reading times collapsing over items (F1), and on item means collapsing over participants (F2). All significant main effects and interactions with $p < .05$ are reported.

Results

Comprehension Task Accuracy

Three out of 44 participants whose task accuracy was below 80% in target sentences and below 85 % in total were removed from further analysis. Among the 41 subjects included in the analysis, average comprehension accuracy was 87%. The average comprehension accuracy in the *match/ null-subject* condition was 89% and 88% in the *match/overt subject* condition. The average comprehension accuracy in the *mismatch/null subject* condition was 88% and 80% in the *mismatch/overt subject* condition.

Self-paced Reading

Residual reading times for *matching overt/null subject* conditions in all regions are shown in Figure 7 and the residual reading times for *mismatching overt/null subject* conditions in all regions are shown in Figure 8.

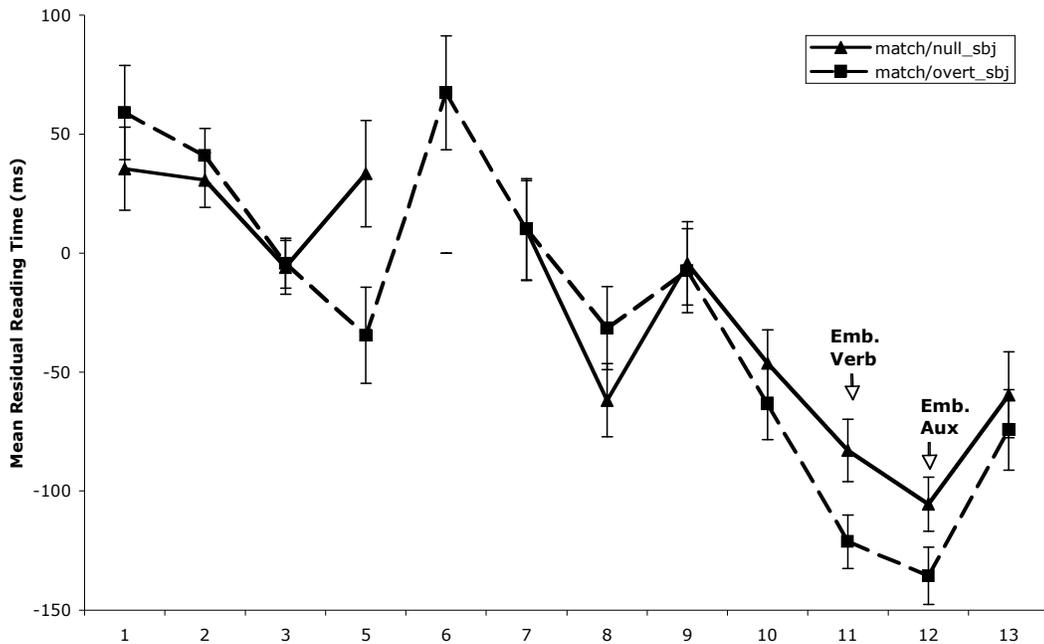


Figure 7: Mean residual reading times per region for the overt subject and null subject matching conditions⁷⁵.

Maitanek₁ ez₂ dio₃ ikasleari₅ zuzendariak₆ ikastaroko₇ txostena₈ oso₉ presaka₁₀
 NP-ERG₁ neg₂ aux-ditrans₃ NP-DAT₅ NP-ERG₆ mod₇ NP-ABS₈ mod₉ adv₁₀
 idatzi₁₁ duela₁₂ esan₁₃
 V₁₁ aux-trans₁₂ V₁₃

The 2x2 ANOVA yielded the following results.

Regions 1-10 (prior to the embedded verb): A 2x2 ANOVA showed a significant main effect of the factor *match* in regions 2 ($F_1(1,23)=9.23$, $MSe=128583$, $p<.01$; $F_2(1, 40)=5.31$, $MSe=89301.2$, $p<.05$) which corresponds to the negative particle, region 3 ($F_1(1,23)=15.89$, $MSe=295479$, $p<.001$; $F_2(1, 40)=14.77$, $MSe=409481$, $p<.001$) the main auxiliary, region 5 ($F_1(1,23)=4.61$, $MSe=541242$, $p<.05$; $F_2(1, 40)=11.22$, $MSe=1008460$, $p<.05$) the dative-marked NP, and region 7

⁷⁵ For analysis reasons, the dative NP was aligned in the two different conditions so that it will always correspond to the same region (= region 5).

($F_1(1,23)=4.57$, $MSe=384308$, $p<.05$; $F_2(1, 40)=6.84$, $MSe=654691$, $p<.05$) the noun modifier.

Pairwise comparisons within each level of the *match* revealed no significant differences in the reading times in the overt and null subject conditions in all regions prior to region 3 (the main auxiliary), and in regions 6 (ergative NP), 7 (noun modifier), 8 (absolute NP), 9 (adverbial modifier) and 10 (adverb).

At region 5, the NP dative following the auxiliary, there was a significant effect in the *matching null subject condition*.

Regions 11 & 12 (embedded verb *idatzi* and its auxiliary *duela*): a significant effect of *embedded-subject* was found ($F_1(1,23)=5.00$, $MSe = 133320$, $p<.05$; $F_2(1, 40)=11.67$, $MSe=244644$, $p<.05$). The effect was due to the longer residual reading times in the null-subject conditions than in the overt-subject conditions (-159 ms vs. -221 ms). However, the interaction of *match* with *embedded-subject* did not reach significance (both F 's < 2, p 's > .1). The lack of an interaction of the two factors is likely to be due to the presence of the main effect of *embedded-subject*.

We ran pairwise comparisons within each level of factor *match*. This analysis yielded no significant effects in the *mismatch* pair. However, there was a significant main effect of *embedded-subject* in the *match* pair ($F_1(1,23)=4.79$, $p<.05$; $F_2(1, 40)=4.93$, $p<.05$). The effect was due to longer reading times at the auxiliary in the null subject condition than for the overt subject condition (residual reading times -82 ms vs. -121 ms respectively).

In region 13, the main verb *esan*, there were no significant differences between the reading times in the two matching conditions ($F_s < 1$).

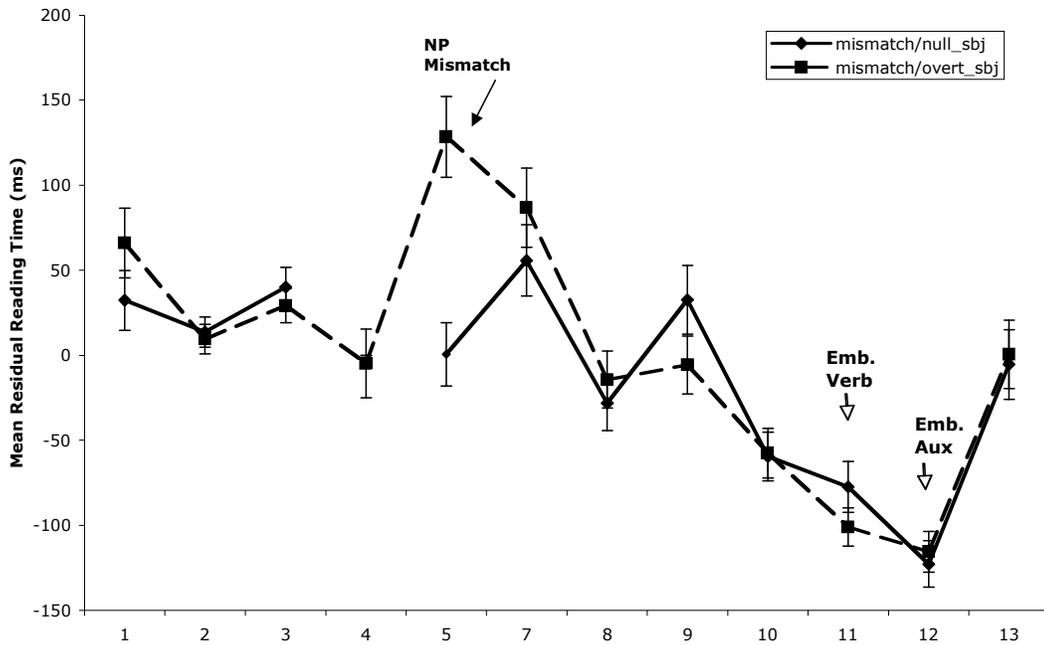


Figure 8: Mean residual reading times in milliseconds per region for the mismatching null and overt subject conditions⁷⁶.

Maitanek₁ ez₂ du₃ zuzendariak₄ **ikasleari**₅ ikastaroko₇ txostena₈ oso₉ presaka₁₀ idatzi₁₁
 NP-ERG₁ neg₂ aux-trans₃ NP-ERG₄ NP-DAT₅ mod₇ NP-ABS₈ mod₉ adv₁₀ V₁₁
 diola₁₂ esan₁₃
 aux-ditrans-Comp₁₂ V₁₃

The pairwise comparison for the *mismatching conditions* showed the following results.

At all regions prior to region 5, the dative NP, there were no significant differences between the reading times in the overt subject and null subject conditions (all $F_s < 1$).

⁷⁶ For analysis reasons, the dative NP was aligned in the two different conditions so that it will always correspond to the same region (= region 5).

In region 5, the noun phrase *ikasleari*, there was a significant main effect of *subject* presence ($F_1(1,23)=6.04$, $p<.05$; $F_2(1, 40)=18.62$, $p<.01$). The effect was due to longer reading times for the overt subject condition than for the null subject condition. The noun phrase *ikasleari* in the former condition was read on average 128 milliseconds slower than in the latter.

In regions 6 through 13, there were no significant differences between the reading times in the two mismatching conditions ($F_s < 1$).

5.5. General Discussion

The goal of this study was to investigate whether the rich morphological information of auxiliaries in Basque can help the parser in making predictions about the existing structure and anticipate the argument structure of the clause final verb in the absence of the lexical form of the verb. Of particular interest was to see if the agreement mismatch between the auxiliary and the following NP could provide cues for building a clause boundary and avoiding a garden-path. In what follows I will summarize the experimental results and then I discuss the implications of these results with respect to processing architectures that I have previously discussed in chapter 1 and at the beginning of this chapter.

One of the effects of interest was the slowdown observed in the overt subject mismatching condition at the dative case-marked noun phrase (*ikasleari* ‘student’) relative to the reading time at the same region in the null subject mismatching condition. This result is partly expected because the sequence of words is not the

same prior to this region (“*Maitanek ez du zuzendariak ikasleari...*”[NP-ERG Neg aux-trans NP-ERG NP-DAT]).

This means that the parser notices the agreement mismatch between the auxiliary and the following NP. The question then is how the parser uses this indirect cue to construct an embedded sentence and avoid the garden path seen in the matching conditions. Should the parser make no attempt to use this cue, there would be a reading time difference expected at the embedded clause, but the fact that there is no reading time difference means that the parser is using this cue effectively. How does the parser go from a mismatch to avoiding a garden-path and what type of parser is required to make this indirect connection is the question that I will try to answer in section 5.5.2.

This region also showed a weak significant effect in the *matching null subject condition*. This result is not expected since the sequence of words is the same up to this region (“*Maitanek ez dio ikasleari...*”[NP-ERG Neg aux-trans NP-DAT]). We take this result to be caused by experimental conditions and not by the experimental design.

The result of central interest in this experiment is the lack of garden-path effect at the embedded auxiliary in the mismatching conditions relative to the garden-path detected in the matching conditions. A shorter reading time was observed at the embedded verb and the verbal auxiliary (regions 10 and 11) in the overt subject matching condition relative to the reading time for the same regions in the null subject matching condition.

The fact that Basque speakers do not show a garden-path at the embedded verb and auxiliary in the mismatching conditions suggests that they are using the agreement mismatch cue and that they are building an embedded clause as a consequence of the mismatch. Unless an incremental parser that anticipates structure is considered, this evidence could not be explained under models that lack anticipatory structure building. These models would predict that the parser would not do anything useful out of the mismatch and it would still fall into the garden-path.

We take the reading time difference at the embedded verb and auxiliary in the matching conditions to reflect a garden-path effect i.e. the parser had to reanalyze the previously made commitments. In the case of the null subject matching condition, the linear sequence of words (and their case array) prior to the embedded verbal cluster matches the predictions established by the fronted ditransitive auxiliary and the parser presumably considers both NPs as part of a single ditransitive clause. The monoclausal structure would be that in (13), which is a perfect sentence in Basque.

- (13) Maitanek ez dio ikasleari ikastaroko txostena oso presaka idatzi.
M-erg neg aux-ditrans NP-DAT mod NP-ABS mod ADV V-ditrans
 ‘Maitane has not written the document for the student in a hurry’

The verbs in the embedded clause are all optionally ditransitive, so even if all the verbs that were used at the embedded sentence can be used both ditransitively and monotonitively, there is a preference for a monotonitive interpretation of these verbs attested by a corpus count of Basque⁷⁷. From the 14 verbs that were used in the embedded clause, 2 verbs were not frequent enough to appear in the corpus (*ureztatu*

⁷⁷ The Basque corpus from which this has been taken is the one belonging to *IXA-taldea* and *Egunkaria*. Thanks to Aitziber Atutxa for making these data available to me.

‘to water’ and *prestatu* ‘to get ready/to prepare’), 1 verb shows a ditransitive interpretation preference (*eman* ‘to give’) and 1 is completely equi-biased for either monotransitive or ditransitive interpretation (*bidali* ‘to send’). There are therefore 10 verbs that show a preference for the monotransitive interpretation⁷⁸. This might be the reason why subjects slow down one word earlier than we would expect. The slowdown at the embedded verb in the null subject matching condition could therefore be due to a preference to have a verb that is biased for a ditransitive reading. The other possible explanation for this slowdown at the embedded verb could be that the parser is puzzled by the fact that there are still two more regions left to process and it might not consider the sentence to be finished at this point.

It is not until the embedded auxiliary appears that it is clear to the parser that the embedded verb (e.g. *idatzi* ‘to write’) is used monotransitively in the embedded sentence and that the sentence is biclausal. At this point the parser has to infer that the previous dative and accusative arguments do not belong within the same sentence and it is forced to reanalyze. This would be the reason for why there is a slowdown at the embedded auxiliary in the null subject matching condition.

This evidence also indicates that the parser is trying to attach each new word incrementally to the syntactic structure being built, and is using any available local information to connect and interpret as much input as possible. If the parser did not attempt to connect the structure previously processed, there would be no reason to expect a reading time difference at this point.

⁷⁸ These 10 verbs are: *atera* ‘to take (out)’, *egin* ‘to do/make’, *ekarri* ‘to bring’, *erakutsi* ‘to show’, *erosi* ‘to buy’, *idatzi* ‘to write’, *konpondu* ‘to fix’, *lapurtu* ‘to steal’, *ordaindu* ‘to pay’, *saldu* ‘to shell’

The experiment in this study tested if the rich agreement encoded in Basque auxiliaries is used by the parser to make predictions about upcoming structure. I tested such predictions in two types of context in which the presence of a structural prediction would lead to a garden-path effect or to early anticipation of an embedded clause. For this purpose, sentences were tested where I manipulated the agreement match/mismatch between the auxiliary and the NP that follows it. I predicted that if there was a mismatch between the auxiliary and the NP the parser would be able to infer a clause boundary before encountering the verbal head and avoid a garden-path, whereas if there was a match between the auxiliary and the NP the parser would analyze the sentence as mono-clausal and it would have to reanalyze when it is clear that the sentence is bi-clausal.

When the auxiliary mismatched the following NP the manipulation in the presence of the embedded subject resulted in the following effects. The dative marked NP was read faster in the null subject condition relative to the same NP in the overt subject condition. In the null subject condition, the NP clearly mismatches with the morphological information in the auxiliary. We take this agreement mismatch cue to be helping the parser in doing the inference that there is a clause boundary. This is contrasted with the condition where the NP following the auxiliary is an Ergative-marked NP, which we take to be a much more direct cue in Basque to determine that there is a clause boundary (see section 5.2.1).

The auxiliary NP matching manipulation had the following effects: the embedded verb and auxiliary were read faster in the overt embedded subject condition relative to the same verb and auxiliary in the null embedded subject

condition. This result was taken as evidence for the fact that the parser attaches every new word into the input and uses the rich agreement information to guide parsing, not needing to wait for the verbal head for confirmation. This creates the possibility of being led down the garden-path. We found this to be the case.

5.5.1. Reanalysis in Basque

5.5.1.1 On-line derivation of the Matching Condition

In this section, I will consider the online derivation of the sentence in (11b) assuming a strong prediction-based left-corner parsing model (discussed in chapter 1, section 1.2.1). The structure for this condition is illustrated incrementally, where the thick lines represent the parsed structure, the dotted lines reflect the predicted structure, the brackets reflect the predicted nodes, the shaded boxes represent the upcoming input, the arrows represent the choices for where in the structure the new input could be interpreted and the star means the interpretation cannot be constructed.

The matching condition in (11b), repeated here as (14), is a case of misanalysis as a mono-clausal structure, requiring subsequent reanalysis as a bi-clausal one. The brackets {} indicate the structure that is temporarily misparsed and the parentheses [] indicate the right analysis for the sentence.

(14) *Matching condition*

{Maitanek ez dio ikasleari [txostena idatzi}
 {M_i-ERG neg **aux-ditrans student-DAT** [pro_i document-ABS write}
 duela] esan
 aux-trans-C] say.

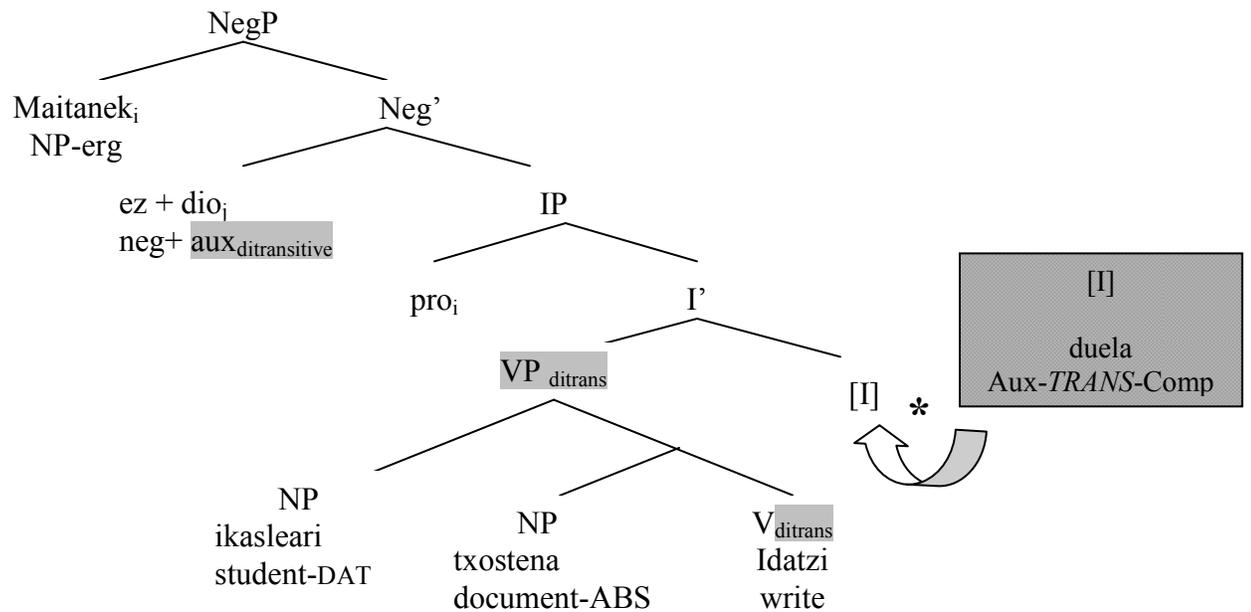
‘Maitane hasn’t told the student that she has written the document’

In this sentence the auxiliary is ditransitive, so the parser projects both the accusative and the dative argument of the verb and the verbal head. When the next word in the input becomes available, the expectation for a dative marked NP is fulfilled. The next word in the input is the absolutive NP which is entered into the structure as the second argument projected by the auxiliary. Then, the embedded verb is the next word and at this point the parser takes this verb to be related to the fronted auxiliary. This yields a single clause analysis.

The parser can only see that a biclausal structure is required at the embedded auxiliary. Then the parser has to revise its initial expectations and restructure the sentence. Consider the online derivation of sentence (14) at this specific point, assuming an incremental parser with a strong predictive mechanism.

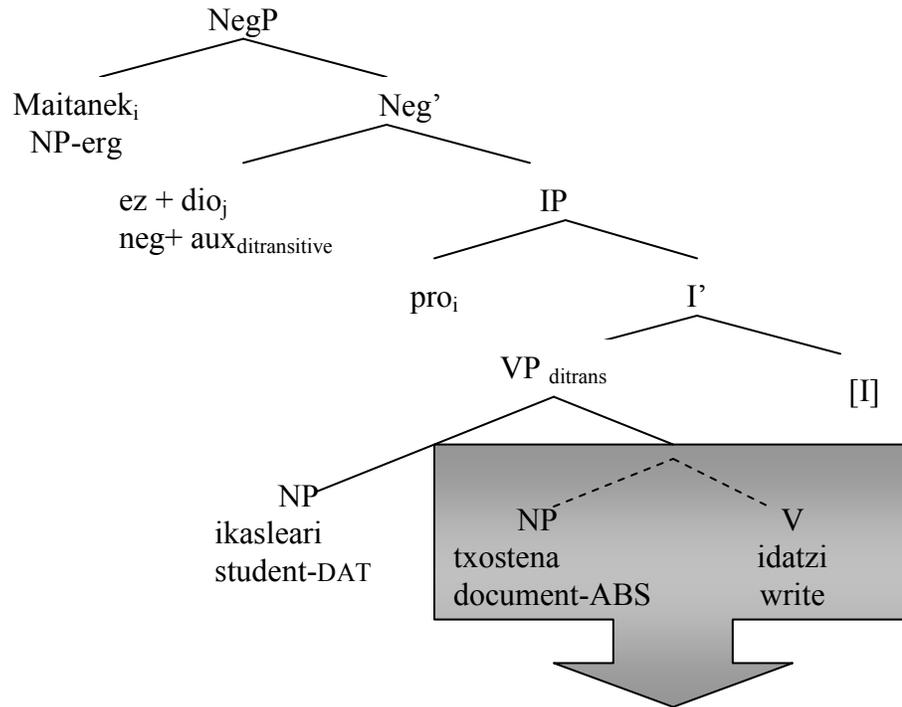
Step 1: the embedded auxiliary has to be included in the ongoing derivation of the sentence but considering the current monoclausal analysis of the sentence, this auxiliary cannot be inserted under the head of IP because it would yield an ungrammatical sentence. This is illustrated in the structure in (15)

(15)



Step 2: The monotransitive morphology of the auxiliary forces the parser to restructure the distribution of the arguments within the VP as shown in the structure in (16). The question here is how the parser figures out the appropriate way to restructure the sentence and which arguments should be reanalyzed as belonging to a different clause and why it doesn't just treat it as an error and can recover from the garden-path by restructuring the sentence.

(16)



If the auxiliary that the parser has encountered is monotransitive, it means the preceding verb 'idatzi' (to write) is used monotransitively and that it has to be lowered in the tree. Furthermore, the absolutive NP has to be the argument for the monotransitive verb that needs to be lowered because the option of having an absolutive argument and an embedded clause in the same sentence is not possible, as shown in the sequence in (17).

- (17) *Maitanek ez dio ikasleari txostena [idatzi
 *M-ERG_i neg aux-ditrans student-DAT document-ABS [pro_j write
 duela] esan
 aux-trans-Comp] V

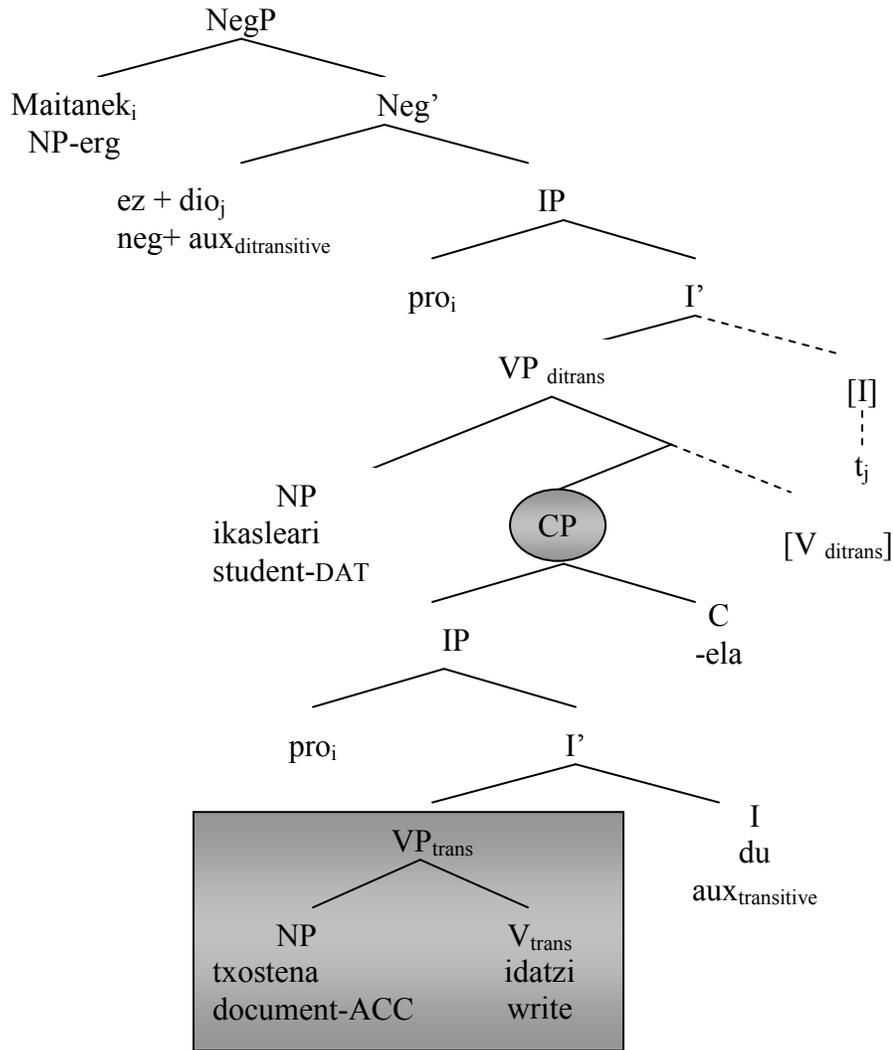
Whenever there is a ditransitive sequence that has to be reanalyzed in a declarative clause, Basque speakers analyze the overt absolutive argument as a

complement of the embedded clause instead of positing a null object within the embedded clause because a main clause overt absolutive argument can't overlap with an embedded clause.

Step 3: The parser has to postulate an embedded sentence where only the absolutive NP and the verb to write (*idatzi*) would be part of the VP⁷⁹ as shown in (18). The next question then is how the parser postulates the embedded sentence and how costly it is to do it.

⁷⁹ Whether this restructuring of the sentence is costly for the parser or not is discussed in section 5.5.1.2 in comparison with some Japanese cases that had been discussed in the literature.

(18)

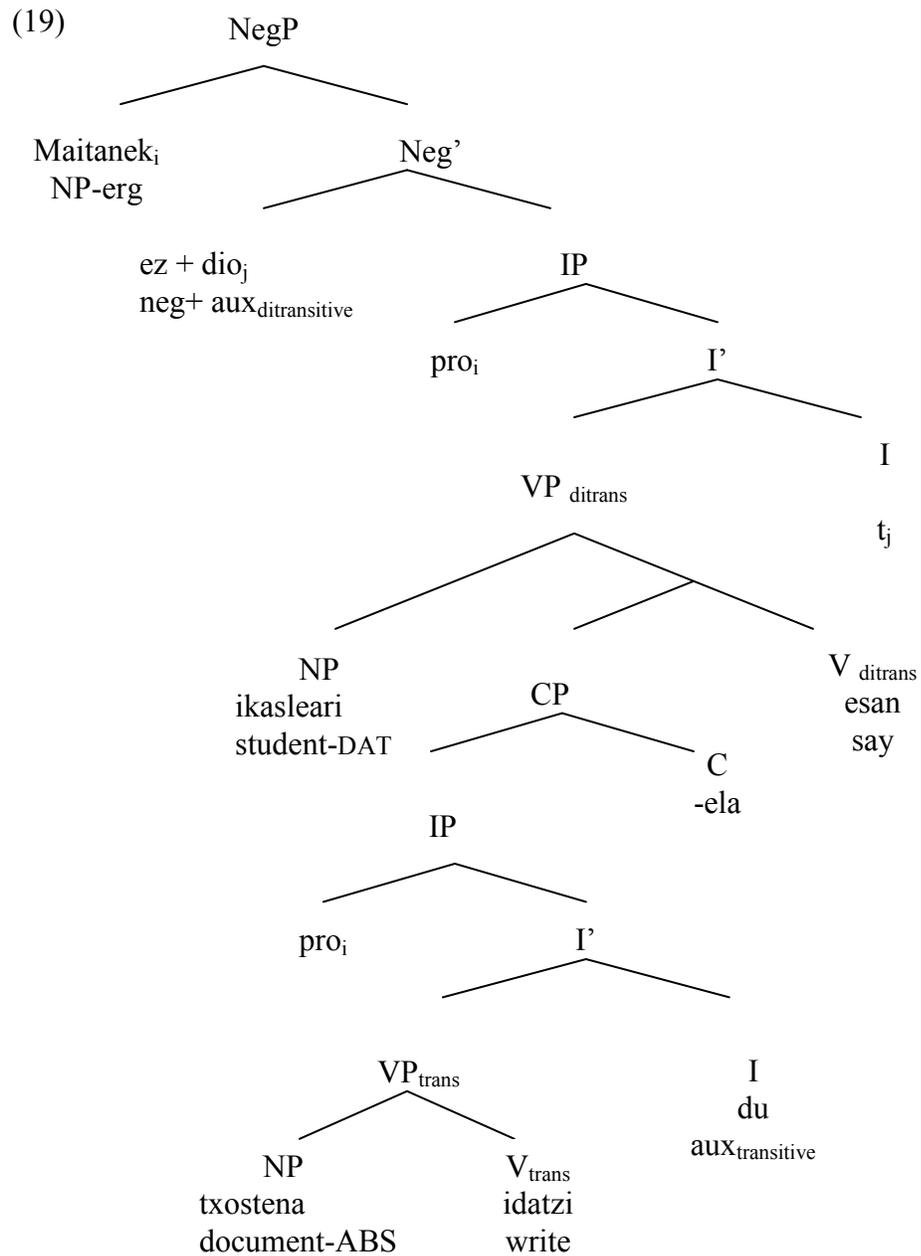


The postulation of a CP is not difficult because the embedded auxiliary introduces the complementizer. The extra steps that the parser needs to take in restructuring the sentence is positing a null subject for the embedded clause⁸⁰ and

⁸⁰ Positing a null subject coreferent with the main subject is considered to be not costly. There seems to be an asymmetry between positing a null subject and positing a null object in parsing (Miyamoto 2003). See the discussion (section 3.2.8) and general discussion (section 3.3) of the Spanish experiment 2 in chapter 3 for the same presupposition that considers the postulation of a null subject which is coreferent with the main subject not to be costly.

maintaining the expectation for the main verb active so that the main clause can be interpreted and all the arguments can be integrated into the structure.

In (19), I show the complete structure for the matching condition in (14).



The control condition for the matching condition in (14) was (20), where the embedded subject was overt.

- (20) *Matching-overt subject condition*
Maitanek_i ez dio ikasleari [zuzendariak txostena
M_i-ERG neg aux-ditrans student-DAT [director-ERG document-ACC
idatzi duela] esan.
write aux-trans-C] say.
‘Maitane hasn’t told the student that the director has written the document’

In contrast, in the control condition (20), the ergative case marking on the NP director (*zuzendariak*) is a clear marking for the embedded sentence. Therefore, in this condition the absolutive marked NP and the verb to write (*idatzi*) would be initially analyzed as part of the embedded clause. Since there is no reanalysis in this case, we did not expect any effect at the embedded auxiliary in this condition, in contrast with the reading time for the same word in the target condition (14).

5.5.1.2 Reanalysis in Basque compared to Reanalysis in Japanese

One of the results of this experiment is the garden-path at the embedded verb and auxiliary in the matching conditions, whose on-line derivation has been discussed in the previous section. It is interesting to contrast the restructuring that has to be carried out in the garden-path cases in Basque with the examples discussed in the literature involving the reanalysis of relative clauses in Japanese (Weinberg 1993, Inoue & Fodor 1995, Mazuka & Itoh 1995, Gorrell 1995, Sturt & Crocker 1996), where depending on the restructuring that needs to be done, the garden-path is expected to have a different cost associated to it.

Previous studies have focused on the restructuring involved in the processing of relative clauses in Japanese. Mazuka & Itoh (1995) specifically discussed the contrast between two types of garden-path sentences in Japanese that derive in different reanalysis cost for the parser, depending on the number of arguments that have to be displaced and on the kind of operation that needs to be performed to restructure⁸¹. One of the basic contrasts they discussed is reflected in (21).

- (21) a. Yamashita-ga yuuzin-o hoomonshita kaisya-de mikaketa
Y-Nom friend-Acc visited company-LOC saw
 ‘Yamashita saw his friend at the company he visited’
- b. Yamashita-ga yuuzin-o hoomonshita siriai-ni tegami-o kaita
Y-Nom friend-Acc visited acquaintance-Dat letter-Acc wrote
 ‘Yamashita wrote a letter to an acquaintance who visited his friend’

The example in (21a) will be reanalyzed as (22), where the parser has to raise the subject and object and posit two empty categories for these two arguments in the embedded clause.

- (22) Yamashita-ga yuuzin-o [Ø_{nom} Ø_{acc} hoomonshita] kaisya-de mikaketa
Y-Nom friend-Acc visited company-LOC saw
 ‘Yamashita saw his friend at the company he visited’

The restructuring of this sentence would proceed in the following way assuming that the parser is considering a monoclausal analysis of the sentence until the verb *to visit* (*hoomonshita*) is processed. The accusative noun phrase *friend* (*yuuzin*) has to be raised out of the VP where the parser would have thought it belonged (the verb ‘*to visit*’ *hoomonshita*) to the object position of the main verb *to see* (*mikaketa*).

⁸¹ There are two views on the literature on this. On the one hand, there is the view that contrasts lowering of structure in the syntactic tree vs. raising of structure in the tree as being associated to a different reanalysis cost (Weinberg 1993, Gorrell 1993). On the other hand, the other view contrasts double displacement vs. single displacement of arguments as the reason for reanalysis being more or less costly respectively (Sturt & Crocker 1996).

Therefore, this would be a case of double displacement and raising. When reanalysis requires restructuring by raising the arguments in the structure as in (22), recovering from the garden-path is supposed to be more costly (Weinberg 1993, Gorrell 1995, Sturt & Crocker 1996).

On the other hand, sentence (21b) would be reanalyzed as (23), where the parser has to lower the subject and postulate an empty category for the subject in the embedded clause.

- (23) Yamashita-ga [\emptyset nom yuuzin-o hoomonshita] siriai-ni
Y-Nom friend-Acc visited acquaintance-Dat
 tegami-o kaita
letter-Acc wrote
 ‘Yamashita wrote a letter to an acquaintance who visited his friend’

The restructuring of this sentence would proceed in the following way assuming that the parser is considering again a monoclausal analysis. The verb *visited* (*hoomonshita*) and the accusative argument *friend* (*yuuzin*) have to be “lowered” and an empty category has to be postulated at the embedded sentence for a subject. Therefore, this would be a case of single displacement and lowering. When reanalysis requires restructuring by lowering arguments in the structure as in (23), the garden-path is not supposed to cause serious processing difficulty.

As discussed in the literature on garden-path recovery, input requiring alteration or deletion of primary structural relations leads to processing difficulty. Particularly “raising” the structure that needs to be reanalyzed is more costly than “lowering” it, irrespective of the number of arguments that are lowered or raised⁸².

⁸² This last claim is debateable as reflected in the discussion of the Basque example at hand.

Taking the previous facts into consideration, the Basque example in (24) taken from the matching null subject condition resembles the Japanese example in (23) in that it is also a case of “lowering”. However, it is a case of double displacement⁸³ that in Japanese garden-path literature is considered to be costly, as seen in (22). The parser has to lower the absolutive marked NP *txostena* (*document*) and the embedded verb *idatzi* (*to write*) and posit an empty subject in the lowered node.

- (24) Maitanek_i ez dio ikasleari [Ønom txostena idatzi
M_i-ERG neg aux-ditrans student-DAT [document-ABS write
duela] esan.
aux-trans-Comp] say.
 ‘Maitane hasn’t told the student that she/he has written the document’

Based on previous arguments for Japanese reanalysis examples, we would predict that the garden-path in sentence (24) would not cause serious processing difficulty and the reanalysis cost would be low because it is a case of lowering structure. This seems to be the case since Basque speakers recover easily from the garden path as shown in the lack of reading time difference at the main verb region (region 12) following the auxiliary in the matching null subject condition of the experimental items. However, the garden-path reported for Basque does not seem to be affected in terms of reanalysis cost by a double displacement of argument as seems to be the case in Japanese. Therefore, Basque raises an interesting contrast with the previous studies.

⁸³ Notice that the displacement of arguments in this case causes the split of the arguments within a VP (the dative NP and the accusative NP).

5.5.2. How to build an embedded clause from a mismatch

5.5.2.1 On-line derivation of Mismatching condition

In this section, I will consider the online derivation of the sentence in (11a) repeated here as (25), a case of agreement mismatch.

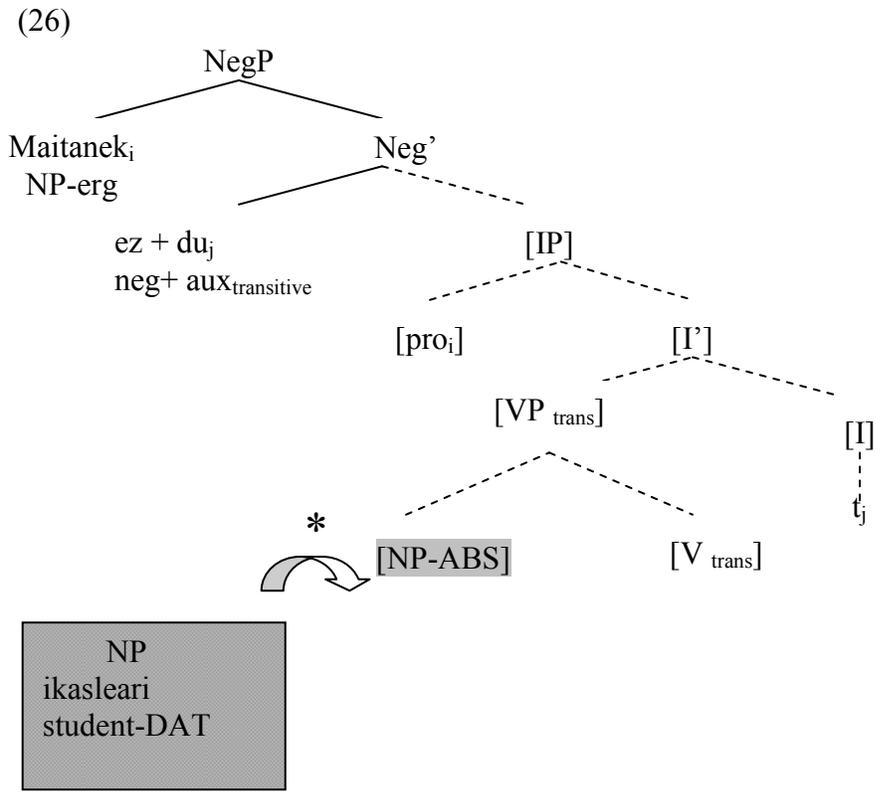
- (25) a. *Mismatching condition*
[Maitanek_i ez du [ikasleari txostena idatzi
[M_i-ERG neg **aux-trans** [pro_i **student-DAT** document-ABS write
diola] esan.]
aux-ditrans-C] say.]
'Maitane hasn't said that she has written the document to the student'

In what follows, the structure for this condition is illustrated incrementally and represented in the same way as that of (11b) in section 5.5.1.1.

Step 1: After the main subject and negation have been processed, the auxiliary is processed and its subcategorization information allows projection of the verb and its argument slots, which in this case is monotransitive. Therefore, we would only expect an absolutive marked NP. The predicted nodes are shown between brackets in the structure in (26)⁸⁴.

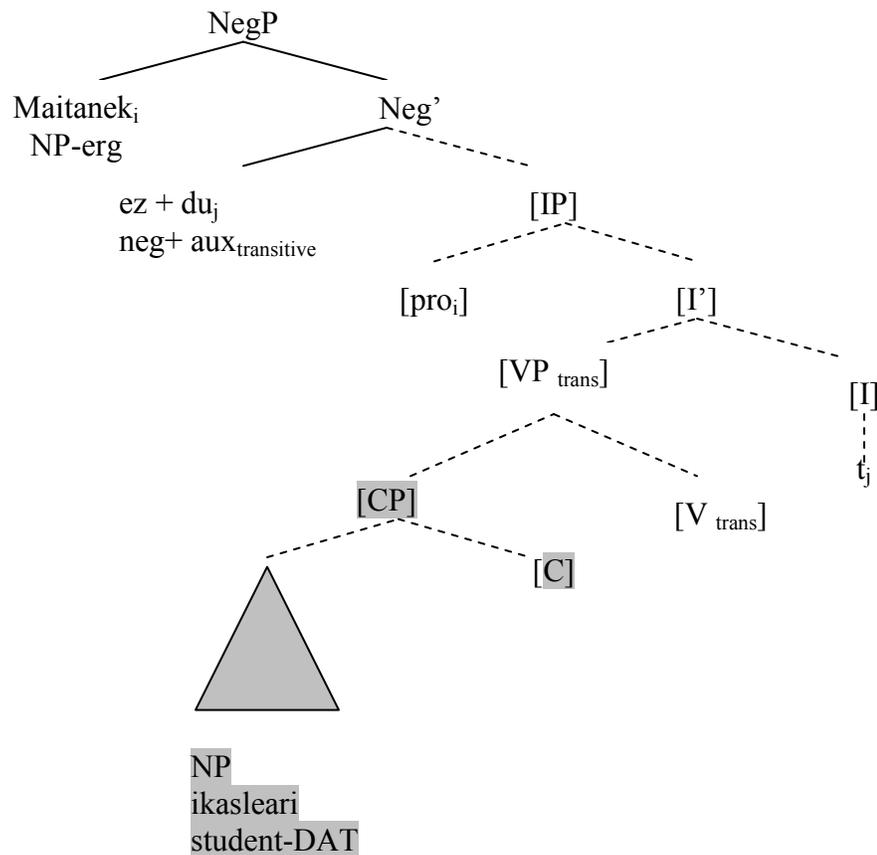
Step 2: The dative-marked NP has to be introduced in the current structure. Its case does not match the predictions established by the auxiliary to have an absolutive marked NP. At this point, the parser must inhibit the step of projecting a new argument slot for the dative NP in the main clause.

⁸⁴ See previous footnote, for the assumptions I make regarding the structure of Negative sentences in Basque.



Step 3: the parser needs to figure out at this point how to accommodate the NP and it needs to conclude from the mismatch that the NP-dative is one of the constituents within an embedded clause. The biggest challenge for the parser is to figure out that there is a whole clause in the absolutive argument slot as illustrated in the structure in (27).

(27)



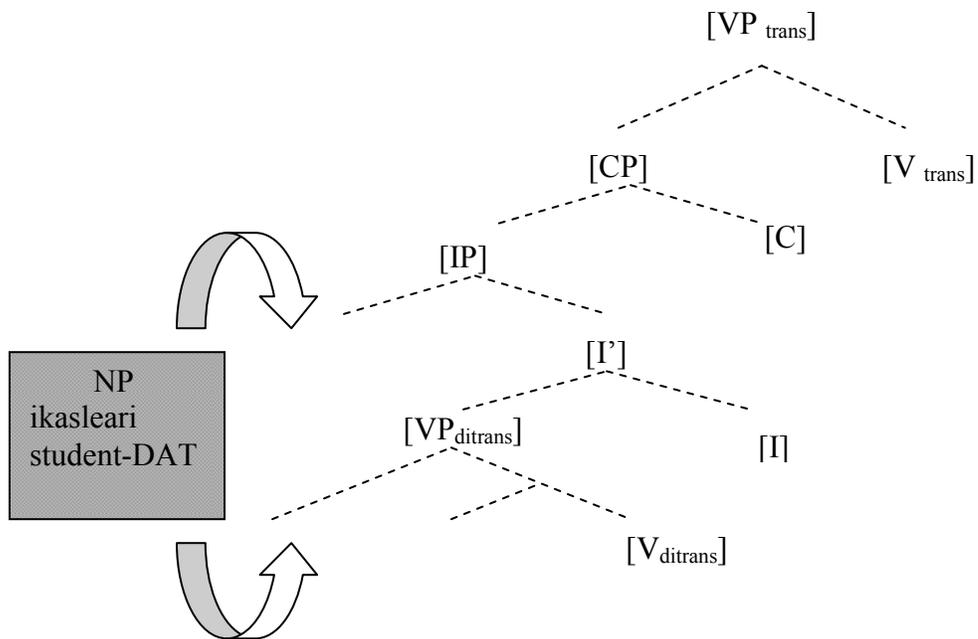
So that the sentence is parsed without any problem, the parser has to be able to infer that there is a whole CP in the absolutive argument and that the dative NP is contained within that CP⁸⁵.

Step 4: once the parser has figured out that there is an embedded clause, it needs to find out the role that the dative-marked NP is playing within that embedded sentence. The dative-marked NP could be either the dative argument of a ditransitive verb phrase where the subject is the main sentence antecedent (*Maitane* in this case)

⁸⁵ The mechanisms employed by the parser in order to connect the dative NP to the CP will be discussed in more detail in the next section.

or the subject of the embedded clause⁸⁶ as illustrated in the structure in (28). We take the first choice, where the NP-dat is part of the VP, to be the most plausible one because that is the most frequent occurrence of datives in Basque⁸⁷ and because it might be easier to presuppose that the subject of the embedded clause has the same antecedent as in the main clause⁸⁸. How this readjustment of the structure is done will be further discussed in the next section.

(28)



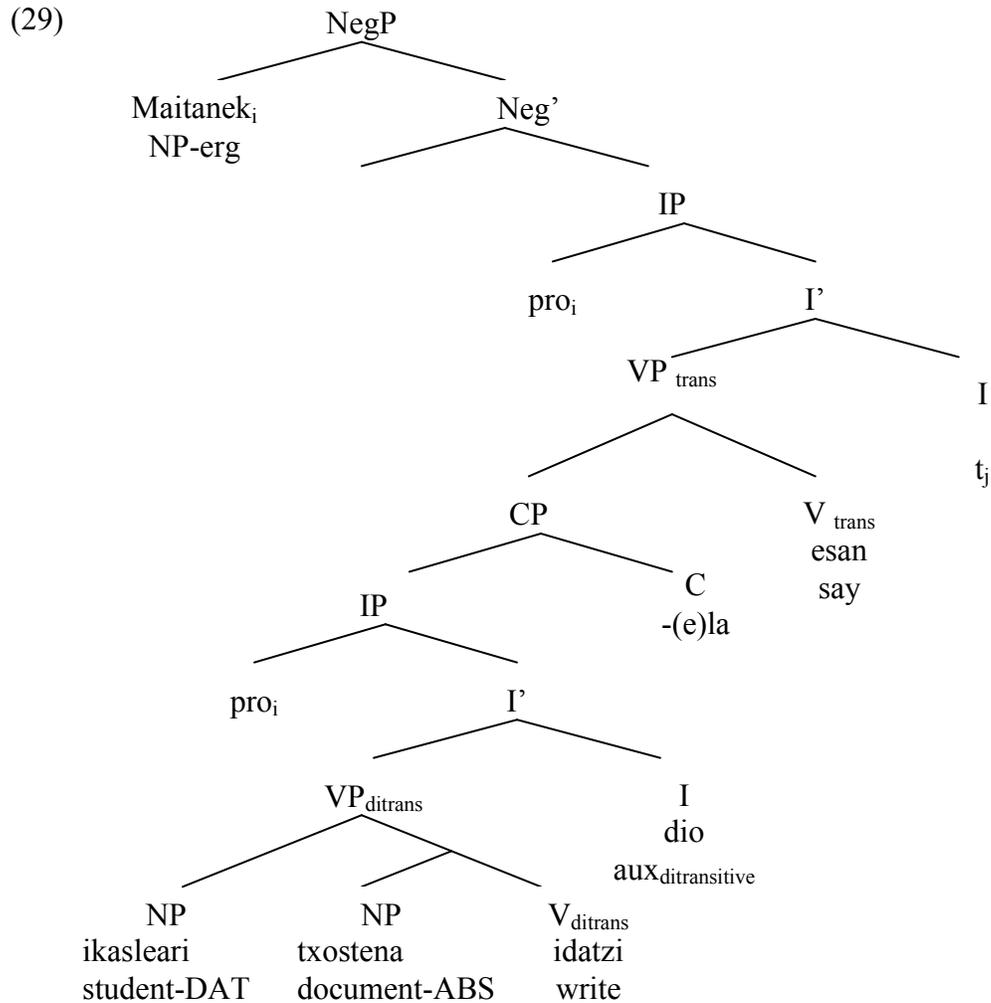
⁸⁶ The embedded sentence could have a dative marked subject.

(i). Maitanek ez du [ikasleari pelikula gustatu zaiola] esan.
M-ERG neg aux student-DAT movie-ACC like aux-Comp say
‘Maitane has not said that the student liked the movie’

⁸⁷ This statement is based on an intuition, I do not have data to support this.

⁸⁸ As in the discussion (section 3.2.8) of the Spanish experiment 2 reported in chapter 3, we take the option of presupposing a null subject with a previously introduced antecedent to be the most economical for the parser rather than having to introduce a new element in the discourse.

So if as we expect, the parser is using the indirect cue provided by the mismatch to avoid the garden-path in the embedded clause, the complete structure for (25) should be as in (29).



To create a control for the previous condition, we matched this sentence with the condition in (30) where the embedded subject was overt and marked with ergative case.

- (30) *Mismatching-overt subject condition*
 [Maitanek ez du [zuzendariak ikasleari txostena
M-ERG neg aux-trans [director-ERG student-DAT document-ACC
 idatzi diola] esan.]
write aux-ditrans-C] say.]
 ‘Maitane hasn’t said that the director has written the document to the
 student’

Here the ergative case marking on the NP *director* is a clear marking for the embedded sentence and it should enter the structural derivation as the subject of the embedded sentence. In contrast to (25), when the attachment of the embedded clause involves the creation of an embedded complement, the complement has already been constructed as soon as the ergative-marked NP is processed.

5.5.2.2 Building an embedded clause from a mismatch

In this section I will concentrate on the result derived from the mismatch of the auxiliary and the following NP within the mismatching conditions. As I have discussed in section 5.3, there needs to be a strong predictive mechanism that allows the parser to infer that there is an embedded clause from the fact that there is a mismatch so that the garden-path in the embedded clause is prevented. In contrast to studies in Chinese (Hsu 2006) where Chinese speakers enter into a parsing breakdown from encountering a classifier mismatch, Basque speakers⁸⁹ recover from the mismatch and furthermore, they are able to use the information provided in this mismatch together with that of the grammar of Basque to project structure in advance and prevent a future garden-path.

⁸⁹ As we have discussed in section 5.2.3. Japanese speakers also recover from the mismatch (Yoshida 2006).

In what follows, I will show how the parser can create an embedded clause from the NP-dat in Basque and I will discuss what kind of parser is required to account for the results of the Basque experiment without falling into structure over-generation. As I will discuss, a more powerful parser than the left-corner parser (Jonhson-Laird 1983, Resnik 1992, Stabler, 1994, Abney & Johnson 1999, Schneider 1999, Aoshima 2003) is required to generate the Basque mismatching conditions since the prediction based on the lexical information provided by the input, which in this case is a dative NP, is not enough to infer a CP from this dative NP. The parser has to be able to infer that there is a whole CP in the absolutive argument of the main verb and that the dative NP is contained within that CP, otherwise we would expect parsing breakdown and the experimental results show that this is not the case. It is also important to note that this is where the two kinds of parsing models I have discussed before would make different predictions. Under the non-anticipatory structure building parser, there is no way to generate all this structure so that the sentence can be parsed, therefore, there would be no creation of an embedded sentence and the agreement violation would be simply left unresolved until the decisive verbal information appears. Under an expectation-based parser, the mismatch would be the trigger to realize that there is something that has to be changed and the parser will try to figure it out from the moment the mismatch happens. A different thing is when does the parser create the embedded sentence. That is something that cannot be determined.

Moreover, the parser that will generate the structures in this experiment also needs to have restrictions on structure generation so that it is not always expected to build an embedded clause for every parsing problem that arises.

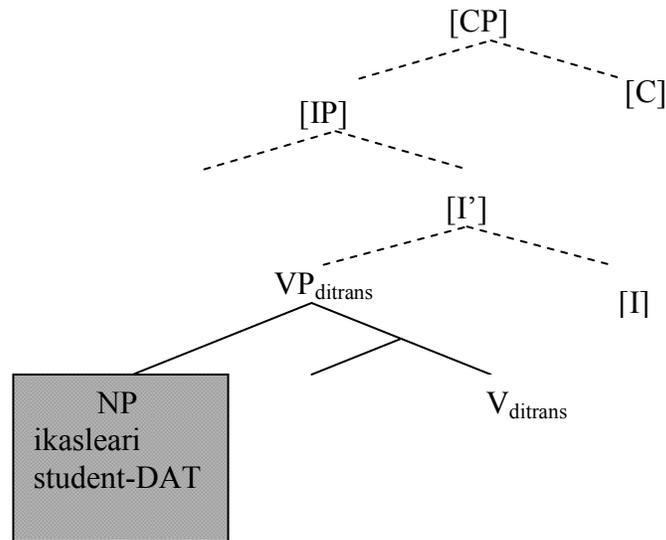
Taking up again the discussion of the on-line derivation of the mismatching condition in the previous section, the parsing of the sentence proceeds normally until the parser notices the mismatch between the monotransitive auxiliary and the dative NP that follows it and it realizes that the NP-dative cannot be directly attached to the current structure (Step 2 in the derivation in previous section, example (26)).

The auxiliary requires an ergative and an absolutive argument to license it. The requirement for the ergative NP has already been fulfilled with the NP-erg *Maitanek* and the expectation is to have an absolutive argument, which could either be an absolutive NP or a complement. The assumption will be that the absolutive argument will preferably be an NP and not a sentential complement so that the argument can be locally licensed and because that will favor the monoclausal analysis of the sentence (as discussed in chapter 1, section 1.1.2). The next word in the input is the dative marked NP, which cannot satisfy the prediction established by the agreement information in the auxiliary. The parser then has to reconsider the structure of the sentence, which up to this point was mono-clausal, and infer that the dative argument is part of the absolutive complement that the auxiliary requires. The question at this point in the derivation is how an incremental parser can connect the word in the input, a dative-marked NP, to the current structure by creating an embedded sentence (a CP in this case) which will be in the absolutive argument slot of the main verb and that contains the dative-marked NP. (The structure in example

(27) in the previous section shows how far the parser's knowledge can go at this point).

Assuming that the parser is taking the dative-marked NP to be part of the verb phrase and that within the embedded clause the dative is in the VP⁹⁰, the verbal phrase and its head can be projected from the presence of the dative NP, but to get the CP projection which is the maximal projection for the embedded clause⁹¹, an intermediate IP needs to be projected first. The tree in (31) illustrates the structure that can be projected based on the dative NP and the one that needs to be projected to get the rest of the sentence right.

(31)



⁹⁰ As discussed in step 4 of the derivation of the mismatching condition in section 5.5.2.1, we make the assumption that the dative-marked NP is considered to be the dative argument of a ditransitive verb phrase where the subject is the main sentence antecedent. This is because we take the occurrence of dative arguments as part of a verbal phrase to be the more frequent in Basque. The second reason to make this assumption is because presupposing that the subject of the embedded clause has the same antecedent as in the main clause might be the less costly option.

⁹¹ This is because embedded clauses in Basque are always introduced by a clause complementizer.

The question then is how the parser can connect the VP with the CP and more relevantly, how can it construct the intermediate IP projection for that. This is the step where the parsing of Basque makes a difference with respect to other languages, since the IP is not something that can be inferred from the case information⁹² provided in the dative, nor from the word order or context⁹³. No additional inference can be made from the dative case marking in the NP besides the fact that it belongs inside a VP that has ditransitive features. Therefore, there needs to be an extra presupposition so that the sentence is parsed and the dative is connected to the rest of the structure.

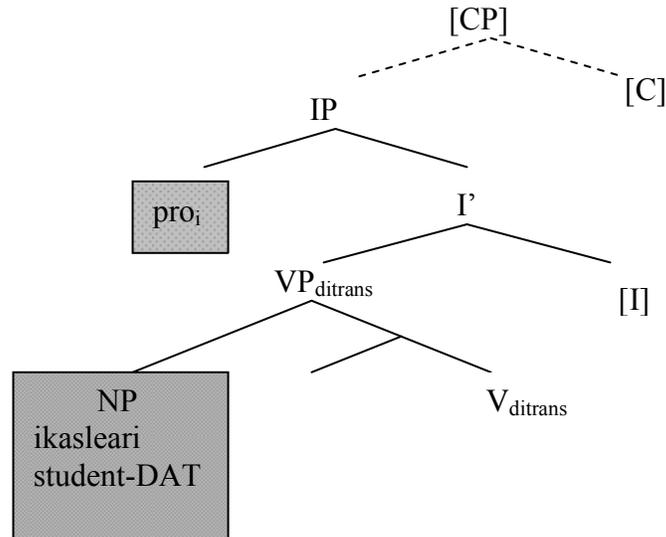
It seems the only way the derivation of the sentence can proceed and not breakdown is by inserting a null subject in the embedded clause that can connect the VP with the IP. This null subject needs to be automatically posited as soon as it is assumed that the dative is part of a VP. If the dative is not the subject of the embedded sentence, then the sentence needs to have a null subject so that the dative can be part of the verb phrase. Basically, once the dative is taken to be part of the VP, it needs to be assumed that there is a null subject and a *pro* needs to be posited. Moreover, the insertion of the null subject in the structure cannot be done before it is seen that the dative NP cannot be connected to the current structure. Somehow, the insertion needs to be done after the input needs restructuring as a last resort operation. By last resort I mean that the insertion of the embedded sentence (and the required heads that need to be projected for that matter) only takes place when there is no other way to connect the words in the input and when even though this is the case, the

⁹² As is done in Japanese RC classifier mismatch cases (Yoshida 2006)

⁹³ As is done in Chinese RC classifier mismatch cases (Hsu 2006)

parser still does not break down. The structure in (32) shows this *pro* insertion, which would be done after the NP-dat cannot be connected to the IP.

(32)



Finally, once the IP has been constructed after the insertion of the null subject in the structure, the construction of the CP is required. The fact that embedded sentences are always introduced by a clause-final complementizer in Basque automatically predicts a CP. Therefore, the crucial step in the derivation of this sentence is the creation of the IP node.

As I have discussed, there are also a couple of assumptions that need to be taken into account when deriving the on-line structure of the sentence. First, the dative argument will preferably be taken to be part of a VP in Basque and, second, a null subject is only inserted in the existing structure when there is evidence that the dative is part of the VP.

The last question is how can we prevent structure over-generation and restrict the insertion of the embedded clause only to those cases that require it. Overgeneration can be avoided if the embedded sentence insertion is a last resort operation. In the Basque example I am analyzing, the embedded sentence needs to be built because the prediction of the main sentence auxiliary to have a monotransitive verbal head has not been fulfilled and it won't be fulfilled unless the dative NP is connected to the ongoing structure. This forces the parser to connect the dative to the ongoing derivation even if the parser cannot directly infer how to connect it. As shown in the structures in (26) and (27) in the previous section, the prediction from the main sentence auxiliary is still on hold while the embedded CP is constructed and the parser figures out what to do with the NP-dative.

Moreover, this structure building and the projection of heads ahead can only be possible under parsing architectures that allow a powerful projection of structure in advance of bottom-up information about the head of phrases (Crocker 1994, Schneider 1999, Lombardo & Sturt 2002). In the case I am analyzing, there is no direct evidence from the input as to how any of the heads can be predicted (e.g. not for the CP or IP). The information on how to predict these heads can only be derived from the requirements intrinsic to the grammar of the Basque language.

Finally, I compare the prediction that is necessary for this experiment with the prediction required for studies on Japanese and Chinese classifier mismatch. With respect to the Japanese cases discussed in section 5.2.3, we see that in the Basque examples there is a need for more inference at the point of the mismatch partly due to the fact that there is more than one alternative as to what could be the role of the

dative in the sentence. In Basque, there is not such clear evidence in the mismatch region (dative NP) as in the Japanese, where the nominative marker at the mismatch region suggests the construction of an embedded clause. In the Basque cases, the knowledge to build an embedded clause only comes from the fact that the dative NP needs to be attached to the current structure, and for that, it is most likely to be attached in the structure as an argument of the VP. In Basque there is an extra step required to infer the embedded clause, which comes as a last resort operation, as a by-product of not being able to attach the dative NP. Moreover, once the embedded clause is posited, it is not clear for a Basque speaker what kind of embedded clause it is, whereas for Japanese speakers there is only one possible structure, the Relative Clause.

With respect to the Chinese cases discussed in (Hsu 2006), Basque differs in that speakers do not enter into a parsing breakdown as Chinese speakers do even if at the point of the mismatch they might not know what to do with the mismatch yet. The range of possibilities does not make them fail to parse these structures. On the other hand, Basque and Chinese have in common the fact that at the point of the mismatch the structure is ambiguous and the postulation of the embedded clause is not immediately derivable from case information as in Japanese. Chinese differs from Basque in that it seems to require more extra steps to go from the input to the predicted structure because, even if the classifier sets a prediction for an upcoming noun and this should promote a creation of an embedded clause, the parser does not know what to conclude from the mismatching noun with ambiguous case features and it breaks down. eation of an embedded clause, the parser does not know what to

conclude from the mismatching noun with ambiguous case features and it breaks down.

5.6. Conclusions

Results of the online experiment suggest that information available prior to verbal heads is being used by the parser to build syntactic structure and that the parser tries to accommodate every new word in the input to the existing tree description. Evidence for this comes from the lack of effect at the most embedded auxiliary in the mismatching conditions.

The lack of effect in the mismatching conditions at the embedded auxiliary also provides evidence that the morphological information available in the auxiliary is used to guide parsing and to predict upcoming material in the structure. This together with the slow down in the mismatch region for these conditions, points to the use of the agreement information encoded in the auxiliary as an indirect cue that induces the parser to posit an embedded clause and avoid an upcoming garden-path.

In contrast to the lack of garden-path in the mismatching conditions, the presence of the garden-path in the matching conditions shows that whenever the indirect cue of the mismatch is not present, the parser cannot get to build an embedded clause ahead of time and cannot avoid the garden-path.

In sum, this experiment shows how the combination of rich agreement and case marking can be used to anticipate structure and guide syntactic parsing. The match/mismatch of the agreement information encoded in Basque auxiliaries with the

case information in the adjacent NP provides the parser with a cue to predict upcoming structure and to avoid in some cases future misanalyses of the sentence. Additionally, this experiment provides evidence for an incremental parser that makes use of every available piece of information to predict structure and that does not wait for the confirmation provided by heads to combine the arguments encountered on the way.

Appendix 5-A: Data items for online reading experiment on basque auxiliary fronting

item 1

- 1a. Maitanek ez dio ikasleari ikastaroko txostena oso presaka idatzi duela esan.
? Maitane al da ikasleari txostena presaka idatzi duela esan ez diona? Y
- 1b. Maitanek ez dio ikasleari zuzendariak ikastaroko txostena oso presaka idatzi duela esan.
? Maitane al da ikasleari zuzendariak txostena presaka idatzi duela esan ez diona? Y
- 1c. Maitanek ez du ikasleari ikastaroko txostena oso presaka idatzi diola esan.
? Maitane al da ikasleari txostena presaka idatzi diola esan ez duena? Y
- 1d. Maitanek ez du zuzendariak ikasleari ikastaroko txostena oso presaka idatzi diola esan.
? Maitane al da zuzendariak ikasleari txostena presaka idatzi diola esan ez duena? Y

item 2

- 2a. Gorkak ez dio amari festarako erreserba interneten bidez egin duela aipatu.
? Ama al da erreserba interneten bidez egin duena? N
- 2b. Gorkak ez dio amari anaiak festarako erreserba interneten bidez egin duela aipatu.
? Ama al da erreserba interneten bidez egin duena? N
- 2c. Gorkak ez du amari festarako erreserba interneten bidez egin diola aipatu.
? Ama al da erreserba interneten bidez egin duena? N
- 2d. Gorkak ez du anaiak amari festarako erreserba interneten bidez egin diola aipatu.
? Ama al da erreserba interneten bidez egin duena? N

item 3

- 3a. Jonek ez dio gaixoari azterketa medikoa arreta gutxiz egin duela esan.
? Egia al da Jonek azterketa medikoa arreta gutxiz egin duela? Y
- 3b. Jonek ez dio gaixoari medikuak azterketa medikoa arreta gutxiz egin duela esan.
? Egia al da medikuak azterketa medikoa arreta gutxiz egin duela? Y
- 3c. Jonek ez du gaixoari azterketa medikoa arreta gutxiz egin diola esan.
? Egia al da Jonek azterketa medikoa arreta gutxiz egin duela? Y
- 3d. Jonek ez du medikuak gaixoari azterketa medikoa arreta gutxiz egin diola esan.
? Egia al da medikuak azterketa medikoa arreta gutxiz egin duela? Y

item 4

- 4a. Igonek ez dio sukaldariari jatetxerako janaria aurreko egunean egin duela azaldu.
? Egia al da Igonek janaria egun berean egin duela? N
- 4b. Igonek ez dio sukaldariari mutikoak jatetxerako janaria aurreko egunean egin duela azaldu.
? Egia al da mutikoak janaria egun berean egin duela? N
- 4c. Igonek ez du sukaldariari jatetxerako janaria aurreko egunean egin diola azaldu.
? Egia al da Igonek janaria egun berean egin duela? N
- 4d. Igonek ez du mutikoak sukaldariari jatetxerako janaria aurreko egunean egin diola azaldu.
? Egia al da mutikoak janaria egun berean egin duela? N

item 5

- 5a. Amaiak ez dio andreari itxasondoko etxea oso merke saldu duela kontatu.
? Itxasondo etxea izan al da oso merke saldu dena? Y
- 5b. Amaiak ez dio andreari eraikuntza-enpresak itxasondoko etxea oso merke saldu duela kontatu.

? Itxasondo etxea izan al da oso merke saldu dena? Y

5c. Amaiak ez du andreari itxasondoko etxea oso merke saldu diola kontatu.

? Itxasondo etxea izan al da oso merke saldu dena? Y

5d. Amaiak ez du eraikuntza-empresak andreari itxasondoko etxea oso merke saldu diola kontatu.

? Itxasondo etxea izan al da oso merke saldu dena? Y

item 6

6a. Xabierrek ez dio gaixoari eguerdiko janaria gatzik gabe prestatu duela azaldu.

? Azaldu al du Xabierrek eguerdiko janaria gatzik gabe prestatu duela? N

6b. Xabierrek ez dio gaixoari erizainak eguerdiko janaria gatzik gabe prestatu duela azaldu.

? Azaldu al du Xabierrek erizainak eguerdiko janaria gatzik gabe prestatu duela? N

6c. Xabierrek ez du gaixoari eguerdiko janaria gatzik gabe prestatu diola azaldu.

? Azaldu al du Xabierrek eguerdiko janaria gatzik gabe prestatu duela? N

6d. Xabierrek ez du erizainak gaixoari eguerdiko janaria gatzik gabe prestatu diola azaldu.

? Azaldu al du Xabierrek erizainak eguerdiko janaria gatzik gabe prestatu duela? N

item 7

7a. Aitziberrek ez dio lagunari enpresako kotxea berandu samar konpondu duela esan.

? Enpresako kotxea izan al da berandu samar konpondu dena? Y

7b. Aitziberrek ez dio lagunari mekanikariak enpresako kotxea berandu samar konpondu duela esan.

? Enpresako kotxea izan al da berandu samar konpondu dena? Y

7c. Aitziberrek ez du lagunari enpresako kotxea berandu samar konpondu diola esan.

? Enpresako kotxea izan al da berandu samar konpondu dena? Y

7d. Aitziberrek ez du mekanikariak lagunari enpresako kotxea berandu samar konpondu diola esan.

? Enpresako kotxea izan al da berandu samar konpondu dena? Y

item 8

8a. Ederrek ez dio emakumeari museorako margolana asegurririk gabe bidali duela kontatu.

? Aseguruaz bidali al du Ederrek margolana? N

8b. Ederrek ez dio emakumeari artistak museorako margolana asegurririk gabe bidali duela kontatu.

? Aseguruaz bidali al du artistak margolana? N

8c. Ederrek ez du emakumeari museorako margolana asegurririk gabe bidali diola kontatu.

? Aseguruaz bidali al du Ederrek margolana? N

8d. Ederrek ez du artistak emakumeari museorako margolana asegurririk gabe bidali diola kontatu.

? Aseguruaz bidali al du artistak margolana? N

item 9

9a. Naiarak ez dio epaileari dibortziorako agiria baimenik gabe erakutsi duela erran.

? Dibortziorako agiria al da baimenik gabe erakutsi dena? Y

9b. Naiarak ez dio epaileari abokatuak dibortziorako agiria baimenik gabe erakutsi duela erran.

? Dibortziorako agiria al da baimenik gabe erakutsi dena? Y

9c. Naiarak ez du epaileari dibortziorako agiria baimenik gabe erakutsi diola erran.

? Dibortziorako agiria al da baimenik gabe erakutsi dena? Y

9d. Naiarak ez du abokatuak epaileari dibortziorako agiria baimenik gabe erakutsi diola erran.

? Dibortziorako agiria al da baimenik gabe erakutsi dena? Y

item 10

10a. Maialenek ez dio poliziarri manifestaldiko argazki-filma legez kontra lapurtu duela aitortu.

? Poliziak lapurtu al du argazki-filma? N

10b. Maialenek ez dio poliziarik kazetariak manifestaldiko argazki-filma legez kontra lapurtu duela aitortu.

? Poliziarik lapurtu al du argazki-filma? N

10c. Maialenek ez du poliziarik manifestaldiko argazki-filma legez kontra lapurtu diola aitortu.

? Poliziarik lapurtu al du argazki-filma? N

10d. Maialenek ez du kazetariak poliziarik manifestaldiko argazki-filma legez kontra lapurtu diola aitortu.

? Poliziarik lapurtu al du argazki-filma? N

item 11a

11a. Julenek ez dio gaixoari beharrezko odola ebakuntza aurretik eman duela kontatu.

? Julene izan al da odola eman duela kontatu ez duena? Y

11b. Julenek ez dio gaixoari gizonak beharrezko odola ebakuntza aurretik eman duela kontatu.

? Julene izan al da gizonak odola eman duela kontatu ez duena? Y

11c. Julenek ez du gaixoari beharrezko odola ebakuntza aurretik eman diola kontatu.

? Julene izan al da odola eman duela kontatu ez duena? Y

11d. Julenek ez du gizonak gaixoari beharrezko odola ebakuntza aurretik eman diola kontatu.

? Julene izan al da gizonak odola eman duela kontatu ez duena? Y

item 12

12a. Aingeruk ez die ikaskideei ariketen erantzuna emailaren bidez bidali duela aipatu.

? Aipatu al die Aingeruk ikaskideei ariketen erantzuna emailaz bidali duela? N

12b. Aingeruk ez die ikaskideei irakasleak ariketen erantzuna emailaren bidez bidali duela aipatu.

? Aipatu al die Aingeruk ikaskideei irakasleak ariketen erantzuna emailaz bidali duela? N

12c. Aingeruk ez du ikaskideei ariketen erantzuna emailaren bidez bidali diela aipatu.

? Aipatu al du Aingeruk ikaskideei ariketen erantzuna emailaz bidali duela? N

12d. Aingeruk ez du irakasleak ikaskideei ariketen erantzuna emailaren bidez bidali diela aipatu.

? Aipatu al du Aingeruk irakasleak ikaskideei ariketen erantzuna emailaz bidali diela? N

item 13

13a. Amuitzek ez dio emakumeari ospakizunerako gonbidapena hilabete hasieran bidali duela erran.

? Hilabete hasieran bidali al du Amuitzek gonbidapena? Y

13b. Amuitzek ez dio emakumeari enbaxadoreak ospakizunerako gonbidapena hilabete hasieran bidali duela erran.

? Hilabete hasieran bidali al du enbaxadoreak gonbidapena? Y

13c. Amuitzek ez du emakumeari ospakizunerako gonbidapena hilabete hasieran bidali diola erran.

? Hilabete hasieran bidali al du Amuitzek gonbidapena? Y

13d. Amuitzek ez du enbaxadoreak emakumeari ospakizunerako gonbidapena hilabete hasieran bidali diola erran.

? Hilabete hasieran bidali al du enbaxadoreak gonbidapena? Y

item 14

14a. Alaznek ez dio etorkinari ezinbesteko dokumentazioa hurrengo egunerako prestatu duela aipatu.

? Etorkinak prestatu al du dokumentazioa? N

14b. Alaznek ez dio etorkinari administrazioak ezinbesteko dokumentazioa hurrengo egunerako prestatu duela aipatu.

? Etorkinak prestatu al du dokumentazioa? N

14c. Alaznek ez du etorkinari ezinbesteko dokumentazioa hurrengo egunerako prestatu diola aipatu.

? Etorkinak prestatu al du dokumentazioa? N
14d. Alaznek ez du administrazioak etorkinari ezinbesteko dokumentazioa hurrengo egunerako prestatu diola aipatu.
? Etorkinak prestatu al du dokumentazioa? N

item 15

15a. Enarak ez die politikoei argazki konprometitua ezkutuko bileran atera duela esan.
? Enara izan al da politikoei argazki konprometitua ezkutuko bileran atera duela esan ez diena? Y
15b. Enarak ez die politikoei kazetariak argazki konprometitua ezkutuko bileran atera duela esan.
? Enara izan al da politikoei kazetariak argazki konprometitua ezkutuko bileran atera duela esan ez diena? Y
15c. Enarak ez du politikoei argazki konprometitua ezkutuko bileran atera diela esan.
? Enara izan al da politikoei argazki konprometitua ezkutuko bileran atera diela esan ez duena? Y
15d. Enarak ez du kazetariak politikoei argazki konprometitua ezkutuko bileran atera diela esan.
? Enara izan al da kazetariak politikoei argazki konprometitua ezkutuko bileran atera duela esan ez duena? Y

item 16

16a. Zigorrek ez dio aitonari kontzerturako txartela legez kontra saldu duela erran.
? Zinemarako txartela saldu al du Zigorrek? N
16b. Zigorrek ez dio aitonari musikariak kontzerturako txartela legez kontra saldu duela erran.
? Zinemarako txartela saldu al du musikariak? N
16c. Zigorrek ez du aitonari kontzerturako txartela legez kontra saldu diola erran.
? Zinemarako txartela saldu al du Zigorrek? N
16d. Zigorrek ez du musikariak aitonari kontzerturako txartela legez kontra saldu diola erran.
? Zinemarako txartela saldu al du musikariak? N

item 17

17a. Jokinek ez dio idazleari liburu berria argitaletxe famatura bidali duela esan.
? Berria al da argitaletxe famatura bidali izan den liburua? Y
17b. Jokinek ez dio idazleari merkataritza-agenteak liburu berria argitaletxe famatura bidali duela esan.
? Berria al da argitaletxe famatura bidali izan den liburua? Y
17c. Jokinek ez du idazleari liburu berria argitaletxe famatura bidali diola esan.
? Berria al da argitaletxe famatura bidali izan den liburua? Y
17d. Jokinek ez du merkataritza-agenteak idazleari liburu berria argitaletxe famatura bidali diola esan.
? Berria al da argitaletxe famatura bidali izan den liburua? Y

item 18

18a. Aintzane ez dio ahizpari matrikularako dirua atzerriko bankura bidali duela aipatu.
? Krediturako dirua izan da atzerriko bankura bidali izan dena? N
18b. Aintzane ez dio ahizpari gurasoek matrikularako dirua atzerriko bankura bidali dutela aipatu.
? Krediturako dirua izan da atzerriko bankura bidali izan dena? N
18c. Aintzane ez du ahizpari matrikularako dirua atzerriko bankura bidali diola aipatu.
? Krediturako dirua izan da atzerriko bankura bidali izan dena? N
18d. Aintzane ez du gurasoek ahizpari matrikularako dirua atzerriko bankura bidali diotela aipatu.
? Krediturako dirua izan da atzerriko bankura bidali izan dena? N

item 19

- 19a. Almikek ez dio jabeari hilabeteko errenta oso berandu ordaindu duela kontatu.
 ? Errenta izan al da Almikek sasoiz ordaindu ez duena? Y
- 19b. Almikek ez dio jabeari ikasleak hilabeteko errenta oso berandu ordaindu duela kontatu.
 ? Errenta izan al da ikasleak sasoiz ordaindu ez duena? Y
- 19c. Almikek ez du jabeari hilabeteko errenta oso berandu ordaindu diola kontatu.
 ? Errenta izan al da Almikek sasoiz ordaindu ez duena? Y
- 19d. Almikek ez du ikasleak jabeari hilabeteko errenta oso berandu ordaindu diola kontatu.
 ? Errenta izan al da ikasleak sasoiz ordaindu ez duena? Y

item 20

- 20a. Katalinek ez dio amonari lorategiko landarea oso gutxitan ureztatu duela azaldu.
 ? Sarritan ureztatu al du Katalinek amonaren landarea? N
- 20b. Katalinek ez dio amonari lorezainak lorategiko landarea oso gutxitan ureztatu duela azaldu.
 ? Sarritan ureztatu al du lorezainak amonaren landarea? N
- 20c. Katalinek ez du amonari lorategiko landarea oso gutxitan ureztatu diola azaldu.
 ? Sarritan ureztatu al du Katalinek amonaren landarea? N
- 20d. Katalinek ez du lorezainak amonari lorategiko landarea oso gutxitan ureztatu diola azaldu.
 ? Sarritan ureztatu al du lorezainak amonaren landarea? N

item 21

- 21a. Koldok ez dio ikasleari gomendiozko gutuna epemugatik kanpo bidali duela erran.
 ? Bidali al du Koldok gutuna epemugatik kanpo? Y
- 21b. Koldok ez dio ikasleari idazkariak gomendiozko gutuna epemugatik kanpo bidali duela erran.
 ? Bidali al du idazkariak gutuna epemugatik kanpo? Y
- 21c. Koldok ez du ikasleari gomendiozko gutuna epemugatik kanpo bidali diola erran.
 ? Bidali al du Koldok gutuna epemugatik kanpo? Y
- 21d. Koldok ez du idazkariak ikasleari gomendiozko gutuna epemugatik kanpo bidali diola erran.
 ? Bidali al du idazkariak gutuna epemugatik kanpo? Y

item 22

- 22a. Eneritzek ez dio emaztegaiari ezkontzarako lore-sorta modu desegokian ekarri duela aipatu.
 ? Emaztegaia izan al da ezkontzarako lore-sorta modu desegokian ekarri duena? N
- 22b. Eneritzek ez dio emaztegaiari umeak ezkontzarako lore-sorta modu desegokian ekarri duela aipatu.
 ? Emaztegaia izan al da ezkontzarako lore-sorta modu desegokian ekarri duena? N
- 22c. Eneritzek ez du emaztegaiari ezkontzarako lore-sorta modu desegokian ekarri diola aipatu.
 ? Emaztegaia izan al da ezkontzarako lore-sorta modu desegokian ekarri duena? N
- 22d. Eneritzek ez du umeak emaztegaiari ezkontzarako lore-sorta modu desegokian ekarri diola aipatu.
 ? Emaztegaia izan al da ezkontzarako lore-sorta modu desegokian ekarri duena? N

item 23

- 23a. Elixabetek ez die alabei Gabonetarako oparia Asiako bidaian erosi duela esan.
 ? Egia al da Elixabetek ez diela alabei esan oparia Asiarako bidaian erosi duela? Y
- 23b. Elixabetek ez die alabei aitak Gabonetarako oparia Asiako bidaian erosi duela esan.
 ? Egia al da Elixabetek ez diela alabei esan aitak oparia Asiarako bidaian erosi duela? Y
- 23c. Elixabetek ez du alabei Gabonetarako oparia Asiako bidaian erosi diela esan.
 ? Egia al da Elixabetek ez duela esan oparia Asiarako bidaian erosi duela? Y
- 23d. Elixabetek ez du aitak alabei Gabonetarako oparia Asiako bidaian erosi diela esan.
 ? Egia al da Elixabetek ez duela esan aitak oparia Asiarako bidaian erosi duela? Y

item 24

24a. Markelek ez dio epaileari hilketako froga epaiketa aurretik erakutsi duela kontatu.

? Kontatu al du Markelek froga epaiketa aurretik erakutsi duela? N

24b. Markelek ez dio epaileari abokatuak hilketako froga epaiketa aurretik erakutsi duela kontatu.

? Kontatu al du Markelek abokatuak froga epaiketa aurretik erakutsi duela? N

24c. Markelek ez du epaileari hilketako froga epaiketa aurretik erakutsi diola kontatu.

? Kontatu al du Markelek froga epaiketa aurretik erakutsi duela? N

24d. Markelek ez du abokatuak epaileari hilketako froga epaiketa aurretik erakutsi diola kontatu.

? Kontatu al du Markelek abokatuak froga epaiketa aurretik erakutsi duela? N

CHAPTER 6: CONCLUSIONS

The main goal of the work in this dissertation has been to look for pre-verbal structure building effects in languages with different configurations by means of using different pre-verbal cues in order to show that syntactic structure can be both interpreted and anticipated before the verbal head.

In Chapter 2, I have examined various kinds of left-peripheral phrases and I have argued that CLLDs and clitic-less left-dislocations have the same syntactic behavior. I concluded that the fact that CLLDs resemble unbounded dependencies that do not involve clitic pronouns at the bottom of the chain suggests that the particular nature of these constructions is not directly related to the clitic pronoun but to the referentiality of the fronted phrase.

In Chapter 3, I demonstrated that the active search mechanism is also triggered in long-distance dependencies involving clitic pronouns. I have shown that CLLD constructions in Spanish require the presence of a clitic pronoun to be interpreted and that the left-dislocation can already be interpreted at this clitic pronoun without requiring the confirmation from the verbal head. I concluded that the active search mechanism is a more general architectural mechanism of the parser that is triggered in all kinds of long-distance dependencies, regardless of whether the search is triggered for a gap or a pronoun.

In Chapter 4, I have confirmed that the active search mechanism for clitic pronouns is triggered in Galician, where the clitic requirement is stronger than in

Spanish. The experimental evidence shows that this active search for clitic pronouns is also apparent in other Romance languages besides Spanish, suggesting that it is a general strategy used when processing CLLD constructions. The findings in this chapter and in Chapter 3 support evidence reported on the literature on backward anaphora (van Gompel & Liversedge 2003, Sturt et al. 2004, Kazanina 2005, Kazanina et al. 2006, Aoshima et al. 2006), where the search for an antecedent is triggered when a pronoun is encountered in the sentence. Based on the results and results from studies on backwards anaphora, one can conclude that the active search mechanism is a general mechanism of the parser applicable to any kind of long-distance dependency and cross-linguistically attested.

In Chapter 5, I have shown that the use of indirect cues such as a mismatch between the morphology of an auxiliary and the following NP helps the parser in creating expectations about the upcoming input and interpreting material incrementally before the verb is processed. In addition, the experimental findings have shown that morphological mismatch can assist the parser in avoiding a garden-path. I argued that in order to account for the results a parser with a powerful projection is required.

Overall, the experimental findings in this work support incremental parsing models that try to connect every new word in the input to the existing syntactic structure and that use every available cue that the configuration of the language provides to interpret the material encountered in the input. In addition, I have shown that in order to do an incremental parsing of structures and in order to interpret arguments immediately, there is a need for a parser with a strong predictive

mechanism. This predictive mechanism should allow the parser to make inferences that go beyond the information provided by the lexical heads in the input assisted by the knowledge of the grammar of each specific language.

Future research will be required to test if clitic-less dislocations would behave similarly to CLLDs in parsing terms if it were the case that they are derived by the same syntactic constraints. If there were no on-line distinction between the dependencies that are interpreted through pronouns and dependencies that are interpreted through gaps, then one would expect to find no distinction with respect to how these dependencies are processed.

Furthermore, the nature of the trace involved in CLLD and clitic-less constructions will have to be specified since it does not exactly behave like wh-dependencies. As discussed in chapter 4, this would depend on the kind of chain and operator that is involved in CLLD and clitic-less dependencies.

Finally, in order to test the predictions of the parser suggested in chapter 5, I would like to further investigate the role that the rich agreement encoded in auxiliaries plays in Basque and extend the results to dependencies where there is a search initiated by an uninterpreted element.

Bibliography

- Abeillé, A., Godard, D., Miller, P. & Sag, I. (1998). "French bounded dependencies". In Balari, S. & Dini, L. (eds.), *Romance in HPSG*. CSLI Publications, Stanford.
- Abney, S. (1989). "A computational model of human parsing". *Journal of Psycholinguistic Research*, 18: 129-144.
- Abney, S. P. & Johnson, M. (1991). "Memory requirements and local ambiguities of parsing strategies". *Journal of Psycholinguistic Research* 20: 233-250.
- Agouraki, Y. (1992). "Clitic Left Dislocation and Clitic Doubling: a Unification". In *UCL Working Papers in Linguistics*, 4: 45-70.
- Alonso-Ovalle, L., Clifton, C., Fernandez-Solera S. & Frazier L. (2002a). "Null vs. Overt Pronouns and the Topic-Focus Articulation in Spanish". *Journal of Italian Linguistics*, 14: 2, 151-169.
- Altmann, G. & Steedman, M. (1988). "Interactions with context during human sentence processing". *Cognition*, 30: 191-238.
- Álvarez, R., Regueira, X.L. & Monteagudo, H. (1986). *Gramática Galega*. Galaxia, Vigo.
- Anagnostopoulou, E. (1997). "Clitic Left Dislocation and Contrastive Left Dislocation". In Anagnostopoulou, E., Van Riemsdijk, H. & Zwarts, F. (eds). *Materials on Left Dislocation*. John Benjamins, Philadelphia.
- Aoshima, S. (2003). *The Parsing and Grammar of Wh-dependencies*. PhD Dissertation, University of Maryland, College Park.
- Aoshima, S., Phillips, C. & Weinberg, A. (2003). "Theoretical implications of the parsing of Japanese wh-scrambling constructions". In Garding, G. & Tsujimura, M. (eds.), *Proceedings of the 22nd West Coast Conference on Formal Linguistics*, 29-42.
- Aoshima, S., Phillips, C. & Weinberg, A. (2004) "Processing long-distance dependencies in a head-final language". *Journal of Memory and Language*, 51:23-54.
- Aoshima, S., Yoshida, M. & Philips, C. (submitted). "Incremental processing of coreference and binding in Japanese".

- Aoun, J, Choueri, L. & Hornstein, N. (2001). "Resumption, Movement and Derivational Economy". *Linguistic Inquiry*, 32: 371-402.
- Aoun, J. & Li. Y. A. (2003). *Essays on the Representational and Derivational Nature of Grammar*. MIT Press, Cambridge.
- Bader, M. (1994) "Syntactic-function ambiguities". *Folia Linguistica*, 28:5-66.
- Bader, M. & Lasser, I. (1994). "German verb-final clauses and sentence processing: evidence for immediate attachment". In Clifton, C., Frazier, L. & Rayner, K. (eds.), *Perspectives on Sentence Processing*. LEA, NJ.
- Bader, M. & Meng, M. (1999). "Subject-object ambiguities in German embedded clauses: an across-the-board comparison". *Journal of Psycholinguistic Research*, 28:121-143.
- Balari, S. (1998). "Pronouns, variables, and extraction in HPSG". In Balari, S. & Dini, L. (eds.), *Romance in HPSG*. CSLI Publications, Stanford.
- Barber, H. & Carreiras, M. (2005). "Grammatical gender and number agreement in Spanish: an ERP comparison". *Journal of Cognitive Neuroscience*, 17:1, 137-153.
- Berwick, R. & Weinberg, A. (1984). *The Grammatical Basis of Linguistic Performance*. MIT Press, Cambridge.
- Betancort, M., Carreiras, M. & Acuña-Fariña, C. (2005). "Processing controlled PROs in Spanish". *Cognition*.
- Bever, T. G. & McElree, B. (1988). "Empty categories access their antecedents during comprehension". *Linguistic Inquiry*, 19: 35-45.
- Bever, T. G. & Sanz, M. (1997). "Empty categories access their antecedents during comprehension: unaccusatives in Spanish". *Linguistic Inquiry*, 28: 69-91.
- Bleam, T. (1996). "Clitic Left-Dislocation, Quantification, and Quirky A-bar Movement". Invited Talk, Universitat Autònoma de Barcelona, November 21, 1996.
- Bleam, T. (1999). *'Leista' Spanish and the Syntax of Clitic Doubling*. PhD Dissertation, University of Delaware.
- Boland, J.E., Tanenhaus, M.K., Garnsey, S.M. & Carlson, G.N. (1995). "Verb argument structure in parsing and interpretation: evidence from wh-questions". *Journal of Memory and Language*, 34:774-806.

- Bordelois, I. (1986). "Parasitic Gaps: Extensions of Restructuring". In Bordelois, I., Contreras, H. & Zagona, K. *Generative Studies in Spanish Syntax*. Foris Publications, Dordrecht.
- Campos, H. (1991). "Silent objects and subjects in Spanish". In Campos, H. & Martínez-Gil, F. (eds.), *Current Studies in Spanish Linguistics*. Georgetown University Press, Washington DC.
- Campos, H. & Zampini, M. (1990). "Focalization Strategies in Spanish". *Probus*, 2: 47-64.
- Carminatti, M. N. (2002). *The processing of Italian subject pronouns*. PhD Dissertation. University of Massachusetts, Amherst, MA.
- Casielles-Suarez, E. (2004). "The Syntax-Information Structure Interface: Evidence from Spanish and English". *Outstanding Dissertations in Linguistics*. Routledge, New York.
- Cecchetto, C. & Chierchia, G. (1999). "Connectivity in Dislocation Constructions and the Syntax/Semantics Interface". In Blake, S., Kim, E. & Shahin, K. (eds.). *Proceedings of WCCFL XVII*. CSLI publications, Stanford University.
- Cecchetto, C. (2000). "Syntactic or Semantic Reconstruction? Evidence from Pseudo-Clefts and Clitic Left-dislocation". In Cecchetto, C., Chierchia, G. & Guasti M.L. (eds.). *Semantic Interfaces*. CSLI publications, Stanford University.
- Cecchetto, C. (2000). "Doubling Structures and Reconstruction". *Probus*, 12:93-126.
- Chomsky, N. (1981). *Lectures on Government and Binding*. Foris Publications, Dordrecht.
- Chomsky, N. (1982). *Some Concepts and Consequences of the Theory of Government and Binding*. MIT Press, Cambridge.
- Chomsky, N. (1986). *Barriers*. MIT Press, Cambridge.
- Christianson, K., Hirotani, M. & Frazier, L. (2005). "Processing English left-dislocations and topicalizations". Poster presented at AMLaP 2005, Ghent, Belgium.
- Cinque, G. (1977). "The Movement Nature of Left Dislocation". *Linguistic Inquiry* 8: 397-411.
- Cinque, G. (1990). *Types of A'-Dependencies*. MIT Press, Cambridge.

- Clahsen, H. & Featherston, S. (1999). "Antecedent priming at trace positions: evidence from German scrambling". *Journal of Psycholinguistic Research* 28: 415-437.
- Clifton, C. & Frazier, L. (1989). "Comprehending sentences with long distance dependencies". In Carlson, G.N. & Tanenhaus, M. (eds.), *Linguistic Structure in Language Processing*, 273-317.
- Clifton, C. & De Vincenzi, M. (1990). "Comprehending sentences with empty elements". In Balota, D., Flores d'Arcais, G. B., & Rayner, K. (eds.), *Comprehension processes in reading*, Erlbaum, Hillsdale, NJ. 265- 284.
- Contreras, H. (1991). "On resumptive pronouns". In Campos, H. & Martínez-Gil, F. (eds.), *Current Studies in Spanish Linguistics*. Georgetown University Press, Washington DC.
- Costa, M.A. & Hub Faria, I. (2005). "Word order and focus effect on European Portuguese sentence processing". Poster presented at AMLaP 2005, Ghent, Belgium.
- Crain, S. & Fodor, J. (1985). "How can grammars help parsers?". In Dowty, D., Karttunen, L. & Zwicky, A. M. (eds.), *Natural Language Parsing. Psychological, Computational, and Theoretical Perspectives*. Cambridge University Press, Cambridge, 94-128.
- Crocker, M. (1994). "On the nature of the principle-based sentence processor". In Clifton, C., Frazier, L. & Rayner, K. (eds.). *Perspectives on sentence processing*. Erlbaum, Hillsdale, NJ, 245-266.
- Crocker, M. (1996). *Computational Psycholinguistics*. Kluwer, Dordrecht.
- Culicover, P.W. (2001) "Parasitic Gaps: a History". In Culicover, P.W. & Postal, P. (eds.), *Parasitic Gaps*. MIT Press, Cambridge.
- De Rijk, R. (1969). "Is Basque an S.O.V. language?". *Fontes Lingae Vascom*, 3: 219-351.
- De Vincenzi, M. (1991). *Parsing Strategies in Italian*. Kluwer, Dordrecht.
- Demestre, J., Meltzer, S., Garcia-Albea, J.E. & Vigil, A. (1999). "Identifying the null subject: evidence from event-related brain potentials". *Journal of Psycholinguistic Research*, 28:293-312.
- Dobrovie-Sorin, C. (1990). "Clitic Doubling, Wh-Movement, and Quantification in Romanian". *Linguistic Inquiry* 21: 351-397.

- Dobrovie-Sorin, C., Bleam, T. & Espinal, T. (in press). "Bare Nouns, number and types of incorporation". In Tasmowski, L. & Vogeleeer, S. (Eds.) *Non-definiteness and plurality*. John Benjamins: Amsterdam.
- Dominguez, A., Cuertos, F. & Segui, J. (1999). "The processing of grammatical gender and number in Spanish." *Journal of Psycholinguistic Research*, 28:485-498
- Fodor, J. (1978). "Parsing strategies and constraints on transformations". *Linguistic Inquiry*, 9:427-473.
- Franco, Jon. (1993). *On object agreement in Spanish*. PhD Dissertation, USC.
- Frazier, L. (1987). "Syntactic processing: evidence from Dutch". *Natural Language and Linguistic Theory*, 5:519-560.
- Frazier, L. & Clifton, C. (1989). "Successive cyclicity in the grammar and the parser". *Language and Cognitive Processes*, 4:93-126.
- Frazier, L. & Flores D'Arcais, G. B. (1989). "Filler driven parsing: a study of gap filling in Dutch". *Journal of Memory and Language*, 28:331-344.
- García-Mayo, P. (1992). "Pro as parasitic gap licenser". In Smith, E. & Zephir, F. (eds.). *Proceedings of the 1992 Mid-America Linguistics Conference and Conference on Siouan-Caddoan Languages*. University of Missouri, Columbia.
- García-Mayo, P. & Kempchinsky, P. (1994). "Finiteness in Romance vs. English Parasitic Gap Constructions". In Mazzolla, M. (ed.). *Issues and Theory in Romance Linguistics: Papers from the Linguistics Symposium in Romance Languages XXIII*. Georgetown University Press, Washington DC.
- Garnsey, S.M., Pearlmutter, N.J., Myers, E. & Lotocky, M.A. (1997). "The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences". *Journal of Memory and Language*, 37.
- Gazdar, G., Klein, E., Pullum, G. & Sag, I. (1985). *Generalized phrase Structure Grammar*, Basil Blackwell Publisher, Oxford.
- Gibson, E. (1991). *A computational theory of human linguistic processing: Memory limitations and processing breakdown*. PhD. Dissertation, Carnegie Mellon University.
- Gibson, E. & Hickok, G. (1993). "Sentence processing with empty categories". *Language and Cognitive Processes*, 8:147-161.
- Gibson, E., Hickok, G. & Schütze, C. (1994). "Processing empty categories: a parallel approach". *Journal of Psycholinguistic Research*, 23 (5):381-405.

- Gorrell, P. (1993). "Evaluating the direct association hypothesis: a reply to Pickering and Barry". *Language and Cognitive Processes*, 8:129-146.
- Gorrell, P. (1995). *Syntax and parsing*. Cambridge University Press, Cambridge.
- Gorrell, P. (1998). "Syntactic analysis and reanalysis in sentence processing". In Fodor, J.D. & Ferreira, F. (eds.), *Reanalysis in sentence processing*. Kluwer, Dordrecht. 201-245.
- Grimshaw, J. (1990). *Argument Structure*. MIT Press, Cambridge.
- Grohmann, K. K. (2000). *Prolific peripheries: a radical view from the left*. PhD Dissertation. University of Maryland, College Park.
- Hernanz, M. L. & Brucart, J.M. (1987). *La sintaxis*. Editorial Crítica, Barcelona.
- Heycock, C. & Kroch, A. (1999). "Pseudocleft connectedness: implications for the LF interface level". *Linguistic Inquiry* 30: 365-397.
- Hornstein, N., Lasnik, H. & Uriagereka, J. (2006) "The Dynamics of Islands: Speculations on the Item of Movement". Ms. University of Maryland, College Park.
- Hsu, C. N., Hurewitz, F. & Phillips, C. (2006). "Contextual and syntactic cues for head-final relative clauses in Chinese". The 19th CUNY Conference on Sentence Processing.
- Hsu, C. N. (2006). *Issues in Head-final relative clauses in Chinese: Derivation, Processing and Acquisition*. PhD Dissertation, University of Delaware.
- Hualde, J.I. & Ortiz de Urbina, J. (2003). *A Grammar of Basque*. Mouton de Gruyter, Berlin.
- Iatridou, S. (1995). "Clitics and Island Effects". *University of Pennsylvania Working Papers in Linguistics*, 2: 11-37.
- Inoue, A. (1991). *A comparative study of parsing in English and Japanese*. PhD Dissertation, University of Connecticut.
- Inoue, A. & Fodor, J. (1995). "Information-paced parsing of Japanese". In Mazuka, R. & Nagai, N. (eds.). *Japanese Sentence Processing*. Erlbaum, Hillsdale, NJ.
- Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness: Cognitive science series (6)*. Harvard University Press, Cambridge.

- Just, M., Carpenter, P. & Wooley, J. (1982). "Paradigms and processes in reading comprehension". *Journal of Experimental Psychology: General*, 111: 228-238.
- Kamide, Y. & Mitchell, D.C. (1992). "Incremental Pre-head attachment in Japanese Parsing". *Language and Cognitive Processes*, 14: 631-662.
- Kaplan, R.M., & Zaenen, A.E. (1989). "Long-distance dependencies, constituent structure, and functional uncertainty". In Baltin, M. & Kroch, A.S. (eds.), *Alternative concepts of phrase structure*. University of Chicago Press, Chicago. 17-42.
- Kato, M. & Raposo, E. (2005). *Topicalization in European and Brazilian Portuguese*. Talk presented at the University of Maryland Colloquium,
- Kayne, R. (1991). "Romance Clitics, Verb Movement, and PRO". *Linguistics Inquiry*, 22: 647-686.
- Kayne, R. (1994). *The Antisymmetry of Syntax*. MIT press, Cambridge.
- Kazanina, N. (2005) *The acquisition and processing of backwards anaphora*. PhD Dissertation, University of Maryland.
- Kazanina, N., Lau, E., Liberman, M., Yoshida, M. & Phillips, C. (submitted) "The Effect of Syntactic Constraints on the Processing of Backwards Anaphora".
- Lafitte, P. (1944). *Grammaire Basque*. Navarro Labourdin Littéraire.
- Laka, I. (1990). *Negation in Syntax: on the Nature of Functional Categories and Projections*. PhD dissertation, MIT.
- Lasnik, H. & Stowell, T. (1991). "Weakest Crossover". *Linguistic Inquiry*, 22: 687-720.
- Lee, M-W. (2004). "Another look at the role of empty categories in sentence processing (and grammar)". *Journal of Psycholinguistic Research*, 33: 51-73.
- Levin, B. (1993). *English Verb Classes and Alternations. A Preliminary Investigation*. The University of Chicago Press, Chicago.
- Lombardo, V. & Sturt, P. (2002) "Incrementality and Lexicalism". In Merlo, P. & Stevenson, S. (eds.), *The lexical basis of sentence processing*, John Benjamins, Amsterdam.
- MacDonald, M. (1989). "Priming effects from gaps to antecedents". *Language and Cognitive Processes*, 4 (1):35-56.

- MacDonald, M.C. (1994) "Probabilistic constraints and syntactic ambiguity resolution", *Language and Cognitive Processes*, 9:692-715
- MacDonald, M.C., Pearlmutter, N.J. & Seidenberg, M.S. (1994). "The Lexical Nature of Syntactic Ambiguity Resolution". *Psychological Review*, 101:4.
- Maia, M. (1996). "The Comprehension of Object Anaphora in Brazilian Portuguese". In Parodi, C., Quicoli, C., Saltarelli, M. & Zubizarreta, M.L. (eds.), *Aspects of Romance Linguistics: Selected papers from the Linguistics Symposium on Romance Languages XXIV March 10-13*. Georgetown University Press, Washington, DC, 93-311.
- Maia, M. (1997). "The Processing of Object Anaphora in Brazilian Portuguese". *Recherches Linguistiques de Vincennes*, 26:151-172.
- Manzini, M.R. (1998). "Syntactic Approaches to Cliticization". *Glott International*, 3:3-7.
- Marslen-Wilson, W. (1973). "Linguistic structure and speech shadowing at very short latencies". *Nature*, 244:522-533.
- Mazuka, R. & Itoh, K. (1995). "Can Japanese speakers be lead down the garden-path?" In Mazuka, R. & Nagai, N. (eds.), *Japanese Sentence Processing*. Erlbaum, Hillsdale, NJ.
- McKoon, G., Ratcliff, R. & Ward, G. (1994). "Testing theories of language processing: an empirical investigation of the on-line lexical decision task". *Journal of Experimental Psychology: Learning, Memory and Cognition*, 20:1219-1228.
- McKoon, G. & Ratcliff, R. (1994). "Sentential context and on-line lexical decision". *Journal of Experimental Psychology: Learning, Memory and Cognition*, 20:1239-1243.
- McMahon, E. (2002). "Clitic Left Dislocation in Modern Greek: an Argument for Movement". Ms. Northwestern University.
- McRae, K., Spivey-Knowlton, M.J., & Tanenhaus, M.K. (1998) "Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension". *Journal of Memory and Language*, 38: 283-312.
- Méndez Álvarez, M.J. (2004) *Ortografía da lingua galega: descripción, regras e ditados*. Xerais, Vigo.
- Meseguer, E., Acuña-Fariña, C. & Carreiras, M. (submitted). "Processing ambiguous Spanish se in a minimal chain: the Active Gap Strategy". Ms. Universidad de La Laguna.

- Miller, P. & Sag, I. (1997). "French clitic movement without clitics or movement". *Natural Language and Linguistic Theory*, 15:573-639.
- Miyamoto, E. & Takahashi, S. (2000). "The Processing of wh-phrases and Interrogative Complementizers in Japanese". In Akatuka, N. & Strauss, S. (eds). *Japanese Korean Linguistics*, 10:62-75.
- Miyamoto, E. & Takahashi, S. (2001). "A filled-gap effect without gaps in Japanese". The 14th Annual CUNY Conference on Human Sentence Processing.
- Miyamoto, E. (2002). "Case markers as clause boundary inducers in Japanese". *Journal of Psycholinguistic Research*, 31: 307-347.
- Miyamoto, E. (2003). "Reanalysis and clause boundaries in Japanese as a constraint-driven process". *Language and Speech*, 46(1): 23-52.
- Miyamoto, E. & Nakamura, M. (2003). "Subject/ object asymmetries in the processing of relative clauses in Japanese". In Garding, G. & Tsujimura, M. (eds.), *WCCFL 22 Proceedings*. Cascadilla Press, Somerville, MA, 342-355.
- Miyamoto, E.T., & Takahashi, S. (2003). "Typing Mismatch Effects in the Processing of wh-phrases in Japanese". Ms. Submitted for publication.
- Monachesi, P. (1998). "Decomposing Italian clitics". In Balari, S. & Dini, L. (eds.), *Romance in HPSG*. CSLI Publications, Stanford.
- Mulders, I. (2003). Transparent Parsing: head-driven processing of verb-final structures. PhD Dissertation, Utrecht Universiteit: LOT.
- Nakano, Y., Felser, C. & Clahsen, H. (2002). "Antecedent priming at trace positions in Japanese long-distance scrambling". *Journal of Psycholinguistic Research*, 31:531-571.
- Nicol, J. & Swinney, D. (1989). "The role of structure in coreference assignment during sentence comprehension". *Journal of Psycholinguistic Research*, 8(1):5-19.
- Nicol, J. L. (1993). "Reconsidering reactivation". In Shillcock, G.A.R. (eds), *Cognitive Models of Speech Processing: The second Sperlonga Meeting*. Erlbaum, NJ. 321-347.
- Nicol, J. L., Fodor, J. & Swinney, D. (1994). "Using cross-modal lexical decision tasks to investigate sentence processing". *Journal of Experimental Psychology: Learning, Memory and Cognition*, 20:1229-1238.

- Olarrea, A. (1996). *Pre and Postverbal subject positions in Spanish: A minimalist account*. Unpublished PhD. Dissertation, University of Washington.
- Ortiz de Urbina, J. (1989). *Parameters in the Grammar of Basque*. Foris Publications, Dordrecht.
- Pablos, L. (2003). *Processing long-distance dependencies involving clitic pronouns in Spanish*, Ms. University of Maryland, College Park.
- Phillips, C. (1996). *Order and Structure*. PhD Dissertation, MIT.
- Phillips, C. (2003). "Linguistics and linking problems". In Rice, M. & Warren, S. (eds.), *Developmental Language Disorders: from Phenotypes to Etiologies*. Erlbaum, Mahwah.
- Phillips, C., Kazanina, N., & Abada, S.H. (2005). ERP effects on the processing of syntactic long-distance dependencies. *Cognitive Brain Research*, 22: 407-428.
- Pickering, M. & Barry, G. (1991). "Sentence processing without empty categories". *Language and Cognitive Processes*, 6:229-259.
- Pickering, M. (1993). "Direct association and sentence processing: a reply to Gorrell and to Gibson and Hickok". *Language and Cognitive Processes*, 8:163-196.
- Pollard, C. & Sag, I. A. (1994). *Head-Driven Phrase Structure Grammar*. University of Chicago Press, Chicago and CSLI Publications, Stanford.
- Prichett, B. (1991) "Head position and parsing ambiguity". *Journal of Psycholinguistic Research*, 20: 251-270.
- Pritchett, B. (1992) *Grammatical competence and parsing performance*. The University of Chicago Press, Chicago.
- Raposo, E. & Uriagereka, J. (2005). "Clitic Placement and FP Projection in European Portuguese". In Cinque, G. & Kayne, R. (eds.). *The Oxford Handbook of Comparative Syntax*. Oxford University Press, Oxford.
- Resnik, P. (1992). *Left-corner parsing and psychological plausibility*. Ms. University of Pennsylvania.
- Rivero, M.L. (1980). "On left-dislocation and Topicalization in Spanish". *Linguistic Inquiry*, 11: 363-393.
- Rizzi, L. (1982). *Issues in Italian Syntax*. Foris Publications, Dordrecht.

- Rizzi, L. (1986). "Null objects in Italian and the theory of pro". *Linguistic Inquiry*, 17: 501-557.
- Rizzi, L. (1990). *Relativized Minimality*. MIT Press, Cambridge.
- Rodman, R. (1974). "On Left-Dislocation". *Papers in Linguistics*, 7: 437-466.
- Rohde, D. (2001-2003). *Linger: A flexible platform for language processing experiments*, MIT.
- Ross, J.R. (1967). *Constraints on Variables in Syntax*. Ph.D. Dissertation, MIT.
- Sag, I. A. & Fodor, J. (1994). "Extraction without traces". *Proceedings of the West Coast Conference on Formal Linguistics*, 13: 365-384.
- Schneider, D. A. (1999). *Parsing and incrementality*. PhD Dissertation, University of Delaware.
- Schneider, D. A. & Phillips, C. (2001). "Grammatical Search and Reanalysis". *Journal of memory and Language*, 45: 308-336.
- Spivey-Knowlton, M., Trueswell, J., & Tanenhaus, M. (1993). "Context effects in syntactic ambiguity resolution: discourse and semantic influences in parsing reduced relative clauses". *Canadian Journal of Experimental Psychology: Special Issue*, 47: 276-309.
- Spivey-Knowlton, M. & Tanenhaus, M. (1994). "Referential context and syntactic ambiguity resolution". In Clifton, C., Frazier, L. & Rayner, K. (eds.), *Perspectives on sentence processing*, Erlbaum, Hillsdale, NJ.
- Spivey-Knowlton, M. & Sedivy, J.C. (1995). "Resolving attachment ambiguities with multiple constraints", *Cognition*, 55: 227-267.
- Sportiche, D. (1993). *Clitic Constructions*, Ms., UCLA.
- Stabler, E.P. (1994). "The finite connectivity of linguistic structure". In Clifton, C., Frazier, L. & Rayner, K. (eds.), *Perspectives on sentence processing*, Erlbaum, Hillsdale, NJ. 303-336.
- Steedman, M. (1996). *Surface structure and interpretation: Linguistic inquiry monographs*, 30. MIT Press, Cambridge.
- Steedman, M. (2000). *The Syntactic Process*. MIT Press, Cambridge.
- Stevenson, S. (1993). "Establishing long-distance dependencies in a hybrid network model of human parsing". In Cognitive Science Society (eds.). *Proceedings of the*

- 15th Annual Conference of the Cognitive Science Society. Erlbaum, Hillsdale, NJ, 982-987.
- Stevenson, S. (1994). *A competitive attachment model for resolving syntactic ambiguities in natural language parsing*. PhD Dissertation. University of Maryland.
- Stowe, L. (1986). "Parsing wh-constructions: evidence for on-line gap location". *Language and Cognitive Processes* 1:227-245.
- Sturt, P. & Crocker, M. (1996). "Monotonic syntactic processing: a cross-linguistic study of attachment and reanalysis". *Language and Cognitive Processes* 11: 449-494.
- Sturt, P. & Pickering, M. (1999). "Structural Change and Reanalysis Difficulty in Language Comprehension". *Journal of Memory and Language* 40: 136-150.
- Sturt, P. (2003). "The time-course of the application of binding constraints in reference resolution". *Journal of Memory and Language* 48: 542-562.
- Sturt, P., Lombardo, V. & Betancort, M. (2004). "Evidence against (even mild) delay models of structure building". AMLaP Conference, Aix-en-Provence.
- Suñer, M. (1988). "The Role of Agreement in Clitic Doubled Constructions". *Natural Language and Linguistic Theory* 6: 391-434.
- Suñer, M. & Yépez, M. (1988). "Null Definite Objects in Quiteño". *Linguistic Inquiry* 19: 511-519
- Suñer, M. (1994). "V-movement and the licensing of argumental wh-phrases in Spanish". *Natural Language and Linguistic Theory*, 12: 335-372.
- Suñer, M. (1999). "Clitic Doubling of Strong Pronouns in Spanish: an Instance of Object shift". In Franco, J., Landa, A. & Martín, J. *Grammatical Analyses in Basque and Romance Linguistics*. John Benjamins, Amsterdam.
- Tabossi, P., Spivey-Knowlton, M., McRae, K. & Tanenhaus, M. (1994). "Semantic effects on syntactic ambiguity resolution: evidence for a constraint based resolution process". In Umiltà, C. & Moscovitch, M. (eds). *Attention and Performance XV*. Erlbaum, Hillsdale, NJ.
- Taraban, R. & McClelland, J. (1988). "Constituent attachment and thematic role expectations". *Journal of Memory and Language*, 27:597-632.
- Taraban, R. & McClelland, J. (1990). "Sentence comprehension: a multiple constraints view". In Balota, D., Rayner, K. & Flores d'Arcais, G. (eds.), *Comprehension processes in reading*. Erlbaum, Hillsdale, NJ.

- Tellier, C. (1991) *Licensing theory and French parasitic gaps*. Kluwer, Dordrecht.
- Torrego, E. (1983) "More Effects of Successive Cyclic Movement", *Linguistic Inquiry*, 14: 561-565.
- Torrego, E. (1984) "On Inversion in Spanish and some of its Effects". *Linguistics Inquiry* 15: 103-129.
- Torrego, E. (1995) "On the Nature of Clitic Doubling". In Campos, H. & Kempchinsky, P (eds.). *Evolution and Revolution in Linguistic Theory*. Georgetown University Press, Washington DC.
- Torrego, E. (1998). *The Dependencies of Objects – Linguistic Inquiry Monograph: 34*. MIT Press, Cambridge.
- Traxler, M., & Pickering, M. (1996). "Plausibility and the processing of unbounded dependencies: an eye-tracking study". *Journal of Memory and Language*, 35: 454-475.
- Trueswell, J.C., Tanenhaus, M.K. & Kello, C. (1993). "Verb-specific constraints in sentence processing: separating effects of lexical preference from garden-paths". *Journal of Psycholinguistic Research*.
- Trueswell, J. & Tanenhaus, M. (1994). "Toward a lexicalist approach to syntactic ambiguity resolution". In Clifton, C., Frazier, L., & Rayner, K. (eds.), *Perspectives on sentence processing*. Erlbaum, Hillsdale, NJ.
- Uriagereka, J. (1995) "Aspects of the Syntax of Clitic Placement in Western Romance". *Linguistic Inquiry* 26: 79-123.
- Uriagereka, J. (1995b) "An F Position in Western Romance". In Kiss, K. (Ed.) *Discourse Configurational Languages*. Oxford University Press, Oxford.
- Uriagereka, J. (1999) *On the Syntax of Doubling*. Ms. University of Maryland.
- Uriagereka, J. (2000) "Doubling and Possession". In Gerlach, B. & Grizenhout, J. (eds), *Clitics in Phonology, Morphology and Syntax*. John Benjamins, Amsterdam.
- Uriagereka, J. (2000b) "Multiple Spell-Out". In Epstein, S. & Hornstein, N. (eds.), *Working Minimalism*. MIT Press, Cambridge.
- Uribe-Etxebarria, M. (1994). *Interface Licensing Conditions on Negative Polarity Items: a Theory of Polarity and Tense Interactions*. PhD Dissertation, University of Connecticut.

- Vallduví, E. (1988) "Functional load, prosody, and syntax: left-detachment in Catalan and Spanish." In MacLeod, L., et al. (eds.), *Papers from the 24th Annual Regional Meeting of the Chicago Linguistic Society*, Chicago, 391-403
- Vallduví, E. (1990) *The Information Component*. PhD. Dissertation, University of Pennsylvania.
- Van, J. (1997) "Left-dislocation, Connectedness and Reconstruction". In Anagnostopoulou, E., Van Riemsdijk, H. & Zwarts, F. (eds), *Materials on Left Dislocation*. John Benjamins, Philadelphia.
- Van Gompel, R.P.G. & Liversedge, S.P. (2003) "The influence of morphological information on cataphoric pronoun assignment". *Journal of Experimental Psychology: Learning, Memory and Cognition*, 29:128-139.
- Villasante, L. (1980) *Sintaxis de la oración simple*. Editorial Franciscana, Oñate.
- Xunta de Galicia (1999) *O Galego, Lingua Propia de Galicia*, Comisión Coordinadora para a normalización lingüística.
- Yamashita, H. (1997). "The effects of word-order and case marking information on the processing of Japanese". *Journal of Psycholinguistic Research*. 26: 163-188.
- Yoshida, M., Aoshima, S. & Phillips, C. (2004) "Relative Clause Prediction in Japanese". The 17th CUNY Conference on Human Sentence Processing
- Yoshida, M. (2006). "Relative Clause Prediction in Japanese". Ms. University of Maryland.
- Yoshida, M. (2006). *Mechanisms and Constraints in Long-Distance Dependency Formation*. PhD Dissertation, University of Maryland, College Park.
- Weinberg, A. (1993) "Parameters in the theory of sentence processing: minimal commitment theory goes East". *Journal of Psycholinguistic Research* 22: 339-364.
- Zagona, K. (2000). *The Syntax of Spanish*. Cambridge University Press, Cambridge.
- Zubizarreta, M.L. (1994). "The Grammatical Representation of Topic and Focus: implications for the structure of the clause". *University of Venice Working papers in Linguistics*, 4.