Straight to the Subject. A Unified Analysis of Preverbal and Postverbal Subjects in Broad Focus Constructions in Spanish

by

Yadira Álvarez López

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Department of Spanish and Portuguese
University of Toronto

© Copyright by Yadira Álvarez López 2019
Abstract

This study re-examines subjects, subject distribution and subject properties in Spanish. It includes an experimental prosodic component that brings new empirical evidence to the theoretical discussion. Theoretically, it adopts a composite understanding of subjecthood (Harley, 1995; McCloskey, 1997), and incorporates the notion of argument structure. The result is a less rigid approach to the notion of default word order, which directly contributes to the possibility of a unified analysis of broad focus (BF) structures. Two options of subject realization, the inflectional layer and the left periphery, were assessed with respect to their prosodic features. The results of the analysis showed that there are differences in duration and fundamental frequency (F0) (for some speakers) between subjects that are old information and appear in the left periphery, and BF subjects. Taking into account these findings, differences in subject distribution in BF contexts, and the event-related interpretation of BF contexts, a new analysis—based upon Borer’s (2005b) Event Phrase—is proposed. In this analysis specEP is the locus of preverbal subject realization, and the need to license the event by binding the event variable is the trigger of the movement to this preverbal position. Since specEP can be filled by non-nominative
elements, this analysis can accommodate different word orders, such as those with a preverbal locative or a dative DP and a postverbal nominative subject. Additionally, this analysis also sheds light on the position that postverbal nominative subjects occupy in these structures. This study also examined the syntactic operation of agreement in contexts that have been claimed to display variable number agreement, and demonstrated that such variability is only apparent, as each instance corresponds to a different underlying structure. Finally, issues related to the syntax-phonology interface are discussed in the light of the experimental findings, specifically the type of information that should be available at the interface, the type of syntactic information that is prosodically represented, and the mapping between syntax and prosody.
Acknowledgments

Carrying out this work was only possible due to the support, help, and encouragement of many people, and some very special four-legged critters that made my life better during the PhD. My most profound gratitude—one beyond words—goes to all of them. (What follows is an attempt to phrase this feeling.)

Special thanks go to the members of my committee. To my supervisor, María Cristina Cuervo, I am indebted for the time we spent together discussing and shaping the ideas presented in this work. Her comments and questions—more than one can possibly imagine—directly contributed to improve this work. I could not have asked for a better interlocutor. I am also thankful for Cristina’s continuous effort to make my writing more transparent. Above all, I am grateful to Cristina for teaching me how to be a better researcher, for giving me the space to disagree and, at the same time, for asking me for better arguments to support my ideas. To Laura Colantoni, I am truly thankful for providing me with the academic support that I needed to undertake the experimental component of this work. Coming from “the other side” (theoretical syntax), this was a major step, one that without her help and guidance would have been impossible. Laura’s moral support, kindness, and friendship were also crucial; for those I will always be grateful too. To Ana Teresa Pérez-Leroux, who has been there for me at every step of this process, goes my warmest gratitude. I feel extremely fortunate for having her unreserved support and constant encouragement, and for giving me all the time I asked for to talk about this and other topics. Finally, I thank Ana for caring so much about me, and for helping me when I needed it the most.

I would also like to thank Olga Fernández-Soriano, whose research has been a source of inspiration for me, for agreeing to be the external appraiser, and for her encouraging comments about this work. Equally, I want to thank Arsalan Kahnemuyipour for his careful reading of this dissertation, valuable comments, encouragement, and thought-provoking questions. It was a privilege to have them both in my committee.

I would not have been able to do this work without my husband’s help and support. He is behind all the Praat scripts, and the wonderful R scripts that allowed me to run the stats
very easily. He was also the first one to hear my ideas about structures, variables, scope, and generic and existential interpretations. Because of his support, help, kindness, understanding, and huge amount of patience, I was able to get to this point.

I owe thanks to Professor Yolanda Iglesias, and to my sister and my brother-in-law, who kindly recorded all the stimuli for the two versions of the experiment. To my participants, I will always be grateful for their time and voices. I would also like to thank Sara Solís, our Graduate Administrator, for all her help, moral support, and kindness during the past years. My deepest gratitude goes to Professor Elizabeth Cowper, who gave me the tools, and the support and encouragement to do generative syntax at the early stages of the PhD. To my professors in Havana, thank you for opening the door into this world of linguistics. To Professor José Gomáriz, thank you for encouraging me to apply to graduate school.

I am extremely grateful to my friends for their support and help. Specially, I would like to thank Tanya Battersby for her friendship, encouragement, and our very long conversations during all these years. I am also very thankful to Tanya for kindly proofreading several parts of this manuscript. (All errors are mine.) Thanks a lot to Matt Patience and Olivia Marasco for their friendship, and for discussing with me some of my experimental results. Additionally, I would like to thank Matt for giving English-sounding titles to my tables and figures. To Anna Limanni, Yasaman Rafat, Natalia Mazzaro, Nelson Dario González, Isabel Cascante, and Aga Bijos, thank you for being there for me all these years. To Gaby Klassen, thank you for your kindness. To Juan, Cris, and Dania, thank you for believing in me and for your unlimited support.

To my furry friends: Pancho, Stella, Poppy, Mars, Kali, Oli, and Moosh, goes my fondest gratitude, because life is better with barks and meows!

Finally, my heartfelt thanks go to my family for their encouragement to pursue the path of learning and their moral support: to my parents for showing me since I was very young that knowledge is freedom, to my grandmother, who is 91 years old and has always valued education more than anyone I’ve ever met, to my sister for caring so much and checking on me every day. I dedicate this work to Ernesto and Moosh, with love.
Table of Contents

Acknowledgments ........................................................................................................................................ iv

Table of Contents ....................................................................................................................................... vi

List of Tables .............................................................................................................................................. ix

List of Figures .......................................................................................................................................... xii

List of Appendices .................................................................................................................................. xiv

Chapter 1 Introduction ............................................................................................................................... 1

Chapter 2 Subject matters: subject category and subject placement .............................................. 6
  2.1. Subjects, their properties and representation in the grammar .............................................. 6
      2.1.1. Subjects and subject properties ............................................................................. 6
  2.2. Towards a subject definition ................................................................................................. 14
      2.2.1. Subjects as a derived notion, and subject positions. .................................... 15
  2.3. Factors that determine subject placement in languages with variable word order 17
      2.3.1. Information structure ....................................................................................... 17
      2.3.2. Argument structure ......................................................................................... 20
          2.3.2.1. SV(O), VS, and DatVS ...................................................................... 21
  2.4. Final remarks ......................................................................................................................... 23

Chapter 3 Previous studies on Spanish preverbal subjects .......................................................... 25
  3.1. Spanish preverbal subjects: different analyses, different structures, same language? ...................................................... 26
  3.2. Spanish preverbal subjects as LD elements ........................................................................... 27
      3.2.1. LD proposals: general claims ............................................................................ 28
      3.2.2. Syntactic evidence of the LD status of preverbal subjects ............................. 29
          3.2.2.1. Distributional evidence ......................................................................... 29
      3.2.2. Interpretational effects: scope, quantifier interpretation, and binding 32
      3.2.3. LD proposals: similarities, differences, and general overview ..................... 33
      3.2.4. Examining some of the evidence of the LD proposal .................................... 36
          3.2.4.1. Distributional facts: on A&A’s subject-verb placement facts ............. 36
          3.2.4.2. O&T (1999)’s ellipsis contexts: What do they really mean? ............ 39
  3.3. Spanish preverbal subjects as elements of the IP layer .................................................................. 48
      3.3.1. IP proposal: general theoretical claims .............................................................. 48
      3.3.2. IP proposal: syntactic evidence ............................................................................ 48
          3.3.2.1. Subject-adverb-verb distribution ............................................................. 48
          3.3.2.2. Scope ........................................................................................................ 49
          3.3.2.3. Binding ................................................................................................. 51
          3.3.2.4. Agreement ............................................................................................ 52
          3.3.2.5. Bare facts .............................................................................................. 55
          3.3.2.6. Subject positions .................................................................................. 57
      3.3.3. IP proposal: general comments .......................................................................... 62
  3.4. Spanish preverbal subjects: a sentential analysis ......................................................................... 63
      3.4.1. Subjects of predication: syntactic evidence ....................................................... 64
      3.4.2. General comments about the sentential analysis of preverbal subjects .... 69
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.5. Final remarks</td>
<td>70</td>
</tr>
<tr>
<td>Chapter 4 When prosody meets syntax: on the syntax-phonology interface and previous research on Spanish prosody</td>
<td>74</td>
</tr>
<tr>
<td>4.1. Theoretical background</td>
<td>75</td>
</tr>
<tr>
<td>4.2. Syntactic proposals of sentential stress</td>
<td>79</td>
</tr>
<tr>
<td>4.3. Information structure and sentential stress</td>
<td>87</td>
</tr>
<tr>
<td>4.4. Syntax-phonology interface and models of sentential stress assignment</td>
<td>94</td>
</tr>
<tr>
<td>4.5. Previous studies on Spanish prosody</td>
<td>97</td>
</tr>
<tr>
<td>4.5.1. Prosodic correlates of focus and topic</td>
<td>98</td>
</tr>
<tr>
<td>4.5.2. The phonology of focus marking</td>
<td>102</td>
</tr>
<tr>
<td>4.5.3. Phrasing</td>
<td>103</td>
</tr>
<tr>
<td>4.5.3.1. Phrasing patterns in Spanish</td>
<td>105</td>
</tr>
<tr>
<td>4.6. Prosodic features of broad focus, focalized and topicalized elements in Spanish</td>
<td>107</td>
</tr>
<tr>
<td>Chapter 5 The sound of subjects: a prosodic analysis of Spanish preverbal subjects</td>
<td>110</td>
</tr>
<tr>
<td>5.1. Study and Hypotheses</td>
<td>111</td>
</tr>
<tr>
<td>5.2. Experimental Design</td>
<td>119</td>
</tr>
<tr>
<td>5.2.1. Sentence design: syntax</td>
<td>120</td>
</tr>
<tr>
<td>5.2.2. Sentence design: information structure</td>
<td>121</td>
</tr>
<tr>
<td>5.2.3. Sentence design: prosody</td>
<td>124</td>
</tr>
<tr>
<td>5.3. Task</td>
<td>125</td>
</tr>
<tr>
<td>5.4. Dialects, participants, testing and data collection</td>
<td>127</td>
</tr>
<tr>
<td>5.4.1. Dialects</td>
<td>127</td>
</tr>
<tr>
<td>5.4.2. Participants</td>
<td>128</td>
</tr>
<tr>
<td>5.4.3. Testing</td>
<td>129</td>
</tr>
<tr>
<td>5.4.4. Data collection</td>
<td>129</td>
</tr>
<tr>
<td>5.5. Data Processing and Analysis</td>
<td>130</td>
</tr>
<tr>
<td>5.5.1. Data Processing</td>
<td>130</td>
</tr>
<tr>
<td>5.5.2. Analysis. Acoustic parameters (measurements and statistical analysis). Phrasing patterns and pitch accents</td>
<td>131</td>
</tr>
<tr>
<td>5.6. Results</td>
<td>134</td>
</tr>
<tr>
<td>5.6.1. Preverbal subjects of transitive verbs</td>
<td>134</td>
</tr>
<tr>
<td>5.6.2. Preverbal non-nominative subjects</td>
<td>144</td>
</tr>
<tr>
<td>5.6.3. Preverbal subjects of unaccusative and unergative verbs</td>
<td>153</td>
</tr>
<tr>
<td>5.6.4. Phrasing patterns and pitch accents</td>
<td>159</td>
</tr>
<tr>
<td>5.6.4.1. Phrasing patterns</td>
<td>159</td>
</tr>
<tr>
<td>5.6.4.2. Peak alignment and pitch accents</td>
<td>168</td>
</tr>
<tr>
<td>5.7. Discussion and conclusions</td>
<td>168</td>
</tr>
<tr>
<td>5.7.1. Prosodic marking</td>
<td>172</td>
</tr>
<tr>
<td>5.7.2. On phrasing and phrasing patterns</td>
<td>174</td>
</tr>
<tr>
<td>5.7.3. On the syntax-phonology interface</td>
<td>174</td>
</tr>
<tr>
<td>Chapter 6 An alternative view on Spanish subjects</td>
<td>179</td>
</tr>
<tr>
<td>6.1. Data and current proposals</td>
<td>179</td>
</tr>
<tr>
<td>6.1.1. Preverbal subjects’ locus and their motivation</td>
<td>181</td>
</tr>
<tr>
<td>6.1.2. Against a predication requirement as movement trigger and a semantic reinterpretation of the EPP in Spanish</td>
<td>185</td>
</tr>
</tbody>
</table>
6.1.2.1. When broad focus statements are not interpreted about the preverbal (overt or covert) element ................................................................. 186
6.1.2.2. Analysis of the predication requirement........................................ 190
6.3. EP and Spanish data............................................................................... 204
  6.3.1. SV(O)............................................................................................... 204
  6.3.2. LocVS............................................................................................. 205
  6.3.3. VS211............................................................................................ 213
  6.3.4. DatVS............................................................................................ 213
    6.3.4.1. A parenthesis on states ................................................................. 216
    6.3.4.2. More on the properties of states .................................................. 218
  6.3.5. Some thoughts on generic and existential XPs in Spanish, and their
distribution .............................................................................................. 227
6.4. Summary ............................................................................................... 230

Chapter 7 On verbal agreement: subject positions and variable agreement. In Spanish? ................................................................. 232
  7.1. Agreement data: preliminarily interpretation and general discussion ........ 234
  7.2. Analysis of the different agreement patterns .......................................... 240
    7.2.1. Agreement and postverbal subjects ................................................ 240
    7.2.2. Ad sensum agreement ................................................................. 247
    7.2.3. SE constructions .......................................................................... 254
  7.3. Final comments .................................................................................... 262

Chapter 8 Conclusion.................................................................................. 263
References.................................................................................................... 269
Appendix A.................................................................................................. 279
List of Tables

Table 2.1. Patterns of subject movement in English, Icelandic, and Spanish

Table 4.1. Prosodic properties of broad, narrow and contrastive focused elements in Spanish

Table 5.1. Differences between sentence topics and background topics (Casielles-Suárez, 2004)

Table 5.2. Types of preverbal DPs, syntactic and information structure contexts

Table 5.3. Types of preverbal DPs, syntactic context, information structure and quantity of tokens

Table 5.4. Ratio paired t-tests results comparing the duration of preverbal subjects in broad focus and topicalized contexts

Table 5.5. Wilcoxon signed-rank tests results comparing the duration of preverbal subjects in broad focus and topicalized contexts

Table 5.6. Ratio paired t-test results comparing the duration of broad focus preverbal subjects and topicalized CLLDs

Table 5.7. Unpaired t-test results comparing the tonal change of broad focus preverbal subjects and topicalized CLLDs

Table 5.8. Wilcoxon rank-sum tests results comparing the tonal change of broad focus preverbal subjects and topicalized CLLDs

Table 5.9. Paired t-tests results comparing the tonal change of preverbal subjects in broad focus and topicalized contexts

Table 5.10. Wilcoxon signed-rank tests results comparing the tonal change of preverbal subjects in broad focus and topicalized contexts
Table 5.11. Paired t-tests results comparing the peak intensity of preverbal subjects in broad focus and topicalized contexts

Table 5.12. Wilcoxon signed-rank tests results comparing the peak intensity of preverbal subjects in broad focus and topicalized contexts

Table 5.13. Unpaired t-tests results comparing the peak intensity of broad focus preverbal subjects and topicalized CLLDs

Table 5.14. Wilcoxon rank-sum tests results comparing the peak intensity of broad focus preverbal subjects and topicalized CLLDs

Table 5.15. Ratio paired t-test results comparing the duration of non-nominative subjects in broad focus and topicalized contexts

Table 5.16. Wilcoxon signed-rank tests comparing the duration of non-nominative subjects in broad focus and topicalized contexts

Table 5.17. Ratio paired t-tests results comparing the duration of broad focus non-nominative subjects and topicalized CLLDs

Table. 5.18. Wilcoxon signed-rank tests results comparing the duration of broad focus non-nominative subjects and topicalized CLLDs

Table 5.19. Paired t-tests results comparing the tonal change of non-nominative subjects in broad focus and topicalized contexts

Table 5.20. Paired t-tests results comparing the tonal change of broad focus non-nominative subjects and topicalized CLLDs

Table 5.21. Wilcoxon signed-rank tests results comparing the tonal change of broad focus non-nominative subjects and topicalized CLLDs

Table 5.22 Paired t-tests results comparing the peak intensity of non-nominative subjects in broad focus and topicalized contexts
Table 5.23 Wilcoxon signed-rank tests results comparing the peak intensity of non-nominative subjects in broad focus and topicalized contexts

Table 5.24 Paired t-tests results comparing the peak intensity of broad focus non-nominative subjects and topicalized CLLDs

Table 5.25. Wilcoxon signed-rank tests results comparing the duration of broad focus preverbal subjects of unergative and unaccusative verbs

Table 5.26. Paired t-tests results comparing the tonal change of broad focus preverbal subjects of unergative and unaccusative verbs

Table 5.27. Wilcoxon signed-rank tests results comparing the tonal change of broad focus preverbal subjects of unergative and unaccusative verbs

Table 5.28. Paired t-tests results comparing the peak intensity of broad focus preverbal subjects of unergative and unaccusative verbs

Table 5.29. Wilcoxon signed-rank tests results comparing the peak intensity of broad focus preverbal subjects of unergative and unaccusative verbs

Table 5.30. Percentages of preverbal DPs phrased as IPs, by information structure context and dialect

Table 5.31. Percentages of DPs phrased as IPs in each syntactic and informational context, by dialect

Table 5.32. Percentages of preverbal DPs phrased as IPs by speaker, syntactic context, and informational context, in the Madrid variety

Table 5.33. Percentages of the different boundary cues used, by dialect

Table 5.34. Percentages of pauses used, by individual speakers
List of Figures

Figure 5.1. Sample Power Point slide used in the experiment

Figure 5.2. TextGrid of “Un hombre revende dulces” produced by a Madrid speaker

Figure 5.3. Scatterplots displaying the paired duration data of preverbal subjects in broad focus and topicalized contexts, by three speakers of each dialect

Figure 5.4. Scatterplots displaying the paired duration data of broad focus preverbal subjects and topicalized CLLDs, by Havana speakers

Figure 5.5. Scatterplots displaying the paired duration data of broad focus preverbal subjects and topicalized CLLDs, by Madrid speakers

Figure 5.6. Box plots and confidence intervals displaying the tonal change of topicalized subjects, topicalized CLLDs, and broad focus subjects, produced by Havana speakers

Figure 5.7. Box plots and confidence intervals displaying the tonal change of topicalized subjects, topicalized CLLDs, and broad focus subjects, produced by Madrid speakers

Figure 5.8. Box plots and confidence intervals displaying the peak intensity of topicalized subjects, topicalized CLLDs, and broad focus subjects, produced by three speakers of each dialect

Figure 5.9. Scatterplots displaying the paired duration data of preverbal non-nominative subjects in broad focus and topicalized contexts, by three speakers of each dialect

Figure 5.10. Scatterplots displaying the paired duration data of broad focus preverbal non-nominative subjects and topicalized CLLDs, by three speakers of each dialect

Figure 5.11. Box plots and confidence intervals displaying the tonal change of broad focus non-nominative subjects, topicalized non-nominative subjects, and topicalized CLLDs, produced by Havana speakers
Figure 5.12. Box plots and confidence intervals displaying the tonal change of broad focus non-nominative subjects, topicalized non-nominative subjects, and topicalized CLLDs, produced by Madrid speakers.

Figure 5.13. Box plots and confidence intervals displaying the peak intensity of broad focus non-nominative subjects, topicalized non-nominative subjects, and CLLDs, produced by three speakers of each dialect.

Figure 5.14. Scatterplots displaying the paired duration data of preverbal subjects of unaccusative and unergative verbs in broad focus contexts, produced by three speakers of each dialect.

Figure 5.15. Box plots and confidence intervals displaying the tonal change of preverbal subjects of unaccusative and unergative verbs in broad focus contexts, produced by Havana speakers.

Figure 5.16. Box plots and confidence intervals displaying the tonal change of preverbal subjects of unaccusative and unergative verbs in broad focus contexts, produced by Madrid speakers.

Figure 5.17. Box plots and confidence intervals displaying the peak intensity of preverbal subjects of unaccusative and unergative verbs in broad focus contexts, produced by three speakers of each dialect.
List of Appendices

Appendix A: Stimuli
Chapter 1
Introduction

This dissertation re-examines subjects, subject distribution, and subject properties in Spanish. This work is done within the minimalist framework of generative grammar; however, the topics addressed as well as the empirical data provided are relevant beyond this theoretical framework. First, the reanalysis of the syntactic data developed here, which allows for a unified treatment of broad focus structures, contributes to a better understanding of some of the differences in word order not linked to information structure in Spanish. Second, the new experimental data that this study offers adds to the prosodic description and characterization of informationally marked elements in Spanish, specifically left-dislocated old information topics.

A property of Spanish subjects is that they may appear in different positions in the string, as in (1a-c). In these cases, depending on the informational properties of the elements and the context, word order varies. For instance, (1a) is felicitous in a broad focus context, where all the information is new, or in a context in which the VP is new information. (1b), on the other hand, corresponds to the context in which the object DP el libro ‘the book’ is old information, and the subject is new information. Finally, (1c) can be used in a context in which both the subject and the object DPs are old information, and the time in which the writing process took place is new information.

(1) a. María escribió el libro el año pasado.

   Maria wrote the book the year last

b. El libro lo escribió María el año pasado.

   The book cl<sub>ACC/MS</sub> wrote María the year last

c. María, el libro, lo escribió el año pasado.

   Maria the book cl<sub>ACC/MS</sub> wrote the year last
‘Maria wrote the book last year.’

Variation in subject distribution, however, is not limited to different contexts of information structure, as it is also attested in broad focus contexts. In such cases, the position in which the subject is realized seems to be dependent on verb type or argument structure (2-4).

(2) Llegaron estudiantes.

arrived students

‘Students arrived.’

(3) A Juan le gusta el chocolate.

Juan$\text{_{DAT}}$ cl$_{\text{DAT}}$ like$_{3S}$ the chocolate

‘Juan likes chocolate.’

(4) En el parque juegan niños. (Example based on Suñer (1982))

in the park play kids

‘Kids play in the park.’

This behaviour and the difference that it represents when compared to languages like English have made Spanish subjects, especially preverbal ones, the topic of an extensive debate for the past thirty years. Generally, most studies take English subjects as the model and compare Spanish subjects to determine their properties. One of the most debated issues is whether or not in Spanish there is a structural preverbal subject position analogous to the English one.

Even though some of the studies about Spanish subjects look at comparable data, their conclusions regarding the position that preverbal subjects occupy are dissimilar. Some propose that preverbal subjects appear in the left periphery (Alexiadou & Anagnostopoulou, 1998; Ordóñez & Treviño, 1999), while others argue that they appear in the inflectional layer (Goodall, 1999; Suñer, 2003; Ortega-Santos, 2008). Yet, there is
a different type of analysis that proposes that constituents other than nominative DPs, such as dative and accusative DPs, and locative phrases, have subject-like properties and behave as such with respect to movement to the preverbal field and interpretation (Masullo, 1992; Fernández-Soriano, 1999a, 1999b; Cuervo, 1999, 2003, 2011; Ortega-Santos, 2008). These studies agree on the idea that these elements also appear in the IP layer, although some of them point out that the syntactic properties of the position that they occupy are not equivalent to those of the English one.

Given the complexity of subject distribution in Spanish, these analyses are characterized by their heterogeneity. This has an impact not only on their conclusions about the positions that preverbal subjects occupy, but also their explanations regarding movement to the preverbal field, and the syntactic operations of verbal agreement and nominative case checking/assignment in this language. None of them taken alone can offer a complete account of the full distribution of subjects in Spanish (preverbal and postverbal) in broad focus structures. Moreover, the options proposed as the mechanism that triggers movement to the preverbal field, predication and the EPP, face both theoretical and empirical challenges. Thus, the aim of this work is twofold: 1) to contribute to the dialogue about the properties of preverbal subjects, by providing new empirical and theoretical evidence and 2) to propose a unified analysis of Spanish data that can account for subject distribution and interpretation in broad focus structures.¹

Although the main focus of this study is the syntax of Spanish subjects, it differs from previous work as it employs an experimental component to test the hypotheses that derive from the syntactic analyses. Taking into account the relationship between syntax and phonology, and the prosodic properties of Spanish, in which features like focus are prosodically encoded (Zubizarreta, 1998; De la Mota, 1998; Face 2001, 2002), I hypothesize that topicalized elements should be different from elements that constitute

¹ The analysis developed in this work corresponds to subjects that are full DPs, and does not include pronominal subjects, which occupy a different position in Spanish. See Fernández-Soriano (1999c) and Luján (1999), and references therein, for a characterization of overt pronominal subjects in Spanish, their distribution and syntactic properties.
new information. In order to obtain empirical evidence that could help us to disambiguate the syntactic status of preverbal subjects in Spanish, this study assesses the two possibilities of subject realization, the left periphery (CP) and the inflectional layer (IP), with respect to their prosodic features. Specifically, duration, fundamental frequency, and intensity in preverbal nominative and non-nominative subjects in broad focus contexts are compared to the same features in topicalized elements, such as left dislocated-doubled clitic phrases and left dislocated nominative and non-nominative subjects.

Theoretically, the current analysis departs from previous analyses by taking a different approach with respect to the notion of default word order. While it is true that SVO is the default word order of transitive verbs in Spanish, this is not the case for unaccusative and psychological verbs of the piacere class (Contreras, 1976; Belletti & Rizzi, 1988; Masullo, 1992; Cuervo, 1999, 2003; Ortega-Santos, 2008). By taking a more inclusive approach, in the sense that SVO is not assumed as the default option against which all other structures need to be compared, this work looks for the common element that underlies all broad focus structures. In doing so, it goes beyond the EPP and predication as movement triggers to the preverbal field, as these prove to be insufficient or too restrictive when evaluated in the light of the Spanish data. It proposes the specifier of the Event Phrase (EP) (Borer, 2005b) as a locus of preverbal subject realization, and the need to license the event by binding the event variable as the trigger of the movement to this preverbal position. Since specEP can also be filled by non-nominative phrases, this analysis has the advantage of accommodating different word orders attested in broad focus contexts in Spanish. Other syntactic phenomena, such as the distribution and interpretation of bare nouns and strong DPs in Spanish, also show a correlation with the syntactic requirements of specEP. Finally, the way in which the syntactic operations of nominative case checking/assignment and verbal agreement take place in this analysis is compatible with minimalist theoretical assumptions.

This work is organized as follows. Chapter 2 addresses the notion of subject as a syntactic category within the generative framework, and presents Spanish subject data in the contexts that will be analyzed. In Chapter 3, I present, review, compare, and evaluate the different analyses about preverbal subjects in Spanish. Special emphasis is placed on the discussion of issues that have not been previously addressed. Chapter 4 constitutes the
link between syntactic and prosodic analyses. The first part of the chapter discusses the syntactic theories about the syntax-phonology interface in the generative framework, and the second part provides an inventory and a discussion of the studies done in Spanish on the phonetic and phonological properties of focalized elements. Chapter 5 comprises the experimental part of this work, which is a prosodic analysis of preverbal elements in two conditions of information structure in terms of their duration, fundamental frequency, intensity, and phrasing patterns. The data analyzed comes from the varieties of Spanish spoken in Havana and Madrid. The chapter starts with the hypotheses that derive from the two main syntactic proposals about preverbal subject realization. Next, I describe the experiment, experimental design, data collection, data analysis and results. The chapter concludes with a discussion of the findings from a prosodic perspective. Taking the results from a syntactic perspective, Chapter 6 presents a new analysis of subjects (preverbal and postverbal) in Spanish, building on Borer’s (2005b) Event Phrase proposal. In Chapter 7, I examine the operation of verbal agreement in contexts that have been claimed to represent variable agreement associated with differences in subject positions, and demonstrate that such variability is only apparent. Chapter 8 offers a summary of the work, and a discussion from a more general standpoint of the experimental results and the syntactic analysis that this work proposes.
Chapter 2
Subject matters: subject category and subject placement

This chapter examines the subject as a syntactic category, subject types, and subject positions within the generative grammar. It discusses the main issues related to the subject category and its representation in the grammar. At the same time, these notions are assessed in the light of Spanish data. In doing so, special attention is paid to the elements that play a role when it comes to subject placement in this language. The chapter concludes with some topics that any theory about subjects and subject positions in Spanish should be able to account for.

2.1. Subjects, their properties and representation in the grammar

The notion of subject underlies most theories about grammar and conforms our basic representation of grammatical elements; however, the definition of the subject is far from trivial. Syntactically, it is not always clear the type of element that plays such role. For instance, it is standardly accepted that in the case of transitive and unergative verbs the external argument of the verb is the subject, whereas in the case of unaccusative verbs the complement of the verb is the subject. Semantically, its definition is not very clear either. While some thematic roles such as agent and experiencer have been linked to subjects, it is the case that subjects can also be patients and themes, roles typically associated with objects. Yet, there are a number of features—most of them unique—that characterize the subject.

2.1.1. Subjects and subject properties

Most grammatical theories work with the general intuition that the subject has a significant status in the sentence. This idea develops from the different properties that subjects have, which make them more salient with respect to other grammatical elements. One such property is the position that the subject occupies in the sentence in some languages, along with the need of its overt realization. There are languages like English in which there is a requirement that every sentence must have an overt subject, and such element must occupy a specific position, examples in (1).
(1)  
  a. Mary brought the book.
  b. The boy sang.
  c. The train arrived.
  d. The students were asked to leave.

Another distinctive characteristic of subjects is that they are generally nominals, as the sentences in (1) illustrate. However, some analyses have argued that other types of elements can also behave as subjects. For instance, in languages like Spanish, it has been proposed that locatives and datives can function as subjects (Torrego, 1989; Masullo, 1992; Fernández-Soriano, 1999a, 1999b; Cuervo 1999, 2003; Ortega-Santos, 2008), based on their syntactic and semantic properties. (2) and (3) are examples of this.

(2)  
A Marcos le gusta la música coral.

Marcos$_{DAT}$ cl$_{DAT}$ like$_{3S}$ the music choral

‘Marcos likes choral music.’  
(Masullo, 1992, p.116, (1))

(3)  
En esta casa falta café.

in this house miss$_{3S}$ coffee

‘Coffee is missing in this house.’  
(Fernández-Soriano, 1999, p.105, (4a))

Predication is another feature that has been associated with subjects. The idea is that there is a relation between the subject and the rest of the sentence, the predicate, by which the predicate states something about the subject. Interestingly, in examples (2) and (3), the preverbal elements that are considered subjects occupy the first position of the string as traditional subjects in languages like English do. This suggests that predication may be the by-product of a structural configuration not necessarily linked to the nominative element.
Another property of the subject is that it takes part in various syntactic operations. The agreement relationship that the verb establishes with it may be the more visible one, as the sentences in (4) illustrate.

(4) a. Mary sings.
   b. The students sing.
   c. My book is on the table.
   d. Ten books are on the table.

The following scheme, based on McCloskey’s (1997) analysis of subjects, summarizes some of the previous subject properties, and adds others. According to McCloskey, it is precisely the existence of these features that confers prominence to the subject with respect to other grammatical elements.

Subject properties according to McCloskey (1997, pp.197-198)

- They are typically formally marked—positionally or morphologically.
- Every clause must have a subject. This is debatable, but the same has not been claimed for any other syntactic position in any language.
- It is the characteristic bearer of certain kinds of semantic roles (AGENT, CAUSE, EXPERIENCER).
- They are almost always nominal.
- Subjecthood is central in the system of promotion of nominals (derived subjects: passives, subject-to-subject-raising, unaccusative advancement, etc.)

All previous subject properties have been captured by different grammatical theories with various degrees of success. In generative grammar, various generalizations and mechanisms have been developed in order to incorporate them. One example is the Extended Projection Principle (EPP) (Chomsky, 1982), which was formulated within the Principles and Parameters framework, and is still maintained in the minimalist framework. This principle states that every sentence must have a subject, and it should occupy the specifier of the Inflection Phrase (IP).

The EPP has generated extensive debate cross-linguistically because not all languages behave alike with respect to its requirements. In other words, in some languages subjects do not pattern alike English subjects with respect to their distribution. Spanish, for
instance, is one such language. Subjects can be found in different positions in Spanish, as (5) and (6). Furthermore, as (7) shows, subjects do not have to be overtly realized. In addition to this, in meteorological constructions the verb shows the default third person singular morphology but there is no overt subject, (8), as opposed to languages like English and French, in which an expletive appears in the subject position, (9) and (10) respectively. ², ³

(5) María cantó.

Maria sang₃s

‘Maria sang.’

(6) Llegó Juan.

arrived₃s Juan

‘Juan arrived.’

(7) Llegó.

arrived

‘She/ he/ it arrived.’

(8) Llueve.

rains

‘It rains.’

(9) It rains.

² There is one variety of Spanish found in the Dominican Republic that uses the overt expletive *ello* in some of these contexts.

³ For the purposes of this discussion it is not relevant whether the pronoun found with meteorological verbs in English is a pure expletive or has a semi-argumental status, as Bolinger (1973), and others have suggested. The point is that, despite the status of this element, there is no overt equivalent in languages like Spanish, at least not in the majority of its varieties.
Il pleu.

expl. rain_{3S}

‘It rains.’

The EPP is not only a problem for Spanish and other Romance languages; it also shows some exceptions in English. Locative inversion contexts with unaccusative verbs, (11), and there insertion contexts, also with this type of verbs, (12), are examples of unexpected behaviours. In both cases, the element that receives the theta role from the verb and enters into the agreement relationship with it is found postverbally, while a locative element occupies what seems to be the canonical subject position. Even some unergative verbs in English can appear in a similar configuration to that of unaccusatives, as (13) shows.

(11) In the backyard entered a man.

(12) There arrived three students.

(13) On the third floor worked two young women called Maryanne Thomson and Ava Brent. (Borer, 2005b, p. 287, (67a))

Although these examples may be claimed to be somehow marginal in terms of their frequency, their existence poses a question to the across-the-board implementation of the EPP. We can hypothesize, based on these cases that, in addition to the EPP, there is some other principle—even in EPP prominent languages, like English—that interacts with the EPP. Taking into account the non-uniform behaviour of subjects in different languages, some researchers have proposed that the EPP is not universal (Borer, 1986; among others), or that it can be derived from other principles (Borer, 2005b; Bošković, 2007, and references therein). In any case, sentences with preverbal locatives constitute evidence that non-nominative elements, specifically locative elements, can appear in preverbal position even in an EPP prominent language like English.

Transitive expletive constructions in Icelandic, (14), and other Scandinavian languages are also a problem for the EPP. In this type of structures even though there is an overt
subject, this element does not move to the preverbal position, which is occupied by an expletive.

(14) *það hefur einhver bórðað epli.*

there has someone eaten an apple (McCloskey, 1997, p. 215, (43b))

These structures support the idea that there may be other principles that are playing a role when it comes to movement into the preverbal position in different languages and, to some extent, in a language like English, if we take into account the previous English data (11-13). In Icelandic, it is the organization of the information with respect to the discourse the main factor that regulates subject movement.

The agreement mechanism has also been closely related to or defined with respect to the subject given the role that this element plays in it. Based on languages like English, it has been proposed that agreement can take place both locally and via long distance through the operation of Agree. In both cases, it always involves the closest nominal in the syntactic domain of the verb.

Since the agreement mechanism is dependent on the subject and subject position, in pro-drop languages and languages with variable word order, this mechanism has been the source of ample debate. For instance, in Spanish and Greek, some researchers have proposed that agreement is never local (Masullo, 1992; Alexiadou & Anagnostopoulou, 1998; Fernández-Soriano 1999a; among others).5

Some of the contexts that have been claimed to be problematic for the possibility of local agreement in pro-drop languages are those that include null subjects. The standard solution in this case proposes a null element, *pro*, that bears the same features of the argument and sits in the specifier of IP (Rizzi, 1990; among others), basically replicating

4 Several works have argued that Agree should be the only option in the grammar. See Boeckx (2008) for arguments in favour of this view, and Boeckx (2006) for a survey of agreement systems crosslinguistically. See also Chapter 7 of this dissertation for an analysis of Spanish data.

5 Camacho (2006) suggests that both possibilities, local and non-local agreement, exist in Spanish, but they depend on the type of verb and the type of subject.
the English scenario. A question that has been raised against this type of analysis is the case of postverbal subjects (Ordóñez & Treviño, 1999) and the possibility of the existence of two elements with the same syntactic features.

Additionally, in Spanish, some researches have noted differences in agreement patterns that seem to be linked to the position in which the subject is realized (Fernández-Soriano, 1999a; Camacho, 1999; Martínez, 1999; Suñer, 2003). Specifically, it has been suggested that variable agreement is only possible when the subject is postverbal (Fernández-Soriano, 1999a; Camacho, 1999; Martínez, 1999), as (15-17) illustrate.

(15) Pero el viejo rey estaba acusado de anglofilo y en la corte existía el favoritismo y la corrupción.

But the old king was accused of being an anglophile, and in the court existed the favouritism and the corruption.

‘But the old king was accused of being an anglophile, and there were favouritism and corruption in his court.’ (Camacho, 1999, p. 2650)

(16) Vino/Vinieron mi hermana y mi cuñado.

came3S/came3PL my sister and my brother-in-law

‘My sister and my brother-in-law came.’

(17) Ahora ya no se suele/suelen llevar esas faldas.

now anymore not SE use3S/use3PL to wear those skirts

‘No one wears those skirts anymore.’ (Martínez, 1999, p. 2772)

On the other hand, when the subject is realized in preverbal position, only one option is possible, (18) and (19).

(18) Mi hermana y mi cuñado *vino/Vinieron.

*
my sister and my brother-in-law *came\textsubscript{3S}/ came\textsubscript{3P}

(19) Esas faldas ya no se *suele/ suelen llevar ahora.

those skirts anymore not SE use\textsubscript{3S} / use\textsubscript{3PL} to wear now

(Martínez, 1999, p. 2772)

Left dislocated subjects have also been claimed to be able to trigger different agreement patterns than regular subjects (Suñer, 2003), as (20) and (21), respectively, illustrate.

(20) El jurado, presionados unos y apocados los más, declararon

the jury under pressure some and fainthearted the majority declared\textsubscript{3P}

innocente al terrorista.

innocent DOM-the terrorist

‘The jury, under pressure some and fainthearted the majority, declared the terrorist innocent.’

(Martínez, 1999, p. 2767, (188a))

(21) *El jurado declararon inocente al terrorista.

the jury declared\textsubscript{3P} innocent DOM-the terrorist

(Martinez, 1999, p. 2767, (188b))

Thus, the behaviour of these constructions has been taken as an indication that there may be differences in the agreement mechanism that depend on the position in which the subject is realized in Spanish.

It is also worth noting that in Spanish there is a widespread tendency for the verb to enter into an agreement relationship with other elements than grammatical subjects in contexts that have been claimed to lack a subject altogether. This type of agreement can be found
in existential constructions (22), periphrastic weather constructions (23), and other impersonal constructions (24), in all dialects.\textsuperscript{6, 7}

(22) \begin{quote} Habían \ muchas animales. \end{quote}

\begin{quote} there.were\textsubscript{3PL} many animals \end{quote}

‘There were many animals.’

(23) \begin{quote} Hacen unos \ días preciosos. \end{quote}

\begin{quote} Make\textsubscript{3P} some days beautiful \end{quote}

‘Days are beautiful.’

(24) \begin{quote} Van ya \ dos horas que vino. \end{quote}

\begin{quote} go\textsubscript{3P} already two hours that came\textsubscript{3S} \end{quote}

‘It’s been two hours since he/she came.’

2.2. Towards a subject definition

Given the syntactic and semantic properties of subjects and their lack of homogeneity as a syntactic category, in the generative tradition subjects are not considered a primitive of the grammar (Harley, 1995; McCloskey, 1997; among others). Within this framework, the concept of subject is understood as a derived notion. In other words, the element that functions as the subject encompasses a number of relevant features that make it the subject. A direct consequence of this proposal is that having a single element fulfilling all subject properties is just one of the possible options. Nonetheless, there has been a tendency to identify the subject with the element that bears a certain subset of these

\begin{itemize}
\item[\textsuperscript{6}] Some of these cases of agreement are highly sanctioned by prescriptive traditional grammars. However, numerous examples can be found even in newspapers.
\item[\textsuperscript{7}] Impersonal constructions also display the default third person singular agreement.
\item[(i)] Había/habían muchas animales.
\begin{quote} there.is/were many animals \end{quote}
\end{itemize}
features. It would seem that some of these properties are more prominent in what can be called the scale of subjecthood.

2.2.1. Subjects as a derived notion, and subject positions.

Treating the subject as a non-homogeneous category allows accommodating cross-linguistic data. In this sense, McCloskey’s (1997) seminal analysis is the most articulated proposal. His analysis represents a deconstruction of the subject category. His main idea is that a number of features that are linked to the element that functions as the subject in languages like English are distributed among various elements in other languages, i.e., these properties do not necessarily converge in the same element.\(^8\) In other words, whereas in English properties like position, movement into the specifier of IP and agreement are linked, in most cases, to a single element, in other languages the element that occupies the preverbal position and triggers agreement on the verb may not the same. McCloskey uses data from topic prominent languages, such as Icelandic, to illustrate differences in subjects’ distribution. In this type of languages, movement to the preverbal position is reserved to the element that is the topic of the sentence.

Interestingly, there is another type of languages like Spanish in which subject properties in some contexts seem to converge in a single element, while in other contexts they appear to be distributed between two elements. Yet, in other contexts, some of these properties seem to be absent altogether. Taking a less radical look at the Spanish data, we can say that it is not clear how some of these mechanisms operate in some syntactic contexts. With respect to movement, for instance, in contexts such as unaccusative constructions and psychological predicates with preverbal dative experiencers, the element that triggers agreement on the verb appears postverbally in broad focus statements. On the other hand, in transitive constructions, the nominative subject appears

\(^8\) The same idea is proposed in Harley (1995).
in a preverbal position.\footnote{There are exceptions to both of these claims: subjects of unaccusative constructions sometimes appear in preverbal position, and subjects of transitive verbs can appear in postverbal position in contexts that include a locative in the preverbal field. These cases are discussed in detail in Chapter 6.} Table 2.1 summarizes subject movement properties in three different languages.

Table 2.1. Patterns of subject movement in English, Icelandic, and Spanish

<table>
<thead>
<tr>
<th>Language</th>
<th>Obligatory movement of the subject</th>
<th>Movement of a constituent to a topic position</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>yes</td>
<td>cleft const.</td>
</tr>
<tr>
<td>Icelandic</td>
<td>variable</td>
<td>yes</td>
</tr>
<tr>
<td>Spanish</td>
<td>variable/?</td>
<td>variable</td>
</tr>
</tbody>
</table>

Another important point of McCloskey’s (1997) proposal is that subjecthood is a composite and derivational concept. In his theory, there is no distinction between underived (25) and derived subjects, (26) and (27). A welcome consequence of this approach is that subjects do not have to be animate agents (26) and (27), or the bearers of certain thematic roles.

(25) Mary sang.

(26) Mary/the train arrived.

(27) The radio broke.

Additionally, in his analysis there is no theory of what kind of nominal can be a subject apart from the argument linking. Although McCloskey does not include cases of arguments that are non-nominative but seem to occupy one of the subject positions within IP, and only mentions locative inversion in English as an exception, his work is a good starting point to consider when analyzing similar constructions in languages like Spanish.
2.3. Factors that determine subject placement in languages with variable word order

As already noted, the EPP—in EPP prominent languages like English—has been claimed to be responsible for subject movement to the preverbal position, specifically to specIP. In this instance, subject movement is the result of a syntactic requirement at the sentence level. At the other end of the spectrum, we find topic prominent languages, for which it has been proposed that the EPP, as it holds in English, does not apply (McCloskey, 1997), and only elements that bear a topic feature move to the preverbal field. In this case, subject movement is linked to the organization of the information at a discursive level (Information Structure). In both instances movement is determined by a principle that always applies. In the middle, in a sort of gray area, we find languages like Spanish that appear to share properties of both groups.

2.3.1. Information structure

In languages with variable word order such as Spanish information structure is one of the factors that determines the realization of the elements in different positions. These differences in realization loci are later mapped onto the linear order, as the following examples of the same sentence and the contexts in which they are felicitous show, (28-30).

Context: What happened?

(28) María compró un libro.

Maria bought a book

‘Maria bought a book.’

Context: Who bought the book?

(29) El libro lo compró María.

The book clACC/MS bought3S Maria

‘The book, it was Maria who bought it.’
Context: What did Maria buy?

(30)  
a. María compró un libro.

Maria bought a book

b. Un libro compró María.

a book bought Maria

‘Maria bought a book.’

Thus, information structure is one of the factors that should be taken into account when analyzing the distribution of subjects in Spanish. The definition of information structure that will be used in this work corresponds to Lambrecht (1996).

Information structure is that component of the sentence grammar in which propositions as conceptual representations of states of affairs are paired with lexicogrammatical structures in accordance with the mental states of interlocutors who use and interpret these structures as units of information in given discourse contexts. (Lambrecht, 1996, p. 5)

It is important to note that the organization of the information does not depend on the elements themselves, but on the representation of those elements on the speaker’s mind and their use in the discourse. Although the definition makes reference to both, mental states and discourse use, for the purposes of this work only the use of the elements in the discourse will be taken into account, for it is supposed to be the final realization of the mental representations.

There are two broad notions that have been used to classify elements according to their informational properties that are central to the organization of information. These notions correspond to the distinction between new and old. This division splits elements into those that are introduced for the first time in the discourse and those that have already
been mentioned. These two categories, *new information* and *old information*, are linked to the grammatical notions of *focus* and *topic*, respectively. For instance, in (29) the DP *María* is new information, thus it is the focus. The same can be said about the DP *un libro* in (30a), whereas the DP *María*, in this case, is old information. Note that in both cases new information (focus) appears in the last position of the string, while old information appears first. This behaviour is in line with the crosslinguistic tendency that both types of informationally marked elements display: old information appears first in the discourse, while new information tends to appear later. This tendency, however, can be overridden, as (30b) illustrates. Finally, the relationship between old and new information and the notions of topic and focus is not univocal and has to be determined for each particular case.

While we generally tend to think about the notions of topic and focus related to a particular constituent, there are contexts in which all the information is new. These are out-of-the-blue sentences or all-new-information sentences. In these contexts all elements have the same informational status. In these all-new-information contexts—such as in (28)—we can still find various word orders, as SV (31), VS (32) and OVS (33). These differences cannot be attributed to information structure, and are associated to variation in argument structure, as discussed next.

(31) María cantó.

Maria sang

‘Maria sang.’

---

10 The notions of topic and focus are very broad and both accept more fine-grained characterizations. For the present discussion, however, they are adequate.

11 There are more complex typologies of information structure than the ones mentioned here; however, for the purpose of the current discussion they are sufficient. The idea was in this case was to illustrate how word order in Spanish is affected by the informational properties of the different elements.

12 This does not entail that the same word order cannot be associated with more than one context of information structure. As (29) and (30a) illustrate, SVO is used in two different contexts. (29) represents a broad focus context, while in (30a) the subject is old information, thus a possibility could be that it is a topicalized constituent in the left dislocated periphery. These examples demonstrate that there are certain cases that are ambiguous between the default word order and some other context of information structure. Yet, there is a dimension in which they could be different: their prosodic realization. Chapter 5 explores some of these contexts from the prosodic perspective.
Based on the differences in the distribution of subjects in all-new information contexts, I propose that there is another factor that can be linked to the realization of subjects in different positions in Spanish, namely the different syntactic frames (argument structures) that a given verb can participate in. It is worth stressing that this claim only applies to broad focus contexts. Moreover, this suggestion does not entail that the final positions in which subjects appear may be analyzed based only on verbs’ argument structures. Nonetheless, the different configurations in which verbs enter should be the starting point when it comes to analyzing different word orders in contexts that can be considered neutral from the perspective of the information structure.

Following this line of thinking, the hypothesis is that in languages that exhibit variable word order, such as Spanish, and to a less extent in languages with fixed word order, such as English, the argument structure of a given verb can influence the position in which the subject appears. This idea is by no means new. The literature on word order in Spanish and Italian often mentions the distinction between unergative and unaccusative verbs, linking them to SV and VS word orders, respectively. (See Contreras (1976) for Spanish.)

Less frequently, some authors talk about the unmarked word order that psych verbs generate: Dat-V-S (Belletti & Rizzi, 1988; Masullo, 1992; Cuervo 1999, 2003; Ortega-Santos, 2008).

Taking into account argument structure, there are three main scenarios that need to be considered when it comes to subjects and subject realization in Spanish. The relevant
criterion to determine such groups is the standard realization of the subject in broad focus contexts. The idea is that the default word order that we find in such cases may be linked to the different argument structures of their verbs, given the correlation among the different typologies and the different positions in which subjects appear.

The hypothesis about the impact of argument structure on word order is not only plausible, but attractive. Note that since most analyses assume that Spanish is an SVO language, anything that deviates from this pattern should be considered an exception. As a result, they miss an important property of the different word orders linked to argument structure differences: the fact that they all correspond to broad focus structures. Under the hypothesis that argument structure influences word order, it may be possible to look for the common element that underlies these structures, and to offer a unified treatment of the different word orders that broad focus contexts display in Spanish without the need of treating structures with postverbal nominative subjects as exceptions.

The work developed here takes a constructionist approach (Borer, 2005b), in which possible syntactic structures are the main constraint in building verbal meanings and verb types. As a consequence, even though terms such as verb type are used, they should be understood within this approach. Keeping this in mind, there are three relevant scenarios in Spanish regarding word order that are linked to verb types: transitive/unergative constructions, unaccusative constructions and psych verbs constructions.

2.3.2.1. SV(O), VS, and DatVS

The first scenario corresponds to SV(O) word order. In this case, the subject is realized in a preverbal position and the verb appears in second position. These are transitive and unergative verbs. The realization of an object after the verb depends on the type of event, and the amount of information that is given. In this case, the element that functions as the subject is merged in the specifier position of a projection above the verbal phrase (VP), such as vP, Chomsky’s denomination, but it can be a voice phrase à la Kratzer, or any other projection that licenses external arguments. (See Chapter 6 for other possibilities.)
Unaccusative verbs constitute the second scenario. In this instance, the element that is realized as the subject is merged as a complement of the verb (Perlmutter, 1978). In Spanish, subjects of unaccusative verbs can appear both preverbally and postverbally, with a tendency of postverbal subjects to appear in broad focus contexts when there is no modification. Even though SV order is also available, its use has been linked to a different interpretation. While VS is associated with a thetic statement, an SV sentence is said to be interpreted about the subject.

Based on the difference between subjects of unaccusative and transitive/unergative verbs in Spanish, it is possible to separate cases in which a specific word order may be due to information structure such as (34) from others in which it is due to argument structure (35).

Context: Who brought the book?

(34) El libro lo trajo María.

the book \text{cl}_{\text{ACC}} brought Maria

‘The book, it was Maria who brought it.’

Context: What happened?

(35) Llegó María.

arrived Maria

‘Maria arrived.’

---

13 This is known as unaccusative advancement, and the term comes from relational grammar (Perlmutter, 1978).

14 Even in English, a language in which the subject appears in the first position of the sentence, this type of construction shows some alternation regarding the realization of its arguments. Unaccusative verbs accept what is called locative inversion construction, and \textit{there} insertion, as illustrated in (i) and (ii), respectively.

(i) Here comes the sun.
(ii) There arrived three men.

15 The same has been proposed for languages like Italian (Cinque, 1993; among others). I return to this idea in Chapter 6.
The third scenario is comprised by psych predicates of the *piacere* class (36), first described by Belletti and Rizzi (1988) in Italian. These verbs always take an experiencer and a theme as their arguments. In this case, the theme triggers agreement on the verb, and bears nominative case, while the experiencer bears dative morphology. The non-marked word order of these constructions in Spanish is Dat-V-S (36) (Masullo, 1992; Cuervo 1999, 2003; Ortega-Santos, 2008).

(36) A Rebeca\textsubscript{i} le\textsubscript{i} gusta el chocolate.

\textit{Rebeca\textsubscript{DAT}} cl\textsubscript{DAT} like\textsubscript{3S} the chocolate

‘Rebeca likes chocolate’

2.4. Final remarks

This chapter examined the main concepts related to subjects and subjection with special emphasis on Spanish data. As we saw, Spanish subjects show a great degree of flexibility regarding the positions in which they appear in the structure if we take the linear word order as an indicator of these differences. This flexibility, however, depends on various factors.

Generally, most studies recognize that information structure has an impact on the distribution of the elements in languages with variable word order. We went a step further and suggested that in addition to information structure, argument structure also has an impact on word order, but this can only be seen in all-new-information sentences.

With this hypothesis in mind there are three stages to consider when analyzing word order, and specifically subjects’ position, in a language like Spanish. From bottom up, the argument structures in which verbs participate, as they seem to set different patterns that are later reflected on linear word order. In addition to this, there could be other factors that also play a role. Within the minimalist framework, elements move from the positions in which they are merged in order to check strong features. This movement has been linked to requirements such as the EPP in languages like English. Although in the case of Spanish it is not clear the principles that are present and how they operate, such principles cannot be excluded; thus they would be the second step in the analysis. Finally, the way
the speaker organizes the information in the discourse is also relevant when it comes to determine the positions in which subjects appear.

Keeping all this in mind, there are some fundamental topics that any analysis of Spanish subjects should cover. The first point is related to subject position(s) and subject movement to the preverbal field. Since subjects do not always occupy the preverbal position, we are left to solve the puzzle of subject movement to the preverbal field. The other side of the coin is the position that preverbal subjects occupy in this language. Additionally, since subjects can appear in different positions, how many of these positions exist in a Spanish?, and where are they in the structure? Another topic is agreement and agreement mechanisms. Given the claim that there are differences in agreement patterns that depend on the position in which the subject is realized, do we need two versions of the agreement operation in Spanish or such differences can be explained in a different way? Finally, how is case assigned in Spanish? Note that the postverbal realization of subjects represents a problem for case assignment theories since it has been claimed that case, unlike agreement, should be checked/assigned in a spec-head configuration, and not through Agree (Bošković, 2007, and references therein). The next chapter reviews and discusses the most relevant analyses of preverbal subjects in Spanish paying special attention to these questions.
Chapter 3
Previous studies on Spanish preverbal subjects

As shown in Chapter 2, subjects in Spanish display great flexibility regarding the position in which they appear in the string. Generally speaking, this possibility derives from the different loci that they occupy in the structure, which is later mapped onto the linear order. This flexibility, however, does not entail that the different word orders can be understood as equivalent. The realization of a given subject in different positions is linked, most likely, to a different interpretation of such element with regard to topic or focus properties, as some of the positions in which they appear are related to these syntactic features.\(^{16}\) In other words, not every word order is suitable for every context. In addition to information structure, there are other factors that also play a role when it comes to subject placement in Spanish. For instance, the different syntactic frames (argument structures) in which verbs appear are the first stage to consider for there seems to be a correlation between them and the different word orders in all-new-information contexts.

While these considerations may seem obvious, it is important to acknowledge them as some analyses evaluate facts in isolation and equate contexts with identical word orders that correspond to different underlying structures. Finally, although variability in subject placement can be found both preverbally and postverbally in Spanish, this chapter focuses on the preverbal field. My main interest is to engage in this area of the research for there is still disagreement among the various theories that analyze preverbal subjects. The aim of this chapter is twofold: to do an inventory of the syntactic proposals that account for the position(s) in which Spanish preverbal subjects appear, and to evaluate

\(^{16}\) Some researchers such as López (2009) have argued that the notions of topic and focus should not be considered primitives since they cannot provide a proper description of the elements in terms of their syntactic behaviour/distribution. Although I do agree with the fact that a characterization in terms of features like anaphoric, non-anaphoric, etc., provides a more accurate description, for the purposes of this chapter and the present discussion this is not relevant. I come back to issues related to this topic in Chapter 4.
and discuss some of their claims. The chapter concentrates on the issues that are more relevant in the analyses which will allow us to develop an argument. Special attention is paid to points that have not been discussed previously.

3.1. Spanish preverbal subjects: different analyses, different structures, same language?

Given the variability that the subject category displays in Spanish, there are different syntactic analyses about the position in which preverbal subjects are hosted in this language. Differences in the analyses are so great that in some cases the proposals differ to the extent of being mutually exclusive. Although there is no agreement among researchers, what seems to be true for Spanish is that there is not just one position where preverbal subjects can be found. Another conclusion seems to be that the subject category is more complex in Spanish than it is in languages like English or French. Furthermore, we can also say that these analyses if taken in isolation, assuming that all of them were suitable, offer at best an incomplete picture of Spanish subjects, preverbal subjects in particular. A recurrent problem, which, to some extent determines their contradictory results, is that the data that most analyses employ is only partial. That being the case, their accounts propose solutions that explain facts that only concern their piece of data. While working with some data should not be a problem, it is also true that such accounts cannot offer a comprehensive explanation of a given phenomenon. Another issue is that some of the analyses treat their data in isolation, in the sense that they only focus on the final syntactic position that corresponds to a particular word order. This is a problem because by doing so they exclude previous steps in the derivation that may be crucial to explain differences among some of these subjects, and/ or subject’s positions.

Within the great pool of work on the topic of preverbal subjects in Spanish, two lines of research can be distinguished based on the type of analysis that they propose. The first one focuses on the position of the subject in the structure, while the other one (although it also considers the structural position) focuses on the sentential role of this element.

There are further differences among the analyses that comprise the first line of research. Specifically, there are two mutually exclusive proposals. The first proposal argues that the position in which preverbal subjects are hosted in Spanish is different than the
traditional subject position in languages like English. Its central claim is that Spanish preverbal subjects are always left dislocated (LD) elements. The main proponents of this analysis are Alexiadou & Anagnostopoulou (1998, henceforth A&A) and Ordóñez & Treviño (1999, henceforth O&T). From now on I refer to their analysis as the LD proposal. The alternative proposal states that Spanish preverbal subjects, or at least some of them, occupy the same position as their English counterparts, namely Spec IP (Goodall, 1999; Suñer, 2003; Camacho, 2006; Ortega-Santos, 2008). I refer to this analysis as the IP proposal.

The second line of research, which includes what may be considered a semantic or interpretational component, links the realization of some preverbal elements, among them non-nominative subjects, to sentential requirements (Masullo, 1992; Fernández-Soriano, 1999a, 1999b; Cuervo, 1999, 2003; Ortega-Santos, 2008). These researchers argue that the main characteristic of some preverbal elements in Spanish is that the sentence states or predicates something about them. In their analyses the notion of subject is broader than the one found in the analyses that only consider the syntactic position of nominative subjects, for they include not only canonical nominative subjects, but also elements that seem to function as subjects, such as some locatives and dative experiencers, in the line of McCloskey (1997). Finally, according to this analysis, these preverbal non-nominative elements are realized in the IP layer.

3.2. Spanish preverbal subjects as LD elements

This section presents the central claims of the LD proposal, along with some of the syntactic evidence upon which A&A (1998) and O&T (1999) build their analyses. It reviews some of the specifics of two theories, and the main differences in their implementation. A general overview of the LD analysis is provided at the end alongside some critiques that the IP proposal’s works do not cover.

17 Camacho (2006) represents a mixed approach.
3.2.1. LD proposals: general claims

Both theories, A&A (1998) and O&T (1999), propose that in Spanish preverbal subjects are located in the left periphery of the sentence. Although the mechanisms that they employ and their implementation are different, their main claim is equivalent: Spanish preverbal subjects are always topicalized elements.

A&A note that subjects in Greek and Romance can appear in different positions in the string. They point out that this possibility exists in different types of predicates alike (1-3), and is not restricted to matrix clauses.

(1) efige o Petros  
left Peter  
‘Peter left.’

(2) epekse o Petros  
played Peter  
‘Peter played.’

(3) ektise i Maria to spiti  
built Mary house  
‘Mary built the house.’

(A&A 1999, p. 495 (7a-c))

They focus on the SVO and VSO alternation, which is central to their argument, and propose that VSO cannot be analyzed along the same lines as transitive expletive constructions in Icelandic, because VSO does not involve an expletive in Romance and Greek. They argue that if there were an expletive in Romance and Greek, preverbal subjects in SVO sentences would behave alike English subjects regarding their syntactic properties. Using syntactic evidence from different contexts, they conclude that this is not the case. Given that VSO does not include an expletive, Greek and Romance qualify as
non/ weak EPP languages (based on Chomsky (1995)). This, in addition to the fact that specifiers are projected only when a strong nominal feature forces the merging or movement of an XP, and subjects in these languages do not seem to bear such feature (taking into account the different positions in which they appear), led A&A (1998) to conclude that the canonical subject position within IP does not project in Greek and Romance.

O&T’s (1999) analysis challenges “the typical assumption that in a pro-drop language like Spanish covert as well as overt subjects occupy a preverbal position at Spell Out in which their case and agreement properties are satisfied” (p. 39). In order to demonstrate that this is not the case, they offer examples in which preverbal subjects pattern with LD objects in contexts of ellipsis, quantifier extraction, and preverbal quantifier interpretation.

3.2.2. Syntactic evidence of the LD status of preverbal subjects

The main body of evidence that A&A (1998) and O&T (1999) provide to demonstrate that preverbal subjects in null subject languages (NSLs), and particularly in Spanish, do not occupy the same locus as their English counterparts comes from differences regarding subject-verb relation/position, subject-verb-adverb placement, quantifier interpretation, and scope, between subjects in these two types of languages.

3.2.2.1. Distributional evidence

A&A (1998) show that a number of adverbs can intervene between the subject and the verb, as (4) illustrates. The same behaviour, as they point out, is attested in Spanish.

(4) O Petros xtes meta apo poles prospathies sinandiese ti Maria

Peter yesterday after from many efforts met Mary


Assuming that the verb raises overtly to T, and considering the case of French, another V-raising language, in which adverbs cannot intervene between the subject and the verb, A&A propose that there is no spec-head relation between these two elements in Greek
and Romance. Moreover, since temporal adverbs, which are not supposed to appear
within the IP layer, may appear between the subject and the verb, they conclude that these
elements must be in different maximal projections. As a result, the subject should be
higher than TP in the structure, and specIP does not project in these languages.

Further evidence of the lack of the canonical subject position in Romance, according to
A&A (1998), comes from the following examples in Spanish, (5a-b).

(5)  a. Temprano salía Julia de casa.
    early left\textsubscript{IMPF/3S} Julia of home

    ‘Julia was leaving/ used to leave home early.’

    early Julia left\textsubscript{IMPF/3S} of home


A&A claim that (5b) is ungrammatical because the subject and the verb compete for the
same preverbal position. They use (5b)’s ungrammaticality as evidence that the specifier
of T does not project. To further account for (5b)’s ungrammaticality, they also propose
that multiple left dislocations are not possible in Spanish.

Additional evidence that preverbal subjects are instances of LDs comes from parallel
behaviour these elements present with respect to non-focused fronted constituents such as
objects and PPs. Specifically both instances can be separated from the verb by series of
adverbs, (6a) and (6b).

(6)  a. O Petro xtes meta apo poles prospathies sinandise ti Maria

    Peter yesterday after from many efforts met Mary

    ‘After many efforts, Peter met Mary yesterday.’

b. sto parko xtes meta apo poles prospathies sinandiese
in the park yesterday after from many efforts met

Peter Mary

‘After many efforts, Peter met Mary yesterday in the park.’


Comparable evidence is provided by O&T (1999) regarding the similar behaviour found between LD direct and indirect objects and preverbal subjects in contexts of VP ellipsis in Spanish. They propose that since all these elements can be remnants of VP ellipsis, they occupy the same structural position in the CP layer. Sentences (7-9) illustrate this.

(7) A Pía le dio unos libros y a Sara también
to Pía cl\_DAT gave\_3S some books and to Sara too

[le dio unos libros]

[cl\_DAT gave\_3S some books]

‘He/ she gave some books to Pía and also to Sara.’ (O&T, 1999, p. 42 (12))

(8) Unos libros le dio a Pía y unos cuadros también
some books cl\_DAT gave\_3S to Pía and some paintings too

[le dio a Pía]

[cl\_DAT gave\_3S to Pía]

‘He/ she gave some books to Pía and some paintings too.’

(O&T, 1999, p. 42 (11))

(9) Él le dio unos libros a Pía y Pepe también
he c\text{DAT} gave_{3S} some books to Pía and Pepe also

\[ \text{le} \quad \text{dio} \quad \text{unos libros a Pía} \]

\[ \text{c\text{DAT} gave}_{3S} \text{ some books to Pía} \]

‘He gave some books to Pía and so did Pepe.’  \hspace{1cm} (O&T, 1999, p. 42, (10))

3.2.2.2. Interpretational effects: scope, quantifier interpretation, and binding

With respect to interpretational effects, both works provide examples that show that preverbal subjects in Romance and Greek do not pattern with their English counterparts. For instance, A&A (1998) argue that in Greek quantifiers and indefinites in preverbal subject position have unambiguous scope, differently from the English ones that are ambiguous. Thus (10a) only allows wide scope reading, whereas a postverbal subject, as in (10b), allows both narrow and wide scope reading. A&A suggest that the unambiguous scope of Greek preverbal subjects can be explained based on the fact that they do not occupy an argumental position, as no reconstruction effect takes place once elements move into A’ positions.

(10)  
\begin{enumerate}
\item a. Kapios fitis stihiotetise karthe arthro.
\end{enumerate}

\begin{align*}
\text{some student} & \quad \text{filed} & \quad \text{every article} & \quad \text{some} & > & \text{every}
\end{align*}

‘Some student filed every article.’

\begin{enumerate}
\item b. stihiotetise kapios fitis karthe arthr
\end{enumerate}

\begin{align*}
\text{filed} & \quad \text{some student} & \quad \text{every article} & \quad \text{some} & > & \text{every, every} & \text{> some}
\end{align*}


A&A also offer examples of binding contexts as evidence of the A’ properties of preverbal subjects in Romance. They suggest, based on Sola (1992), that overt personal pronouns can only be construed as bound variables when they appear in postverbal position, as (11) shows. When the subject pronoun is preverbal, as (12), there cannot be co-reference between the pronoun and the subject of the matrix clause because the alternation null/ overt exists.
3.2.3. LD proposals: similarities, differences, and general overview

Both proposals develop in an attempt to explain the differences between pro-drop languages and non pro-drop languages (and the differences between different types of pro-drop languages, in A&A (1998)’s case) regarding subject placement, subject and verb movement, and EPP checking, and explore these issues within the minimalist framework. In a sense, the ultimate goal of A&A’s theory is to accommodate cross-linguistic variation in terms of parameters. Their theory is more articulated and extensive than O&T (1999)’s, as it does not only include Spanish, but also Greek, Celtic, and Arabic. For the same reason, some of the claims that they make do not entirely describe Spanish subjects. On the other hand, O&T’s analysis has the advantage of being more specific, since it is based on Spanish data only.

As already noted, both works argue that in NSLs preverbal subjects are instances of left dislocations, as they pattern with LD elements, such as LD objects in some syntactic environments, and they are different from English subjects. Both proposals conclude that in Spanish (also in Greek and other Romance languages in the case of A&A (1998)), the syntactic operation of agreement between the subject and the verb does not take place in a Spec-head configuration. This goes in favour of the elimination of the Specifier of IP or AgrS, as possible subject locus, which is a central claim of both theories.

There are significant differences between two analyses as well. In A&A’s proposal preverbal subjects are always base-generated in the left periphery of the sentence, whereas in O&T’s, they are generated within the vP as adjuncts. This difference is
important in terms of the predictions that derive from it. O&T’s analysis predicts that the unmarked word order in Spanish is VSO, fact that is problematic given the uses, meaning and frequency of this word order. In the first chapter I talked about unmarked word orders in Spanish, we saw that in cases in which all the information is new and there is no emphasis on any given element, the type of argument structure in which the verb enters seems to set or have an impact on the final word order. It is worth pointing out that there may be other factors involved in rendering the final word orders, but we have not explore this possibility so far. Keeping in mind this limitation, it is clear that one of the orders not attested in that inventory is precisely VSO, which would be the default word order in Spanish, according to this theory. More importantly, neither of the two proposals can account for out-of-the-blue statements in which all the information is new, as previous analyses (Goodall, 1999; Suñer, 2003; Ortega-Santos, 2008) have also noted.

Further differences between the two proposals relate to the specificities of the mechanisms by which EPP, case, and agreement are checked. A&A (1998) claim that the main distinction between pro-drop and non pro-drop languages lies on the differentiation of the mechanism by which the EPP is checked. In their theory, this mechanism does not include movement or insertion of an XP in Greek and Romance languages, but the raising of the verb, as its morphology “includes the requisite nominal feature” (A&A, 1998, p. 491). They also propose that in Greek and Romance, unlike Arabic and Celtic, the specifier of TP is not licensed. This way they account for further differences found among languages that belong to the pro-drop group.

In their proposal non-referential (expletive) pro does not exist. A&A (1998) conclude that there is no pro in Greek and Romance after examining SV(O)/ ExpletiveVS(O)

---

18 Please note that this is one of the instances in which my approach differs from previous analyses of Spanish subjects, as I do not consider that there is just one unmarked word order in this language, but different ones that specifically correlate to different argument structures. This approach has the advantage that it allows us to not consider cases such as postverbal subjects of unaccusative verbs or postverbal subjects of psych predicates as mere exceptions. In addition to this, it offers the opportunity to provide a more complete analysis of the subject category in this language because it includes more cases, which could ultimately contribute to determining the principles that operate in the language.

19 Even O&T (1999) recognize this as a problem and state that VSO is not the more neutral order in the discourse. The reader can find more of this discussion in their work.

20 In their proposal EPP refers to checking a nominal feature of AGR, in the sense of Chomsky (1995).
alternations in Germanic and transitive expletive constructions in Icelandic, and comparing them to VSO order in Greek and Romance. They suggest that since no Definiteness Effects (DE) are attested in Romance and Greek in unaccusative constructions, neither in transitive constructions with postverbal subjects in Greek, it follows that VSO order does not involve an expletive in these languages, and as result there is no evidence to postulate the existence of expletives in Romance languages and Greek. With respect to referential pro, they do not discard the possibility that it exists, but their proposal does not depend on it, since all the information that this null pronoun bears is encoded in the morphology of the verb.

O&T (1999), on the other hand, based on the idea that agreement can be considered a clitic, propose the elimination of AgrS as a functional projection that hosts subjects in Spanish. In their analysis, agreement is a clitic that absorbs theta role and case. As a result, the realization of overt subjects is seen as the same phenomenon as clitic doubling, which, as they accurately point out, is quite common in Spanish. Thus the information related to person and number is encoded in a clitic like morpheme located in inflection. O&T’s motivation to disjoint this information from the DP is based on the fact that some plural DPs can show various possibilities with respect to agreement: first person, second person, and third person plural; therefore they conclude that agreement depends on a pronominal like item, and not on the DP itself (13). This way they justify how the co-reference is possible based on subject agreement and not on the subject DP, (14) and (15)

(13) Los estudiantes tienen/tenéis/tienen mala memoria.

    the students   have-1P/-2P/-3P    bad memory

    ‘We students/ You students/ The students have bad memory.’

    (O&T, 1999, p. 58, (68a-c))

(14) *[Los estudiantes]i salimos de la reunión después de que los ellos acusaran.

    the students, left-1P of the meeting after of that them accused-3P

    (O&T, 1999, p. 59, (71a))
Since agreement takes place without the need of an overt subject, case checking or agreement cannot motivate subject movement to a preverbal position. As a result, subject movement is justified by discourse features. A direct consequence of this theoretical take is that preverbal subjects in Spanish always move to the left periphery of the sentence, for being this layer the one that relates to discourse properties. O&T propose that referential pro does not exist in Spanish, as all its information is encoded in the agreement morpheme.

As will be seen, most of the syntactic evidence that the LD proposal provides to support the left dislocated status of Spanish preverbal subjects is debunked by the IP analysis. But before moving onto that, some comments about part of the evidence that the LD proposal offers that the IP analysis does not cover are in order.

3.2.4. Examining some of the evidence of the LD proposal

3.2.4.1. Distributional facts: on A&A’s subject-verb placement facts

A&A (1998) begin their analysis by looking at distributional facts. They examine alternations in word order such as SVO and VSO, along with sentences that contain adverbs in the preverbal field in order to find out the subject’s syntactic requirements. The first thing worth noting about their examples is that they do not make any distinction among the different structures that display the same word order. In other words, their analysis assumes that a given word order is possible regardless of the conditions in which it may be acceptable. This premise undermines any conclusion that derives from their analysis given that it presupposes that factors such as information structure and, possibly, argument structure do not play a role when it comes to word order variation, when the opposite is standardly accepted, especially in the case of information structure in
languages with variable word order. As a result, A&A end up treating alike sentences that may not be the same in terms of their syntactic features/structure.

The following sentences, which are a translation into Spanish of their Greek examples, demonstrate that indeed they are possible in Spanish. However, the conditions in which they are used are different, which suggests that their syntactic structures may not be equivalent. (16) could be the answer to a question like “What happened?” Example (17) is acceptable as the answer to the question “Who played?” Lastly, (18) is very hard to imagine as anything else but a question in Spanish, or some type of focalization.

(16) Salió Pedro.

left Pedro

‘Pedro left.’

(17) Jugó Pedro.

played Pedro

‘Pedro played.’

(18) Construyó María la casa.

built María the house

‘Maria built the house.’

In addition to the previous examples whose behaviour in Greek they just extrapolate to Spanish, some of their examples in Spanish are not appropriate to back up their syntactic claims based on distributional facts either. A particularly illustrative case is (5a-b).

(5) a. Temprano salía Julia de casa.

early leftIMPF/3S Julia of home

‘Julia was leaving/ used to leave home early.’
b. *Temprano Julia salía de casa

   early Julia left\textsubscript{IMPF/3S} of home


A&A (1998) argue that (5b) is ungrammatical because the subject competes with the adverb for the same preverbal position. They further propose that in Spanish, differently from Italian and Greek, multiple left dislocations are not allowed. (5), however, does not exemplify any of these two claims. First, (5a) and (5b) depict unaccusative structures. A possibility may be that since subjects of unaccusative verbs are generated as the complement of the verb, there is indeed some kind of competition when it comes to movement to the preverbal field in Spanish (in general) and particularly so in this syntactic environment. It has been noted that there is a difference in the interpretation of sentences (19) and (20).\textsuperscript{21} (19) is suitable in out-of-the-blue contexts, while (20) is said to be interpreted about the subject.

(19) Llegó Juan.

   arrived Juan

   ‘Juan arrived.’

(20) Juan llegó.

   Juan arrived

   ‘Juan arrived.’

This difference is clearly seen if we add an adverb, as (21). In this case the neutral word order is SVAdv, while other possibilities are just odd in out of the blue contexts.

(21) Juan llegó temprano.

\textsuperscript{21} The same argument has been proposed for languages like Italian (Cinque, 1993, among many others).
Juan arrived early

‘Juan arrived early.’

Their conclusion about the impossibility of multiple left dislocations in Spanish is also erroneous, as (22) illustrates. Their mistake originates because they assume that (5b) is ungrammatical across-the-board, without taking into account that it is only so under a broad focus stress pattern. In other words, it is not the case that (5b) is ungrammatical, but it requires a specific intonation that derives from the informational properties of its elements, which confirms that information structure is relevant. As (22) shows, nothing prevents any given sentence in Spanish from having more than one topic.

(22) Esta mañana, a María, nadie la ha visto.

‘No one has seen María this morning.’

Taking into account (22), a sentence like (5b) should be acceptable even under the left dislocated analysis, as its ungrammaticality cannot be explained based on the topic character of its subject, and/or the impossibility of multiple left dislocations in this language.

For all previous reasons, (5b) does not constitute evidence to support any of the claims based on distributional facts in Spanish. If anything, it argues more about the specificities of unaccusative structures in this language. Moreover, Ortega-Santos (2008) notes the so-called incompatibility of subjects and adverbs in the same projection cannot be maintained. He argues that based on the assumption that adverbs enter the structure as specifiers, along with the standard assumption that there could be multiple specifiers, nothing prevents a subject and an adverb to appear within any given projection.

3.2.4.2. O&T (1999)’s ellipsis contexts: What do they really mean?

O&T use the parallel behaviour between LD direct and indirect objects and preverbal subjects in contexts of VP ellipsis in Spanish as evidence of the LD status of preverbal subjects. Since Spanish allows for all these different types of syntactic elements to be
remnants of VP ellipsis, a plausible hypothesis is that all of them have something in common, which, according to O&T, is the fact that they occupy the same structural position in the left periphery. The relevant examples, (23-28), follow.

(23) A Pía le dio unos libros y a Sara también [le dio unos libros] to Pía cl gave_{3S} some books and to Sara too [cl gave_{3S} some books]

‘She/he gave some books to Pía and so did Sara.’

(O&T, 1999, p. 42, (14))

(24) Unos libros le dio a Pía y unos cuadros también [le dio a Pía] some books cl gave_{3S} to Pía and some paintings too [cl gave_{3S} to Pía]

‘She/he gave some books to Pía and some paintings too.’

(O&T, 1999, p. 42, (11))

(25) Él le dio unos libros a Pía y Pepe también [le dio unos libros a Pía] he cl gave some books to Pía and Pepe also [cl gave some books to Pía]

‘He gave some books to Pía and so did Pepe.’

(O&T, 1999, p. 42, (10))

(26) A ti te admitirán en Harvard pero es probable DOM you cl will.admit in Harvard but is probable

que a tu amiga no [la admitan^{22} en Harvard]

that DOM your friend not [cl admit in Harvard]

---

^{22} In their example the tense is the synthetic future *admitirán* ‘will.admit’, but this context generally uses present of the subjunctive mode.
‘You will be admitted to Harvard, but it is probable that your friend will not.’

(O&T, 1999, p. 42, (16))

(27) María no sabe ruso pero es posible que Luis sí [sepa ruso]

Maria not knows Russian but is possible that Luis yes [know\textsubscript{SUBJ} Russian]

‘Maria does not know Russian, but it is possible that Luis does.’

(O&T, 1999, p. 42, (15))

(28) Luis no sabe traducir pero yo conozco a una alumna

Luis not knows to translate but I know \textsubscript{DOM} a student

que sí [sabe traducir]

who yes [knows to translate]

‘Luis does not know how to translate, but I know a student who does.’

(O&T, 1999, p. 42, (17))

(23) and (24) include LD objects, whereas (25) includes a preverbal subject. Sentences (26-28) show that the same behaviour is also attested in cases that include island contexts.

Even though preverbal subjects and LD objects behave alike in the previous sentences with respect to ellipsis, the conclusion of extending the LD status of these subjects to all preverbal subjects in Spanish is erroneous. It is the case that preverbal subjects and LD objects behave alike in the previous examples, but this is only so given the properties of these constructions. All previous examples constitute clear instances of focalized or topicalized contrastive elements in preverbal position, and the ones that include subjects are no exception. This is expected since the contexts that make this type of ellipsis possible include polarity particles such as sí ‘yes’, no ‘not’, tampoco ‘neither’, and también ‘too’. These polarity particles create a contrast between the element in the first clause and its counterpart in the following one where part of the VP is elided. One of the
properties of the left periphery is that elements that bear a special informational status such as topic, focus, contrastive focus, etc., check their features there. For this reason, given the context that these polarity particles trigger, and the preverbal position that these elements occupy, it is safe to assume that all these preverbal elements are the left periphery, therefore it is only expected that they behave as such. Nevertheless, this does not suffice to propose that all preverbal subjects in Spanish are left dislocated. In other words, these examples confirm that it is possible for a subject in Spanish to appear in the left periphery, but this cannot be generalized to all preverbal subjects. Furthermore, this type of ellipsis also takes place in contexts that include adjuncts, such as PPs, which have nothing in common with DO, datives or subjects, since they are not arguments.

Based on the same type of contexts, O&T (1999) also test the null hypothesis that overt subjects and referential pro occupy the same position within the IP layer, as represented in (29) and (30).

(29) \[ IP [\text{Subject no/ también/ tampoco/ sí}] \]  
     \[ \text{(O&T, 1999, p. 43, (19a))} \]

(30) \[ CP [\text{DO/IO IP [pro no/ también/ tampoco/ sí]}] \]  
     \[ \text{(O&T, 1999, p. 43, (19b))} \]

Although this may seem like an attractive idea, it fails for two reasons. First, they compare the position of pro and overt subjects in syntactic environments in which pro would never appear. Second, the contexts that they work with require an identity condition between the subject of the first sentence and the second one, which is violated in the example that they provide (31). Compare (31) to their other examples, or to (32), below, a version of (31) in which the identity condition holds.

(31) a. A ti la policía te va a detener,
     \[ \text{DOM you, the police is going to detain} \]
     pero a Pedro el juez no lo va a detener.
     \[ \text{but DOM Pedro, the judge not is going to detain} \]
     ‘The police is going to detain you, but the judge is not going to detain Pedro.’
b. *A ti la policía te va a detener,

\[
\text{DOM you, the police cl is going to detain}
\]

pero a Pedro el juez no.

but DOM Pedro the judge not

(O&T, 1999, p. 43, (20a-b))

(32) A ti la policía te va a detener,

\[
\text{DOM you, the police cl is going to detain}
\]

pero a Pedro no [lo va a detener la policía]

but DOM Pedro, not [cl is going to detain the police]

‘The police are going to detain you, but not Pedro.’

With respect the first problem (possible contexts in which pro appears), note that the previous examples represent ellipsis contexts that include contrastive elements. One of the premises of contrast is that two or more elements are evaluated with respect to each other. A property associated with this type of elements is that they bear specific prosodic features that make them more salient in languages in which contrastiveness is encoded prosodically. Spanish is one such language. Given the combination of these properties it is very hard to contrast two elements without mentioning one of them. The only thing that their example proves is that it is impossible to obtain this type of ellipsis when there is no identity in the elided portion of the two sentences. Moreover, the fact that pro and overt subjects do not have the same syntactic distribution does not constitute evidence that a position within IP does not exist, which is another O&T’s (1999) conclusion. (As will be seen, the IP analysis argues in favour of various subject positions within inflection, and provides evidence for it.)

A similar problem emerges with Camacho (2006)’s data based on this type of ellipsis constructions. The behaviour of this type of LD subjects is the origin of his claim that
preverbal lexical subjects are located in the left periphery in Spanish, whereas other subjects such as pro, _ello_ ‘it’ and _nadie_ ‘no one’ appear in the IP layer. His argument does not hold for the same reasons as O&T’s. Camacho’s examples that include modals and show further differences between LD elements like direct objects and datives, on the one hand, and LD preverbal subjects, on the other hand, do not provide any insight about the position of non-dislocated preverbal subjects. Additionally, some of the syntactic evidence is hard to evaluate, as shown below.

According to Camacho (2006), the following examples show asymmetries in the elided material depending on the type of preverbal element: preverbal lexical subjects and CLLD objects. In (33), the modal cannot be part of the VP, whereas in (34) the modal is part of it. Example (33) posses the extra challenge of being an unaccusative structure; nevertheless, I think that the same behaviour is attested in transitive constructions. Based on (33) and (34), he proposes that preverbal lexical subjects occupy a position higher in the structure within the CP, i.e., more external, than CLLD objects.

(33) Los médicos pueden haber salido, pero Pedro dice que las enfermeras no _[han salido]_

‘Doctors may have left, but Pedro says that the nurses haven’t (left).’

‘*Doctors may have left, but Pedro says that the nurses may not.’

(Camacho, 2006, p. 55, (9a))

(34) A los médicos les puede haber dado permiso, pero Pedro dice que _ DOM_ the doctors  cl can have given permission, but _ DOM_ Pedro says that

a las enfermeras no _[pueden haber dado permiso]_

DOM the nurses not cl canhave given permission
‘Doctors may have been granted permission, but Pedro says that nurses may not
(have been granted permission).’ (Camacho, 2006, p. 55, (9b))

Camacho notes further differences in the behaviour of non-quantified preverbal and
postverbal subjects with respect to their compatibility with a CLLD constituent in
contexts that include modals. He suggests that these differences could be somehow
similar to those argued by A&A (1998) regarding the scope properties of preverbal and
postverbal subjects. Differently from the scope properties, which change depending on
the position in which the subject is realized, in this case the interpretation of the sentence
is compromised. The relevant examples follow.

(35) #Sí, a los pacientes, los médicos residentes los deben atender,

   yes, DOM the patients, the doctors residents cl must attend

   por eso están tan bien.

   for that are so well

   ‘Yes, patients must be taken care of by resident doctors, that’s why they are so
well cared for.’ (Camacho, 2006, p. 56, (10a))

(36) Sí, a los pacientes, los deben antender los médicos residentes,

   yes, DOM the patients, cl must attend the doctors residents

   por eso están tan bien.

   for that are so well

   ‘Yes, patients must be taken care of by resident doctors, that’s why they are so
well cared for.’ (Camacho, 2006, p. 56, 10b)

The previous data is hard to evaluate. (35) may be odd because of the way that the
information is structured and not necessarily because of the status/properties of the
subject. The same type of oddness is obtained when we substitute the preverbal non-
quantified subject for a specific subject which suggests that notions like quantification or specificity may not be the problem in this case, as (37) shows.

(37)  Sí, a los pacientes, la Dr. Smith los debe atender, por eso están tan bien

yes, DOM the patients the Dr. Smith cl must attend for that are so well

‘Yes, Dr. Smith must be taking care of patients that’s why they are so well cared for.’

(37) improves once we add a locative like aquí ‘here’ at the beginning of the sentence, as a topic (38).

(38)  Sí, aquí, a los pacientes, los médicos residentes los deben antender,

yes here DOM the patients the doctors residents clACC must attend

por eso están tan bien.

for that are so well

‘Yes, resident doctors must take care of patients here, that’s why they are so well cared for.’

The same locative appears in (39) and (40), taken from Camacho (2006), meant to illustrate that the problem does not arise when the CLLD is dismissed. Interestingly, both examples degrade once this locative is removed, as (41) and (42) show, which again points to problems of information structure.

(39)  Sí, aquí los médicos residentes deben atender a los pacientes,

yes here the doctors resident must attend DOM the patients

por eso están tan bien.

for that are so well
‘Yes, resident doctors must take case of patients here, that’s why they are so well cared for.’

(Camacho, 2006, p. 56, (11a))

(40) Si, aquí deben atender a los pacientes los médicos residentes,

yes here must attend DOM the patients the doctors resident

por eso están tan bien.

for that are so well

‘Yes, resident doctors must take care of patients here, that’s why they are so well cared for.’

(Camacho, 2006, p. 56, (11b))

(41) Sí, los médicos residentes deben atender a los pacientes,

yes the doctors resident must attend DOM the patients

por eso están tan bien.

for that are so well

(42) # Sí, deben atender a los pacientes los médicos residentes,

yes must attend DOM the patients the doctors resident

por eso están tan bien.

for that are so well

In any case, it is hard to take ellipsis data as conclusive evidence of the status of preverbal subjects in Spanish, primarily, because these contexts correspond to LD subjects, which is valid, but they do not say anything about non-LD subjects. This along with the previously discussed erroneous distributional evidence that A&A (1998) use to build part of their argument weakens the LD proposal. As will be shown in the following section that covers the IP account, most of the evidence that A&A use does not describe Spanish data.
3.3. Spanish preverbal subjects as elements of the IP layer

3.3.1. IP proposal: general theoretical claims

The IP proposal claims that preverbal subjects in Spanish, or at least some of them, are located within the inflectional layer, similarly to their English counterparts (Goodall, 1999; Suñer, 2003; Camacho, 2006; Ortega-Santos, 2008).23 This claim does not deny the possibility of preverbal subjects to appear in the left periphery. It is a well-known fact that in Spanish almost any element can be dislocated, and subjects are no exception (Suñer, 2003). However, this does not necessarily entail that all of them are dislocated (Suñer, 2003). As already noted, some of these researchers argue that a crucial fact in this case is that preverbal subjects in Spanish can appear in contexts that correspond to all-new-information sentences (Goodall, 1997; Suñer, 2003; Ortega-Santos, 2008), which should be impossible if they are located in the left periphery.

It is worth pointing out that all works of the IP account developed as a reaction to the proposal that preverbal subjects are LD elements in Spanish.24 As a result, most of the evidence that they use to assess the status of preverbal subjects overlaps with that of the LD account. This is especially useful as it allows us to more readily compare the syntactic evidence provided by both accounts.

3.3.2. IP proposal: syntactic evidence

3.3.2.1. Subject-adverb-verb distribution

Part of the evidence that A&A (1998) provide to demonstrate that preverbal subjects do not occupy the same position in Greek and Romance as in English comes from the relative positions that subjects, verbs and adverbs occupy with respect to one another. Based on such distributional differences, they conclude that subjects and verbs are not within the same functional projection. Evidence of the lack of spec-head relation between the verb and the subject is shown by the fact as a number of adverbs can intervene between the two, especially the case of temporal adverbs, (6).

23 Only some subjects are in IP in Camacho’s analysis.
(6) a. O Petros xtes meta apo poles prospathies sinandise ti Maria

Peter yesterday after from many efforts met Mary


Suñer (2003), on the other hand, notes that in Spanish adverbs from a different category can intervene between the subject and the verb without forcing a parenthetical reading, as (43) shows. This behaviour, according to Suñer’s analysis, suggests that subjects are not always LD in Spanish.

(43) El maestro probablemente encontrará rápidamente el error.

the teacher probably will.find quickly the mistake

‘Probably the teacher will find the mistake quickly.’ (Suñer, 2003, p. 343, (3a))

3.3.2.2. Scope

Another piece of evidence that A&A (1998) use to support the LD status of preverbal subjects is the scope properties of indefinite and quantifier preverbal subjects in Greek. As already seen, their claim is that indefinite and quantificational preverbal subjects are not ambiguous with respect to their interpretation, while postverbal subjects preserve their ambiguity, like their English counterparts. They argue that this is the case because, differently from languages like English in which preverbal subjects occupy an argumental position, in Greek, the final landing site of these elements is a non-argumental one (A’); thus no reconstruction effects can take place at LF if an element has moved into an A’ position. They conclude that since preverbal subjects pattern alike LD direct objects in Greek regarding their interpretational properties, preverbal subjects are also instances of LDs. They extend their analysis to Spanish.

Suñer (2003) counters this claim by arguing that in Spanish preverbal subjects preserve their postverbal interpretation after moving into the preverbal field. She points out that this fact would be impossible if they had moved into an A’ position. Examples (44a-c) illustrate this ambiguity. According to Suñer, the three instances are ambiguous between the distributive and the quantificational reading; however, the distributive reading, in
which *cada* ‘each’ takes scope over *algún* ‘some’, is more prominent than the quantificational reading for (44a), despite the preverbal position of the subject.

(44) 

a. En la biblioteca departamental, algún estudiante sacó prestado cada libro.

in the library departmental, some student borrowed each book
cada>>algún preferred

b. En la biblioteca departamental, sacó prestado *algún estudiante* cada libro.

in the library departmental, borrowed some student each book

c. En la biblioteca departamental, sacó prestado *cada libro* algún estudiante.

in the library departmental, borrowed each book some student

‘In the departmental library, some student borrowed each book.’

(Suñer, 2003, p. 344, (5a-c))

With respect to the lack of ambiguity of LD objects, Suñer points out that it is only expected given the properties of accusative clitics that are always specific.

Suñer also notes that in Spanish there is not only one option when it comes to the interaction between the position of an indefinite subject and the interpretation of a modal, which is another claim of A&A (1998). This fact goes against the idea that the preverbal subject always takes wide scope over the modal, and, as a result, it can only receive specific interpretation. According to Suñer, both the epistemic and the deontic readings are possible in Spanish independently of the subject position. The correlation that she found was between perfect tenses and the epistemic reading. About examples (45a-c), she argues that the epistemic reading is preferred.

(45) 

a. Un estudiante pudo haber asesinado a su loro. poder>>un

a student could have assassinated DOM his parrot

b. Pudo haber asesinado un estudiante a su loro.
could have assassinated  a student  DOM  his parrot

c. Pudo haber asesinado  a  su loro un estudiante.

could have assassinated  DOM   his parrot   a student

‘A student could have assassinated his parrot.’

(Suñer, 2003, p. 345, (8a-c))

Suñer points out that with other verbal tenses possibilities vary depending on the speakers. Some speakers report the same type of ambiguity, whereas others prefer the deontic reading with preverbal subjects, and both readings with postverbal ones. As she notes, the behaviour described for indefinite preverbal subjects in Greek does not hold in Spanish. In other words, scope properties of preverbal subjects do not constitute evidence of the A’ status of these elements in Spanish.

3.3.2.3. Binding

Based on Montalbetti’s (1986) observation that in null subject languages embedded subject pronouns cannot be bound by a quantifier in the matrix clause since the alternation null/ overt exists, A&A (1998) claim that the dissimilar status of overt and null subjects in these languages is responsible for this syntactic behaviour.25

Suñer (2003) challenges this claim by examining contexts that range from postverbal to preverbal subjects in the embedded clause, examples (50-52). She found that even though postverbal subjects of the embedded clause favour binding, fact that, as she points out, challenges Montalbetti’s observation, preverbal subjects can also be bound by a

---

25 The following examples adapted from Suñer (2003, p. 347) illustrate Montalbetti’s claim.

(i) Muchos estudiantes, piensan que pro, son inteligentes.
   Many students think that are intelligent
   ‘Many students, think that they, are intelligent.’

(ii) Muchos estudiantes piensan que ellos son inteligentes.
   Many students think that they are intelligent
   ‘Many students, think that they, are intelligent.’
quantifier in the matrix clause. These findings also go against A&A’s suggestion that there is a difference in the interpretation of these two types of subjects that derives from the loci that they occupy in the structure. In other words, there is no evidence that null subjects and overt subject pronouns behave differently in this context with regard to their interpretation.

(46)  

a. Todos los jugadores piensan que ellos ganarán la copa.  

All the players think that they will win the trophy

b. Todos los jugadores piensan que ganarán ellos la copa.  

All the players think that will win they the trophy

c. Todos los jugadores piensan que ganarán la copa ellos.  

All the players think that will win the trophy they

‘All the players think that they will win the trophy.’

(Suñer, 2003, p. 347, (15a-c))

Suñer proposes that all previous sentences are ambiguous between the interpretation that establishes a co-reference between the quantifier and the overt pronoun, and the interpretation in which the referent of the overt pronoun is different from the subject of the matrix clause.

3.3.2.4. Agreement

Agreement is another point that has been used to assess the status of preverbal subjects. The LD account argues that since this operation does not take place locally, AgrS must be eliminated as a functional projection in NSLs. As previously seen, this conclusion derives

---

26 Suñer notes that similar facts were found for Catalan in cases of postverbal subjects. She cites Barbosa’s (1997) findings for this language.
from the behaviour/ position of overt preverbal subjects, verbs and adverbs, along with the possibility of null subjects.

Suñer (2003), in a different take, suggests that subject-verb agreement with collective nouns in Spanish indicates differences between preverbal non-LD subjects and LD subjects in Spanish. She uses the following sentences to show that topic subjects accept _ad sensum_ agreement, examples in (47), while non-LD subjects do not, examples in (48). According to this idea, in Spanish, the result of the agreement operation may be different depending on the subject position.²⁷

(47) a. El jurado, Mara nos aseguró que estaban presionados.

the jury₃S Mara us assured that were₃P pressured

‘The jury, Mara assured us that they felt pressured.’

b. La familia, Mara nos aseguró que se pusieron de acuerdo.

the family₃S Mara us assured that SE were₃P of agreement

‘The family, Mara assured us that they reached an agreement.’

c. La escuadra, el general explicó que atacaron a los enemigos.

the squad₃S the general explained that attacked₃P DOM the enemy

‘The squad, the general explained that they attacked the enemy.’

(Suñer, 2003, p. 350, (26a-c))


the jury₃S were₃P pressured

²⁷ It could be argued that differences among these sentences arise from the fact that they have different numerations, thus they cannot be compared. For the moment, I am just presenting Suñer’s (2003) arguments. I discuss the agreement mechanism, and these cases in particular in Chapter 7.
b. *La familia se pusieron de acuerdo.

the family$_{3S}$ SE were$_{3P}$ in agreement

c. *La escuadra atacaron a los enemigos.

the squad$_{3S}$ attacked$_{3P}$ DOM the enemy

(Suñer, 2003, p. 350, (25a-c))

Camacho (2006), whose analysis represents a mixed approach, proposes that agreement can be local or unbound in Spanish. The two possibilities depend on the subject type (as will be seen in 2.3.2.6), and on whether or not a projection is active. In his analysis, verb movement, for instance, renders a projection active. As a result, the subject moves to establish agreement. With respect to the various possibilities that plural DPs display in Spanish, he suggests that there is no need to dissociate the subject DP from the agreement mechanism and postulate a clitic-like morpheme like O&T (1999) propose, when using a more articulated version of the CP layer can solve the ambiguity. Camacho (2006) proposes the structure in (49), based on Poletto (2000).

(49) $\left[\text{TopP Subj DP}_i \left[\text{NegP NumP t}_i \left[\text{Num inf V}_j\right]\left[\text{Hearer P t}_i\left[\text{Hearer t}_j\left[\text{Speaker P t}_j [\text{TP}]\right]\right]\right]\right] \right]$

(Camacho, 2006, p. 64, (33))

This structure represents a more detailed view of the CP layer as it includes functional projections that relate to the discourse, such as the notions of speaker and hearer which make possible to recover any kind of information related to the subject. When no information is available, the default third person is used. In Camacho’s structure, preverbal lexical subjects are in TopP, but this should not be the only option if we take into account the syntactic evidence discussed thus far. Examples (50a-c)) show the implementation of the previous structure.

(50) a. [Los amigos] [SpeakerP pro$_{1PL}$ salimos (1PL.)]

the friends left
b. [Los amigos] [HearerP pro(2PL salisteis (2PL))]

c. [Los amigos] [IP pro(3PL) salieron (3PL)]

(Camacho, 2006, p. 64, (35))

3.3.2.5. Bare facts

Differences in the distribution of bare nouns have also been taken as a proof of the dissimilar status of preverbal subjects and topics by the IP account. It is well known fact that in Spanish preverbal subjects in broad focus contexts cannot be bare nouns. This regularity has been observed in different studies (Suñer, 1982; Torrego, 1989; Casielles-Suárez, 2004).\(^{28}\) The following examples illustrate this fact.

(51) Llegaron invitados.

arrived\(_3\)P guests

‘Guests arrived.’

(52) *Invitados llegaron.

guests arrived\(_3\)P

(53) Aquí duermen animales.

here sleep\(_3\)P animals

‘Animals sleep here.’ (Torrego, 1989)

---

\(^{28}\) The phenomenon is more complex than the differentiation between cases of preverbal and postverbal subjects. Cuervo (2010) also notices that subjects of psych verbs in Spanish cannot be bare nouns, even though they appear in a postverbal position.

(i) A María le gustan las flores.

Mariá\(_{\text{DAT}}\) cl please\(_3\)P the flowers

‘Maria likes flowers.’

(ii) *A María le gustan flores.

Mariá\(_{\text{DAT}}\) cl please\(_3\)P flowers

However, and differently from psych verbs, it is also true that in unaccusative and some unergative constructions the possibility of having a postverbal bare subject exists. I come back to these issues in Chapter 6.
(54) *Animales duermen aquí.  

animals sleep<sub>3P</sub> here

Goodall (1999) compares the impossibility of preverbal subjects in broad focus contexts to appear as bare nouns to the possibility of topics to do so, (55), and argues that these two types of elements cannot be equated given their dissimilar behaviour in this context.  

(55) Yo a él libros no le dejo.

I to him<sub>i</sub> books not cl<sub>i</sub> lend

‘Books, I don’t lend him.’  

(Goodall, 1999, p. 99, (9))

He provides further evidence for the claim that preverbal subjects differ from topics. Based on the idea that bare quantifiers cannot be topics, Goodall argues that if subjects are instances of LD elements they should pattern alike in this respect, (56) and (57).

(56) *A nadie, Juan lo ha visto.

DOM no-one Juan cl<sub>MS</sub> has seen  

(Goodall, 1999, p. 99, (6))

(57) Nadie ha visto a Juan.

no-one has seen DOM Juan

‘No one has seen Juan.’  

(Goodall, 1999, p. 99, (7))

However, (56) is hard to evaluate because it may be ungrammatical not only because it includes a bare quantifier as a topic, but because there is a clitic that is specific that

---

29 These sentences become grammatical once the subject is interpreted as contrastive. In this case, the stress pattern of the sentence is different from a broad focus statement.

30 This type of evidence is based on Casielles-Suárez (1997) (as in Goodall (1999)).
makes impossible the co-reference with the negative bare quantifier.\textsuperscript{31,32} The sentence improves once the clitic is dismissed (58). Further evidence that there is a problem linked to the clitic can be seen by the fact that \textit{nadie} ‘no-one’ is impossible to cliticize when it functions as a DO, (59) and (60).

(58) \begin{align*}
A & \text{nadie}, \quad \text{Juan ha visto}.
\end{align*}
\begin{align*}
\text{DOM} & \quad \text{no-one} \quad \text{Juan has seen} \\
\text{‘Juan hasn’t seen anyone.’}
\end{align*}

(59) \begin{align*}
\text{Juan no} & \quad \text{ha visto a} \quad \text{nadie}.
\end{align*}
\begin{align*}
\text{Juan not} & \quad \text{has seen} \quad \text{DOM} \quad \text{no-one} \\
\text{‘Juan hasn’t seen anyone.’}
\end{align*}

(60) \begin{align*}
\text{*Juan no} & \quad \text{lo} \quad \text{ha visto} \quad (\text{where the clitic} \text{lo} \text{refers to} \text{nadie} \quad \text{‘no one’})
\end{align*}
\begin{align*}
\text{Juan not} & \quad \text{cl}_{\text{ACC/MS}} \quad \text{has seen}
\end{align*}

Leaving aside the case of bare quantifiers in preverbal position, whose syntactic behaviour is not conclusive, the distribution of bare nouns in Spanish does offer a different picture for preverbal subjects and topics. Even though it is not clear what kind of restrictions are playing a role in the distribution of bare nouns in Spanish, it is a fact that the preverbal position in broad focus contexts seems to be forbidden for bare nouns, while this is not the case for topics. Thus we can conclude that preverbal subjects in broad focus statements do not behave like topics in this context.

3.3.2.6. Subject positions

While the LD proposal claims that the specifier of inflection does not project in Spanish, Suñer (2003) proposes that it does not only project, but there may be various positions

\textsuperscript{31} See Suñer (2003) and references therein, for a similar claim regarding the specific interpretation of direct object clitics.

\textsuperscript{32} The same observation is done by O&T (1999) regarding Rizzi’s (1986) example.
within the IP layer for subjects to appear in. Building upon Carinaletti’s (1997) specialization hypothesis, Suñer argues that preverbal subjects in Spanish occupy different positions, as they do not behave alike when it comes to their distribution with respect to the verb. She points out that some subjects show adjacency effects, while others do not. The relevant examples follow, (61a-b).

(61)  

a. Juan/ Él, a mi parecer, es muy simpático.  

STANDARD SPANISH

Juan/ he, according to me, is very nice

‘In my opinion, Juan/he is very nice.’

b. Ello (*a mi parecer) no sería malo estudiar.  

DOMINICAN SPANISH

EXPL according to me, not would hurt to study

(Suñer, 2003, p. 351, (29a-b)).

Based on the previous cases, Suñer concludes that strong subjects like DPs and personal pronouns are in the upper specifier of IP, while weak pronouns like ello and pro appear in a lower specifier. She proposes the structure in (62).

(62)  

[TP {Juan/ él} [XP parenth. [TP {ello/ pro} Vfin

STRONG WEAK

(Suñer, 2003, p. 352, (30c))

Camacho (2006) also explores the idea of various subject positions, and adds new examples with negative quantifier subjects, which, according to his analysis, must also be adjacent to the verb. He contrasts the behaviour of negative quantifier subjects to the behaviour of lexical subjects, postverbal negative quantifiers and CLLDs. The following examples (63-65) illustrate the dissimilar behaviour of preverbal lexical subjects and preverbal negative quantifiers.

(63) Nadie pudo avanzar 3 metros.
no-one could advance 3 meters

‘No one could advance 3 meters.’ (Camacho, 2006, p. 57, (13a))

(64) *Nadie casi pudo avanzar 3 metros.

no-one almost could advance 3 meters (Camacho, 2006, p. 57, (13b))

(65) La tortuga casi pudo avanzar 3 metros.

the turtle almost could advance 3 meters

‘The turtle almost could advance 3 meters.’ (Camacho, 2006, p. 57, (14a))

From the previous examples Camacho concludes that preverbal subjects in Spanish do not occupy a single position. He proposes, like Suñer (2003), that the expletive *ello, and presumably pro should appear within the IP projection, to which he adds preverbal negative quantifier subjects. Among the elements that appear within the IP layer, he finds further differences in their distribution, as they do not behave alike in contexts with adverbs and parentheticals. He proposes that *ello is structurally closer to the verb than negative quantifiers, as (66) and (67) illustrate.

(66) Ello (*a mi parecer) no sería malo estudiar.

*Expl, according to me, wouldn’t hurt to study  (Suñer, 2003, p. 351, (29b))

(67) Ninguno, a mi parecer, pudo avanzar 3 metros.

no-one, according to me, could advance 3 meters

‘As far as I know, no one could advance 3 meters.’ (Camacho, 2006, p. 58, (17))

Even though the conclusion about the existence of different subject positions is a plausible one, example (64) that includes casi and a preverbal negative quantifier subject

33 However, unlike Suñer (2003), Camacho (2006) argues that lexical subjects are found in the left periphery, as already seen.
does not necessarily show adjacency effects. Furthermore, it may be ungrammatical for other reasons than the lack of adjacency between *nadie* and the verb. The following example that includes different types of adverbs shows that it is possible to separate *nadie* from the verb. Although adverb-placing data should be taken carefully, it is the case that the original example includes *casi*, for this reason is only logical to compare the behaviour of other adverbs in the same syntactic context.

(68)  Nadie nunca/ todavía/ ya/ ahora/ ayer/ desafortunadamente
      no-one never/ yet / anymore /now/ yesterday/ unfortunately
      puede/pudo avanzar 3 metros.
      can/could advance 3 meters

(68) shows that adverbs from different classes and with various meanings can intervene between *nadie* and the verb without rendering an ungrammatical sentence. (68) confirms that preverbal negative quantifier subjects can be separated from the verb; thus we can assume that (64) is ungrammatical due to other reasons different than the lack of adjacency between the negative quantifier subject and the verb.

A possibility may be that the ungrammaticality of (64) derives from the combination of the properties of negation and the meaning/ function of *casi*. It is well known fact that negative quantifiers when they appear in preverbal position take scope over the verb and function as negation. In such cases there is no need to use a negative particle such as *no* ‘not’ to convey the negative meaning. Furthermore, the use of such particle makes the sentence ungrammatical, as the following example shows.

(69)  *Nadie no avanzó 3 metros.
      no-one not advanced 3 meters

*Casi*, on the other hand, has similar properties to negation when it takes scope over the verb, but its meaning is richer. In addition to work as negation, when *casi* takes scope over the verb, it adds some type of modification, which in some cases could be equivalent to a type of quantification. Depending on the verb class (activity, accomplishment and
achievement) this property develops from the interaction between casi and the verb, or between casi and the object. In any case, the meaning of sentences that include casi is richer than those that include negation. These characteristics make casi a special kind of negative particle. The following examples illustrate these two properties of casi.

(70) La tortuga casi avanzó 3 metros.

the turtle almost advance 3 meters

‘The turtle almost advanced 3 meters.’

Interpretation: It did not cover the whole distance, but covered most of it

(71) La tortuga no avanzó 3 metros.

the turtle not advanced 3 meters

‘The turtle didn’t advance 3 meters.’

Interpretation: It did not cover the distance; there is no information about the distance

(72) La tortuga casi no avanzó 3 metros.

the turtle almost not advance 3 meters

‘The turtle could almost not advance 3 meters.’

Interpretation: It did cover the distance, but it was with a lot of effort, it took a really long time, etc.

The combination of casi with a preverbal negative quantifier like nadie results in the same type of ungrammaticality as the one shown in (64). This, however, does not constitute evidence of the existence of adjacency effects between nadie and the verb. If there were in fact adjacency effects they should manifest in other types of contexts and, as already seen, this is not the case. However, cases that show differences in the distribution of ello, strong pronouns and nouns support the idea that there could be different subject positions within the IP layer.
3.3.3. IP proposal: general comments

One of the advantages of the IP proposal is that it includes the possibility of preverbal subjects in the IP layer, as well as in the left periphery in Spanish. Given the differences between these two layers, and considering the interpretation of sentences from an information structure perspective, these two possibilities of subject realization seem more adequate in terms of the data that they are able to cover. Specifically, the IP proposal can account for the broad focus interpretation of sentences that include preverbal subjects in Spanish, while the LD proposal cannot. In addition, the IP proposal does not face any problems regarding the prediction of wrong word orders (VSO) as the default option in Spanish, as the LD account (O&T, 1999) does.

In terms of the syntactic evidence that the IP proposal examines, it also constitutes a step forward, both for the contexts that it covers, and its analyses. By examining the same syntactic contexts of the LD proposal, with some variations in the case of adverb types, for instance, the IP proposal shows that some preverbal subjects in Spanish behave as their English counterparts. It also provides further syntactic evidence that suggests that there is a difference between preverbal subjects and left dislocated topics, with respect to the distribution of bare nouns in the two contexts. Moreover, by examining the syntactic behaviour of different types preverbal subjects such as DPs, pro, and the expletive pronoun ello, the IP proposal demonstrates that it is not only possible that a position in IP layer exists in Spanish, but that various positions may be available within this layer.

However, the IP proposal is not without shortcomings. Some of its problems arise from the analyses of the data that they provide which are not always appropriate. Another problem has to do with not addressing some theoretical issues, or addressing them just partially. This, I believe, is the result of the type of approach that characterizes most of these works. Specifically, since the IP account is a reaction to the LD proposal, most researchers focus on demonstrating that the syntactic evidence that the LD account offers is not always correct or the only possibility. By the same token, they overlook, or do not elaborate on how the syntactic mechanisms of case assignment, and agreement, for instance, take place in Spanish. Additionally, there is variation regarding some explanations, and the analyses do not include a complete set of constructions in Spanish.
It is worth stressing, however, that the problem of including some only constructions concerns both accounts. The following approach incorporates some of the less common word orders in Spanish, and tackles some of these issues from a different perspective.

3.4. Spanish preverbal subjects: a sentential analysis

The sentential analysis states that movement to the preverbal position of some elements in Spanish is triggered by a sentence requirement, which has been identified as predication (Masullo, 1992; Fernández-Soriano, 1999a; Cuervo, 1999, 2003; Ortega-Santos, 2008). One of the main characteristics that traditional grammars signal about subjects is that the sentence states (or predicates) something about them. This is the well-known distinction between *thema* and *rhema*. Thus, putting this into context what this line of research suggests is that some elements that occupy a preverbal position in Spanish are interpreted or understood as subjects of predication. These subjects of predication, however, cannot be always equated to the nominative subject, or the element that triggers agreement on the verb, as we shall see. This approach, as already noted, supports McCloskey’s (1997) idea that subject properties do not necessarily convey on a single element, and may be distributed across various elements in some languages.

It is important to point out that while the original proposals (Masullo, 1992; Fernández-Soriano 1999a; Cuervo, 1999, 2003) developed specifically to account for non-nominative elements that appear in the preverbal field and seem to have subject like characteristics, Ortega-Santos’ (2008) analysis explicitly argues in favour of a semantic driven EPP in Spanish that applies to nominative subjects as well.

Unlike the LD and IP accounts, this type of approach takes as a starting point the interpretation of an element with respect to the sentence. This does not mean that the syntax is secondary, but taking an interpretational approach provides these researchers with the necessary tool that allows them to explain the preverbal presence of some elements different than nominative subjects, at the same time that it offers a plausible link between such elements and preverbal nominative subjects.
3.4.1. Subjects of predication: syntactic evidence

The initial analyses (Masullo, 1992; Fernández-Soriano, 1999a; Cuervo, 1999, 2003) hold that some non-nominative elements that occupy the preverbal position in Spanish have subject like properties, and propose that they occupy a position within the IP layer. Their claim is based on the parallel behaviour that these elements and preverbal nominative subjects display regarding their interpretation, and syntactic properties such as idiom formation, raising, extraction from coordinate structures, position in interrogatives, nominalization, etc., as the following examples illustrate.

In raising constructions, for instance, it is the dative or the locative element the one that raises, examples (73-75). This parallels the behaviour of regular nominative subjects. Masullo points out that like nominative subjects datives and locatives in these constructions cannot super-raise, example (76).

(73) Aquí parece sobrar/ faltar/ ocurrir algo
here seems to-be-extra/ to-miss/ to happen something

‘Something seems to be extra/ missing/ happening here.’

(Fernández-Soriano, 1999, p.108, (7a))

(74) A Juan parece habersele ocurrido una idea.
Juan\textsubscript{DAT} seems have.cl.cl occurred an idea

‘It seems that an idea has occurred to Juan.’ (Masullo, 1992, p. 124, (17a))

(75) A Adriana parece gustarle la música coral.

34 Masullo (1992) and Fernández-Soriano (1999a) offer syntactic evidence to demonstrate the fact that these preverbal elements are not in the left periphery, similar to what we saw in the case of the other syntactic analyses for nominative subjects. In the case of Masullo, he shows that dative subjects of psych verbs pattern with nominative agentive subjects regarding word order in wh- questions, quantification, extraction and raising from embedded clauses, adverbial modification (solamente ‘only’), etc. Fernández-Soriano also shows that preverbal locatives pattern with nominative agentive subjects in raising contexts, wh-questions formation, extraction from coordinate structures, nominalizations, their impossibility of being bare nouns, etc.
Adriana\textsubscript{DAT} seems to like cl the music choral

‘Adriana seems to like choral music.’ (Masullo, 1992, p. 123, (16a))

(76) *A Adriana\textsubscript{DAT} parece que Marcos cree que ti le gusta la música coral ti

Adriana\textsubscript{DAT} seems that Marcos believes that cl like the music choral

(Masullo, 1992, p. 123, (16d))

In interrogatives, datives and locatives behave like nominative subjects with respect to subject verb inversion, as (77) and (78) show.

(77) a.* ¿Qué Juan compró?

What Juan bought

b. ¿Qué compró Juan?

What bought Juan

‘What did Juan buy?’ (Masullo, 1992, p. 124, (18a-b))

(78) a. *¿Qué a Juan le gusta?

What Juan\textsubscript{DAT} cl likes

b. ¿Qué le gusta a Juan?

What cl likes Juan\textsubscript{DAT}

‘What does Juan like?’ (Masullo, 1992, p. 124, (19a-b))

Another piece of evidence comes from nominalizations. Fernández-Soriano (1999a) notes that in nominalizations in Spanish the subject and the direct object are preceded by the genitive preposition \textit{de}, while “other internal arguments (including datives) and adjuncts
retain the same preposition they take in the corresponding verbal constructions” (p. 111). (79) and (80) illustrate this.35

(79)  
a. El descubrimiento de América en 1492  
the discovery of America in 1492  
‘the discovery of America in 1492’  (Fernández-Soriano, 1999a, p.111, (12a))  
b. el paseo de Juan por el parque  
the walk of Juan by the park  
‘Juan’s walk on the park’  (Fernández-Soriano, 1999a, p.111, (12c))  
c. la entrega de premios a los ganadores  
the giving of awards to the winners  
‘the award ceremony to the winners’  (Fernández-Soriano, 1999a, p.111, (12d))

She further points out that in the subset of impersonal verbs that she analyses, locatives behave in the same way as the previous subjects.36

(80)  
a. la nevada de/ *en Sevilla  
the snowing of/ *in Seville  
‘the snowing in Seville’  
b. el suceso de/ *en Barcelona  
the event of/ *in Barcelona

35 These examples include some modifications in their translations.  
36 These examples include some modifications in their translations.
‘the event in Barcelona’

c. la entrega de premios *de/ en Barcelona

the presentation of awards *of/ in Barcelona

‘the award ceremony in Barcelona’

(Fernández-Soriano, 1999a, pp. 111-113, (13a-c))

With respect to the preverbal realization of these non-nominative elements, Masullo (1992) explains their presence as a requirement of the sentence to have a subject of predication. In other words, elements move to the preverbal field in order to fulfill a requirement that is related to the interpretation of the sentence at the discursive level. According to his proposal, every sentence must have a subject, but the notion of subject is understood as essentially related to the content of the sentence. Thus, elements found in the preverbal position are interpreted as the logical subject of the sentence. He identifies this property with an aboutness feature.

Since other elements than the traditional nominative subject that triggers agreement on the verb occupy the preverbal position in the IP layer, case is not assigned in a spec-head configuration. As a result, specTP is not a case position. Also, since the element that triggers agreement is not in the preverbal position either, Masullo proposes that specTP is not the position in which person and number features are checked. Hence, movement to such position is strictly motivated by predication. This way he accounts for the presence of non-nominative elements that appear preverbally (81-84). As specTP is a caseless and non-theta position in Spanish, it is a non-argumental one. Interestingly, this conclusion relates this proposal to the LD account, for the latter claims that preverbal subjects are left dislocated because they present A’ properties.

(81) A Marcos le gusta la música coral.

MarcosDAT clDAT please3S the music choral

‘Marcos likes choral music.’ (Masullo, 1992, p. 115, (1))
(82) A Marcos le cayó una piedra encima.

Marcos$_{\text{DAT}}$ cl$_{\text{DAT}}$ fell a rock on top

‘A rock fell on top of Marcos.’ \hspace{1cm} (Masullo, 1992, p. 115, (2))

(83) Aquí se come muy bien.

here cl eat$_{3S}$ very well

‘One eats very well here.’ \hspace{1cm} (Masullo, 1992, p. 115, (4))

(84) Nunca pasa nada en ese pueblo.

never happen$_{3S}$ nothing in that town

‘Nothing ever happens in that town.’ \hspace{1cm} (Masullo, 1992, p. 136, (61))

Fernández-Soriano (1999a) accounts for the presence of non-nominative elements in the preverbal position as a result of the need of fulfilling the EPP requirement. She also proposes a separation of the EPP from case assignment, and phi features checking. Like Masullo (1992), she argues that case and agreement are not checked in a spec-head configuration in Spanish. The following examples show the type of elements that appear in the preverbal position in order to fulfill the EPP (85-87).

(85) Aquí falta café.

here misses coffee

‘Coffee is missing here.’ \hspace{1cm} (Fernández-Soriano, 1999a, p. 103, (1b))

(86) En este impreso consta que eres el responsable.

in this form states that are the responsible.

‘In this form, it states that you are the responsible one.’

\hspace{1cm} (Fernández-Soriano, 1999a, p. 103, (1a))
While Masullo’s (1992) analysis was mainly developed to account for datives subjects of psych verbs that appear in preverbal position, Fernández-Soriano’s (1999a) was done for locatives and also datives in impersonal constructions that also appear preverbally. Both authors extend their analyses to other elements, and end up covering similar kind of data, and using the same type of syntactic evidence. Their explanations regarding the phenomenon itself, i.e., movement to a preverbal position, are also similar.

Similarly, Ortega-Santos (2008) suggests that movement to the preverbal field, specifically to specIP, in Spanish, occurs because of an EPP requirement of the sentence. However, he offers a reinterpretation of the EPP. Differently from more standard versions of the EPP that call for a nominal in T, he proposes that in Spanish the EPP requirement is a semantic one. This way he correlates internal merge (movement) with surface semantics. In his analysis such EPP requirement is present in all sentences alike, whenever the specific conditions arise. For economy principles, in cases in which it is not necessary (he names non-neutral information structure contexts), specIP does not project. He further argues that under this type of analysis, specIP is available not only to nominative subjects but to other elements as well. For instance, based on minimality constraints, he justifies dative experiencers’ movement to such position.

3.4.2. General comments about the sentential analysis of preverbal subjects

The most important contribution of the sentential analysis is the idea that other elements different than the nominative subject also have subject-like properties, and seem to be able to fulfill some kind of sentence/discourse requirement in Spanish, which has been identified as predication. Because of this approach, these researchers can account for a wider range of data that otherwise would be treated as an exception.

However, it is not clear how this predication requirement operates or develops. If predication is a requirement of every sentence in broad focus contexts, and predication
makes the interpretation of the sentence about the element that occupies the preverbal position, what happens in cases in which the sentence is not interpreted about the preverbal element itself but about the whole event or state? It would seem that proposing that predication is the trigger of the movement is too strong as it leaves out this possibility. It is not the case that all sentences with preverbal subjects or subject-like elements in Spanish correspond to a categorical judgement about such elements. There is ambiguity: they can be about the subject or they can be about the event. Moreover, how can we accommodate sentences in which there is no predication such as existential, impersonal, and meteorological constructions? A possibility may be that there could be some other requirement different than predication itself that is present. Nevertheless, the idea that other elements different than the nominative subject may fulfill whatever requirement there might be is quite relevant, because it allows accounting for some facts that otherwise would be left unexplained, or simply considered isolated cases or exceptions. At the same time, this analysis provides a link between nominative subjects and some subject-like elements. I will revisit these issues in Chapter 6 in the light of a different proposal based on Borer’s (2005b) Event Phrase.

3.5. Final remarks

After reviewing the different proposals we are left with a more general idea of subjects and elements that behave as such in Spanish. The diversity of these analyses along with the various approaches that they employ confirms that the subject category in this language is more complex than it is in languages like English and French. However, there are still unsolved issues and, in most cases, too many possible answers to the questions noted in the first chapter. Specifically, the possibilities regarding the syntactic operations of agreement and nominative case checking or assignment are rather abundant. Movement to the preverbal field has also received different explanations, some incompatible with each other, or too restrictive in terms of the predictions that derive from them. In addition, some of the specific claims of these works are difficult to reconcile with the current theoretical framework. Finally, most models only cover some data.
For instance, the works that address the syntactic operation of nominative case checking/assignment (the LD proposal and the sentential analysis) propose that it does not take place in a spec-head configuration in a language like Spanish. This is a problem if we consider valid the evidence that suggests that differently from agreement, case assignment should occur in a spec-head configuration (Bošković (2007) and references therein).

With respect to agreement, there are two possibilities with different specifications, from which, at least, four possible scenarios can be derived. The LD proposal alone offers three of such scenarios. Both A&A (1998) and O&T (1999) agree on the elimination of AgrS or specIP, which entails that this operation never takes place in a spec-head configuration in their analyses. A&A propose that all the information related to agreement is encoded in the verb. This entails that subject’s phi features are part of the numeration. Such features are later realized in the verb. This, according to them, also amounts to checking the EPP in this type of languages. On the other hand, O&T suggest that information related to the subject is encoded in a clitic like morpheme (a type of covert element with phi features), this way they delink such information from the subject DP. As noted, the motivation for disjoining this information from the subject DP develops from the ambiguity that can arise from plurals DPs in Spanish, which can show different person agreement.

Camacho (2006) who represents a mixed analysis (between the LD and the IP proposals) argues that there is no need to postulate a clitic like morpheme in order to account for such cases. Based on Poletto (2000) he proposes that what is needed is more detailed CP layer that includes functional projections that relate to the discourse, such as the notions of speaker and hearer which makes possible to recover any kind of information related to the subject when there is ambiguity. He also proposes that the agreement mechanism can be local or unbound in Spanish. In his analysis, local agree is dependent on whether a projection becomes active for agreement or not. Verb movement renders a projection agreement-active. He also suggests that the type of subject can also determine the type of agreement. For instance, subjects that display locality constraints such as the Dominican expletive ello ‘it’ trigger local agree given the configuration in which they appear.
Finally, Suñer (2003) implicitly assumes that agreement can take place either in a local configuration or long distance when she claims that *ad sensum* agreement is only possible with left dislocated subjects, while subjects in the IP layer reject it. Moreover, it seems that depending on the configuration in which agreement takes place, its results could vary. This is a problem because it entails that depending on the way in which agreement takes place we would have different results. If we take agreement (Agree) to be a universal operation by which a probe selects a goal and matches its properties, there should be no differences regardless of the position of the elements that take part in this operation. I will come back to these issues in Chapter 7.

With respect to movement, things are equally unclear. O&T (1999) propose that movement to the preverbal field occurs due to subjects bearing discourse properties, while in A&A’s (1998) proposal there is no movement since preverbal subjects are base-generated in the left periphery. The IP approach also suggests various motivations for subject movement. Goodall (1999), for instance, maintains the EPP as a feature of T that triggers overt movement, but widens its way of fulfilment by including not only DPs as possible candidates but also PP arguments. In his analysis the same head that checks the EPP feature also bears a *Quantifier* feature. This way he accounts for the fact that preverbal subjects do not co-occur with *wh*-phrases. For Camacho (2006), as already noted, subject movement is linked to verb movement. Lastly, for the sentential analysis movement to the preverbal field occurs because of a predication requirement, which in some proposals is equated with the EPP.

Leaving aside for a moment the multiple options that these accounts propose, there is a less explored route that could be used to test the two main claims regarding the positions that preverbal subjects occupy in Spanish (the IP layer and the left periphery), namely the relationship between syntax and prosody. Since there are two possibilities for preverbal subjects localization, the IP layer or the CP domain, and syntactic notions like topic and focus are also encoded prosodically in languages like Spanish, these proposals make different predictions with respect to the prosodic realization of elements depending on whether they are in the left periphery or the IP layer. If we assume that there is a correspondence between some syntactic categories such as topic and different type of foci and the special prosodic realization of the elements that bear such features in Spanish,
one could hypothesize that the prosodic realization of subjects in the two positions may be different given the syntactic differences of such positions. If this idea is on the right track, the prosodic behaviour of preverbal subjects could help us to disambiguate their syntactic status. Aside from Contreras (1976), Suñer (1982) and Zubizarreta (1998), who does not address subjects in particular but links the role of prosody to the variable word order in Spanish, there has not been any other proposal that tries to associate the theory of syntactic structure to the prosodic realization in Spanish. This is precisely what the experiment carried out in Chapter 5 is about, but before that a review of the theoretical model on which the analysis is based is in order. The next chapter covers the main syntactic theories about the interface between syntax and phonology, as well as the phonetic and phonological works carried out in Spanish that analyze the prosodic realization of elements in different contexts of information structure.
Chapter 4
When prosody meets syntax: on the syntax-phonology interface and previous research on Spanish prosody

As discussed in Chapter 3, there are two syntactic analyses about the localization of preverbal subjects in Spanish: the left dislocated proposal (Alexiadou & Anagnostopoulou, 1998; Ordóñez & Treviño, 1999), and the IP proposal (Goodall, 1999; Suñer 2003; Camacho, 2006; Ortega-Santos, 2008; Masullo, 1992; Cuervo, 1999, 2003, Fernández-Soriano, 1999a). Assuming a theoretical framework like generative grammar in which these analyses are developed, one can expect differences between elements located in the CP and the middle field with respect to not only their syntactic properties, but also their prosodic characteristics. Differences in the prosodic realization of left-dislocated and broad focus elements should arise (at the phonological level) as a manifestation of their syntactic and informational features in languages where prosody is used to encode informational differences. Spanish seems to be one such language, as different types of focalized elements display differences in duration, peak height and alignment, and intensity (Toledo, 1989; De la Mota, 1998; Face, 2001, 2002) with respect to broad focus elements. Taking into account this property of Spanish along with the different predictions that these theories make regarding the prosodic realization of preverbal subjects, a prosodic analysis of preverbal subjects could provide the empirical data that may help us to disambiguate their syntactic status. This is precisely what the study in Chapter 5 aims to do, but before doing that a review of the studies about the syntax-phonology interface is in order.

This chapter reviews and discusses the previous research done in the fields of syntax and prosody at both theoretical and experimental levels. These two types of analyses, although quite different, complement one another very effectively. On the one hand, syntactic analyses propose mechanisms that allow us to derive different sentential stress patterns, taking as a starting point the syntactic material or syntactic structure. On the other hand, prosodic analyses provide a prosodic characterization of the different types of elements, taking into account the syntactic notions of topic and focus.
The chapter is organized as follows. The next section presents a brief outline of the theoretical framework. Sections 4.2 to 4.4 comprise a summary of the research done in the area of syntax regarding the relation between syntax and phonology (prosody), in other words, how this relationship is captured by the grammar. In these sections I examine and discuss various syntactic models that account for this relationship at the interface level. Section 4.5 reviews the experimental studies that have been done in the field of prosody that are used as the starting point for my study, for their goals and their results, as well as for their methodology. The chapter ends with a summary of the prosodic properties of broad focus, focalized and topicalized elements in Spanish.

4.1. Theoretical background

Questions about the status of preverbal subjects in Spanish should be considered from the perspective of the interface between syntax and phonology within minimalist framework. The Minimalist Program (Chomsky, 1995, 2001) “explores the hypothesis that the language faculty is the optimal realization of interface conditions” (Hornstein, Nunes & Grohmann, 2005, p. 14). Since sentences have both form and meaning, which can be understood as the phonological expression of the words and their semantic interpretation, it is expected that “sentential outputs of the grammar ‘interface’ with systems that give them their articulatory and perceptual (A-P) properties and those that provide them with their conceptual and intentional (C-I) characteristics.” (Hornstein, Nunes & Grohmann, 2005, p. 8). Therefore, the only interface levels that appear in this framework are Phonological Form (PF) and Logical Form (LF), which interact with the cognitive systems that are responsible for A-P and C-I characteristics, respectively.

In the architecture of this framework there is a lexicon and a computational system that generates syntactic structures. These syntactic structures are sent to the A-P and C-I systems through PF and LF. I adopt the Y model of grammar, which is widely accepted within this tradition.

The computational system, the syntax, combines elements via two operations, merge and move (internal merge), in order to create complex syntactic objects. Elements that are used in the syntax come from the numeration, an array of items and features that are the material or the starting point of the derivation. These elements are part of the lexicon. All
elements that are used in the derivation have to be part of the numeration. This means that no extra items or features are added in the course of the derivation. In this model, all features have to be checked prior to the moment in which the syntax ships out the material to the PF and LF domains, otherwise the whole process, i.e., the derivation, crashes. Depending on the type of feature, the checking process takes place overtly or covertly. Strong features are checked overtly, and generate movement (internal merge), while weak features are checked covertly.

In this model the syntax feeds both the phonological and the semantic components, PF and LF respectively. As a result, there is no interaction between the PF and LF modules. Furthermore, once elements are shipped to PF and LF there is no communication in the opposite direction. A direct consequence of this is that no syntactic operations are triggered by phonological and/or semantic considerations. Another property of this system is that since all items and features come from the numeration they take part in the syntactic operations. This is relevant when we think about the notions of topic and focus as syntactically encoded features.

With respect to clause structure, I am assuming a structure with a small v (or voice projection), which licenses the external argument in transitive constructions, and above this, the IP and the CP layers. I am working with an articulated version of the CP layer (Rizzi, 1997). It is important to keep in mind that the CP is not only responsible for the type of sentence (declarative, interrogative, etc.), but also interacts with the discourse domain; therefore discourse related properties such as focus and topic are expected to be checked in this layer, either overtly or covertly. 37 I assume, at least, two projections in the CP layer: one related to topic(s), Topic Phrase (TopP), and the other related to focus, Focus Phrase, FocP.

Based on this model of the grammar, we can make some predictions. Since the notions of topic and focus are syntactically encoded (elements enter the derivation with such

37 Note that not all focus types are checked in the left periphery. Informational focus is checked in an intermediate focus projection within the IP layer. Traditionally, this type of focus corresponds to the last position of the string, once the elements are linearized.
features), and we know that all features have to be checked prior to spell-out (the moment in which the syntax ships out the material to PF and LF), then we expect that elements that are marked as [+Topic] and [+Focus] check their features in the CP layer, in the Topic and Focus phrases. Depending on the flavour of these features, weak or strong, they are going to be checked covertly or overtly, respectively. Let us take for instance the following sentence in Spanish.

(1) El vino lo trajo Juan.

the wine cl\textsubscript{ACC/MS} brought\textsubscript{3S} Juan

‘The wine, it was Juan who brought it.’

This sentence pronounced with neutral intonation is felicitous in a context where someone wants to find out who brought the wine. In this case, the DP *el vino* ‘the wine’ would be an example of old information (Lambrecht, 1996; Erteschik-Shir, 2007), which could be equated to a ratified topic in the sense of Hedberg and Sosa (2007), because it is already part of the discourse. As (1) shows, the DP *el vino* ‘the wine’ appears in the first position of the string.\textsuperscript{38} Since this DP has a topic feature to check, it moves to the CP layer to the TopP to check it. We know that this is so because the DP *el vino* ‘the wine’ appears in preverbal position, and transitive structures in Spanish in broad focus contexts display SVO word order, compare (1) and (2).

(2) Juan trajo el vino.

Juan brought\textsubscript{3S} the wine

‘Juan brought the wine.’

In addition to the fronting of the direct object, there is a clitic that bears the same case, gender and number features as the DP, before the verb. One of the characteristics of clitic

\textsuperscript{38} There is cross-linguistic evidence that old information tends to appear first, while new information appears later in the sentence. There are cases in which this tendency is overridden, but this by no means invalidates this generalization.
left-dislocated constructions (CLLD) in Spanish is this clitic doubling feature (Suñer, 1988). Regarding the position of the subject, there are two possibilities. The first one would be an intermediate Focus projection, the default informational focus phrase below the CP layer, which is reserved for new information focus and that usually corresponds to the last element of the string once the elements are linearized. Another possibility, according to Zubizarreta (1998), would be a focus projection to the right. I will not examine any of these possibilities here because they are not relevant to the present discussion, but the interested reader is encouraged to see Zubizarreta (1998).

Going back to the topicalization mechanism, it is a well-known fact that it is available to almost any kind of element in the sentence. The following examples in Spanish include different topicalized elements, (3) and (5), along with their broad focus counterparts, (4) and (6), respectively. Note that in (5) and (6) there are no differences in word order between a sentence that includes a topicalized subject and a broad focus sentence with a preverbal subject. In this case, however, the relevant question would be about the prosodic realization of these elements.

Context: Any news about María?

(3) A María la vi ayer.

DOM Maria clACC FS saw1S yesterday

‘Maria, I saw her yesterday.’

Context: What happened?

(4) Vi a María ayer.
saw$_{1S}$ DOM Maria yesterday

‘I saw Maria yesterday.’

Context: What’s the problem with Roberto?

(5) Roberto nunca piensa antes de hablar.

Roberto never think$_{3S}$ before talking

‘Roberto talks without thinking.’

Context: What happened?

(6) Roberto nunca piensa antes de hablar.

Roberto never think$_{3S}$ before talking

‘Roberto never thinks before talking.’

Since the syntax ships out the material to the articulatory system, which then realizes it phonetically, a long-standing question that generative grammar has aimed to answer concerns the relation between syntax and phonology, in other words, how these two domains interact. In the following sections, I summarize and discuss some of the main theoretical models of the interface between syntax and phonology.

4.2. Syntactic proposals of sentential stress

There are several proposals that explore the link between syntax and prosody taking a syntactic approach to this matter.  As will be seen, there is a structural progression from the first theories, which employ not only syntactic information but also some kind language specific phonological stress rule, to theories that rely only on structural facts.

---

41 Prosody and phonology are considered synonyms in this context, and are used interchangeably in this and subsequent chapters.
From the beginning of generative grammar, one of its quests has been to explain how stress is assigned at the sentence level. The goal has been to propose a mechanism of stress assignment that accounts for the regularity of sentential stress patterns found across sentences. The first attempt at doing so was Chomsky and Halle’s (1968) Nuclear Stress Rule (NSR henceforth), which aimed to capture the fact that in English the most prominent word is sentence final. The NSR and subsequent rules and models discussed in this section only apply to broad focus contexts, i.e., contexts in which all the information has the same status. After reviewing the proposals that deal with broad focus, I refer to some of the proposals that work with contexts in which one or more elements are informationally marked in section 3.3.

The NSR proposes a mechanism that takes syntactic constituents as its primary source and applies in a cyclic manner, assigning stress to the different elements of the syntactic constituents in order to generate the desired result of main stress or prominence on the final stressed word. This rule assumes that all content words bear stress at the word level. The procedure applies cyclically, taking syntactic constituents and assigning stress to each accented word. It always assigns primary stress to the rightmost word of each constituent, and reduces the stress by one on the other(s) word(s) of the constituent. In the following example, it can be seen that the first cycle applies to the VP, assigning primary stress to the object and reducing it by one on the verb.

\[
\text{[}_5\text{The cat } \text{VP [ate the cookie]}\]

\[\text{VP [eat the cookie]}
\]

\[2 \quad 1\]

The next cycle includes the subject. The rule takes into account the previous cycle, and assigns primary stress to the object again, reduces the stress on the verb by one, and assigns stress to the subject. The output of this rule is primary stress on the direct object, and secondary stress on the subject.

\[
\text{[}_5\text{The cat } \text{VP [eat the cookie]}\]
\]

\[2 \quad 3 \quad 1\]
While it is a fact that the NSR takes syntactic constituents as its primary elements, it also applies a specific English phonological rule of stress assignment: it assigns primary stress to the rightmost word of the syntactic constituents which means that it cannot be considered entirely syntactic. Another problem that this rule faces is that there are some contexts in English that do not display prominence on the final word. Unaccusative and passive constructions are cases in which it has been claimed that the main stress goes on the subject or first element of the sentence instead of on the final element, in broad focus contexts.\(^42\)

In order to accommodate these differences in stress patterns, various actualizations of the NSR were developed over the following years within the different frameworks of generative grammar. These proposals also take syntactic constituents and apply either stress-specific phonological rules, similar to the NSR, or some kind of parameterization, in order to capture language-specific properties that allow them to generate the different stress patterns found across languages. However, most of them are too powerful since they predict stress patterns that are not attested cross-linguistically (Kahnemuyipour, 2009). In addition to this, since they apply language-specific phonological rules, they cannot account for any kind of regularity even in languages that behave alike, which is an undesired result of any theory that intends to capture cross-linguistic regularities (Kahnemuyipour, 2009).

In the nineties, a more structural approach was developed based only on syntactic information. Cinque’s (1993) model is an example of this. Cinque’s proposal is based on the idea that all the information that is necessary to derive the unmarked stress pattern of a phrase comes from the syntax; therefore there is no need to postulate any language-specific phonological rule.\(^43\) His goal was to develop a null theory of phrase stress, as he called it. Furthermore, Cinque’s aim was to propose a model that could account for different languages with the same stress pattern, and also for different languages with different stress patterns. He based his model primarily on data from three languages:

\(^{42}\) It should be noted that there is not a unified opinion about the prosodic behaviour of unaccusative and passive derived subjects.

\(^{43}\) This applies to neutral focus declaratives or out-of-the-blue statements.
German, English and Italian, although he also includes examples in other languages that can be analyzed along the same lines.

Unlike previous models, Cinque’s model relies only on syntactic information, final surface structure and syntactic structure, and the Head Parameter to generate the different stress patterns found across these languages. According to Zubizarreta (1998), Cinque’s proposal was meant to capture a basic generalization about nuclear stress (NS): “NS falls on the most embedded element on the recursive side of the tree” (p. 85), which can vary depending on the head-initial or head-final parameter.

The main contribution of Cinque’s model is that it is the first instance of a stress assignment theory that can explain why there is such limited variation with respect to sentential stress patterns across languages (Kahnemuyipour, 2009). In addition to this, it does not suffer from the overgeneralization problem (Kahnemuyipour, 2009). However, Cinque’s system faces some theoretical problems when it is evaluated in light of new versions of the generative framework since it relies on the head parameter, which was dispensed with in subsequent work. Another issue for this model is that it does not preform well with sentences that have complex subjects, for instance subjects with various layers of embedding, as it predicts that the most embedded element is the one that receives more prominence. Another problem that has been noted by Kahnemuyipour (2009) about Cinque’s system is that it cannot generate the stress pattern of unaccusative sentences without relying on the notion of information structure. Cinque’s approach in this case seems inadequate because, in principle, these constructions are potential instances of broad focus contexts. However, while it is true that Cinque uses the term information structure, his explanation actually points to differences in the argument structure of unaccusative and transitive verbs, which are later realized in terms of information structure in some languages. These differences along with other language-dependent syntactic properties such as obligatory movement of the subject to a preverbal position in English, and movement optionality in languages like Italian and Spanish, in which such movement is linked, among other possible things, to the information structure of the sentence, make Cinque’s claim less objectionable. It is a fact that there is a mismatch between the stress patterns of unaccusative structures in English and languages
like Italian and Spanish, and any system that tries to generate a stress assignment mechanism for both groups will probably run into problems in one way or another.

Another proposal of the nineties is Zubizarreta (1998). One of the most valuable features of the theory that she develops is that it offers a mechanism that accounts for prosodic facts that seem to be related to word order and the notion of focus in two different groups of languages: Germanic and Romance. Her work is particularly relevant because it is partially based on Spanish data. The other Romance language is French. On the Germanic side, it includes English and German. German data was a problem for previous models of the NSR, among other things, because they could not accommodate the fact that verb final transitive sentences display sentential stress on the object and not on the verb, which is the final element (Kahnemuyipour, 2009). Regarding the significance of her proposal, it is important to stress that it not only tries to explain the type of data that was problematic for other theories, but also explores the possibility that nuclear stress is relevant to the identification of focus in Germanic and Romance languages.

Zubizarreta’s theory covers two different sets of data: broad focus and informationally marked contexts. This section discusses her proposal for broad focus contexts, while informationally marked contexts are analyzed in the following section. Her model for broad focus contexts also takes the NSR as a starting point and modularizes it into two rules. The first one is sensitive to asymmetric c-command and the other one is sensitive to selectional properties. These two rules operate depending on the language. Both rules are active in the Germanic group, whereas in the Romance group, the selectional rule does not operate and only the hierarchical version of the NSR does. She notes that in Romance sentential stress is always sentence final. In the case of Spanish, this final position is privileged for new information. This feature together with the free word order property that this language displays is the reason why there is a redistribution of constituents in cases in which the new information does not coincide with the last constituent in the primary word order.

One of the main criticisms that Zubizarreta’s proposal has received is that it is redundant (Kahnemuyipour, 2009). In her theory, there are two rules that are used to generate the stress pattern of Germanic SVO sentences. In addition to that, there is the Romance group
which displays the same prosodic pattern in these types of sentences and only one rule applies. Again, we are faced with another example of a model that tries to generate the stress patterns of languages with different syntactic requirements, and it is only able to do so by proposing two sets of rules. Leaving aside this problem for a moment, what is interesting in Zubizarreta’s theory, in my opinion, is that it constitutes a step forward with respect to Cinque’s (1993), since it can describe different stress patterns based on rules that take into account different syntactic properties, and this is a key issue. Her second rule that refers to the selectional ordering of the constituents points to the existence of something else playing a role in the computation of the nuclear stress, besides the already known relevance of hierarchical structure. The next two proposals discussed in this section capture this structurally dependent factor by incorporating the notion of phases.

Going back now for a moment to the case of passive and unaccusative constructions, there is a piece of structural information that is missing from the previous models of sentential stress. It is the case that subjects of both unaccusative and passive constructions are generated as complements of the verb. Although Cinque (1993) makes some claims about the differences between subjects of unaccusative and transitive verbs in terms of information structure/structural position in the tree, this syntactic information is not formally captured in his model. It is Legate’s (2003) proposal the first one to integrate this piece of information.44

Legate’s proposal is developed within the minimalist phase-based framework of the grammar and the multiple spell-out system that derives from it.45 She argues that VPs are

44 Legate’s (2003) work is not developed as a stress assignment theory; therefore, there are some details that are not fully worked out. To some extent it can be said that her work is done in the opposite direction. She takes the stress pattern of unaccusative and passive sentences as an argument to prove that not only CPs and vPs are phases, but so are VPs. In other words, she uses the stress pattern of these sentences as phasehood tests.

45 In a phase-based account, when there is a phase in the derivational process, the elements that constitute the syntactic complement of the Head of the Phase are sent off to PF, for phonological realization, and to LF, for logical interpretation. This type of procedure allows sending syntactic material in chunks to the two other systems instead of sending it all together at the end of the derivation. What is crucial in this system is the notion of phases itself, and that it is precisely what is different in the model that Legate (2003) proposes: unaccusatives and passives constitute phases in the same fashion as CPs and transitive vPs. The latter two are what most scholars consider phase-inducing categories, while unaccusatives and passives are usually regarded as non-phase-inducing categories.
phase-inducing categories based on the prosodic properties that subjects of unaccusative and passive constructions display. Since the subjects that appear in these constructions bear main stress, she proposes that their prosodic behaviour is due to the fact that they retain the stress that they receive in their base-generated position after moving to the preverbal field.

Kahnemuyipour (2009) criticizes Legate’s proposal because it potentially predicts that every element that moves from a stress bearing position into a position outside the phase should retain the stress that it receives in its original position. This, however, does not seem to be the case for wh-phrases and topicalized objects, which are considered to be unstressed elements, as he points out. Although a valid observation, I think the issue is not as straightforward as it seems. A pertinent question at this point would be about the differences among all these elements, and the final position that they occupy in the structure. Regarding position, if we compare subjects of passives and unaccusatives, on one hand, to wh-phrases and topicalized elements, on the other hand, we find that there is a difference between them. In the first case we have elements that do not belong to the CP layer, whereas wh-phrases and topicalized objects belong to this layer, the former as focalized elements and the latter as topics in the topic phrase. A possible consequence of the differences in their syntactic distribution could be some type of prosodic differentiation.

It is important to acknowledge that in this case we are faced with two different contexts of information structure. Let us leave aside for a moment wh-questions, and focus on subjects of unaccusatives and passives versus topicalized elements. It is a standard assumption that in English, subjects of unaccusative and passive constructions move to the specifier of IP in order to check the EPP, whereas topics move to the left periphery. The first context qualifies as a broad focus context in which just the NSR applies. On the other hand, topicalization is a process that involves movement to the left periphery based on the specific syntactic properties of the given element. One possible difference between these two contexts is that while in the broad focus scenario only one rule of stress assignment applies, the NSR, in the informationally marked context the computation of the stress may be different due to the fact that more than one rule could apply. This is precisely what happens in the case of focalization, as will be seen once we cover
Zubizareta’s (1998) and Kahnemuyipour’s (2009) models for these contexts. These researchers propose that in contexts that include a focused element, there is an interaction between the NSR and another rule (some type of focus stress rule). The outcome of the interaction between the two rules modifies the stress pattern of the sentence. In other words, the stress assigned by the focus rule overrides the initial stress pattern that the NSR assigns and, as a consequence, the prosodic realization of the elements also varies. With this stress-assigning mechanism in mind one can only wonder what happens in cases in which there is topicalization. Is there any specific rule that applies in these contexts that also modifies the original output of the NSR? If this were the case, the potential problem that emerges from Legate’s (2003) proposal simply goes away.

Another question relevant to this discussion is how different sentential stress patterns compare across languages; take for instance differences in stress patterns of unaccusative sentences in languages like English and Spanish. What do they tell us about such languages? Why is there a difference in prominence among these subjects and subjects of transitive verbs in some languages but not in others? More importantly, are these differences in prominence relevant to the syntax? While I do not intend to provide an answer to these questions, I believe that they can give us a broader perspective about the language system, and maybe help us to understand some specific prosodic manifestations within a bigger context. I revisit these questions in Chapter 5 in light of the data from my experiment.

The final proposal on sentential stress in broad focus constructions included in this section is Kahnemuyipour (2009). His model also offers an exclusively syntactic theory of sentential stress assignment. It takes phases and the multiple spell-out system as a starting point, similar to Legate (2003). However, differently from Legate’s proposal, in his model only transitive VPs, i.e. vPs, and CPs are considered phases. According to Kahnemuyipour’s theory, “sentential stress is assigned at the phase to the highest phonologically non-null element of the spelled out constituent” (2009, p. 68). This way he accounts for main stress on the direct object in transitive constructions. His theory uses the notions of asymmetric c-command and dominance to determine the highest element.
Kahnemuyipour’s proposal can generate the stress patterns of transitive, unaccusative and passive constructions alike without having to rely on the notions of information structure and/or base-generated positions, using only the syntactic structure and the multiple spell-out system architecture. In addition to this, his theory does not face any empirical problems with wh-phrases and topics. At the same time, it accommodates data from different languages, including German and Persian, which proved to be a problem for previous theories. This proposal, however, cannot capture the behaviour of Spanish unaccusative constructions, which do not display main stress on the subject (Zubizarreta, 1998), in contrast to English. A pertinent question at this point is why subjects of unaccusative constructions in English and Spanish display different stress patterns, as noted before. One possibility is that, if Kahnemuyipour’s proposal is on the right track, given the differences related to the obligatory movement of the subject to the preverbal field in English and its optionality in Spanish, when there is indeed movement in unaccusative constructions in Spanish, the final landing site of the derived unaccusative subject is not the same as their English counterparts. A possible hypothesis would be that the differences in the stress patterns in these cases would be the result of structural differences in these two languages regarding the positions that these elements occupy in the upper layers of the structure.

4.3. Information structure and sentential stress

This section includes two syntactic models of sentential stress assignment that do not correspond to broad focus contexts. In other words, contexts in which one or more elements are informationally marked or more prominent than the others. For the sake of clarity, I refer to these contexts as informationally marked versus out-of-the-blue or broad focus contexts.

There is a general consensus that informationally marked contexts show specific prosodic properties. These contexts display more variability (and also freedom) regarding their prosodic patterns when compared to broad focus contexts. This statement is true for

---

46 We already saw that the Legate’s (2003) potential problem is not as much so if we treat topicalized items within their specific informational context, and not as if they belonged to a broad focus construction.
languages that mark different types of foci prosodically.\textsuperscript{47} English and Spanish are examples of this type of language. Take for instance contrastively focused or narrow focused elements in the (b) version of the following sentences, along with a possible context in which they are felicitous (a). Capital letters are used in this case to indicate that the element is pronounced with a special intonation when compared to the prosodic realization of same element in a broad focus context.

(7) a. I think John brought his cat to the party.

b. No, he brought his DOG.

(8) a. Creo que Juan trajo refrescos.

\begin{align*}
\text{believed$_{1S}$ that Juan brought$_{3S}$ soft drinks} \\
\text{‘I believe that Juan brought soft drinks.’}
\end{align*}

b. No, Juan trajo VINO.

\begin{align*}
\text{no Juan brought$_{3S}$ wine} \\
\text{‘No, he brought wine.’}
\end{align*}

(9) a. Conociste a Luisa en mi fiesta, ¿no te acuerdas?.

\begin{align*}
\text{met$_{2S}$ DOM Luisa at my party, not you remember} \\
\text{‘You met Luisa at my party, don’t you remember?’}
\end{align*}

b. No, a MARÍA la conocí en tu fiesta, y a Luisa en casa de Manuel.

\begin{align*}
\text{no DOM Maria \text{cl}_{ACC} meet$_{3S}$ at your party and DOM Luisa \text{at home of Manuel} } \\
\text{‘No, I met Maria at your party, and Luisa at Manuel’s place.’}
\end{align*}

\textsuperscript{47} In this case, focus should be understood as narrow or contrastive focus, not as information focus, which is the focus that appears in broad focus contexts that can refer to one single word, but also to various constituents and even the whole sentence.
According to the different theories of stress assignment, the prosodic marking displayed by the elements of the previous examples is the result of an interaction between the stress pattern of the given sentence in the broad focus context and the existence of a focalized constituent with special prosodic properties. This is particularly noticeable in cases in which the linear order in both cases is the same, (7b) and (8b). The other case in which narrow or contrastive focus is marked via movement to the left periphery, sentence (9b), may also display prosodic marking in languages in which this option is available. If this is the case, the left dislocated element may show special prosodic properties as well. Spanish seems to be one such language, in which both movement and prosody are used to indicate the syntactic property of the focused element in the left periphery, as has been shown by some prosodic studies.\textsuperscript{48} The stress patterns of sentences (7b) and (8b) are examples of the interaction between the broad focus stress pattern of the sentence and the focus stress of their focalized elements. It is precisely the result of this interaction what syntactic models such as Zubizarreta (1998) and Kahnemuyipour (2009), among others, account for.\textsuperscript{49}

One common feature that these two proposals share is that they assume a correspondence between the primary stress of the sentence and the focalized constituent. In other words, the focalized constituent always includes the more prominent element of the sentence. Generally speaking, these models combine the NSR and another rule that refers to the focalized element. During the process of stress derivation, these two rules interact, and their interaction generates the stress pattern of the informationally marked sentence. Kahnemuyipour (2009), for instance, combines his Focus Stress Rule with his version of the NSR in order to generate the stress pattern of sentences that contain a focalized element. In his model, both rules apply in a phase-based manner.

\textsuperscript{48} See Domínguez (2004).
\textsuperscript{49} There are differences in the type of approach that different models assume regarding this interaction. Some models take sentential stress as a given, and from it they derive a mechanism that generates the special stress of the focused element. Other proposals take the focus structure as the starting point, and from there they derive sentential stress.
Focus Stress Rule: at the phase HP, mark a focussed subconstituent $C$ to receive focus stress. (Kahnemuyipour, 2009, p.129)

In Kahnemuyipour’s model, the two rules apply independently and their only interaction arises in the phonetic realization of the elements that are marked. There is no particular order in which the rules must apply. Since both rules apply, we end up with elements bearing focus stress and sentential stress. Sometimes the same element may bear both types of stress. Based on the correspondence noted between focus and prosody, the element marked for focus stress receives higher prominence in the sentence. In other words, focalized elements are always more prominent than non-focalized elements. Focalized elements are also more prominent than elements marked with sentential stress.

An interesting prediction that derives from Kahnemuyipour’s model, as he notes, is that there is a difference between sentences with narrow focus on the last element, and the same sentence in a broad focus context. Let us take for instance an SVO structure in a language like Spanish or English with a narrow focused direct object, and the same structure in a broad focus context. Based on his model, in one case the direct object bears focus stress, while in the other case it bears sentential stress. Since focus stress is different than sentential stress, there should be a difference in the prosodic realization of the direct object in these two contexts. Although this prediction is not borne out by Spanish data,\(^{50}\) and in the case of English the data seems to be inconclusive as Kahnemuyipour (2009) points out, there are languages in which the differentiation happens. Cinque (1993,) provides an example of a Polish DP that illustrates this fact, (10).

\begin{align*}
\text{(10)} & \quad \text{a. [waznosc [komunikacj[i SAmochodowej]]]} \\
& \quad \text{b. [waznosc [komunikacj[i samochoDOwej]]]} \\
& \quad \text{‘the importance of traveling by car’ Cinque (1993, p. 260, footnote 23, i (a-b))}
\end{align*}

---

\(^{50}\) See Zubizarreta (1998) and Domínguez (2004).
According to Cinque, only (b) can be understood as a broad focus statement. In this case, any of the bracketed elements can be interpreted as new information. On the other hand, (a) can only be interpreted with focus on the last element, i.e., only the last element is focused. The difference in this case arises from stress placement in both constructions. According to Cinque’s consultant, standardly, Polish places stress on the penultimate syllable of the most embedded constituent. When stress is changed to the first syllable of the word, as in (a), the element is interpreted as focused.

While I think it is interesting that a difference like one this exists, it seems that its actualization in the prosody has more to do with the specific inventory of pitch accents and/or their different combinatorial possibilities that a given language may have, or the prosodic cues that languages use to encode syntactic differences, than with the syntax itself. Let us consider another instance in which a syntactic theory aims to distinguish between different types of foci. Zubizarreta (1998) argues that there is a difference between contrastive focus and narrow focus with regard to their syntactic treatment and their prosodic realization. Her proposal has been criticized on the basis that such differentiation is not relevant (Kahnemuyipour, 2009). The main issue here is that some theories suggest that there are no semantic differences between these two types of foci. It is a fact, however, that there is a difference in the prosodic realization of these two types of foci in Spanish, and it seems that in Italian as well, as per Zubizarreta’s (1998) claims. The prosodic behaviour of these two languages with respect to the different realization of these two types of focus, along with the previous prediction of Kahnemuyipour’s (2009) model, raise an important question about what type of information should be considered part of the syntax and what information is just a by-product of a given language system that coincides with some semantic properties, but not relevant to the syntax itself.

Zubizarreta’s theory of stress assignment for informationally marked contexts is also based on the Focus Prosody Correspondence Principle. As already noted, this principle

51 There are several studies in Spanish that show that there are prosodic differences between narrow and contrastive focus (Face, 2001, 2002).
states that in a sentence the constituent that is marked for focus must contain the most prominent word. Based on this correspondence she proposes the Focus Prominence Rule according to which given two sisters, the one that is marked [+F] is more prominent than the other one. Her model also works with her modularized version of the NSR, which, as the reader probably recalls, has two versions that apply depending on the type of language. In her model, as well as in Kahnemuyipour’s, all rules apply. In cases in which there is contradiction in the output of the rules, different languages use different mechanisms to resolve it. For English and French Zubizarreta proposes that non-focused material is considered invisible; thus the focalized material is always more prominent. Technically speaking, this mechanism allows placing narrow and contrastive focus stress on any element of the sentence in any given position, while deaccenting the rest of the material. In Spanish and Italian, on the other hand, all material is prosodically visible, which leads her to propose that defocalized material moves to a higher position in the structure and the focalized element remains in a position where it can receive stress by the NSR. This is what she calls p-movement, in which “p” stands for prosodic.

Zubizarreta’s (1998) proposal for languages like Spanish and Italian has attracted much criticism because it seems to be incompatible with the Y-model of the grammar standardly assumed in the minimalist program (Ortega-Santos, 2008; Kahnemuyipour, 2009). The fact that a syntactic operation such as movement is motivated by prosodic requirements, i.e., the need for the focalized element to coincide with a specific position so it can receive stress in it, seems conflicting with the principle that the relationship between syntax and prosody goes only in one direction: from syntax to prosody. According to this property of the system, no syntactic operation should be triggered by prosodic considerations.

While at first sight this looks like a classic example of PF tampering with the syntax, I believe that this is not the case. The first thing worth noting is that Zubizarreta’s (1998) proposal aims to capture the fact that in Spanish there is a redistribution of the elements of the sentence when something different than the last element bears some type of informational focus. In other words, while some languages use deaccenting (i.e., the non-focused material becomes metrically invisible (Zubizarreta, 1998)) and have across-the-board freedom to assign stress to any element, other languages use movement. Whether
or not the position in which the narrow focused element ends up being—due to movement of the non-focalized material to another position, and probably its own movement to a focused position—is privileged in terms of stress assignment, it is a fact that cannot be ignored. However, motivating movement by the need for a given element to appear in a stressed position is not entirely necessary. The proposal could capture the same facts by stating that non-focalized material moves into a given position to check its features (let us assume some type of topic position because it is old information), and the focalized material moves to an intermediate focus phrase which has been claimed to exist (this is the default focus phrase which the NSR aims to capture). The results of such movements in terms of stress assignment are the same without compromising the proposal’s validity in the minimalist framework by suggesting that the redistribution takes place in order to align a given element and a position, based on the stress properties of the position. This is actually what Zubizarreta’s proposal intends to convey when she states that “p-movement feeds the NSR (…) Therefore its point of application depends on where the NSR applies. (…) I suggested that the NSR applies at the end of the syntactic derivation. This implies that p-movement (…) must also apply in the syntax” (p.141). Since this type of movement only takes place when the element is narrow focused, one can speculate that in Spanish narrow focus has a strong feature, therefore any element that bears it should check it overtly before spell-out. In other words, the element has to move into a specific syntactic position in order to check the given feature. This way, movement would be obligatory. By the same token, this would be the default mechanism in this language to convey narrow focus.52

52 There is debate with respect to contrastively focused elements, and their differences with narrow focus in Spanish. It is a well-known fact that contrastively focused elements can appear in any position in Spanish and receive stress there, i.e., they can stay in situ. This property of contrastively focused elements by no means invalidates Zubizarreta’s (1998) proposal about narrow focused elements. Furthermore, contrastively focused elements can also move to the left periphery and receive a special stress there. The difference between these two foci with respect to movement and also stress patterns is what Zubizarreta’s proposal aims to capture.
4.4. Syntax-phonology interface and models of sentential stress assignment

To sum up, the goal of all previous models is to account for the relationship between syntax and phonology at the interface level within the generative framework of grammar. They do so by proposing mechanisms that allow them to derive sentential stress from the syntactic structure and/or the syntactic properties of the different elements of a given utterance. One of the main features of the minimalist framework is that the syntax feeds the phonological component. Another important feature is that, in this framework, notions like topic and focus are syntactic properties; thus they are computed through the derivation. This is valuable because, as we have seen, these notions are relevant to the computation of stress as well.

One of the advantages that this theory of the grammar has is that it can propose a mechanism that describes the interface between two modules of the grammar, syntax and phonology, without having to rely on stipulations. Since the main elements that can have an impact on stress assignment, at least primary stress, are syntactically encoded, the information for the computation of sentential stress comes from the syntax.\(^{53}\)

One of the shortcomings, however, is that different theories of sentential stress, in the spirit of incorporating particular properties of a given language, may be proposing differences that may not be relevant to the syntax itself. I think this is a complex issue and given the fact that we do not have all the information regarding all languages and how these properties are encoded, there are two possibilities.

\(^{53}\) Some of these theories also try to derive the secondary stress of the sentence: the second most prominent element in the sentence. However, the case of secondary stress is more complex than primary stress, as there are various factors that intervene in its computation. Some models such as Chomsky and Halle’s (1968) Nuclear Stress Rule and Kahnemuyipour (2009) predict secondary stress on the subject of transitive constructions. Other theories such as Zubizarreta (1998) do not deal with secondary stress. The reason is that, as Zubizarreta correctly points out, secondary stress not only depends on the syntax. She argues that, unlike primary stress, which is linked to the notion of focus in the phrase, secondary stress cannot be predicted from the syntactic structure. She mentions various studies within the phonological tradition that demonstrate that secondary stress can be moved around the sentence depending on factors such as syllabic quantity, constituents size, etc. The interested reader is encouraged to see Zubizarreta (1998) for details.
First, it could be the case that in fact these systems are incorporating differences among sub-types of elements that are not relevant to the syntax, for instance, differences among types of foci, as we previously saw. One way to solve this problem is to propose that the only two categories of information structure that are primitives of the syntax are the notions of topic and focus (Erteschik-Shir, 2007), and that their various subclassifications can be derived from features that those elements may also bear. In the same vein, another possibility would be to propose that differences among different types of foci come from the actualizations of these two notions in the pragmatic context in which a given utterance is evaluated, since it has been suggested that some of those distinctions come from the implicatures that arise from a given context (Büring, 2007).54 The problem with this theoretical move is that some of the prosodic properties of those elements would have to have been derived after an evaluation of the context in which they appear. This extra step complicates the process, and would require an additional communication between the two modules. Furthermore, the different behaviour regarding the locus that these elements occupy would be a bigger problem for this option, since the syntax would have no way to handle differences in their distribution. Therefore, the idea of having only two primitives with extra features is the preferable option at this point. It also offers a more adequate treatment of the phenomenon of focus as a continuum, in addition to the fact that it complies with the minimalist spirit.

The second interpretation of the differences among sub-types of elements could be that they are indeed part of the grammar. It could be the case that there are semantic differences among these elements that are relevant to their distinction, and those semantic differences translate into different syntactic elements as well (Zubizarreta, 1998). One of the arguments that could be put forth to prove the validity of the syntactically encoded version of the differences among these elements is that in some languages they display distinct prosodic realizations. In Spanish, for instance, there are differences between the prosodic realizations of contrastive and narrow focus. Another example is Polish, which distinguishes between narrow and broad focus (Cinque, 1993). Unfortunately, the same

---

54 Büring (2007) points out that although information structure categories do not affect the truth conditions of utterances, they do affect aspects of meaning such as implicatures and felicity conditions.
prosodic argument goes in the other direction, since there are languages that do not distinguish among some of the different sub-types. There is Spanish, again, which does not distinguish between narrow and broad focus when it refers to the element that occupies the last position of the string (the opposite case of Polish). Furthermore, English seems to be another language that behaves alike, as there is no conclusive data that proves that these contexts are prosodically different.

I believe that in this particular instance prosody may not be the best indicator, as there are different possibilities that go in both directions even with the same type of elements depending on the language. As stated before, these types of prosodic differences have more to do with the specific prosodic properties of a given language, than with actual differences among syntactic elements and/or categories. The fact that in some cases they coincide with specific sub-types is something that cannot be ignored, but proposing syntactic differences based on the prosodic behaviour is at best uncertain at this point. A more reasonable approach to this matter is to explore a syntactic analysis that includes two primitives only: focus and topic, and try to accommodate the different behaviours of the different elements within it in terms of different features: [+/-Contrastive], [+/-Informational], etc. The next section, which summarizes the prosodic analysis of different types of informationally marked elements, illustrates the wide variation regarding their prosodic marking, which corroborates the idea that at this point it is more accurate to propose parallels that only relate to the two primitives.

Although these syntactic proposals do not say much about topicalized elements, they can be analyzed along the same lines as focalized elements based on this theory of the grammar. One of the reasons why topics are scarcely mentioned in these analyses is that, in most cases, they are considered old information, and have been claimed to be unstressed.\textsuperscript{55, 56} Thus, in principle, there is no need to generate any special mechanism of

\textsuperscript{55} The unstressed claim is valid for languages in which the phonological mechanism of deaccenting is available. English has been suggested to be one such language. In the case of Spanish, deaccenting was not considered to be possible, and this is one of the reasons why Zubizarreta (1998) proposes a connection between word order in Spanish and the expression of focus. Even though deaccenting may not be possible in Spanish (some experimental studies have challenged this claim), or at least may not be the default mechanism, one can expect that elements that are old information do not behave like elements that are new
stress assignment for elements that do not bear any stress. However, the prosodic analysis of topics can also be informative in terms of differences among topic and non-topic elements, such as broad focus elements. Also, as we have already seen, topicalization may have an impact in the computation of sentential stress as well.

4.5. Previous studies on Spanish prosody

Among the studies that investigate the prosodic properties of elements that encode the notions of topic and focus we find a wide range of variation if we consider their goals. Although in all cases these studies look at the prosodic properties of utterances and/or words, their research spectrum goes from acoustic and/or phonological descriptions of prosodic cues to signal broad, narrow or contrastive focus, to the possible relation between syntax and prosody. In order to facilitate the summary and discussion of all these works and their findings, I divide the studies in two groups.

The first group comprises works that provide a description, either phonetic or phonological (or both) of the intonational properties of a given element, more in the line of information structure analyses from a prosodic point of view. The second group includes works that seek to investigate the possible relationship between syntactic

_Information in terms of their prosodic characteristics. As a general observation we can say that there seems to be a trade off between word order and the deaccenting mechanism. Languages with more rigid word order favour deaccenting, while languages that display a freer word order may not use it. It is clear, however, that these two extremes can be stretched out: English allows some variation in word order and Spanish seems to also allow deaccenting as per Face’s (2003) findings. According to Domínguez’s (2004) data, deaccenting in Spanish is available to postfocal material only when the elements are presupposed. Even though there may be common trends among languages, I believe that its properties are linked to the specific prosodic features of a given language. Part of the idea of topics being unstressed elements comes from the fact that some topics (old information topics) are mainly pronouns, which substitute for items that have been already mentioned in the discourse. We know that pronouns are unstressed unless they are contrastive; therefore, topics are unstressed. However, there are other types of topics that are not necessarily pronouns, and might not bear the same prosodic properties. (See Hedberg and Sosa (2007) for a prosodic characterization of topics in spontaneous speech in English.) Another important issue when talking about deaccenting is that, depending on the position that the element occupies in the string, it may be realized differently. In German, for instance, when topics are realized in prefocal position they display shorter duration and a flatter F0 contour than the same elements in the same position in a broad focus context. On the other hand, when topics are realized in a postfocal position, they display about the same duration and F0 pattern as the same element in the same position in a broad focus context. See Féry and Kügler (2008), and Kügler (2008) for details about topics in German.

In this case, the research is mainly based on phrasing patterns and their possible correspondence with syntactic constituents._
constituents and prosody, specifically with regard to phrasing. This division does not entail that the analyses in the second group are different than those carried out in the first group with respect to their methods (although they may differ), but their research questions are. In addition, it is important to note that the first group of studies also mentions syntactic phenomena as focus constructions are part of the syntax, but most of them do not integrate a syntactic analysis.

4.5.1. Prosodic correlates of focus and topic

In this section, I focus on the group of works that investigate the prosodic correlates of narrow and contrastive focus in Spanish. Since my study includes the analysis of topicalized elements, we expect to find some type of similarity between their prosodic realization and the realization of contrastive or narrow focused elements, which also bear specific informational properties. This statement should not be understood as if the prosodic characteristics of topicalized and focused elements are the same, but more in the line that there should be prosodic differences between the realization of elements that bear discourse-related properties such as focus and topic, on one hand, and the prosodic realization of elements in broad focus constructions, on the other.

The most general finding of this group of studies is that there is a difference between narrow and contrastive focus constructions and broad focus constructions. Another important finding, however, is that there is not a unified strategy in Spanish to signal narrow and contrastive focus across dialects, and even within the same dialect there is variation, which makes this phenomenon very complex in terms of its prosodic analysis.

One of the first experimental studies that investigates the realization of narrow focus in Spanish, Toledo (1989), shows that intensity is the most important prosodic cue to indicate it. The dialect under research was the Spanish spoken in Buenos Aires, and the variables that it looked at are fundamental frequency (F0), intensity and duration. This study was an acoustic analysis of sentences in which the focal element was determined by contextual variation. The material was obtained through a reading task and included 18 different questions with their respective answers. Speakers were recorded reading a context and a question about it, and answering the question with a sentence that was also provided. Speakers were able to decide how to realize prosodically the different elements
based on the given contexts. There were a total of three male speakers that participated in the study.

Results show the increase of intensity to be the most systematic prosodic parameter to signal narrow focus in this dialect. Pauses and duration are not prominent cues, and differences in the F0 are only significant for two of the speakers. Although a relatively small study, Toledo (1989) sets an important starting point for this line of research, and gives a good overview of the complexity of the prosodic system in Spanish. As this study does not control for verb types, it has a significant shortcoming with respect to the syntax. Since it treats constructions that are not syntactically equivalent as if they were, part of its results could be affected by this treatment.

Studies done in Central Peninsular Spanish that looked at the realization of narrow and/or contrastive focus show different results. De la Mota (1998) and Face (2000, 2001, 2002) found that duration is affected in both cases of narrow and contrastive focus. In these studies duration is the most prominent cue that speakers used to signal these types of foci. In Face (2001), there is a lengthening of the stressed syllable in narrow focused words, which basically corroborates De la Mota’s (1998) findings.

These studies also show variation across speakers when it comes to signalling focus. De la Mota found that F0 peaks in focused words were higher than those found in the same words in broad focus constructions. In contrast, Face (2001) shows that differences in F0 peaks are not statistically significant between narrow and broad focus constructions. This contradiction is solved by Face (2002) who does a more refined analysis and notes that there are two strategies to mark narrow focus in Spanish. The first strategy is peak alignment: the peak is realized within the stress syllable in focused words vs. its postonic realization in broad focus contexts. In this case, the height of the peak does not vary with respect to broad focus, only its location does. The second strategy is to increase the height of the peak, but the peak itself is realized in a postonic position. This is very interesting

58 This late peak alignment is the same configuration that appears in broad focus words. The height of the peak, however, is different.
because we see a trade-off between height and alignment in these two strategies that signal narrow focus.

In addition to this, Face (2001) found that intensity is not a cue used in this dialect, differently from Buenos Aires Spanish. This last result points to the existence of cross-dialectal prosodic differences in Spanish. Overall, however, these results support the idea that despite dialectal differences, informationally marked elements are prosodically signalled.

Another study done in Central Peninsular Spanish is Domínguez (2004), who also examines different types of focus and their prosodic realization. Her work includes an analysis of preverbal subjects in broad focus sentences and preverbal contrastively focalized subjects, as well as other elements in the left periphery. The data in this case also comes from a reading task. This work is particularly interesting because it establishes a direct dialog with the syntax with respect to the mechanisms of focus signalling at the same time that it explores their different prosodic correlates.

Domínguez (2004) found that preverbal subjects in Spanish, regardless of their informational status, bear the highest peak in the sentence. In addition to this, preverbal subjects that are contrastively focused show an increase in the F0 height with respect to preverbal subjects that appear in broad focus contexts. On the other hand, fronted focalized objects do not show the highest peak of the sentence. The author explains the differences in the prosodic realization between contrastively focused subjects and fronted focalized objects based on the different positions that they occupy in the structure. She argues that unlike contrastive initial subjects, which appear in the same syntactic position

---

59 This result is in line with the attested crosslinguistic trend of downstepping or declination in both tonal and non-tonal languages (Ladd 2008; Nooteboom 1997; Prieto 1997, among many others). In a broad focus utterance, the general trend is for the fundamental frequency to drop over time, which means that every time that a pitch accent is realized it is going to be lower than the previous one. This tendency, of course, can be modified depending on the type of utterance. There is no declination in some types of questions, for instance, and when there is a contrastively focused element in the middle of the string, its pitch accent is going to be higher than the previous one, precisely because it is something very salient, as it is not expected. However, in the case of broad focus utterances, the trend for the F0 is to be lower at every subsequent point in time. Therefore it is generally expected that the first pitch accent of the utterance is the highest one.
as preverbal broad focus subjects, fronted objects appear in initial position due to the result of focus movement. She suggests that since the canonical word order is altered in the case of fronted objects there is no need for prosodic marking, differently from contrastive initial subjects which display the same word order as broad focus preverbal subjects, and receive special prosodic marking in situ.60

Domínguez’s results regarding contrastive focus are expected. Contrastive focus is the most salient case; therefore anything that is compared to it should display less prosodic prominence. This goes in line with what previous studies have already found, and confirms the idea that there is a gradation in focus marking from the most salient case to the least salient one. The puzzle concerns to focused initial objects, which do not display the highest F0 of the sentence despite being in the first position of the string and being focalized. This prosodic behaviour of fronted objects points to the idea of a possible trade-off between word order and prosody, which seems likely in a language with the possibility of flexible word order. An interesting scenario to analyze would be one that included a contrastive fronted object. In this case, there would be movement to the preverbal field and probably prosodic marking as well. A pertinent question that derives from it is about the mechanisms of focus signalling, presumably dependent on focus type on this model.

With respect to left dislocated elements, Domínguez (2004) found that there is no pause between them and the rest of the sentence, which she points out as an interesting result because of the tendency to represent these elements separated with a comma from the rest of the utterance in Spanish. Based on this result, she proposes that the whole utterance is pronounced as a single intonational phrase. This last claim should be taken cautiously because there are other cues that can signal a phrase boundary such as pitch reset (Frota et al., 2007), final syllable lengthening (Face, 2002; Frota et al. 2007), the presence of boundary tones, etc. that the author does not take into account. Furthermore, pauses are

60 This supposes different mechanisms of focus marking depending on the focus type. According to this proposal, narrow focused elements need to move to check their features, while contrastively focused elements stay in situ and receive prosodic marking in that position. A direct consequence of this theoretical take is that preverbal subjects in broad focus constructions and contrastive focalized initial subjects appear in the same syntactic position.
not common in laboratory speech, as most of the studies on intonation suggest (Frota et al., 2007; Face, 2001, 2002, 2003; Prieto, 2007). Consequently, linking phrasing exclusively to pauses would limit the results to one of the possibilities. Unfortunately, the study does not provide data regarding the behaviour of the F0 of left-dislocated elements. Another finding of this work is that in Spanish neutral declaratives (i.e., broad focus constructions) do not differ prosodically from sentences with narrow focus on the last element with respect to their F0. This behaviour is captured in Zubizarreta’s (1998) proposal. Although this is a pilot study, for both the amount of tokens and the number of participants, its results suggest some important trends about Spanish prosody and the realization of discourse marked elements that should be further tested.

4.5.2. The phonology of focus marking

At the phonological level there are several proposals regarding focus marking in Spanish. Some of these proposals can be found in the studies previously mentioned, Face (2001, 2002), De la Mota (1998), Domínguez (2004), but also in Hualde (2002), Niebert (2000), among others. Most of these scholars have pointed out that in cases of narrow focus the F0 peak tends to be realized within the stressed syllable of the focused word, a phenomenon labeled as early alignment, in opposition to late peak alignment, which is the realization of the peak after the stressed syllable. Late peak alignment is commonly found in broad focus contexts.61

There are two theories that try to explain early peak alignment. The first one proposes the existence of a different pitch accent in cases of narrow focus L+H*, Face (2001), in contrast to the pitch accent used in broad focus constructions L*+H. The other proposal, Hualde (2002), states that there is just one pitch accent but in cases of narrow focus there is an intermediate boundary after the focused word which forces the peak to be realized within the stressed syllable. Face (2002) points to the possibility that multiple accounts arise because there are multiple prosodic strategies for marking focus in Spanish. The

---

61 Unlike most of studies, Domínguez (2004) reports a rate of 40% of early peak alignment in broad focus contexts.
existence of these different strategies, according to Face, is the main cause of disagreement among previous studies.

Face (2002) identified two strategies for marking narrow focus. The first strategy consists in the realization of the peak within the stressed syllable (early peak alignment). In this case the height of the peak is no different than peaks in broad focus words, but the position in which it is realized changes. The second strategy is to increase the height of the peak, but the peak itself is realized postonically, in the same position as it would be realized in broad focus. Changes in height were also a cause of disagreement among Face’s (2001) results and De la Mota’s (1998). Doing an item analysis allowed Face to tease apart the two strategies, and to further corroborate some of the previous results that have been proposed for this language in this context.

4.5.3. Phrasing

Next I review studies that focus on the relationship between syntactic and prosodic domains. Among other things, this type of research aims to determine whether the domains of intonational phrasing are dependent on phonological or syntactic considerations. One of the starting points of this type research are the different phonological phenomena that seem to be influenced or affected by syntactic contexts, such as French liaison, sandhi rules in various languages, vowel deletion, etc.

Several theories try to capture the relationship between syntax and prosodic domains. Elordieta (2007) notes that there are strictly syntactic proposals, like the Direct Reference Theory (Kaisse, 1985), which states that syntactic relations such as c-command constrain the domains in which external sandhi rules apply, while other proposals represent a mixed approach. Elordieta (1997), for instance, combines syntax and morphology. In his theory phonological domains are derived from morphosyntactic feature-checking relationships. Another important model within this field of research is the Prosody Hierarchy Theory (Selkirk, 1980; Nespor & Vogel 1986), which establishes different prosodic levels that are mapped from morphosyntactic structure. In this theory, there is

62 De la Mota (1998) reports an increase of peak height in narrow focused words in Spanish as compared to broad focus.
not an isomorphic correspondence between syntactic structure and prosodic domains: there are prosodic levels such as the syllable and the foot that do not have syntactic counterparts, and others, like the phonological word, do not always coincide with one syntactic element. Furthermore, syntactic constituents do not determine the domains of the application of phonological rules, as Elordieta (2007) points out. Finally, more recent theories like Dobashy’s (2003) combine phonological phrasing and the multiple spell-out system of the minimalist framework. In this theory, phonological phrases are mapped from syntactic phases (Elordieta, 2007). Based on the architecture of the minimalist framework, this theory proposes that the subject is mapped together with C in SVO sentences. Thus the prediction is that in all languages that display SVO word order, the subject will be an independent phonological phrase.

Just by briefly reviewing some of the main theories of this field, we see that there is variation regarding not only the framework that they develop, but also the theoretical stance that they assume and, as a consequence, the results that derive from it. It is important to note that all these theories display various degrees of success with respect to the type of data that they intend to account for, and their ample variety is precisely due to the need to explore different possibilities in order to cover different types of data. Equally important is that most of these theories recognize that extralinguistic factors such as speech rate can also intervene in phrasing strategies. Other factors, such as syllabic quantity and the weight of the constituents, can also influence phrasing patterns. One thing that is evident at this point is that the relationship between syntax and prosodic domains is not clear, and by no means is one-to-one.

Before getting into the description of Spanish phrasing patterns, I would like to illustrate this last point by some conclusions based on the one-to-one correspondence approach. Specifically, I refer to some works that establish parallels between phrasing pattern(s) and syntactic positions of preverbal subjects in languages like Spanish, European Portuguese and Italian. These studies (Elordieta, Frota & Vigário, 2005; Feldhausen, 2014) propose matches between phrasing and syntactic positions, but do not offer principled theoretical
explanations for their conclusions. They simply link the different phrasing patterns of these languages to the two competing syntactic proposals about subjects, i.e., the left dislocated proposal (CP layer) and the IP proposal (inflectional layer). Their conclusions, however, should be taken cautiously. If we look at the bigger picture, and analyze the results from all these languages, or the variation found within the same languages in order to have a comprehensive view of this type of analysis, their conclusions cannot be maintained.

Elordieta, Frota & Vigário (2005) claim that in a language like Spanish, preverbal subjects are located in the CP layer because they constitute an independent phrase. On the other hand, since in European Portuguese subjects are phrased with the verb, they propose that in this language preverbal subjects are located in the IP layer. Their conclusions seem less strong when we add data from a different variety of these same languages. It is the case that Northern European Portuguese phrase their subjects as independent elements the same way that Spanish does (D’Imperio et al., 2005). What would these results mean? Do these different phrasing patterns reflect different structures? Moreover, how can we accommodate this type of variation within the syntax?

Despite the impossibility of establishing a direct correspondence between phrasing and syntactic constituents and/or syntactic positions, the decision to include phrasing in this summary is twofold. First, it is an important component of the intonational patterns of any given language and consequently should be part of any analysis that aims to describe the intonational properties of a language and their elements. Second, the decision not to use it as a test of syntactic status of preverbal subjects and other preverbal elements, as has previously been done in this field, had to be justified.

4.5.3.1. Phrasing patterns in Spanish

Although there are a number of studies that look at phrasing patterns in Spanish, in this section, I only include works that analyze phrasing patterns in broad focus constructions.

---

63 Elordieta, Frota & Vigário (2005) use an optimality approach in terms of alignment and ranking; however, this does not explain the reasons behind the connection between each type of phrasing pattern and a given syntactic structure. They only associate phrasing patterns with syntactic proposals.
D’Imperio et al. (2005) explored the relation between intonation and prosodic phrasing by looking at three factors: syntactic branching, i.e., the number of syntactic phrases or constituents, prosodic branching, i.e., the number of prosodic words, and constituent length, i.e., the number of syllables. They found that in Spanish the most common phrasing pattern across all conditions was (S)(VO). In their corpus, length does not seem to be a decisive factor that influences phrasing strategies. With respect to branching effects, they found that prosodic branching does not play a role, whereas syntactic branching appears to be a major factor. This result is the opposite of what the authors found for Catalan, Neapolitan Italian, Standard European Portuguese and North European Portuguese in the same study. D’Imperio et al.’s (2005) general result about phrasing corroborates Elordieta, Frota & Vigário (2005) findings regarding phrasing, as they also propose (S)(VO) as the most common pattern in Spanish. However, it contradicts Nibert (2000) who found (SV)(O) to be the most extended pattern. The differences among these results may be due to several factors, ranging from differences in the type of stimuli (experimental design) to the speech rate that the speakers used, as most of these experiments only asked the speakers to use a normal speech rate but did not control for it.

Contrary to D’Imperio et al. (2005), Prieto (2007) argues that the prosodic structure of Spanish cannot be accounted for based only on syntactic information, for there are other prosodic constraints, such as the size and the balance of phrase constituents, that also play a role when it comes to phrasing strategies. She concludes that length has an impact in phrasing decisions as well. Prieto also finds different phrasing possibilities for SVO sentences in Spanish in her corpus. Her results show that (S)(VO), (SV)(O) and (SVO) are all possible phrasing patterns, although the last possibility is only found in utterances pronounced at a fast speech rate. Her analysis is also based on a reading task, but the participants were instructed to use three different speech rates at different moments. This fact made possible to include speech rate as a variable in the analysis.

Frota et al. (2007), another study on phrasing, examine the phonetics and phonology of intonational phrasing. Similarly to other studies about the prosodic characteristics of focus types, Frota et al. seek to find the phonetic and phonological boundary cues that are present when there is a phrase in order to propose a typology for each of the five Romance languages they examine. For their analysis, they use the same corpus of the
Romance languages intonational phrasing project used in D’Imperio et al. (2005), Elordieta, Frota & Vigário (2005), and others, but in their analysis they only included cases that did not pose any doubt in their classification. In other words, they use cases that were unambiguously classified as having a clear phrasing boundary. At the phonetic level, they identify the boundary cues and their frequency, while at the phonological level they provide descriptions of such boundaries by using nuclear accents plus their possible combination with the two dominant contours: continuation rise (H) and sustained pitch (!H).

Frota et al. (2007) propose that a phrasing boundary must show one or more of the following characteristics: pitch reset after the boundary, boundary marked by either a High (H) or Low (L) tone, pause after the boundary, and preboundary lengthening. Their data shows that Spanish uses a continuation rise contour and a High boundary tone as main cues to signal a phrase boundary, at rates of 88.4% and 99.3%, respectively. Pitch reset was also found in 76% of the utterances, while preboundary length was present in 40% of their corpus. Pauses were less frequently used, only 28.2% of the time, as well as Low tone. Similar results related to pauses and Low tones were also found for other Romance languages. Again, speech rate could have been a factor that influenced the results regarding pauses, as this was a reading task and speakers were asked to use a normal speech rate, but the researchers did not control for it afterward. With respect to nuclear accents, the most commonly found in Spanish is L*+H.

It is worth pointing out that all these studies have looked at laboratory speech. While this is not a problem per se, and in fact it may be the only way to obtain certain information, it is also true that they do not entirely reflect what actually happens in spontaneous speech. Nonetheless, as Face (2003) notes, there is no clear division between spontaneous speech and laboratory speech; rather, they should be seen as part of a continuum.

4.6. Prosodic features of broad focus, focalized and topicalized elements in Spanish

To summarize the general features of Spanish prosody, I focus on the three contexts under discussion: broad focus, focalized, and topicalized constructions. Several studies have shown that constructions with focalized constituents (narrow and contrastive foci)
are prosodically distinguishable from broad focus constructions. Focalized elements present differences in duration, F0 height, intensity and peak alignment when compared to broad focus elements. These prosodic properties are distributed across the different dialects examined. Furthermore, there are also differences among speakers when it comes to using these strategies. Different speakers may use different strategies, and even the same speaker may use different strategies or a combination of them to mark the elements. However, the existence of these prosodic features corroborates the idea that some syntactic properties such as different types of foci are prosodically encoded. Based on this, and also on our theory of the grammar, we expect that elements marked with topic features also show some prosodic differences when compared to broad focus elements.

In addition to the previous differences between broad focus constructions and focalized constructions, another characteristic of broad focus constructions is their various phrasing patterns. The most extended phrasing pattern in Spanish for broad focus utterances is (S)(VO) according to the majority of the studies, but it is also possible to find (SV)(O) and (SVO), depending on factors such as the size and balance of the phrase constituents, and speech rate. In most cases, phrasing is signalled mainly by a high boundary tone and a continuation rise.

Regarding phrasing in topicalized constructions, contrary to what was commonly believed, in most cases, there is no pause between the fronted element (left dislocated element) and the rest of the utterance (Domínguez, 2004; Feldhausen, 2016). However, the dislocated element constitutes a separate intonational unit with respect to the other part of the sentence (Feldhausen (2016) contra Domínguez (2004)). In most cases, its phrasing as an independent phrase is marked by a rising nuclear pitch accent (continuation rise) and a high edge tone (high boundary tone) (Feldhausen, 2016), like subjects in broad focus utterances.

Interestingly, the prosodic description of topics in Spanish is not very detailed. Most data is limited to the phrasing patterns that they display, i.e., little is known about duration,

---

64 There are two contexts in Spanish that are indistinguishable: broad focus statements and sentences with a narrow focused last element. See Zubizarreta (1998) and Domínguez (2004).
fundamental frequency and intensity as compared to elements in broad focus constructions. To provide a detailed prosodic description of topics in Spanish is precisely one of the goals of my study. Table 4.1 summarizes the prosodic differences among different types of informationally marked elements.

**Table 4.1. Prosodic properties of broad, narrow and contrastive focused elements in Spanish**

<table>
<thead>
<tr>
<th></th>
<th>Intensity (stressed syllable)</th>
<th>Duration (stressed syllable)</th>
<th>F0 peak (stressed syllable)</th>
<th>Peak Alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Broad Focus</td>
<td></td>
<td></td>
<td></td>
<td>Late (Madrid)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Early (40%, Alicante)</td>
</tr>
<tr>
<td>Narrow focus</td>
<td>Increases (with respect to BF, Buenos Aires)</td>
<td>Lengthening (with respect to BF, Madrid)</td>
<td>Increases (with respect to BF) (Mixed results: Madrid. It is also significant for some Buenos Aires speakers)</td>
<td>Early (Madrid)</td>
</tr>
<tr>
<td>Contrastive focus</td>
<td>Lengthening (with respect to BF, Madrid)</td>
<td>Increases (with respect to BF, Madrid and Alicante)</td>
<td>Early (Madrid)</td>
<td></td>
</tr>
<tr>
<td>Topics</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>
Chapter 5
The sound of subjects: a prosodic analysis of Spanish preverbal subjects

This chapter presents the experimental study of the prosody of preverbal subjects in Spanish. This work is intended to test the predictions that derive from the two syntactic proposals that locate Spanish preverbal subjects in different domains of the clause. According to the first theory, preverbal subjects are considered topics in the left periphery (LD proposal) (Alexiadou & Anagnostopoulou, 1998; Ordóñez & Treviño, 1999), whereas, in the other case, they are regular subjects that appear within the inflectional layer (IP proposal) (Goodall, 1999; Suñer, 2003; Ortega-Santos, 2008).

These two contradictory syntactic analyses constitute excellent testing material from a prosodic perspective. If we consider the fact that these two layers have different functional properties, in addition to the claim that the prosodic realization of the elements that appear in them in Spanish is also different, then the prosodic realization of preverbal subjects could be a cue to disambiguate their syntactic status. The prosodic claim is particularly significant with regard to elements that occupy the left periphery of the sentence which are often associated with special intonational properties. For instance, in Spanish, there is a generally accepted idea that there is some prosodic marking between the left dislocated element and the rest of the sentence. Zagona (2002) notes that some “left dislocated constituents are followed by a juncture and, optionally, by a pause” (p. 212).  

The pause as signaling element in this case is a simplification of the issue; however, it captures the important intuition that there is a difference between the left dislocated element and the rest of the utterance. This difference seems to be strong enough that it carries into the written language where it is often represented by a comma. Even though the idea of the comma (as a pause) seems to be widely extended and used in the syntactic field, most experimental studies on prosody show that pauses are not common, at least not in laboratory speech, in this context. Furthermore, pauses are not the only way to convey the information that there is difference between the left dislocated element and the rest of the utterance, and, as already pointed out, they are in fact the least frequent one according to studies on Spanish prosody. More information about this topic can be found in the sections of previous studies on prosody and
Following this line of thinking, and assuming that there is a relationship between syntax and prosody—which is particularly visible with regard to items that encode discourse related information such as focus and topic—preverbal subjects should resemble the prosodic realization of left dislocated elements, if it is the case that they are left dislocated elements. If, on the other hand, they are located within the inflectional layer, and bear no discourse properties, they should not pattern with left dislocated elements. Clearly, there are several factors that have to be taken into account when comparing different types of elements, but the previous reasoning is meant to illustrate the course of analysis that is followed in this chapter. Details about the way this is implemented are given in the section about experimental design.

The chapter is organized as follows. Section 5.1 presents an outline of the study and the hypotheses tested. The following sections, 5.2 and 5.3, detail the experimental design. Section 5.4 provides an explanation of the data collection process, dialects under research and participants. Section 5.5 describes the processes of data processing and analysis. Section 5.6 presents the results. The chapter ends with a discussion of the results from a prosodic perspective.

5.1. Study and Hypotheses

This study compares preverbal subjects to left dislocated topicalized elements. The goal of this comparison is to bring new empirical evidence to the theoretical discussion by testing the predictions that derive from the two syntactic analyses, the IP and the LD proposal. At a more general level, the results from this analysis will also contribute to a better understanding of the interface between syntax and phonology. Finally, but equally important, these results will provide a better characterization of the prosodic features of topicalized elements in the left periphery in Spanish.

In order to obtain the relevant data, the study examines different types of preverbal subjects across two conditions of information structure: broad focus and topicalization. Subject classification is derived from the argumental frames in which verbs participate, in the results section in the following chapter. In the particular case of my data, pauses are not a relevant marker of left dislocated elements.
and they correspond to three different types verbal configurations. The first group comprises subjects of transitive and ditransitive verbs. The second group consists of non-nominative subjects (Cuervo 1999, 2003, Masullo 1992). The third group contains preverbal subjects of unaccusative verbs.

The reason to include the information structure distinction regarding subjects derives from the claim that Spanish preverbal subjects are topics. It is important to point out, however, that none of the two accounts that state this specifies the type of topics that they refer to. Based on the syntactic evidence that they provide, i.e., the type of syntactic elements that are used to show similarities with respect to preverbal subjects, one can infer that they refer to old information, background, or ratified topics (Lambretch 1996, Erteschik-Shir 2007, among others). Moreover, the fact that they claim that the information that preverbal subjects in Spanish provide is not essential, and one of the reasons that they provide is that it can be recovered from the verb, also points to the idea that these elements are already part of the discourse. Consequently, by including two contexts of information structure, this study covers two possible scenarios: preverbal subjects as background topics and preverbal subjects as something different from background topics.

The idea of preverbal subjects being something different than background topics deserves further comment. Various studies on topics point to the classification of preverbal subjects in some languages as default topics and topics of the sentence. This is the case for English (Erteschik-Shir, 2007; Ortega-Santos, 2008) and Spanish (Masullo, 1992; Casielles-Suárez, 2004; Ortega-Santos, 2008), for instance. In this type of classification, topics are not necessarily old as the most extended definition of this term proposes. The

66 In Spanish datives of psych verbs pattern with regular subjects with respect to some of their syntactic and semantic properties, which makes them different from regular datives of ditransitive verbs (Masullo 1992; Cuervo, 1999). One of their most salient characteristics is their preverbal position as the default word order in broad focus contexts. See Chapter 3 for a more detailed characterization of this type of datives.

67 I am abstracting away here from the fact that this type of analysis rules out broad focus constructions with preverbal subjects, something that seems highly unlikely in any given natural language.

68 The literature on topics and the syntactic mechanism of topicalization shows that this is a very complex phenomenon and varies vastly across languages. The interested reader is encouraged to see Erteschik-Shir (2007) for an informative summary and coverage of this matter.
common characteristic that these analyses single out is that the sentence states something about these preverbal elements.

Casielles-Suárez (2004), based on the examination of topic phrases in Spanish and English, offers a typology of these two types of topical elements: sentence topics, which she equates to preverbal subjects, and background topics, which can be any element of the sentence. Table 5.1 summarizes the different properties that characterize these two types of topical elements.

Table 5.1. Differences between sentence topics and background topics (Casielles-Suárez, 2004)

<table>
<thead>
<tr>
<th>Sentence Topics</th>
<th>Background Topics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Restricted to a unique element</td>
<td>Not restricted to a unique element</td>
</tr>
<tr>
<td>Correlate with sentence initial position</td>
<td>Do not correlate with a particular sentence position</td>
</tr>
<tr>
<td>Seem to be restricted to discourse referents</td>
<td>Seem not to be restricted to discourse referents</td>
</tr>
<tr>
<td>Not necessarily discourse old</td>
<td>Necessarily discourse old</td>
</tr>
<tr>
<td>Not necessarily unaccented</td>
<td>Necessarily unaccented</td>
</tr>
</tbody>
</table>

Casielles-Suárez’s interpretation of preverbal subjects as sentence topics is similar to Masullo’s (1992) proposal about some non-nominative elements that occupy the preverbal position in Spanish, which he labels as subjects of predication, and share some properties with traditional subjects. Both proposals, although meant to account for different syntactic data, point to the fact that there is a distinct property of some elements that occupy a specific preverbal position that cannot be equated to the traditional notion of topic as old information. The current study also incorporates this distinction: broad focus subjects would be the so-called “topics” of the sentence, while topicalized preverbal subjects are old information topics. While I am not committing myself to the specifics of these proposals, I believe that the distinction between these two types of topics points to an important differentiation that might be among the cues to explain part of the behaviour of Spanish preverbal subjects.

---

69 This is also similar to Fernández-Soriano’s (1999a) proposal for preverbal locatives and datives.
After covering the two variables that the study includes (subject type and information structure), let us examine the different comparisons that derive from them. The first comparison comprises subjects of transitive verbs. In this case, subjects of transitive verbs in a broad focus context are compared to topicalized elements such as left dislocated direct objects and datives, as well as to the same subjects of transitive verbs in a topicalized construction. The same procedure is followed with preverbal non-nominative subjects. Taking into account their linear order, the position of preverbal subjects and preverbal non-nominative subjects in broad focus contexts is the same as in the topicalized construction, i.e., in both constructions they appear in the first position of the string, which makes their prosodic comparison suitable. Sentences (1), (2) and (3), and (4), (5) and (6) illustrate the two groups. Topicalized elements appear underlined.

(1) El hombre revende dulces.

A man resells3S sweets

‘A man resells sweets.’

(2) El hombre revende dulces.

A man resells3S sweets

‘A man resells sweets.’

(3) Al hombre le vende dulces.

the manDAT clDAT sells3S sweets

‘She sells sweets to the man.’

(4) A Juan le gusta el dinero.

JuanDAT clDAT likes3S the money

‘Juan likes money’

(5) A Juan le gusta el dinero.
Juan\textsubscript{DAT} cl\textsubscript{DAT} likes\textsubscript{3S} the money

‘Juan likes money.’

(6) A Juan le gasta el dinero.

Juan\textsubscript{DAT} cl\textsubscript{DAT} spends\textsubscript{3S} the money

‘He/She spends Juan’s money.’

Finally, preverbal derived subjects of unaccusative verbs are contrasted to preverbal subjects of unergative verbs, as (7) and (8) exemplify. In this group, both types of subjects only appear in the broad focus condition.

(7) Juan Marcel creció.

Juan Marcel grew up

‘Juan Marcel grew up.’

(8) Juan Marcel comió.

Juan Marcel ate

‘Juan Marcel ate.’

The goal of the comparison in the first two groups is straightforward: to find out if preverbal nominative and non-nominative subjects in broad focus constructions and left dislocated elements in topicalized constructions (such as topicalized preverbal nominative and non-nominative subjects, topicalized datives and direct objects) pattern alike or differently with respect to their prosodic properties. The last comparison that contrasts preverbal derived subjects of unaccusative verbs to preverbal subjects of unergative verbs seeks to investigate the possible prosodic differences or similarities in Spanish between a derived subject, based generated in the complement position and then moved to the
preverbal field, and subjects of unergative verbs.\(^70\) Table 5.2 summarizes the types of syntactic elements that the study includes, along with the informational properties of the contexts in which they appear.

*Table 5.2. Types of preverbal DPs, syntactic and information structure contexts*

<table>
<thead>
<tr>
<th>Information Structure</th>
<th>Syntactic Structure</th>
<th>Transitive &amp; ditransitive verbs</th>
<th>Psych verbs &amp; unaccusative predicates (non-nominative Subjects)</th>
<th>CLLD</th>
<th>Unaccusative/Unergative</th>
</tr>
</thead>
<tbody>
<tr>
<td>BROAD FOCUS</td>
<td>S</td>
<td>n-NomS</td>
<td></td>
<td>S (unacc)</td>
<td>S (unerg)</td>
</tr>
<tr>
<td>TOPICALIZATION</td>
<td>S</td>
<td>n-NomS</td>
<td>DO &amp; Dat</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Let us now consider the hypotheses and predictions that can be derived from the syntactic contexts in Table 5.2, under the assumption that there is a position in the inflectional layer in Spanish that is available for elements that function as sentential subjects in broad focus contexts. Unlike English, this position is not limited to the nominative element that triggers agreement on the verb, as it is possible to find in it other elements that display subject like properties, in line with McCloskey’s (1997) idea of subjecthood.

The first hypothesis of our study originates from the two competing syntactic theories about preverbal subjects in Spanish: subjects in the left periphery (LD proposal), and subjects in the inflectional layer (IP proposal). This hypothesis is based on the idea that preverbal subjects in Spanish can appear in the IP (contra the LD proposal).\(^71\)

\(^70\) The reader should remember that in the case of English these two types of subjects are claimed to display different prosodic properties (Legate, 2003; Kahnemuyipour, 2009), although the evidence in this case is not conclusive. On the other hand, in Spanish, this differentiation does not seem to hold (Zubizarreta, 1998).

\(^71\) This hypothesis does not rule out the possibility of subjects appearing in the left periphery when they bear specific informational features that are checked there.
**Hypothesis 1:** If preverbal subjects of transitive verbs in broad focus constructions are in the IP layer, then they will display different prosodic properties than topicalized elements in the left periphery, such as CLLD and topicalized subjects. Specifically, topicalized elements will display less duration, a flatter F0 contour, and less intensity than elements in the broad focus condition.

Given that in Spanish the syntactic category of focus has also prosodic manifestations, it is likely that syntactic differences between topics and non-topics are also mapped onto the prosody. In other words, since left dislocated elements bear a distinct topic feature that is responsible for their movement to the left periphery and preverbal subjects of transitive and ditransitive verbs in broad focus contexts do not carry such a feature (thus they stay in the inflectional layer), the prosodic realization of both groups will differ.

Since contrastively or narrow focused elements in the left periphery in Spanish display special prosodic properties which differentiate them from broad focus elements, as several studies have shown, topicalized elements in the left periphery are expected to behave in the same line. In this case, based on results of previous studies such as Féry and Kügler (2008) and Kügler (2008) for German, and also on the general idea that topicalized elements being old information are generally unstressed, topicalized elements are expected to show less duration than the same elements in a broad focus context. However, it may be the case that other processes, such as lengthening of the last syllable of the word in cases in which the element constitutes an independent phrase, could affect duration as well. In addition to this, and also based on Féry and Kügler (2008) and Kügler (2008), topicalized elements may display a flatter F0 contour with respect to their broad focus counterparts, as greater differences in F0 are related to more saliency of the element. Finally, there could be differences in intensity, although in previous studies of different varieties, this parameter has only been found to be relevant in the variety of Spanish spoken in Buenos Aires.

The second hypothesis refers to preverbal non-nominative subjects and mirrors the previous one. This hypothesis derives from the syntactic proposal that argues that non-nominative subjects of psychological verbs and some unaccusative predicates have
subject like properties, which makes them different from CLLDs (Masullo, 1992; Fernández-Soriano, 1999a, 1999b; Cuervo 1999, 2003; Ortega-Santos, 2008).

One of the arguments that have been put forth in order to differentiate non-nominative subjects from CLLDs is that the default word order of structures that include non-nominative subjects is n-NomSVS, whereas the preverbal realization of CLLDs is the result of a transformation. Specifically, these analyses have argued that preverbal datives of ditransitive verbs are the result of a topic movement of the dative element to the left periphery of the sentence, whereas preverbal dative experiencers of psych verbs appear preverbally without bearing any special informational property. The same has been proposed for non-nominative subjects of other unaccusative predicates. In other words, these non-nominative subjects appear preverbally in broad focus contexts. Additionally, they pattern like regular nominative subjects in various syntactic environments. The question that we aim to answer is whether or not these syntactic differences translate into prosodic differences as well.

**Hypothesis 2**: If preverbal non-nominative subjects of psych verbs and other unaccusative predicates in broad focus contexts are in the IP layer, then they will display different prosodic properties than topicalized CLLDs and topicalized non-nominative subjects of psych verbs and other unaccusative predicates. Particularly, topicalized elements will display less duration, flatter F0 contour and less intensity than elements in the broad focus condition.

These predictions are also based on the findings of Féry and Kügler (2008) and Kügler (2008). However, as for Hypothesis 1, there is a possibility that if left dislocated elements constitute independent phrases, their duration gets affected in the other direction, given the tendency of final syllable lengthening that several studies have found in these cases.

The third and last hypothesis refers to the group that comprises two different types of subjects: derived subjects of unaccusative verbs and subjects of unergative verbs. In this group the sentences appear in the broad focus condition only, and both subjects in preverbal position.
**Hypothesis 3:** There will be no prosodic differences between preverbal subjects of unaccusative verbs and unergative verbs in broad focus constructions. The prosodic features of duration, fundamental frequency and intensity that these subjects display should be alike in the two syntactic contexts.

Although this scenario examines two types of preverbal subjects: subjects based generated as complements of the verb and moved to a preverbal position, and subjects merged in the specifier position of the vP, it is expected that both types of preverbal subjects present similar prosodic realization because there are no differences in their informational status.

This hypothesis arises from the theoretical assumption that both types of subjects in broad focus contexts can appear in the inflectional layer in Spanish. The rationale behind this is that even though subjects of unaccusative verbs are based generated as the complement of the verb, their movement to the preverbal field is not necessarily linked to any special information status of the given element. The movement to a preverbal position of subjects of unaccusative verbs can be classified as an “aboutness” or predication type of movement, in line with Masullo (1992) and Ortega-Santos’ (2008) proposals. This is the same type of movement that has been proposed for nominative subjects of transitive and unergative verbs, and non-nominative subjects of psych verbs and some unaccusative predicates. This seems to be the case since preverbal subjects of unaccusative constructions are possible in broad focus contexts without triggering any special interpretation regarding their informational properties aside from predication, and/or any special prosodic distinction (Zubizarreta, 1998). As in the case of subjects of other types of verbs, this does not exclude the possibility that these subjects can appear in the left periphery due to these elements bearing special informational properties.

### 5.2. Experimental Design

There are two syntactic variables that the study includes: verb type and information structure. With respect to verb type, the experiment comprises four groups: transitive/ditransitive verbs, psychological verbs and other unaccusative predicates (with preverbal non-nominative subjects), unaccusative and unergative verbs. In terms of information structure, there are two contexts: broad focus constructions and topicalized constructions.
The next paragraphs detail how the different contexts are constructed, syntactically and prosodically.

5.2.1. Sentence design: syntax

The first group, sentences with transitive/ditransitive verbs, compares preverbal subjects to left dislocated elements, which can be either datives or direct objects. In this case, SVO or SVAdv sentences, as shown in (9) and (11) respectively, are compared to sentences with a left dislocated dative (10) or direct object (12) with clitic doubling.

(9)  Un hombre revende dulces.
    ‘A man resells sweets.’

(10) Al hombre le vende dulces.
     ‘She sells sweets to the man.’

(11) Los barcos entren temprano.
     ‘The boats train early.’

(12) Los barcos los guardan temprano.
     ‘Boats are put away early.’

The second group contrasts preverbal non-nominative subjects of psych verbs and other unaccusative predicates to CLLDs of ditransitive verbs.

(13) A Juan le gusta el dinero.
Juan\text{\textsubscript{DAT}} cl\text{\textsubscript{DAT}} \text{likes}_3\text{S} \text{the money}

‘Juan likes money.’

(14) A Juan le gusta el dinero.

Juan\text{\textsubscript{DAT}} cl\text{\textsubscript{DAT}} \text{spends}_3\text{S} \text{the money}

‘He/She spends Juan’s money.’

The third and last group corresponds to sentences with preverbal subjects of unergative verbs (15), which are compared to preverbal subjects of unaccusative verbs (16).

(15) Los estudiantes de Marta leyeron.

the students of Marta read\text{\textsubscript{3P}}

‘Marta’s students read.’

(16) Los estudiantes de Malta vinieron.

the students of Malta came\text{\textsubscript{3P}}

‘The students from Malta came.’

5.2.2. Sentence design: information structure

The other syntactic variable is information structure. This is particularly relevant because the informational properties of the context affect its prosodic realization. Note the differences in the prosodic realization of the following examples (17-20).

Context: What happened?

(17) The cat ate the cookie. \text{BROAD FOCUS}

Context: Who ate the cookie?

(18) The cat ate the cookie. \text{INFORMATIONAL FOCUS OR BROAD FOCUS REALIZATION}

(19) THE CAT ate the cookie. \text{EMPHASIS ON THE SUBJECT}
Consequently, if we want to compare sentences prosodically in order to derive possible syntactic differences from their prosodic properties, it is necessary to control the context in which the sentence appears.

The experiment includes two different contexts: broad focus and topicalization. The syntactic contexts in which we are interested are not equivalent in terms of information structure as we are comparing preverbal subjects, presumably the result of broad focus structure, to left dislocated elements, which correspond to specific topic constructions. One way to circumvent this is by having two versions of the context of the preverbal subject sentences. In the first case they appear in a broad focus construction, which corresponds to the “What happened?” type of question in (21), while in the second context the subject is topicalized. In this case, the question includes the subject, making it a topic in the answer, as in (22), similarly to one of the possibilities of generating left dislocated constructions. The third question of the group includes either an indirect dative or direct object, which appears as the first DP in the answer, as a left-dislocated topic.

(23)

**BROAD FOCUS CONTEXT**

¿Qué pasa? ‘What happens?’

(21) Un hombre revende dulces.

   a man resells sweets

   ‘A man resells sweets.’

**TOPICALIZED CONTEXT** (Old information topic, subject included in the question)

¿Qué revende el hombre? ‘What does the man sell?’

(22) El hombre revende dulces.

   the man resells sweets
‘The man resells sweets.’

TOPICALIZED CONTEXT (Old information topic, dative included in the question)

¿Qué le vende al hombre? ‘What does she sell to the man?’

(23) Al hombre le vende dulces.

the manDAT clDAT sells sweets

‘She sells sweets to the man.’

This procedure allows us to compare sentences that are not equivalent regarding their informational properties by comparing them indirectly. The same approach is used in the comparison of datives of psychological verbs and datives of ditransitive verbs.

In the case of unaccusative verbs, on the other hand, there are only broad focus constructions, which are compared to broad focus constructions of unergative verbs. Table 5.3 summarizes all syntactic contexts with their respective information structure and the quantity of tokens per context.

Table 5.3. Types of preverbal DPs, syntactic context, information structure and quantity of tokens

<table>
<thead>
<tr>
<th>Information Structure</th>
<th>Syntactic Structure</th>
<th>Transitive &amp; ditransitive verbs</th>
<th>Psych verbs &amp; unaccusative predicates (non-nominative Subjects)</th>
<th>CLLD</th>
<th>Unaccusative/Unergative</th>
</tr>
</thead>
<tbody>
<tr>
<td>BROAD FOCUS</td>
<td>S (15 tokens)</td>
<td>n-NomS (9 tokens)</td>
<td></td>
<td>S (unacc) (8 tokens)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>S (unerg) (8 tokens)</td>
<td></td>
</tr>
<tr>
<td>TOPICALIZATION</td>
<td>S (15 tokens)</td>
<td>n-NomS (9 tokens)</td>
<td>DO &amp; Dat (24 tokens)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
5.2.3. Sentence design: prosody

Since the study aims to compare the prosodic realization of different syntactic structures in each of the three groups, there are a number of factors that have to be taken into account as they can have an impact in the prosodic realization of a given string. For instance, syllabic quantity, type of syllable, stress position, and even sound types—voiceless vs. voiced—have an effect on the prosodic realization of any given string. Thus, if we are to compare two sentences prosodically in order to derive possible syntactic differences we have to control all possible variables that can lead to external variation, i.e. variation that does not depend on the syntax itself. For this reason, sentences that are compared have the same syllabic quantity. In addition to this, the position of stressed syllables, and syllable types in the string is equivalent as well, see (26) to (28). Note that all of these sentences have 8 syllables, with stress on syllables 2, 5 and 7. Every attempt was made to control for syllable type, i.e. heavy vs. light; however, there are cases in which it was not possible to find sentences that met the necessary syntactic requirements and also maintained the same syllabic structure. Syllables and vowels that are compared, however, are always identical. In each group there are always two sentences that are exactly the same: the broad focus version and the topicalized version of the preverbal subject sentence.

(26)  \textit{Un hom bre re ven de dul ces.}  
\begin{tabular}{cccccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
\end{tabular}  

‘A man resells sweets.’

(27)  \textit{El hom bre re ven de dul ces.}  
\begin{tabular}{cccccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
\end{tabular}  

‘The man resells sweets.’

(28)  \textit{Al hom bre re ven de dul ces.}  
\begin{tabular}{cccccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
\end{tabular}  

‘She sells sweets to the man.’
Likewise, I also tried to control for sound types, specifically, voiced vs. voiceless sounds in order to avoid chunks of non-visible F0. Sometimes this was impossible because, again, syntactic restrictions on verb types made it very hard to control for both.

5.3. Task

To test the hypotheses, I designed a controlled production task. The goal was to simulate a communicative situation by giving participants an audio-recording of a context and an image related to it, followed by a question about that context. Participants are also given the words that should be used in the target answer, in the order that they should appear in the string. Words are given in a list form (vertically), different from a sentence-like form (horizontally). In addition to this, the verb is not conjugated. This added task incorporates an extra processing step for the speakers, as they have to conjugate the verb according to the tense of the question they are asked. This extra processing step potentially reduces the effect of a reading task placing the answer on the side of the regular speech variant.

In order to facilitate the task and engage the participants, I used a Power Point presentation, in which each slide includes a picture related to the situation presented, an audio-recording with a story, and at the end of the audio-recording a question about it. The slide also includes the specific words that participants have to use on the right side of the screen in a vertical manner with a tenseless verb, as figure 5.1 shows. Words do not appear on the screen until the question is formulated so participants do not become too familiar with them beforehand. Using a Power Point presentation allows me to combine images and sounds, at the same time that it allows participants to control the speed rate at which they chose to see the slides.

---

72 Following Jessop (2012).
A controlled production task has an advantage over a reading task since it allows recreating situations that resemble spontaneous speech by engaging speakers and, as a consequence, favouring more natural like answers. It is important to note that even though this type of task is not ideal in the sense that it does not reflect spontaneous speech, it allows us to obtain the type of answers that we are looking for in order to carry out the analysis. This is the same reason why previous studies, as seen in the previous chapter, used reading tasks. It is a fact that there are multiple ways to answer a question, and oftentimes speakers can do so by uttering a single word, for instance. This is the reason why an induced production task, although preferable in the sense that it would be closer to spontaneous speech, was ruled out. The last observation that I would like to make regarding the decision of using a controlled production task is that the target answers are not the only possible ones. As I needed not only complete sentences, but also sentences with words located in specific positions in the string in order to compare them later, it was necessary to give speakers the words that they were supposed to use, in the order they should appear.

The audio-contexts used in the experiment were recorded by speakers of the dialects that the study examines; thus there were two versions of the recordings, one for Madrid and one for Havana. Spanish dialects differ crucially, among other things, in their intonation; therefore, reader selection was done in order not to influence or change any pattern of a given dialect by presenting participants with intonational strategies used in another dialect. In addition to this, the context plays an important role in the creation of a more
realistic environment that would also favour a more spontaneous like answer. Contexts were presented in a random order.

5.4. Dialects, participants, testing and data collection

5.4.1. Dialects

The study examines two dialects: the Central Peninsular Spanish variety spoken in Madrid, Spain, and the Caribbean variety spoken in Havana, Cuba. The motivation for including these two dialects is twofold. The first reason is that even though there are significant differences between the two varieties regarding not only their phonetic inventory, but also their intonational and morphosyntactic properties, the topic under research belongs to the core of the grammar; therefore, it should be represented in both.

Regarding intonational differences between these two dialects, “Yes or no” questions, for instance, are realized with a rising contour in the Peninsular variety whereas in the Caribbean they use a falling pattern at the end (Sosa, 1999). There are also syntactic characteristics such as overt subject realization and non-obligatory verb-subject inversion in questions in the Caribbean variety that crucially differ from the peninsular variety. However, and even though the two varieties may use different prosodic strategies to encode information related to discourse items such as topic constructions, there should be a similar behaviour in both cases, as we are investigating something that should be the same across dialects, namely the position of the preverbal subject in broad focus constructions, and left dislocated elements.

The second reason for selecting these two dialects is that the central Peninsular Spanish variety has been one of the most studied Spanish dialects regarding some of its prosodic characteristics; hence it represents a good opportunity to compare our results to those that already exist in other prosodic areas. At the same time, this study will also contribute to a more detailed description of the properties of this dialect. With respect to the Havana variety the situation is quite the opposite as the number of studies is very limited. This analysis, therefore, is an important contribution to the description of this variety.

At a descriptive level, the goal of this study is to provide further characterization of these two dialects in terms of their prosodic strategies. At a theoretical level, the study seeks to
investigate the possible relationship between syntax and prosody, and how it could be represented and accounted for within the grammar. It also aims to obtain prosodic evidence to evaluate the syntactic theory that claims that not all preverbal subjects in Spanish are located in left periphery of the sentence. More broadly, the objective is to explore the idea that cross-dialectal differences are not represented in the core grammar of the language.

5.4.2. Participants

There are a total of 10 participants in the study, 5 from each dialect. In order to minimize variation and facilitate data analysis the sample only includes women, ranging from 21 to 35 years of age. To be eligible to participate, their educational level had to be at least grade 12, although the majority of them had the equivalent of a bachelor degree completed at the time the testing was done.

I recruited participants in both Havana and Madrid. In Cuba, recruitment took place through personal networks. In Madrid, participants were recruited at a campus cafeteria at the Universidad Complutense de Madrid around lunchtime.

In both cases, individuals were invited to participate in a study about word order variation in Spanish. At that moment they were informed that the study would last approximately one hour, and that they could withdraw from it at any point without any consequences whatsoever. They were also told that they would be compensated for their participation with 20 CAN dollars, a standard amount for a one-hour linguistics study. They were given a letter in which all this information was stated, as well as the researcher contact information along with the contact information of the ethics committee of the University of Toronto in case they wished to know more about the project. In case they agreed to participate, they were also given a consent form that they had to sign the day the testing was done.

In order to participate in the study, individuals had to be speakers of the dialect and have lived in the area for the past 20 years. They also had to be monolinguals, and those who knew a second language could not have lived for more than two months in a country in which the second language was spoken. This was not a problem in Cuba as most people
have not lived in any other country, but in Spain it was a decisive factor in the recruitment because most Europeans participate in exchange programs to learn a second language, and as part of it they travel and live in foreign countries for extended periods of time.

5.4.3. Testing

The testing session lasted approximately one hour, divided in four intervals of 10 minutes of recording time with breaks of five minutes in between. At first, and to help the participants to feel more comfortable with the researcher and get familiarized with the equipment they were going to use, we carried out an informal conversation about favourite festive celebrations. Participants were asked to talk about their preferred celebrations including Christmas, New Year’s Eve, birthdays, etc. Although this was a topic to break the ice, they were told that they could talk about anything they liked, and oftentimes they introduced new topics of their interest. These conversations lasted between five and ten minutes, and they were also recorded. Even though they do not constitute part of the analysis of the present work, they could be used in the future if we wished to extend this research to spontaneous speech.

After the conversations, participants were introduced to the experiment. Following the initial explanation, they were asked to do a sample trial of three contexts in order to make sure they understood the task. Participants were able to hear the context and the question as many times as they wanted or needed. Since the experiment involved listening to a context and a question about it, participants were recommended to wear headphones in order to minimize any distraction from background noise. I decided to do this after a pilot test in which I noticed that participants were easily distracted by noises, and oftentimes had to replay the recordings. Introducing the use of headphones helped to cut down the testing time because, in most cases, participants did not need to replay the contexts.

5.4.4. Data collection

Testing was carried out on site in both cases: Havana and Madrid. In both cases I tried to find a quiet place in which the recordings could be done without getting too much exposure to environmental noise, although these were not soundproof. In Havana I did so
in my family house, and in Madrid I used an empty study room on campus. In both cases, however, some background noise was present, especially in the recordings from Cuba.\textsuperscript{73}

As participants were doing the task, I had a sheet with all target answers in the same order they appear in the experiment. In cases in which they did not produce the target sentence I made a note about it, and at the end of the whole session, I ask them to repeat those contexts. The purpose of this was to preserve the amount of tokens per condition in order to be able to carry out the statistical analysis. The average of repetitions was six sentences per participant. Participants were told in those cases that due to some background noise it was necessary to repeat those contexts, as I did not want to influence their answers by giving them clues. In most cases, they produced the target sentence during the repetition.

Participants were recorded using a Marantz PMD661 Compact Flash/MP3 Wav Recorder and a unidirectional condenser microphone, placed about 30 cm from the speaker’s lips. The sampling rate used in the recordings was 44.1kHz. The data was stored on memory cards and then transferred to a computer where it was processed using PRAAT 5.3.63 (Boersma and Weenink, 2014).

5.5. Data Processing and Analysis

5.5.1. Data Processing

The first step of the data processing was to convert all sound files to mono. Secondly, all sound files were resampled to 22.1Hz, and then normalized due to some background noise, especially in the recordings from Cuba. After that, the data was analyzed acoustically and qualitatively. The acoustic analysis was carried out using PRAAT.

In order to do the acoustic analysis the material was annotated. For this purpose I used TextGrids with two tiers. The first tier corresponds to the general syntactic division of the sentence: preverbal determiner phrase, labelled as DP, and the predicate, labelled as VP. The second tier contains the stressed syllables of the preverbal noun and the verb, along

\textsuperscript{73} For this reason the recordings had to be normalized.
with the last syllable of the preverbal DP, which corresponds to the last syllable of the preverbal noun in most cases. These were labelled as ssN, ssV and DP-Is, respectively. Figure 5.2 shows an example of a TextGrid of the sentence “Un hombre revende dulces” ‘A man resells sweets’, in the broad focus context produced by a Madrid speaker. In the figure I added a third tier that includes the words. Although the segmentation includes both DP and VP, most of the acoustic analysis carried out for this work was done in the DP, as this is the most relevant part for this study.

*Figure 5.2. TextGrid of “Un hombre revende dulces” produced by a Madrid speaker*

For the qualitative analysis, various parameters were recorded. These parameters were taken from Frota et al. (2007) and they correspond to the characterization of the F0 in the final part of the DP as well as the information related to the peak alignment of the noun and the verb. In this part of the analysis, I also used some of the acoustic measurements in order to make some decisions while labeling. This qualitative information was recorded in an Excel sheet for each speaker.

5.5.2. Analysis. Acoustic parameters (measurements and statistical analysis). Phrasing patterns and pitch accents

For the comparison of the elements, three acoustic parameters were considered: duration, fundamental frequency (F0) and intensity.
Duration was measured in the stressed syllable of the noun (ssN) of the DP, and also in the last syllable of the DP (DP-ls). As noted in the previous chapter, duration is one of the most important cues in Spanish, at least in Peninsular Spanish, when it comes to signalling different types of foci.

For the F0, two values were measured: the F0 at the beginning of the stressed syllable of the noun, and the F0 peak. The first value was obtained automatically using a PRAAT script that measures the F0 at the first boundary of the syllable. After obtaining this value, I manually checked those that were zeroes and those that seemed to be outliers, which could have been due to jumps in fundamental frequency, creaky voice, etc. In these cases, whenever it was possible, the automatically generated value was replaced by the first F0 value available. Cases of creaky voice were discarded in the statistical analysis. In the case of the F0 peak, the procedure was done manually. The first step was locating the peak, and then measuring it. This information was recorded in an Excel sheet.

The purpose of including two F0 measurements (beginning and peak) was to obtain the derived value of the tonal change associated with the stressed syllable. The tonal change was calculated by subtracting the F0 of the beginning of the syllable from the F0 peak. One of the advantages of using the derived measurement of tonal change instead of the absolute value of the peak is that one can compare relative values without having to take into account the F0 baseline for each sentence. The goal in this case is to see if one context, broad focus or topicalization, favours a bigger tonal change than the other. Based on the idea that old information topics are unstressed, it is expected to find a flatter F0 contour, which means less tonal change, in the case of topics.

Intensity, the third acoustic parameter, was measured in the stressed syllable of the noun as well. In this case, I worked with the maximum intensity value within this segment. This value was obtained automatically using a PRAAT script.

As the experiment was specifically designed to compare the same element in two different conditions, paired statistical analyses were performed in most cases. When this was not possible, unpaired analyses were done. This was only the case in the tonal change comparison between preverbal subjects and CLLDs, where not all the tokens were the identical. In each case, the Shapiro-Wilk normality test was performed before submitting
the data to either parametric or non-parametric analyses. The alpha level selected was .05 in all cases.

For the statistical analysis of duration, I worked with the ratio of the duration of the element in the two different conditions. The ratio was the appropriate choice, instead of absolute values, because we are looking at an effect, either increase or decrease. For intensity, I compared the maximum intensity associated with the stressed syllable of the element in the two different conditions. Further details about the procedures followed in each instance are provided in the results section.

For the statistical analysis, the first step was to consider each speaker individually. The intention was to find out the strategies that were relevant for each speaker. Although some domains are more stable than others—for instance the use of duration is more stable than the use of the fundamental frequency—an individual analysis was the preferable choice, as the results of a dialectal analysis could be affected or masked by individual differences. After the individual analyses were done, in cases in which it was applicable and/or necessary, dialectal analyses were carried out.

For the qualitative analysis different parameters were considered. Since I was also interested in exploring the possible differences or similarities in phrasing patterns that are associated with each context, broad focus and topicalization, different parameters that could signal phrasing were analyzed. These parameters were taken from Frota et al. (2007), and correspond to the description of final stretch of the F0 in the preverbal field. They were used to determine whether a boundary was present between the preverbal DP and the VP. The parameters that this study looks at are: the realization of the pre-boundary stretch (continuation rise or sustained pitch), High (H%) or Low (L%) tone signalling the boundary, pitch reset after the boundary, and pauses.

The preboundary stretch was described as: continuation rise and sustained pitch. Boundary Tones were labeled as: H%, L%, and other. Pitch resets were considered present when there was a drop of the F0 to the speaker’s baseline or when the difference between the first peak and the second peak of the sentence was equal or larger than 1 erb. Since pitch distance is more relevant to the perception of changes than absolute pitch values, I decided to work with the relation between the two peaks. The value of 1 erb was
determined based on the combination of two proposals for reliable discrimination, 3 semitones and 1.5 semitones (Nootboon, 1997). Pauses were considered present when there was a silence gap that lasted more than 150 ms. As in Frota et al.’s (2007) study, in cases in which there was more than one of these elements present in an utterance, it was considered that there was a boundary. In such cases, the element before the boundary (the preverbal DP) was classified as an independent phrase (IP).

The information regarding peak alignment, i.e., the realization of the peak either within the stressed syllable or after it, was also recorded. The reader should recall that in Peninsular Spanish there is a tendency to realize the peak within the stressed syllable in contexts of narrow focus, as opposed to its postonic realization in broad focus contexts. Since there is no information regarding the peak realization in contexts of topicalization, I decided to include it in this analysis.

5.6. Results

5.6.1. Preverbal subjects of transitive verbs

According to Hypothesis 1, it is expected that preverbal subjects of transitive verbs in broad focus constructions would differ prosodically from topicalized elements, such as preverbal topicalized subjects of transitive verbs and CLLDs. The prediction was that topicalized elements should display less duration than broad focus elements. In addition to this, topicalized elements should also show a flatter F0 contour and have less intensity with respect to broad focus elements, due to the fact that they were previously mentioned in the discourse. For the sake of clarity, I repeat an example of the contexts to which this hypothesis refers. Underlined elements are old information.

WHAT HAPPENED?

(29)  El empleado midió la tela. SUBJECT (transitive verbs)-BROAD FOCUS

   the employee measured the fabric

   ‘The employee measured the fabric.’

WHAT DID THE EMPLOYEE MEASURE?

(30)  El empleado midió la tela. SUBJECT(transitive verbs)-OLD INFORMATION
the employee measured the fabric

‘The employee measured the fabric’

WHAT DID SHE GIVE TO THE EMPLOYEE?

(31) Al empleado le dio la tela. CLLD-OLD INFORMATION

the employee DAT cl gave the fabric

‘He/She gave the fabric to the employee.’

The scatterplots in Figure 5.3 illustrate the comparison of the duration of preverbal subjects in two conditions of information structure, broad focus and topicalization, for six speakers, three from each dialect. In the graphs, the points represent paired observations, and the axes represent the two conditions. The blue line marks the boundary where observations display the same duration in the two conditions. The fact that the majority of the points are on the right side of the blue line, which corresponds to the broad focus condition, indicates that elements in this condition display a longer duration than their topicalized counterparts. Tables 5.4 and 5.5 summarize the results of the ratio paired t-tests and Wilcoxon signed-rank tests for the comparison of the duration of preverbal subjects in the two conditions of information structure. As can be seen, differences in duration are statistically significant for all the speakers of the two dialects.

Figure 5.3. Scatterplots displaying the paired duration data of preverbal subjects in broad focus and topicalized contexts, by three speakers of each dialect*

*The values in the axis scales are logarithms of the duration.
Table 5.4. Ratio paired t-tests results comparing the duration of preverbal subjects in broad focus and topicalized contexts

<table>
<thead>
<tr>
<th>Dialect</th>
<th>Speaker</th>
<th>M</th>
<th>95% CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Havana</td>
<td>H6C</td>
<td>1.19</td>
<td>1.11 - 1.28</td>
<td>&lt; .001</td>
</tr>
<tr>
<td></td>
<td>H7C</td>
<td>1.18</td>
<td>1.14 - 1.23</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Madrid</td>
<td>H2M</td>
<td>1.12</td>
<td>1.08 - 1.17</td>
<td>&lt; .001</td>
</tr>
<tr>
<td></td>
<td>H4M</td>
<td>1.12</td>
<td>1.05 - 1.19</td>
<td>.002</td>
</tr>
<tr>
<td></td>
<td>H5M</td>
<td>1.13</td>
<td>1.08 - 1.18</td>
<td>&lt; .001</td>
</tr>
<tr>
<td></td>
<td>H6M</td>
<td>1.13</td>
<td>1.06 - 1.20</td>
<td>&lt; .001</td>
</tr>
<tr>
<td></td>
<td>H7M</td>
<td>1.09</td>
<td>1.05 - 1.14</td>
<td>&lt; .001</td>
</tr>
</tbody>
</table>

Table 5.5. Wilcoxon signed-rank tests results comparing the duration of preverbal subjects in broad focus and topicalized contexts

<table>
<thead>
<tr>
<th>Dialect</th>
<th>Speaker</th>
<th>Mdn</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Havana</td>
<td>H1C</td>
<td>1.15</td>
<td>.011</td>
</tr>
<tr>
<td></td>
<td>H2C</td>
<td>1.10</td>
<td>&lt; .001</td>
</tr>
<tr>
<td></td>
<td>H3C</td>
<td>1.14</td>
<td>.013</td>
</tr>
</tbody>
</table>

The scatterplots in figures 5.4 and 5.5 suggest that the comparison between broad focus preverbal subjects and CLLDs shows a similar trend, but in this case differences are statistically significant for two speakers only: H3C (t(4) = 2.91, p = .044) and H5M (t(4) = 3.34, p = .029), as the results of the Ratio paired t-tests summarized in Table 5.6 show. Note that in this context only five pairs were used in the comparison, given that they were the only ones that contained the same words (figures 5.4 and 5.5). A dialectal analysis was carried out to see whether the quantity of tokens was a factor in this case. The results of the ratio paired t-tests reveal that the quantity of tokens was likely the problem (Havana: t(23) = 3.85, p < .001; Madrid: t(22) = 3.35, p = .003).
Figure 5.4. Scatterplots displaying the paired duration data of broad focus preverbal subjects and topicalized CLLDs, by Havana speakers

*The values in the axis scales are logarithms of the duration.

Figure 5.5. Scatterplots displaying the paired duration data of broad focus preverbal subjects and topicalized CLLDs, by Madrid speakers

*The values in the axis scales are logarithms of the duration.
Table 5.6. Ratio paired t-test results comparing the duration of broad focus preverbal subjects and topicalized CLLDs

<table>
<thead>
<tr>
<th>Dialect</th>
<th>Speaker</th>
<th>M</th>
<th>95% CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Havana</td>
<td>H1C</td>
<td>1.07</td>
<td>0.94 - 1.21</td>
<td>.224</td>
</tr>
<tr>
<td></td>
<td>H2C</td>
<td>1.13</td>
<td>0.91 - 1.39</td>
<td>.163</td>
</tr>
<tr>
<td></td>
<td>H3C</td>
<td>1.18</td>
<td>1.01 - 1.37</td>
<td>.044</td>
</tr>
<tr>
<td></td>
<td>H6C</td>
<td>1.17</td>
<td>0.89 - 1.53</td>
<td>.181</td>
</tr>
<tr>
<td></td>
<td>H7C</td>
<td>1.08</td>
<td>0.87 - 1.32</td>
<td>.381</td>
</tr>
<tr>
<td>Madrid</td>
<td>H2M</td>
<td>1.07</td>
<td>0.97 - 1.19</td>
<td>.116</td>
</tr>
<tr>
<td></td>
<td>H4M</td>
<td>1.06</td>
<td>0.92 - 1.23</td>
<td>.312</td>
</tr>
<tr>
<td></td>
<td>H5M</td>
<td>1.17</td>
<td>1.03 - 1.34</td>
<td>.029</td>
</tr>
<tr>
<td></td>
<td>H6M</td>
<td>1.08</td>
<td>0.89 - 1.31</td>
<td>.280</td>
</tr>
<tr>
<td></td>
<td>H7M</td>
<td>1.13</td>
<td>0.73 - 1.77</td>
<td>.439</td>
</tr>
</tbody>
</table>

The results from the fundamental frequency analysis are less categorical than those of duration. The specific measurement in this case was the tonal change associated with the stressed syllable of the preverbal noun. The expectation was that elements in the topicalized condition should display less tonal change than elements in the broad focus condition.

Figure 5.6 illustrates the raw data of the tonal change for all of the Havana speakers, along with the calculated confidence intervals. The first thing that can be noticed is that there is great variability among speakers. This variation comes from two sources: inter-speaker variability and intra-speaker variability. Inter-speaker variability can be particularly seen when we compare speaker H7C, who does not use the F0 to distinguish among these contexts, to the rest of the speakers. Note that H7C pattern resembles a flat line. Inter-speaker variability is also visible in the different patterns used by the rest of the speakers. Although there seems to be a trend for topicalized CLLDs to display a bigger tonal change than broad focus preverbal subjects, this tendency is not across the board. While speakers H1C, H2C and H3C show this pattern, speaker H6C shows the opposite trend. The different treatment of CLLDs and TopSs by H3C constitutes an example of intra-speaker variability, as both elements correspond to contexts of topicalization. Finally note that in all cases there is a large overlap of the confidence intervals.

The Madrid speakers’ tonal change data also shows great variation. Figure 5.7 depicts the tonal change raw data along with the confidence intervals for the speakers of this variety. As in the case of the Havana speakers, there is a large overlap in the confidence intervals.
As a result, no pattern seems to emerge in any of the comparisons that arise from these contexts.

Figure 5.6. Box plots and confidence intervals displaying the tonal change of topicalized subjects, topicalized CLLDs, and broad focus subjects, produced by Havana speakers

Figure 5.7. Box plots and confidence intervals displaying the tonal change of topicalized subjects, topicalized CLLDs, and broad focus subjects, produced by Madrid speakers
Unpaired t-tests and Wilcoxon rank-sum tests were conducted to determine whether differences in the tonal change of broad focus preverbal subjects and topicalized CLLDs were significant for any of the speakers of the two dialects. Paired t-tests and Wilcoxon signed-rank tests were used in the case of preverbal subjects in the two conditions of information structure, broad focus and topicalization. The results of the tests revealed that differences in the tonal change of the comparison of broad focus subjects and CLLDs are statistically significant for three speakers of the Havana variety: H1C (t(25) = -2.64, p = .014), H2C (t(25) = -4.01, p < .001), and H3C (t(26) = -2.85, p = .008). For these speakers, CLLDs display bigger tonal change than broad focus subjects. No statistically significant differences were found for the Madrid speakers, as the results summarized in tables 5.7 and 5.8 show.

Table 5.7. Unpaired t-test results comparing the tonal change of broad focus preverbal subjects and topicalized CLLDs

<table>
<thead>
<tr>
<th>Dialect</th>
<th>Speaker</th>
<th>M BFS</th>
<th>SE BFS</th>
<th>M CLLD</th>
<th>SE CLLD</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Havana</td>
<td>H1C</td>
<td>34.59</td>
<td>3.55</td>
<td>50.60</td>
<td>5.01</td>
<td>.014</td>
</tr>
<tr>
<td></td>
<td>H2C</td>
<td>22.19</td>
<td>3.88</td>
<td>40.03</td>
<td>2.33</td>
<td>&lt; .001</td>
</tr>
<tr>
<td></td>
<td>H3C</td>
<td>20.37</td>
<td>3.00</td>
<td>33.38</td>
<td>3.44</td>
<td>.008</td>
</tr>
<tr>
<td>Madrid</td>
<td>H2M</td>
<td>49.43</td>
<td>3.63</td>
<td>58.30</td>
<td>3.35</td>
<td>.084</td>
</tr>
<tr>
<td></td>
<td>H4M</td>
<td>30.28</td>
<td>3.00</td>
<td>33.03</td>
<td>1.36</td>
<td>.416</td>
</tr>
<tr>
<td></td>
<td>H6M</td>
<td>34.56</td>
<td>3.60</td>
<td>37.62</td>
<td>3.35</td>
<td>.541</td>
</tr>
<tr>
<td></td>
<td>H7M</td>
<td>46.07</td>
<td>3.81</td>
<td>57.70</td>
<td>5.84</td>
<td>.113</td>
</tr>
</tbody>
</table>

Table 5.8. Wilcoxon rank-sum tests results comparing the tonal change of broad focus preverbal subjects and topicalized CLLDs

<table>
<thead>
<tr>
<th>Dialect</th>
<th>Speaker</th>
<th>Mdn BFS</th>
<th>Mdn CLLD</th>
<th>W</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Havana</td>
<td>H6C</td>
<td>41.83</td>
<td>30.83</td>
<td>135.5</td>
<td>.089</td>
</tr>
<tr>
<td></td>
<td>H7C</td>
<td>17.73</td>
<td>22.08</td>
<td>81.00</td>
<td>.454</td>
</tr>
<tr>
<td>Madrid</td>
<td>H5M</td>
<td>69.99</td>
<td>67.13</td>
<td>103</td>
<td>.583</td>
</tr>
</tbody>
</table>

Tonal change differences between preverbal subjects in broad focus and topicalized conditions are statistically significant for two speakers of the Havana variety only: H2C (Z = -2.38, p = .017) and H6C (Z = -3.43, p < .001), as the results in tables 5.9 and 5.10 show. Note that differences in these two cases are very small, especially for H2C (Mdn = -5.39). Also note that the pattern is different for the two speakers. It is important to bear in mind that in this instance we are comparing the same DP in two different conditions of
information structure, while the previous case compares two DPs with different syntactic function—preverbal subjects and CLLDs—in two conditions of information structure.

Table 5.9. Paired t-tests results comparing the tonal change of preverbal subjects in broad focus and topicalized contexts

<table>
<thead>
<tr>
<th>Dialect</th>
<th>Speaker</th>
<th>M</th>
<th>SE</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Havana</td>
<td>H1C</td>
<td>-6.18</td>
<td>5.03</td>
<td>.240</td>
</tr>
<tr>
<td></td>
<td>H3C</td>
<td>1.99</td>
<td>3.34</td>
<td>.560</td>
</tr>
<tr>
<td></td>
<td>H7C</td>
<td>0.76</td>
<td>3.43</td>
<td>.823</td>
</tr>
<tr>
<td>Madrid</td>
<td>H2M</td>
<td>-2.40</td>
<td>3.28</td>
<td>.477</td>
</tr>
<tr>
<td></td>
<td>H4M</td>
<td>3.20</td>
<td>3.81</td>
<td>.415</td>
</tr>
<tr>
<td></td>
<td>H5M</td>
<td>2.50</td>
<td>4.73</td>
<td>.605</td>
</tr>
<tr>
<td></td>
<td>H6M</td>
<td>3.73</td>
<td>2.65</td>
<td>.184</td>
</tr>
<tr>
<td></td>
<td>H7M</td>
<td>-7.56</td>
<td>5.33</td>
<td>.182</td>
</tr>
</tbody>
</table>

Table 5.10. Wilcoxon signed-rank tests results comparing the tonal change of preverbal subjects in broad focus and topicalized contexts

<table>
<thead>
<tr>
<th>Dialect</th>
<th>Speaker</th>
<th>Mdn</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Havana</td>
<td>H2C</td>
<td>-5.39</td>
<td>.017</td>
</tr>
<tr>
<td></td>
<td>H6C</td>
<td>15.95</td>
<td>&lt; .001</td>
</tr>
</tbody>
</table>

The first explanation for the imbalanced behaviour between these two contexts could be that the sample does not provide the necessary data. Since there is so much variation, more tokens may be necessary. However, even considering the limitations of this data, a plausible hypothesis is that differences in information structure that are linked to differences in syntactic function, and also correlate with differences in word order have a bigger impact on F0 changes.

A second hypothesis is that differences in the tokens themselves are responsible for this asymmetry. It is a fact that the comparison of preverbal subjects in the two contexts of information structure always includes the same element in two conditions. On the other hand, in the case of subjects and CLLDs only five of the tokens are the same, the rest of the tokens present some phonetic differences. It could be the case that differences in the fundamental frequencies of the segments themselves are influencing these results. This idea, however, proves to be false when we look at the individual patterns that speakers display. It can be noticed in Figure 5.6 that while speakers H1C, H2C, H3C display a bigger tonal change in CLLDs than in broad focus subjects (H1C: BFS Mdn = 35.5,
CLLD $Mdn = 50.6$; H2C BFS $Mdn = 22.2$, CLLD $Mdn = 40$; H3C: BFS $Mdn = 20$, CLLD $Mdn = 33$), speaker H6C displays the opposite pattern (BFS $Mdn = 41.83$, CLLD $Mdn = 30.83$). This asymmetry suggests that the tokens themselves do not have an impact on the F0 behaviour, and only speaker strategies modulate it.

The last possibility is that further characterization is needed for these comparisons. The current analysis does not take into account the different pitch accents that these elements have, as it only looks at their tonal change. Even though the realization of the pitch accent is mostly postonic in this data, two configurations are still possible: L$^{*}$+H and L$^+>H^*$. Thus, it could be the case that the tonal change variation in these two configurations is different. A similar point was raised by Face (2002) regarding the mixed results that different studies had reported about the increase of the F0 in contexts of narrow focus. In that case, only instances in which the realization of the pitch accent coincided with the broad focus configuration, i.e., postonic, the increase in height was relevant. In cases where the pitch accent was realized within the stressed syllable, differences in height were not significant between broad and narrow focused elements. Our sample is less different than Face’s because the pitch is always postonic, yet there could be differences with respect to the height of the peak in the two configurations.

Finally, there are differences between the two dialects with respect to the use of F0 in these contexts. While in Havana F0 is one of the acoustic parameters used to distinguish between some of these structures, the data from Madrid shows that the speakers do not use F0 to differentiate old information from new information in the preverbal field.

Differences in intensity between topicalized and broad focus elements are minimal for all of the speakers of the two dialects, as can be seen in figure 5.8 that illustrates the peak intensity raw data with the calculated confidence intervals for three speakers of each dialect. Paired t-tests and Wilcoxon signed-rank tests were conducted for the comparison of the peak intensity of the stressed syllable of preverbal subjects in two conditions of information structure. As the results summarized in tables 5.11 and 5.12 show, no significant differences were found except for speaker H4M from Madrid ($r(13) = 3.56$, $p = .003$). However, the difference in this case seems rather small (M: 2.07)
Figure 5.8. Box plots and confidence intervals displaying the peak intensity of topicalized subjects, topicalized CLLDs, and broad focus subjects, produced by three speakers of each dialect.

Table 5.11. Paired t-tests results comparing the peak intensity of preverbal subjects in broad focus and topicalized contexts

<table>
<thead>
<tr>
<th>Dialect</th>
<th>Speaker</th>
<th>M</th>
<th>SE</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Havana</td>
<td>H1C</td>
<td>0.60</td>
<td>0.47</td>
<td>.219</td>
</tr>
<tr>
<td></td>
<td>H2C</td>
<td>1.63</td>
<td>1.04</td>
<td>.142</td>
</tr>
<tr>
<td></td>
<td>H6C</td>
<td>0.47</td>
<td>0.76</td>
<td>.548</td>
</tr>
<tr>
<td></td>
<td>H7C</td>
<td>1.46</td>
<td>0.73</td>
<td>.067</td>
</tr>
<tr>
<td>Madrid</td>
<td>H2M</td>
<td>-0.43</td>
<td>0.64</td>
<td>.515</td>
</tr>
<tr>
<td></td>
<td>H4M</td>
<td>2.07</td>
<td>0.58</td>
<td>.003</td>
</tr>
<tr>
<td></td>
<td>H5M</td>
<td>0.01</td>
<td>0.58</td>
<td>.988</td>
</tr>
<tr>
<td></td>
<td>H6M</td>
<td>0.12</td>
<td>0.45</td>
<td>.790</td>
</tr>
<tr>
<td></td>
<td>H7M</td>
<td>-0.37</td>
<td>0.75</td>
<td>.627</td>
</tr>
</tbody>
</table>

Table 5.12. Wilcoxon signed-rank tests results comparing the peak intensity of preverbal subjects in broad focus and topicalized contexts

<table>
<thead>
<tr>
<th>Dialect</th>
<th>Speaker</th>
<th>Mdn</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Havana</td>
<td>H3C</td>
<td>-0.08</td>
<td>.670</td>
</tr>
</tbody>
</table>

Unpaired t-tests and Wilcoxon rank-sum tests were employed in the comparison of the peak intensity of the stressed syllable of preverbal broad focus subjects and CLLDs.
Tables 5.13 and 5.14 summarize the results of these tests. In this case no significant
differences were found for any of the speakers of the two dialects.

**Table 5.13. Unpaired t-tests results comparing the peak intensity of broad focus
preverbal subjects and topicalized CLLDs**

<table>
<thead>
<tr>
<th>Dialect</th>
<th>Speaker</th>
<th>$M_{BFS}$</th>
<th>$SE_{BFS}$</th>
<th>$M_{CLLD}$</th>
<th>$SE_{CLLD}$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Havana</td>
<td>H1C</td>
<td>82.05</td>
<td>0.64</td>
<td>83.06</td>
<td>0.37</td>
<td>.190</td>
</tr>
<tr>
<td></td>
<td>H2C</td>
<td>82.57</td>
<td>0.86</td>
<td>83.95</td>
<td>0.62</td>
<td>.197</td>
</tr>
<tr>
<td></td>
<td>H3C</td>
<td>80.82</td>
<td>0.62</td>
<td>82.31</td>
<td>0.61</td>
<td>.099</td>
</tr>
<tr>
<td></td>
<td>H7C</td>
<td>83.23</td>
<td>0.67</td>
<td>82.81</td>
<td>0.72</td>
<td>.674</td>
</tr>
<tr>
<td>Madrid</td>
<td>H2M</td>
<td>81.45</td>
<td>0.83</td>
<td>81.86</td>
<td>0.80</td>
<td>.723</td>
</tr>
<tr>
<td></td>
<td>H5M</td>
<td>85.18</td>
<td>0.73</td>
<td>84.94</td>
<td>0.35</td>
<td>.773</td>
</tr>
<tr>
<td></td>
<td>H6M</td>
<td>83.00</td>
<td>0.51</td>
<td>83.20</td>
<td>0.45</td>
<td>.776</td>
</tr>
<tr>
<td></td>
<td>H7M</td>
<td>82.13</td>
<td>0.62</td>
<td>83.86</td>
<td>0.63</td>
<td>.061</td>
</tr>
</tbody>
</table>

**Table 5.14. Wilcoxon rank-sum tests results comparing the peak intensity of broad focus
preverbal subjects and topicalized CLLDs**

<table>
<thead>
<tr>
<th>Dialect</th>
<th>Speaker</th>
<th>$Md_{BFS}$</th>
<th>$Md_{CLLD}$</th>
<th>$W$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Havana</td>
<td>H6C</td>
<td>83.77</td>
<td>84.34</td>
<td>80.50</td>
<td>.435</td>
</tr>
<tr>
<td>Madrid</td>
<td>H4M</td>
<td>84.56</td>
<td>83.68</td>
<td>123.00</td>
<td>.265</td>
</tr>
</tbody>
</table>

The intensity results are in line with what has been reported for the Madrid variety, where
this acoustic parameter is not used to mark differences in information structure in
contexts that are more salient than the ones that this experiment analyzes, e.g., narrow
focus (Face, 2000). Based on this data, it can be suggested that the Havana variety does
not use this parameter to signal differences in information structure between old and new
information. However, these results should be taken cautiously because there is no
baseline data. Consequently, to have a better representation of the prosodic properties of
this dialect, more data is necessary.

### 5.6.2. Preverbal non-nominative subjects

Hypothesis 2 refers to preverbal non-nominative subjects, and mirrors the previous one
about preverbal subjects of transitive verbs. It is expected that topicalized preverbal non-
nominative subjects of psych verbs and other unaccusative predicates and topicalized
CLLDs differ prosodically from preverbal non-nominative subjects of psych verbs and
other unaccusative predicates in broad focus constructions. As in Hypothesis 1, it was
predicted that topicalized elements should display shorter duration than broad focus
elements. With respect to their fundamental frequency, a flatter F0 contour is expected for topicalized DPs, when compared to DPs that are new information. Finally, topicalized DPs should have less intensity than DPs that are new information. Next, I repeat an example of the contexts to which this hypothesis refers. Underlined DPs are old information.

WHAT HAPPENED?

(32) A Juan le gusta el dinero. NON-NOMINATIVE SUBJECT-BROAD FOCUS

Juan_{DAT} cl_{DAT} likes the money

‘Juan likes money.’

WHAT DOES JUAN LIKE?

(33) A Juan le gusta el dinero. NON-NOMINATIVE SUBJECT-OLD INFORMATION

Juan_{DAT} cl_{DAT} likes the money

‘Juan likes money.’

WHAT DOES HE/SHE DO TO JUAN?

(34) A Juan le gasta el dinero. CLLD-OLD INFORMATION

Juan_{DAT} cl_{DAT} spends the money

‘He/she spends Juan’s money.’

Note that in the previous sentences the preverbal DP is a dative. The only difference among these datives resides in the type of verbal structure in which they are embedded. In one case they are dative experiencers, as in (32-33), arguments of psych verbs, while in (34) they are part of a ditransitive structure. (The same structural differentiation holds in the case of non-nominative subjects of other unaccusative predicates and ditransitive structures. However, in this case, the interpretation of the non-nominative subject does not always corresponds to a dative experiencer.) Their informational status also varies.
Non-nominative subjects appear in two conditions of information structure, broad focus and topicalization, while CLLDs appear in the topicalized condition only.

The number of tokens per condition was 9. Since all participants produced all sentences, all tokens were employed for the statistical analysis. Duration, tonal change and intensity, as in the previous case, were measured in the stressed syllable of the preverbal DP. Only cases that presented creaky voice were discarded for the F0 analysis.

The scattered plots in Figure 5.9 show the comparison of the duration of preverbal non-nominative subjects in two conditions of information structure, broad focus and topicalization, for three speakers of each dialect. As in the cases examined for Hypothesis 1, the tendency is for old information to show less duration than new information, in accordance with what was predicted. Ratio paired t-tests and Wilcoxon signed-rank tests were conducted to determine whether differences in duration were significant.

*The values in the axis scales are logarithms of the duration.*

The results from these tests, summarized in tables 5.15 and 5.16, show that the duration of topicalized preverbal non-nominative subjects is statistically significantly smaller than
that of broad focus non-nominative subjects for seven speakers—and also patent for speaker H6C \((t(8) = 2.13, p = .066)\). Even though these results are consistent with the predictions about the behaviour of this parameter, they contrast with the findings from the context of preverbal subjects, analyzed for Hypothesis 1, where duration differences were significant for all of the speakers. A possible explanation for the different behaviour between the two comparisons is that the tokens used in these sentences are causing this effect. Some of the DPs employed in the sentences that include datives are Spanish names such as Juan, Julia and Luis. Unfortunately, at that point I did not think about the impact that the variable realization that phonemes like /x/, and /n/ and /s/ in coda may pose, especially in dialects of the Caribbean in which they can be realized as [n], [ŋ] or a nasal vowel, and [s]>[h]>[Ø], respectively. Additionally, since these preverbal DPs constitute independent intonational phrases in most cases, and some of them have only one syllable (or they are oxytone, like Manuel, another name used), there may be two processes taking place in the same syllable. The first one would be a shortening process that reflects the type of information that the item encodes. The other is a lengthening process, which usually occurs in the syllable before the boundary. Given the properties of these constructions, it is likely that in these DPs both processes are taking place at the same time. In the case of the contexts of Hypothesis 1, although this lengthening process is also arguably happening, most of the nouns had two or more syllables; therefore, lengthening takes place in the last syllable of the DP, which is different from the stressed syllable. As a result, there is no interaction between the two processes.

Table 5.15. Ratio paired t-test results comparing the duration of non-nominative subjects in broad focus and topicalized contexts

<table>
<thead>
<tr>
<th>Dialect</th>
<th>Speaker</th>
<th>BF n-NomS vs. Top n-NomS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>M</td>
</tr>
<tr>
<td>Havana</td>
<td>H1C</td>
<td>1.11</td>
</tr>
<tr>
<td></td>
<td>H3C</td>
<td>1.11</td>
</tr>
<tr>
<td></td>
<td>H6C</td>
<td>1.05</td>
</tr>
<tr>
<td></td>
<td>H7C</td>
<td>1.15</td>
</tr>
<tr>
<td>Madrid</td>
<td>H2M</td>
<td>1.09</td>
</tr>
<tr>
<td></td>
<td>H4M</td>
<td>1.18</td>
</tr>
<tr>
<td></td>
<td>H5M</td>
<td>1.13</td>
</tr>
<tr>
<td></td>
<td>H6M</td>
<td>1.08</td>
</tr>
</tbody>
</table>
Table 5.16. Wilcoxon signed-rank tests comparing the duration of non-nominative subjects in broad focus and topicalized contexts

<table>
<thead>
<tr>
<th>Dialect</th>
<th>Speaker</th>
<th>Mdn</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Havana</td>
<td>H2C</td>
<td>1.14</td>
<td>.148</td>
</tr>
<tr>
<td>Madrid</td>
<td>H7M</td>
<td>1.09</td>
<td>.129</td>
</tr>
</tbody>
</table>

The scatterplots for the comparison of the duration of broad focus non-nominative subjects and topicalized CLLDs, shown in Figure 5.10, suggest that the same could be happening in these contexts. Ratio paired t-tests and Wilcoxon signed-rank tests were conducted to see whether differences in duration were significant. The results of these tests, summarized in tables 5.17 and 5.18, respectively, show that the duration of topicalized CLLDs is statistically significantly smaller than that of broad focus preverbal non-nominative subjects for five speakers, and also patent for speaker H1C ($t(6) = -2.32$, $p = .060$). As in the previous case, the variable realization of some of the segments, along with the possibility that the two processes may be taking place in the same syllable are likely the causes of these results.

Figure 5.10. Scatterplots displaying the paired duration data of broad focus non-nominative subjects and topicalized CLLDs, by three speakers of each dialect

*The values in the axis scales are logarithms of the duration.
Table 5.17. Ratio paired t-tests results comparing the duration of broad focus non-nominative subjects and topicalized CLLDs

<table>
<thead>
<tr>
<th>Dialect</th>
<th>Speaker</th>
<th>$M$</th>
<th>95% CI</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Havana</td>
<td>H1C</td>
<td>1.13</td>
<td>0.99 - 1.28</td>
<td>.060</td>
</tr>
<tr>
<td></td>
<td>H2C</td>
<td>1.14</td>
<td>0.92 - 1.40</td>
<td>.186</td>
</tr>
<tr>
<td></td>
<td>H6C</td>
<td>1.09</td>
<td>0.96 - 1.23</td>
<td>.149</td>
</tr>
<tr>
<td></td>
<td>H7C</td>
<td>1.13</td>
<td>1.02 - 1.25</td>
<td>.028</td>
</tr>
<tr>
<td>Madrid</td>
<td>H2M</td>
<td>1.03</td>
<td>0.90 - 1.18</td>
<td>.597</td>
</tr>
<tr>
<td></td>
<td>H4M</td>
<td>1.08</td>
<td>1.01 - 1.15</td>
<td>.024</td>
</tr>
<tr>
<td></td>
<td>H5M</td>
<td>1.10</td>
<td>1.01 - 1.20</td>
<td>.030</td>
</tr>
<tr>
<td></td>
<td>H6M</td>
<td>1.10</td>
<td>1.03 - 1.18</td>
<td>.014</td>
</tr>
<tr>
<td></td>
<td>H7M</td>
<td>1.06</td>
<td>0.90 - 1.25</td>
<td>.405</td>
</tr>
</tbody>
</table>

Table 5.18. Wilcoxon signed-rank tests results comparing the duration of broad focus non-nominative subjects and topicalized CLLDs

<table>
<thead>
<tr>
<th>Dialect</th>
<th>Speaker</th>
<th>$Mdn$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Havana</td>
<td>H3C</td>
<td>1.05</td>
<td>.008</td>
</tr>
</tbody>
</table>

Figures 5.11 and 5.12 illustrate the raw data of the tonal change by speaker, along with the confidence intervals, for Havana and Madrid, respectively. As in the contexts examined for Hypothesis 1, there is large overlap of the confidence intervals in both dialects. However, unlike the contexts analyzed for Hypothesis 1, in this case no pattern seems to emerge for any of the speakers. The results from the paired t-tests and Wilcoxon signed-rank tests—summarized in tables 5.19, 5.20 and 5.21—confirm that there are no statistically significant differences except for speaker H6C ($t(7) = -3.19$, $p = .015$) in the comparison of BF non-nominative subjects vs. CLLDs. Interestingly, the pattern that she employs in this case is different from the constructions that include subjects of transitive verbs. While broad focus non-nominative subjects display less tonal change than topicalized CLLDs ($M = -20.35$), broad focus subjects of transitive verbs display more tonal change than CLLDs (BF $M = 41.83$, CLLD $M = 30.83$, Table 5.8), as well as more tonal change than their topicalized counterparts ($Mdn = 15.95$, Table 5.10).

These results, however, are not unexpected. First, the use of the fundamental frequency to distinguish among old and new information in these two dialects is not across the board, as we saw in the contexts related to Hypothesis 1. Second, it is a fact that we are comparing elements that belong to the same syntactic category in the first position of the string. Although it can be argued that they bear different syntactic functions and occupy
different positions in the structure, i.e., non-nominative subjects versus left dislocated objects, it is possible that this type of information cannot be expressed due to restrictions that prosodic structures impose.

Figure 5.11. Box plots and confidence intervals displaying the tonal change of broad focus non-nominative subjects, topicalized non-nominative subjects, and topicalized CLLDs, produced by Havana speakers

Figure 5.12. Box plots and confidence intervals displaying the tonal change of broad focus non-nominative subjects, topicalized non-nominative subjects, and topicalized CLLDs, produced by Madrid speakers
Table 5.19. Paired t-tests results comparing the tonal change of non-nominative subjects in broad focus and topicalized contexts

<table>
<thead>
<tr>
<th>BF n-NomS vs. Top n-NomS</th>
<th>Dialect</th>
<th>Speaker</th>
<th>M</th>
<th>SE</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Havana</td>
<td>H1C</td>
<td>-6.39</td>
<td>6.03</td>
<td>.324</td>
<td></td>
</tr>
<tr>
<td></td>
<td>H2C</td>
<td>-0.49</td>
<td>7.33</td>
<td>.949</td>
<td></td>
</tr>
<tr>
<td></td>
<td>H3C</td>
<td>-8.14</td>
<td>5.44</td>
<td>.173</td>
<td></td>
</tr>
<tr>
<td></td>
<td>H6C</td>
<td>-13.19</td>
<td>10.66</td>
<td>.251</td>
<td></td>
</tr>
<tr>
<td></td>
<td>H7C</td>
<td>-2.09</td>
<td>2.86</td>
<td>.485</td>
<td></td>
</tr>
<tr>
<td>Madrid</td>
<td>H2M</td>
<td>-0.34</td>
<td>3.22</td>
<td>.918</td>
<td></td>
</tr>
<tr>
<td></td>
<td>H4M</td>
<td>9.16</td>
<td>5.89</td>
<td>.164</td>
<td></td>
</tr>
<tr>
<td></td>
<td>H5M</td>
<td>1.63</td>
<td>3.99</td>
<td>.694</td>
<td></td>
</tr>
<tr>
<td></td>
<td>H6M</td>
<td>-2.58</td>
<td>3.95</td>
<td>.532</td>
<td></td>
</tr>
<tr>
<td></td>
<td>H7M</td>
<td>-6.46</td>
<td>5.59</td>
<td>.281</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.20. Paired t-tests results comparing the tonal change of broad focus non-nominative subjects and topicalized CLLDs

<table>
<thead>
<tr>
<th>BF n-NomS vs. CLLD</th>
<th>Dialect</th>
<th>Speaker</th>
<th>M</th>
<th>SE</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Havana</td>
<td>H3C</td>
<td>-11.85</td>
<td>7.02</td>
<td>.135</td>
<td></td>
</tr>
<tr>
<td></td>
<td>H6C</td>
<td>-20.35</td>
<td>6.37</td>
<td>.015</td>
<td></td>
</tr>
<tr>
<td></td>
<td>H7C</td>
<td>-4.12</td>
<td>2.92</td>
<td>.202</td>
<td></td>
</tr>
<tr>
<td>Madrid</td>
<td>H2M</td>
<td>-4.20</td>
<td>2.70</td>
<td>.163</td>
<td></td>
</tr>
<tr>
<td></td>
<td>H4M</td>
<td>0.905</td>
<td>4.38</td>
<td>.843</td>
<td></td>
</tr>
<tr>
<td></td>
<td>H5M</td>
<td>-15.17</td>
<td>10.4</td>
<td>.187</td>
<td></td>
</tr>
<tr>
<td></td>
<td>H6M</td>
<td>0.639</td>
<td>3.77</td>
<td>.870</td>
<td></td>
</tr>
<tr>
<td></td>
<td>H7M</td>
<td>-10.32</td>
<td>5.38</td>
<td>.096</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.21. Wilcoxon signed-rank tests results comparing the tonal change of broad focus non-nominative subjects and topicalized CLLDs

<table>
<thead>
<tr>
<th>BF n-NomS vs. CLLD</th>
<th>Dialect</th>
<th>Speaker</th>
<th>Mdn</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Havana</td>
<td>H1C</td>
<td>-0.18</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>H2C</td>
<td>-2.70</td>
<td>.297</td>
<td></td>
</tr>
</tbody>
</table>

Finally, as in the contexts examined for Hypothesis 1, intensity seems not to be a prosodic correlate to distinguish between old and new information in these two dialects, as the data in Figure 5.13 illustrates. There seems to be no variation in the peak intensity among broad focus elements and topicalized elements, for any of the speakers.

Paired t-tests and Wilcoxon signed-rank tests were conducted to see if differences in peak intensity among elements in the different conditions were significant for any of the speakers. No statically significant differences were found except for one speaker (H2M:
(8) = 3.48, \( p = .008 \) in the comparison between non-nominative subjects in the two conditions of information structure, as the results summarized in tables 5.22 and 5.23 show. On the other hand, differences in intensity between broad focus non-nominative subjects and topicalized CLLDs were statistically significant for two speakers H1C (\( t(6) = -3.37, p = .015 \)) and H3C (\( t(8) = -2.72, p = .026 \)) (Table 5.24). The tendency is for topicalized CLLDs to show more intensity than broad focus non-nominative subjects. However, it should be noted that differences in these cases, while statistically significant, are very small (H1C, \( M = -2.01 \), and H3C, \( M = -2.35 \)), which is also the case of H2M in the comparison of BF n-NomS and Top n-NomS (\( M = 1.72 \)). Moreover, the fact that differences in intensity are not found in any other context for these speakers suggests that the differences seen in these cases should not be taken as evidence that intensity is used by these speakers. In this case, more data would be necessary to determine whether some of these speakers are using intensity to distinguish among different elements in different contexts of information structure.

Figure 5.13 Box plots and confidence intervals displaying the peak intensity of broad focus non-nominative subjects, topicalized non-nominative subjects, and CLLDs, produced by three speakers of each dialect
Table 5.22 Paired t-tests results comparing the peak intensity of non-nominative subjects in broad focus and topicalized contexts

<table>
<thead>
<tr>
<th>Dialect</th>
<th>Speaker</th>
<th>M</th>
<th>SE</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Havana</td>
<td>H1C</td>
<td>0.73</td>
<td>0.82</td>
<td>.403</td>
</tr>
<tr>
<td></td>
<td>H2C</td>
<td>-0.46</td>
<td>0.76</td>
<td>.564</td>
</tr>
<tr>
<td></td>
<td>H3C</td>
<td>-0.77</td>
<td>0.65</td>
<td>.270</td>
</tr>
<tr>
<td></td>
<td>H6C</td>
<td>0.91</td>
<td>1.04</td>
<td>.407</td>
</tr>
<tr>
<td></td>
<td>H7C</td>
<td>-0.82</td>
<td>0.45</td>
<td>.102</td>
</tr>
<tr>
<td>Madrid</td>
<td>H2M</td>
<td>1.72</td>
<td>0.49</td>
<td>.008</td>
</tr>
<tr>
<td></td>
<td>H4M</td>
<td>0.22</td>
<td>0.63</td>
<td>.730</td>
</tr>
<tr>
<td></td>
<td>H6M</td>
<td>0.07</td>
<td>0.56</td>
<td>.909</td>
</tr>
<tr>
<td></td>
<td>H7M</td>
<td>-1.12</td>
<td>0.72</td>
<td>.160</td>
</tr>
</tbody>
</table>

Table 5.23 Wilcoxon signed-rank tests results comparing the peak intensity of non-nominative subjects in broad focus and topicalized contexts

<table>
<thead>
<tr>
<th>Dialect</th>
<th>Speaker</th>
<th>Mdn</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Madrid</td>
<td>H5M</td>
<td>-0.52</td>
<td>.820</td>
</tr>
</tbody>
</table>

Table 5.24 Paired t-tests results comparing the peak intensity of broad focus non-nominative subjects and topicalized CLLDs

<table>
<thead>
<tr>
<th>Dialect</th>
<th>Speaker</th>
<th>M</th>
<th>SE</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Havana</td>
<td>H1C</td>
<td>-2.01</td>
<td>0.60</td>
<td>.015</td>
</tr>
<tr>
<td></td>
<td>H2C</td>
<td>0.03</td>
<td>0.87</td>
<td>.975</td>
</tr>
<tr>
<td></td>
<td>H3C</td>
<td>-2.35</td>
<td>0.87</td>
<td>.026</td>
</tr>
<tr>
<td></td>
<td>H6C</td>
<td>-0.06</td>
<td>0.36</td>
<td>.882</td>
</tr>
<tr>
<td></td>
<td>H7C</td>
<td>-1.21</td>
<td>0.92</td>
<td>.226</td>
</tr>
<tr>
<td>Madrid</td>
<td>H2M</td>
<td>0.17</td>
<td>0.69</td>
<td>.813</td>
</tr>
<tr>
<td></td>
<td>H4M</td>
<td>-1.07</td>
<td>0.80</td>
<td>.219</td>
</tr>
<tr>
<td></td>
<td>H5M</td>
<td>0.81</td>
<td>1.02</td>
<td>.450</td>
</tr>
<tr>
<td></td>
<td>H6M</td>
<td>1.39</td>
<td>0.89</td>
<td>.158</td>
</tr>
<tr>
<td></td>
<td>H7M</td>
<td>0.24</td>
<td>0.96</td>
<td>.807</td>
</tr>
</tbody>
</table>

5.6.3. Preverbal subjects of unaccusative and unergative verbs

The last comparison examines preverbal subjects generated in two different types of verbal configurations: unaccusative and unergative structures. According to Hypothesis 3, it is expected that these subjects behave alike prosodically since they bear the same informational status and their final position in the structure is arguably the same. We predicted that there would be no differences among the acoustic parameters of duration, fundamental frequency and intensity that these two types of preverbal subjects display.
As in the previous contexts, all speakers produced all sentences, and all tokens were used in the statistical analysis. 8 pairs of sentences were analyzed. Only cases of creaky voice were discarded for the F0 analysis. Examples of the analyzed sentences follow, along with the contexts in which they are uttered and their informational and syntactic statuses.

WHAT HAPPENED?

(15) Los estudiantes de Marta leyeron.  SUBJECT-Unerg- BROAD FOCUS

the students of Marta read

‘Marta’s students read.’

WHAT HAPPENED?

(16) Los estudiantes de Malta vinieron.  SUBJECT-Unacc- BROAD FOCUS

the students of Malta came

‘The students from Malta came.’

The graphs in Figure 5.14 illustrate that both types of subjects behave alike with respect to duration in most cases. The results from Wilcoxon signed-rank tests confirmed that there are no statistically significant differences between the duration of preverbal unaccusative and unergative subjects in broad focus contexts, as the summary of the results in Table 5.25 shows, consistent with the prediction of the behaviour about this acoustic parameter.
Figure 5.14. Scatterplots displaying the paired duration data of preverbal subjects of unaccusative and unergative verbs in broad focus contexts, produced by three speakers of each dialect.*The values in the axis scales are logarithms of the duration.

Table 5.25. Wilcoxon signed-rank tests results comparing the duration of broad focus preverbal subjects of unergative and unaccusative verbs

<table>
<thead>
<tr>
<th>Dialect</th>
<th>Speaker</th>
<th>Mdn</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Havana</td>
<td>H1C</td>
<td>1.00</td>
<td>.109</td>
</tr>
<tr>
<td></td>
<td>H2C</td>
<td>1.00</td>
<td>.383</td>
</tr>
<tr>
<td></td>
<td>H3C</td>
<td>1.00</td>
<td>.547</td>
</tr>
<tr>
<td></td>
<td>H6C</td>
<td>1.00</td>
<td>.129</td>
</tr>
<tr>
<td></td>
<td>H7C</td>
<td>1.00</td>
<td>.820</td>
</tr>
<tr>
<td>Madrid</td>
<td>H2M</td>
<td>1.00</td>
<td>.547</td>
</tr>
<tr>
<td></td>
<td>H4M</td>
<td>1.00</td>
<td>.652</td>
</tr>
<tr>
<td></td>
<td>H5M</td>
<td>1.00</td>
<td>.945</td>
</tr>
<tr>
<td></td>
<td>H6M</td>
<td>1.00</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>H7M</td>
<td>1.00</td>
<td>.820</td>
</tr>
</tbody>
</table>

Regarding fundamental frequency, as can be seen in figures 5.15 and 5.16, which illustrate the raw data of the tonal change with the calculated confidence intervals for the Havana and Madrid speakers, respectively, there seems to be no difference in tonal change between the two subject types. The results from Paired t-tests and Wilcoxon signed-rank tests, summarized in tables 5.26 and 5.27, respectively, confirm that there are
no statistically significant differences between these two subject types for any of the speakers.

**Figure 5.15.** Box plots and confidence intervals displaying the tonal change of preverbal subjects of unaccusative and unergative verbs in broad focus contexts, produced by Havana speakers

**Figure 5.16.** Box plots and confidence intervals displaying the tonal change of preverbal subjects of unaccusative and unergative verbs in broad focus contexts, produced by Madrid speakers
Table 5.26. Paired t-tests results comparing the tonal change of broad focus preverbal subjects of unergative and unaccusative verbs

<table>
<thead>
<tr>
<th>Dialect</th>
<th>Speaker</th>
<th>M</th>
<th>SE</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Havana</td>
<td>H2C</td>
<td>9.62</td>
<td>5.58</td>
<td>.128</td>
</tr>
<tr>
<td></td>
<td>H3C</td>
<td>-2.52</td>
<td>2.60</td>
<td>.365</td>
</tr>
<tr>
<td></td>
<td>H6C</td>
<td>4.61</td>
<td>15.00</td>
<td>.767</td>
</tr>
<tr>
<td></td>
<td>H7C</td>
<td>-3.05</td>
<td>5.45</td>
<td>.591</td>
</tr>
<tr>
<td>Madrid</td>
<td>H2M</td>
<td>1.53</td>
<td>5.08</td>
<td>.772</td>
</tr>
<tr>
<td></td>
<td>H4M</td>
<td>-3.62</td>
<td>3.46</td>
<td>.325</td>
</tr>
<tr>
<td></td>
<td>H5M</td>
<td>5.95</td>
<td>6.45</td>
<td>.383</td>
</tr>
<tr>
<td></td>
<td>H6M</td>
<td>7.19</td>
<td>6.60</td>
<td>.312</td>
</tr>
<tr>
<td></td>
<td>H7M</td>
<td>-4.28</td>
<td>5.00</td>
<td>.417</td>
</tr>
</tbody>
</table>

Table 5.27. Wilcoxon signed-rank tests results comparing the tonal change of broad focus preverbal subjects of unergative and unaccusative verbs

<table>
<thead>
<tr>
<th>Dialect</th>
<th>Speaker</th>
<th>Mdn</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Havana</td>
<td>H1C</td>
<td>9.98</td>
<td>.383</td>
</tr>
</tbody>
</table>

Figure 5.17 illustrates the peak intensity raw data with the calculated confidence intervals for six speakers, three from each dialect. As in the previous cases, there seem to be no differences in peak intensity between the two compared elements for any of the speakers. Paired t-tests and Wilcoxon signed-rank tests results confirm that there are no statistically significant differences in the peak intensity between the two subject types in any of the two dialects, as the results summarized in tables 5.28 and 5.29 show. It would be inadequate, however, to propose that this constitutes evidence that these subjects behave alike. The intensity results simply corroborate that these two dialects seem not to use this acoustic parameter to signal prosodic differences in any of the contexts analyzed in this experiment. As noted before, the same result has been found in the comparison of broad focus and narrow focused elements in Madrid (see Face, 2000).
Figure 5.17. Box plots and confidence intervals displaying the peak intensity of preverbal subjects of unaccusative and unergative verbs in broad focus contexts, produced by three speakers of each dialect.

Table 5.28. Paired t-tests results comparing the peak intensity of broad focus preverbal subjects of unergative and unaccusative verbs

<table>
<thead>
<tr>
<th>Dialect</th>
<th>Speaker</th>
<th>M</th>
<th>SE</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Havana</td>
<td>H1C</td>
<td>0.94</td>
<td>1.14</td>
<td>.439</td>
</tr>
<tr>
<td></td>
<td>H2C</td>
<td>-0.36</td>
<td>1.03</td>
<td>.735</td>
</tr>
<tr>
<td></td>
<td>H3C</td>
<td>0.77</td>
<td>0.71</td>
<td>.317</td>
</tr>
<tr>
<td></td>
<td>H6C</td>
<td>0.22</td>
<td>0.70</td>
<td>.766</td>
</tr>
<tr>
<td></td>
<td>H7C</td>
<td>0.07</td>
<td>0.69</td>
<td>.922</td>
</tr>
<tr>
<td>Madrid</td>
<td>H2M</td>
<td>-0.31</td>
<td>0.57</td>
<td>.600</td>
</tr>
<tr>
<td></td>
<td>H4M</td>
<td>2.33</td>
<td>1.15</td>
<td>.078</td>
</tr>
<tr>
<td></td>
<td>H5M</td>
<td>0.23</td>
<td>0.54</td>
<td>.680</td>
</tr>
<tr>
<td></td>
<td>H7M</td>
<td>-0.72</td>
<td>0.74</td>
<td>.360</td>
</tr>
</tbody>
</table>

Table 5.29. Wilcoxon signed-rank tests results comparing the peak intensity of broad focus preverbal subjects of unergative and unaccusative verbs

<table>
<thead>
<tr>
<th>Dialect</th>
<th>Speaker</th>
<th>Mdn</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Madrid</td>
<td>H6M</td>
<td>-2.29</td>
<td>.250</td>
</tr>
</tbody>
</table>
5.6.4. Phrasing patterns and pitch accents

The last aspects that this study examines are the phrasing patterns of the different syntactic structures, and the pitch accents associated with topicalized elements. Phrasing patterns in this case only refer to the preverbal DP with respect to the rest of the utterance. Thus the specific data that I report is whether these preverbal elements are phrased as independent units. Phasing patterns are analyzed in the context of the three hypotheses of our experiment. I present phrasing results first, and after that some general comments about pitch accents.

5.6.4.1. Phrasing patterns

In order to do the phrasing analysis, sentences were divided according to their informational properties into two categories: broad focus and topicalization. Next, a second grouping was done taking into account the syntactic function of the first element. The following scheme depicts these divisions.

<table>
<thead>
<tr>
<th>Broad Focus</th>
<th>Topicalization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects of transitive verbs</td>
<td>Subjects of transitive verbs</td>
</tr>
<tr>
<td>Non-nominative subjects</td>
<td>Non-nominative subjects</td>
</tr>
<tr>
<td>Subjects of unergative verbs</td>
<td>CLLD</td>
</tr>
<tr>
<td>Subjects of unaccusative verbs</td>
<td></td>
</tr>
</tbody>
</table>

A third division was made taking into account syllabic quantity. The grouping by syllabic quantity corresponds to some of the syntactic divisions of the previous step. Sentences with subjects of transitive verbs, preverbal non-nominative subjects of psych verbs and other unaccusative predicates, and CLLDs have short preverbal DPs that range from 2 to 4 syllables, with an average of 3.25 syllables. Sentences with preverbal nominative subjects of unaccusative and unergative verbs have longer preverbal DPs that range from 5 to 8 syllables, with an average of 5.5 syllables. Hence there are two groups: short DPs (2-4 syllables) and long DPs (5 to 8 syllables).

VPs also have similar lengths once separated into the previous syntactic groups. Transitive predicates range from 4 to 7 syllables, with an average of 5 syllables. VPs of
psych verbs and other unaccusative predicates range from 6 to 7 syllables, with an average of 6.7 syllables. VPs of unaccusative and unergative verbs range from 2 to 3 syllables, with an average of 2.8 syllables. As a consequence, in the case of VPs there are also two groups: long and short. Short VPs correspond to unaccusative and unergative structures with preverbal nominative subjects, while long VPs correspond to transitive, psychological, and other unaccusative predicates. Although I do not report any data on VP phrasing, it is important for sentences that are compared to have the same balance between the DP and the VP because of the effect that syllabic quantity has on phrasing patterns. Accordingly, structures that are compared in terms of phrasing always have the same balance.

Several cues were taken into account to determine whether an element was phrased independently. These cues were taken from Frota et al. (2007), and are summarized below. When one or more cues were present, the element was considered to be an independent phrase (IP).

- Use of a pre-boundary stretch indicating a boundary such as: continuation rise (-H), sustained pitch (-H!), or other
- Use of boundary tones signalling a boundary: high (H%) and low (L%)
- Pitch reset after the boundary
- Use of pauses after the DP

The results from the division based on information structure (broad focus and topicalized contexts) show a strong preference for phrasing the first syntactic constituent as an IP in both dialects (Table 5.30). These results are in line with what has been found for Peninsular Spanish, as most studies that have analyzed broad focus utterances with transitive verbs report that the most common phrasing pattern in this variety is (S)(VO) (D’Imperio et al., 2005; Elordieta, Frota & Vigário, 2005). The same phrasing pattern has also been attested in contexts of different types of left dislocations, such as hanging topics, clitic left dislocations and left dislocations without a resumptive pronoun, according to Feldhausen (2016) for Peninsular Spanish.

Table 5.30. Percentages of preverbal DPs phrased as IPs, by information structure context and dialect

<table>
<thead>
<tr>
<th>Dialect</th>
<th>N BF</th>
<th>% IP BF</th>
<th>N Top</th>
<th>% IP Top</th>
</tr>
</thead>
<tbody>
<tr>
<td>Havana</td>
<td>201</td>
<td>72.64</td>
<td>226</td>
<td>82.74</td>
</tr>
<tr>
<td>Madrid</td>
<td>203</td>
<td>87.68</td>
<td>228</td>
<td>85.09</td>
</tr>
</tbody>
</table>
Although the tendency is to phrase the first constituent independently, our data shows that there are differences between the two dialects, and some differences between the two contexts of information structure in Havana. With respect to topicallyized contexts, both dialects behave in a similar way (Havana: 82.74%, Madrid: 85.09%) ($\chi^2 (1) = 0.34, p = .563$) (Table 5.30). Differences arise when we compare the broad focus condition, in which Madrid shows a higher percentage of DPs phrased as IPs with respect to Havana, 88% vs. 73%, respectively ($\chi^2 (1) = 14.35, p < .001$). It is interesting that in Madrid these two contexts of information structure behave alike with respect to their phrasing patterns (87.68% and 85.09%, for BF and topicalization, respectively) ($\chi^2 (1) = 0.61, p = .434$).

On the other hand, in Havana, although there is a strong tendency for speakers to phrase the first constituent as an IP in both contexts, there are differences between the two. Topicalized contexts display a stronger tendency to have their first element phrased as an IP with respect to broad focus, 82.74% versus 72.64%, respectively ($\chi^2 (1) = 6.33, p = .012$). Even though this first grouping does not take into account syllabic quantity, these results are very similar to those obtained once syllabic quantity is incorporated, as will be seen next.

The results from the analysis that incorporates different syntactic structures and syllabic quantity are more complex. For this part of the analysis, three general groups were established based on the data grouping and the type of elements that are compared. The first group contains DPs with different syntactic function, in different conditions of information structure. The relevant comparisons in this case are the following two: subjects of transitive verbs versus CLLDs, and non-nominative subjects of psych verbs and other unaccusative predicates versus CLLDs. The second group comprises the same elements in two conditions of information structure. In this group subjects of transitive verbs and preverbal non-nominative subjects of psych verbs and other unaccusative predicates are compared in two conditions, broad focus and topicalization. The last group includes different types of syntactic subjects in the same condition of information structure. In this case, the compared elements are subjects of unergative verbs and derived subjects of unaccusative verbs in broad focus contexts.

Table 5.31 shows the percentages of phrasing the first DP as an IP that the different syntactic structures display in each dialect in the two conditions of information structure.
It can be seen that some syntactic contexts seem to show sensitivity to information structure, i.e., they exhibit variation in their phrasing patterns depending on the context of information structure in which they appear, while others stay the same. These tendencies are not across the board, and there is also dialectal variation.

Table 5.31. Percentages of DPs phrased as IPs in each syntactic and informational context, by dialect

<table>
<thead>
<tr>
<th>Dialect</th>
<th>Syntactic Structure</th>
<th>N BF</th>
<th>% BF</th>
<th>N Top</th>
<th>% Top</th>
</tr>
</thead>
<tbody>
<tr>
<td>Havana</td>
<td>Subject (transitive)</td>
<td>70</td>
<td>65.71</td>
<td>69</td>
<td>78.26</td>
</tr>
<tr>
<td></td>
<td>Non-nominal Subject</td>
<td>44</td>
<td>79.54</td>
<td>44</td>
<td>81.82</td>
</tr>
<tr>
<td></td>
<td>CLLD</td>
<td>-</td>
<td>-</td>
<td>113</td>
<td>85.84</td>
</tr>
<tr>
<td></td>
<td>Subject (unaccusative)</td>
<td>32</td>
<td>74.42</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Subject (unergative)</td>
<td>33</td>
<td>75.00</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Madrid</td>
<td>Subject (transitive)</td>
<td>70</td>
<td>71.43</td>
<td>70</td>
<td>71.43</td>
</tr>
<tr>
<td></td>
<td>Non-nominal Subject</td>
<td>45</td>
<td>93.33</td>
<td>45</td>
<td>82.22</td>
</tr>
<tr>
<td></td>
<td>CLLD</td>
<td>-</td>
<td>-</td>
<td>113</td>
<td>94.69</td>
</tr>
<tr>
<td></td>
<td>Subject (unaccusative)</td>
<td>44</td>
<td>95.45</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Subject (unergative)</td>
<td>44</td>
<td>100</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Taking into account the syntactic function of the first element and its informational properties, there are differences between preverbal subjects of transitive verbs in the broad focus condition and topicalized CLLDs. CLLDs show a higher percentage of phrasing the preverbal DP as an IP in both dialects: 65.71% vs. 85.84%, Havana, and 71.43% vs. 94.69%, Madrid (Havana: $\chi^2(1) = 10.25, p = .001$, and Madrid: $\chi^2(1) = 19.19, p < .001$). This result supports our first hypothesis, which states that preverbal subjects in broad focus sentences do not pattern prosodically with CLLDs. The phrasing patterns’ results show that there are differences between these two types of syntactic constructions. On the other hand, there are no differences between preverbal non-nominal subjects of psych verbs and other unaccusative predicates in the broad focus condition and topicalized CLLDs in any of the two dialects, 79.54% vs. 85.84%, Havana, and 93.33% vs. 94.69%, Madrid (Havana: $\chi^2(1) = 0.76, p = .383$, and Madrid: $\chi^2(1) = 0.10, p = .748$) against hypothesis two.

With respect to cases in which the same DP is evaluated in the two conditions of information structure, there is a tendency for subjects of transitive verbs to display higher rates of being phrased as IPs in the topicalized condition than in the broad focus condition in Havana, 65.71% vs. 78.26%, respectively ($\chi^2(1) = 2.71, p = 0.100$). However, this
tendency does not hold in Madrid, which shows the same phrasing pattern for these elements in the two conditions of information structure (71.43%). The behaviour found in the Madrid variety is probably less puzzling than we may think at first, if we consider the previous results from the F0 acoustic analysis in the two dialects. The F0 analysis also showed that there were statistically significant differences in the Havana variety in these contexts for some of the speakers, whereas in Madrid no statistically significant differences were found. Since the parameters for the phrasing analysis are based on the description of the F0 contour, these results are only expected.

Preverbal non-nominative subjects of psych verbs and other unaccusative predicates phrasing patterns, on the other hand, do not vary between the two conditions of information structure in Havana: 79.54%, broad focus condition, vs. 81.82%, topicalization ($\chi^2 (1) = 0.07, p = .787$). The difference observed in Madrid—93.33%, broad focus, vs. 82.22%, topicalization—is not statistically significant ($\chi^2 (1) = 2.59, p = .108$). Moreover, this difference is the result of variations related to one speaker phrasing patterns. An individual analysis reveals that the percentage drop in the topicalized condition in Madrid is linked to one speaker who significantly changes her phrasing pattern in this structure.

Table 5.32 shows the individual percentages of IPs in the two conditions of information structure for preverbal non-nominative subjects of psych verbs and other unaccusative predicates, and subjects of transitive verbs in Madrid. Note the significant change of H4M phrasing patterns in constructions that include non-nominative subjects, 77.78% vs. 35.71%, broad focus and topicalization, respectively, versus the similarity of phrasing patterns in contexts that include preverbal subjects of transitive verbs (35.71%, in both cases). Also note the more stable behaviour that the rest of the speakers show.
Table 5.32. Percentages of preverbal DPs phrased as IPs by speaker, syntactic context, and informational context, in the Madrid variety

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Broad focus</th>
<th>Topicalization</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S (transitive)</td>
<td>n-NomS</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>H2M</td>
<td>14</td>
<td>100</td>
</tr>
<tr>
<td>H4M</td>
<td>14</td>
<td>35.71</td>
</tr>
<tr>
<td>H5M</td>
<td>14</td>
<td>78.57</td>
</tr>
<tr>
<td>H6M</td>
<td>14</td>
<td>78.57</td>
</tr>
<tr>
<td>H7M</td>
<td>14</td>
<td>78.57</td>
</tr>
</tbody>
</table>

Going back to Table 5.31, it can be noticed that, in both dialects, subjects of unaccusative and unergative verbs behave alike (Havana: 74.42% vs. 75.00%, Madrid: 95.45% vs. 100%, unaccusative and unergative, respectively, in Madrid) (Havana: \( \chi^2 (1) = 0.004, p = .950 \); Madrid: \( \chi^2 (1) = 2.05, p = .153 \)). In this case, as in the previous contexts, the tendency is for the preverbal DP to be phrased independently. One more time, these two types of subjects seem to be prosodically indistinguishable in Spanish. Note that the tendency of phrasing the first element as an IP is higher in Madrid than in Havana for both types of subjects (unaccusative verbs: 74.42% vs. 95.45%, unergative verbs: 75.00% vs. 100%) (unaccusatives: \( \chi^2 (1) = 7.57, p = .006 \); unergatives: \( \chi^2 (1) = 12.57, p < .001 \)). As the reader probably recalls these are cases of long preverbal DPs (7 syllables) followed by short VPs (2 syllables).

Although phrasing the first DP as an IP is a trend that has been attested in all cases, a pertinent question at this point is about the differences between these two dialects. While both dialects show high rates of phrasing the first DP independently, in the case of Madrid this tendency seems to be categorical in unaccusative and unergative constructions. Thus, not only the effect of syllabic quantity on phrasing patterns has to be taken into account when proposing any mapping between syntactic structure and phrasing—as phrasing patterns are influenced/determined by syllabic quantity—but also the specific phonotactic features of a given dialect.

With respect to the different boundary cues that speakers use to signal an IP, in this study we found that the realization of the preboundary stretch as a continuation rise (-H) (Havana: BF 64.50%, Top 77.43%; Madrid: BF 81.77%, Top 85.53%) and the presence of a high tone (H%) at the boundary (Havana: BF 61.69, Top 74.11%; Madrid: BF
79.80%, Top 80.26%) are the most common cues for both dialects. Another frequent cue in Madrid is pitch reset (BF 67.00%, Top 61.40%). Pauses are less frequent in both dialects, as can be seen in Table 5.33.

Table 5.33. Percentages of the different boundary cues used, by dialect

<table>
<thead>
<tr>
<th>Dialect</th>
<th>Information Structure</th>
<th>Cont. Rise</th>
<th>Bound. Tone (H)</th>
<th>Pitch Reset</th>
<th>Pauses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>Havana</td>
<td>BF</td>
<td>200</td>
<td>201</td>
<td>201</td>
<td>201</td>
</tr>
<tr>
<td></td>
<td>Top</td>
<td>226</td>
<td>224</td>
<td>226</td>
<td>226</td>
</tr>
<tr>
<td>Madrid</td>
<td>BF</td>
<td>203</td>
<td>203</td>
<td>203</td>
<td>203</td>
</tr>
<tr>
<td></td>
<td>Top</td>
<td>228</td>
<td>228</td>
<td>228</td>
<td>228</td>
</tr>
</tbody>
</table>

It is interesting how topicalized contexts show a higher tendency to have their first element phrased independently, especially in some syntactic environments—e.g., CLLDs vs. preverbal subjects of transitive verbs—and how pauses seem to be playing no role in this distinction.

There is also great variation in terms of the use of pauses by individual speakers as can be seen in Table 5.34, which summarizes the percentages of pause use per speaker in the two conditions of information structure in both dialects. Note that while some speakers use pauses in the two conditions, others do not use pauses at all (H1C, H3C, H7M). Moreover, there is no pattern with respect to the use of pauses in the two conditions of information structure: three speakers decrease their use in the topicalized condition (H2C, H4M, H6M), three speakers increase their use of pauses (H6C, H7C, H5M), and speaker H2M behaves the same in the two conditions.

Table 5.34. Percentages of pauses used, by individual speakers

<table>
<thead>
<tr>
<th>Dialect</th>
<th>Speaker</th>
<th>Broad focus</th>
<th>Topicalization</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>Havana</td>
<td>H1C</td>
<td>40</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>H2C</td>
<td>39</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>H3C</td>
<td>40</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>H6C</td>
<td>41</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>H7C</td>
<td>41</td>
<td>46</td>
</tr>
<tr>
<td>Madrid</td>
<td>H2M</td>
<td>41</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>H4M</td>
<td>41</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>H5M</td>
<td>41</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>H6M</td>
<td>39</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>H7M</td>
<td>41</td>
<td>46</td>
</tr>
</tbody>
</table>
Finally, the percentages of the use of pauses in this study are lower than those reported in previous works (Face, 2003; Domínguez, 2004, Frota et al., 2007). There are two possible explanations for this, and both relate to the type of data. The sentences employed in our experiment can be classified as short sentences, whereas other studies have used longer sentences. It is a well-know fact that syllabic quantity has an impact on different aspects of speech production; therefore differences in utterance length have an impact on the use of pauses as well. The other factor may be differences in elicitation techniques. Previous studies consisted on reading tasks, and this study does not correspond to this type of task. Although in this study participants were given the words that they had to use in a vertical way that depicts the word order that they had to employ, the verb was not conjugated. This intermediate processing step forced speakers to come up with the sentence themselves, which places our sentences more on the side of spontaneous speech. A plausible hypothesis is that reading tasks can potentially have the effect of a more careful pronunciation; thus they could trigger a higher use of pauses.

There is a third possibility that does not pertain to the data, but to individual differences among speakers. It could be the case that some speakers may use pauses more frequently than others. If this is so, most of our participants fall on the side of the spectrum that uses fewer pauses. In this case, it would be a matter of incorporating more individuals to this analysis to see if it is the data that is having an effect or speakers’ preferences. The percentages of pause use per speaker, summarized in Table 5.34, reveal that even with this limited number of speakers there is variation.

Going back to Table 5.33, it can be noticed that an important difference between these two dialects is that pitch reset is more frequently used in Madrid than in Havana, in both conditions of information structure (Madrid: BF 67.00%, Top 61.40%; Havana: BF 26.37% Top 13.71%) (Broad Focus: $\chi^2 (1) = 66.96, p = < .001$; Topicalization: $\chi^2 (1) = 109.93, p = < .001$). This difference could potentially contribute to a more melodic contour in the case of the Madrid variety versus what may sound like a more monotonous tune of the Havana variety. This characterization corresponds to the impression of what some speakers perceive about the two dialects. Moreover, the limited use of pitch reset in the Havana variety may also contribute to less differentiation among the different elements, and therefore to the perception of bigger chunks, which could potentially make
comprehension more difficult, especially among speakers from different dialects. Another impressionistic characterization reported by speakers of other varieties with respect to the Havana variety.

There are also differences between the two dialects regarding the realization of the preboundary stretch as a continuation rise in both contexts of information structure. In this case, we also find higher rates in Madrid (BF: 81.77%, Top: 85.53%) when compared to Havana (BF: 64.50%, Top: 77.43%) (Broad Focus: $\chi^2(1) = 15.32, p < .001$; Topicalization: $\chi^2(1) = 4.93, p = .026$). The use of a High boundary tone (H%) in the broad focus condition is more frequent in Madrid than in Havana as well (79.80% vs. 61.69%, respectively) ($\chi^2(1) = 16.02, p = < .001$). However, in the topicalized condition, both dialects behave alike with respect to the use of high boundary tones (Havana: 74.11%, Madrid: 80.26%) ($\chi^2(1) = 2.43, p = .119$).

One can hypothesize that the use of less tonal movements in the case of Havana contributes to the perception of a faster speech rate in this variety. Perceived pauses are the result of a combination of tonal movements that make the listener perceive a given point in the F0 as a silence, when in fact there is no actual silence (Nootboom, 1997). Thus the lack of tonal movements or their lower frequency may contribute to the perception of a faster speech rate.

Finally, there seems to be a trade off between the use of these prosodic cues and their potential use as markers of different syntactic or informational structures, at least in this piece of data. Differences in the use of some of the cues in the two contexts of information structure in Havana (High boundary tone: 61.69% vs. 74.11%; Continuation rise: 64.50% vs. 77.43%; Pitch reset: 26.37% vs. 13.71%, for BF and Top, respectively) (High boundary tone: $\chi^2(1) = 7.53, p = 0.006$; Continuation Rise: $\chi^2(1) = 8.68, p = .003$; Pitch reset: $\chi^2(1) = 16.35, p = < .001$) contribute to the differentiation of some of the elements and syntactic structures in these two contexts (see results from the tonal change analysis of preverbal subjects of transitive verbs in section 5.6.1). On the other hand, in Madrid, most of these differences are not present due to the use of the different prosodic cues across the board (High boundary tone: 79.80% vs. 80.26%; Continuation rise: 81.77% vs. 85.53%; Pitch reset: 67.00% vs. 61.40%, for BF and Top, respectively) (High
boundary tone: $\chi^2 (1) = 0.01, p = .905$; Continuation Rise: $\chi^2 (1) = 1.11, p = .292$; Pitch reset: $\chi^2 (1) = 1.46, p = .227$.

5.6.4.2. Peak alignment and pitch accents

Peak alignment in topicalized contexts is generally postonic in both dialects. This is not unexpected because topicalized elements being old information are not supposed to bear any type of pitch accent that signals prominence, and late peak alignment is the default option in Spanish. The specific pitch accents that appear in the data are L*+H and L+>H*.

There are also a few instances of topicalized elements in which the peak occurs within the stressed syllable in this sample. These cases correspond to DPs that contain oxytone words that are separated from the rest of the utterance by a pause. Since the occurrence of this pitch accent derives from the combination of the specific phonetic properties of the word and the insertion of a pause, which interrupts the F0 and forces the pitch accent to be realized within the stretch before the silence, the existence of this pitch accent can be classified as a by-product. Thus we can conclude that even though this configuration is found in the data, it is not representative. (This pitch accent is normally used in nuclear position, and it is often linked to focalization. This context does not match any of the cases included in this study.)

Finally, all topicalized elements in this sample, despite being old information, bear pitch accents. As pointed out before, this could be the result of the position that these elements occupy in the string. Since elements in this study always appear in the first position, which is privileged in the sense that it sets the course of the F0 in the rest of the utterance, it is expected that they bear pitch accents.

5.7. Discussion and conclusions

The findings from this study provide further characterization of the prosodic strategies that the dialectal varieties of Spanish spoken in Havana, Cuba, and Central Peninsular Spain employ. Particularly, the acoustic results from the prosodic characterization of topicalized elements, understood as old information, expand the empirical base of our
understanding of informationally marked elements in Spanish. Additionally, the data from the Havana variety is a welcome addition, for this is a less studied dialect.

Overall, these results show that there are prosodic differences between elements that are new information and topicalized elements that are old information. Specifically, DPs marked as old information differ from DPs in broad focus or all-new-information contexts in two acoustic parameters: duration and fundamental frequency. In both dialects, DPs that are old information display less duration than DPs that are new information. This result is in line with what has been found in other languages (Féry & Kügler, 2008; Kügler, 2008). Furthermore, it supports the idea that old information, being already part of the discourse, is less salient. These results also corroborate the cross-linguistic trend that duration is a very stable acoustic correlate used to signal differences in information structure. In Spanish, particularly, duration seems to be the most reliable cue as this and previous studies have shown (De la Mota, 1998; Face, 2001, 2002).  

With respect to the use of the F0, the results that our data provides are less categorical. This is not entirely unexpected given the idiosyncratic use of this parameter. In addition to this, and before addressing the F0 results, there are two factors that need to be considered in this particular instance that can have an impact on the behaviour of the F0: the position the element occupies in the string, and the lack of saliency of old information topics. Regarding the first factor, the DPs analyzed in this study always appear in the first position of the string. While this is not a problem per se, it is a fact that this position is prosodically privileged in the sense that it sets the course of the F0 in the utterance. Given the tendency that most neutral utterances display crosslinguistically, which is for the F0 to be lower in every subsequent point in time, it is generally expected that the first pitch accent of the utterance is the highest one (at least in broad focus statements). This tendency, known as downstepping or declination, is attested in tonal and non-tonal languages alike (Ladd, 2008; Nooteboom, 1997). By the same token, the element that appears in this position is less susceptible to display variations in its fundamental frequency because it is the one that sets the course of the F0. Furthermore, since there is  

74 Except for Buenos Aires, Toledo (1989).
nothing before it to which it could be compared, any change in its F0 is probably more
difficult to interpret. Although old information is supposed to be less salient and changes
in F0 can potentially convey that, one can speculate that the fact that these DPs appear in
the first position makes changes in their F0 less likely to occur.

Despite these facts, results from the behaviour of the F0 raise some interesting points
when analyzed from the perspective of the interface between syntax and phonology. Let
us abstract away from the impact that the previous factors may have on the prosodic
realization of these elements, and also from the cross-linguistic tendency that right
branching languages display, and focus on the behaviour of the F0 in this piece of data.
The following discussion is based on the results from the Havana variety, as no
statistically significant changes in F0 were observed in Madrid. A question that I will
address later is precisely how a theory of the interface of syntax and phonology can
incorporate this dialectal difference.

The first thing worth noting is that changes in F0 to distinguish between old and new
information are not across the board. In other words, not all speakers use F0 to
distinguish between these two contexts of information structure. Furthermore, among
those speakers who show some variations in F0, the strategies used are not the same.
While most speakers use more tonal change in structures that contain old information
elements, one speaker uses the opposite pattern. This issue gets more complicated than
differences in information structure, as the study also includes differences in the syntactic
functions of the elements. As the reader recalls, elements analyzed in this experiment
correspond to different syntactic functions, such as subjects of transitive verbs, CLLDs,
and non-nominative subjects.

Based on the limited data analyzed in this study, one may suggest that changes in F0
could indeed be signalling differences between old information elements (in the left
periphery) and broad focus elements. However, not all old information elements that
move to the left periphery show the same behaviour. Moreover, there seem to be
syntactic environments in which changes in F0 are more likely to occur. Specifically, a
more stable use of F0 seems to be linked to syntactic environments that correlate with
word order variation, in addition to the specific informational status of the elements that
appear in them. Constructions that display a different word order than the same utterance in a broad focus context—as a consequence of having items marked as old information that move to left periphery—show more variation in the F0 of their first element with respect to broad focus constructions. On the other hand, differences in information structure that do not alter the linear word order are not always marked. In other words, elements that are marked as old information and their position in the string also varies are more likely to show F0 changes with respect to broad focus elements that appear in the same first position of the string. This instance can be considered an example of a redundant behaviour of the system, since these elements are marked twice: first, by occupying a different locus in the string (this is the result of syntactic movement), and second, by differences in their F0 when compared to broad focus elements in the same position of the string.

The most significant characteristic of this prosodic behaviour is that, for most speakers, it does not extend to syntactic environments that do not display differences in word order; despite the fact that in these cases there is also syntactic movement to the left periphery. The different treatment of the F0 in these two types of syntactic environments may be the cue to explain its variable use among speakers, and between the two dialects as well. The fact that F0 changes mainly appear in contexts that also involve changes in word order points to the possibility that the use of the F0 in this context may be of secondary order. Since changes in word order already signal syntactic differences among these constructions (topicalization vs. broad focus structures, for instance), one can hypothesize that the use of the F0 would be optional. Furthermore, the fact that there is no trace of this prosodic behaviour in Madrid also supports the idea that the use of the F0 in this case only comes to reinforce the already manifest difference in word order. Other syntactic environments analyzed in this study also support this idea. For instance, there are no differences between preverbal non-nominative subjects of psych verbs and other unaccusative predicates in the two conditions of information structure, or between broad focus preverbal non-nominative subjects and old information CLLDs. Moreover, differences among preverbal subjects of transitive verbs are only relevant for two Havana speakers, and the magnitude of the tonal change is very small.
We can conclude that the use of the F0 offers a sharp contrast when compared to that of
duration. Additionally, the disparity among its use in the different syntactic environments
constitutes a problem when we consider the type of syntactic information that is
prosodically represented. However, despite the different behaviour of these two acoustic
parameters, and dialectal differences among their use, the findings of this study suggest
that grammatical properties that are common within a given language are represented in
all dialects.

5.7.1. Prosodic marking

One of the most extended claims about the characterization of old information is that it is
unstressed. As noted in the previous chapter, the notion of old information being
unstressed probably derives from the behaviour of pronouns, which in some contexts are
instances of old information and are said to be unstressed. Although the property of being
unstressed has not been defined, one of the possibilities that comes to mind is the
existence of a specific pitch accent that signals old information. Other options are the
absence altogether of pitch accents, or the realization of pitch accents in a less salient way
in comparison to other elements that bear the same pitch accents in other contexts of
information structure. With respect to the latter, for instance, there could be some
variation in the acoustic realization of old information elements, such as a lower
fundamental frequency associated with their peak when compared to the same element in
broad focus context. However, the unstressed notion has proved to be more complex than
the qualitative and quantitative behaviour of the fundamental frequency, as duration is the
main cue that signals the distinction between elements that are old information and
elements that are part of broad focus utterances in these two dialects.

Another related topic to revisit in the light of this data is the deaccenting mechanism.
According to Face (2003) and Domínguez (2004), deaccenting is possible in Spanish.75
Old information elements, which are already part of the discourse, would be good
candidates to be deaccented, since that would be another way to be less salient. This

75 Domínguez (2004) proposes that deaccenting is available in Spanish in cases of topicalization when the
element appears in postfocal position.
option, however, does not seem to be available in cases in which the element appears in the first position of the string.

The differentiation in the prosodic behaviour of elements based on the position that they occupy in the string is by no means new. Féry and Kügler (2008) and Kügler (2008) found that in German there are differences in the realization of topics depending on the position in which they appear with respect to focus. They found that when topics are realized in prefocal position they display shorter duration and a flatter F0 contour in comparison to the same elements in a broad focus context. When topics are realized in a postfocal position, they display about the same duration and F0 pattern as the same element in a broad focus context. Based on the data obtained in our experiment, in which old information topics are in prefocal position, Spanish topics behave similarly to German topics in prefocal position with respect to duration. F0 results, again, are more complex. It remains to be tested the behaviour of old information topics in postfocal position in Spanish with respect to duration.

Finally, it can be concluded that there are no specific pitch accents associated with old information in Spanish, at least not when the element occupies the first position in the string. The pitch accents associated with elements classified as old information in this data are L*+H and L+>H*, which are instances of late peak alignment. These pitch accents are also found in contexts of new information and even in contexts of narrow focus in Spanish. In the case of narrow focus, however, there is an increase in the fundamental frequency associated with the peak of the narrow focused element, with respect to the same element in the broad focus condition (Face 2002). A question that needs to be answered is about the exact location of the peak in this data. Although its realization is postonic, it could be the case that it is displaced with respect to broad focus. In other words, the peak in cases of topicalization could be realized later than in broad focus contexts. This is just a hypothesis, as it could be the case that, since topicalized elements display shorter duration, the location of their peak will not vary significantly with respect to broad focus.

---

76 Hedberg and Sosa (2007) also propose that there is no specific pitch accent to mark old information in English based on their analysis of spontaneous speech.
5.7.2. On phrasing and phrasing patterns

Given the inexact nature of the relationship between phrasing and syntax, as discussed in Chapter 4, phrasing results can only be taken as one more feature of a given syntactic environment. Therefore our approach to phrasing is a descriptive one, in which phrasing patterns are only considered within the contexts of information structure, specifically the notions of old information and broad focus.

Overall, our results confirm that in Spanish there is a tendency to phrase the first constituent independently regardless of its informational properties and syntactic function. In both dialects, this tendency was attested in cases of subjects of transitive verbs, preverbal non-nominative subjects, CLLDs, and preverbal subjects of unaccusative and unergative verbs alike, in both broad focus and topicalized contexts. This tendency, however, is higher in constructions that correspond to topicalizations in most syntactic environments. These findings suggest that structural differences between these two types of constructions contribute to the differentiation of their phrasing patterns.

It was also observed that syllabic quantity has an effect on phrasing. This was attested in the context of unaccusative and unvergative verbs, which have long subjects. These long subjects are largely phrased independently in both dialects. However, in the case of Madrid the effect of syllabic quantity seems to be categorical. An interesting scenario to consider would be the one about the phrasing pattern that these constructions would display if both elements were short. My guess in this case is that both, the preverbal DP and the verb, would be phrased as a single intonational phrase. If this result were confirmed, it would entail that phrasing patterns are equally, or even more, susceptible to syllabic quantity, prosodic weight (Frota et al. 2007), etc., than to actual syntactic constituents/structure. Additionally, based on the differences between these two dialects regarding phrasing patterns in some environments, what can be said about the mapping between syntactic structure and phrasing patterns?

5.7.3. On the syntax-phonology interface

In this section I focus on the examination of the following two topics from the perspective of the syntax-phonology interface in the light of our experimental findings: 1)
the elements that should be included in a theory of information structure at the syntax-phonology interface, 2) the type of syntactic information that is prosodically encoded.

With respect to the first topic, the answer is to some extent straightforward. As the results from our experiment confirm, old information elements that appear in the left periphery show prosodic differences in duration when compared to broad focus elements in preverbal position in Spanish. Some topicalized elements also show differences in their fundamental frequency with respect to broad focus elements. These findings suggest that topics or a [topic] feature should be comprised in a theory of information structure at the interface of syntax and phonology. Since [topic] is a syntactic feature, elements obtain this specification in the syntax. Because this type of syntactic information seems to be relevant to the prosodic realization of such elements, at least in languages like Spanish, this feature should be visible at the syntax-phonology interface. At that point, the feature is expressed by the relevant prosodic cues of a given language. In Spanish, for instance, this is done by decreasing the duration of the topicalized element, at least when it appears in prefocal position, and in some cases by F0 changes as well.

Further support for the need to include a topic feature at the interface comes from the potential behaviour of old information contrastive topics in the left periphery, whose prosodic realization may resemble that of contrastively focalized elements. If this were the case, these topics may have an impact on sentential stress. However, unlike focalization, topicalization has not been included in the computation of sentential stress. As a side note, we can add that the behaviour of the different types of focused and topicalized elements supports analyses such as López (2009), who argues that the notions of topic and focus may not be appropriate for a syntactic and prosodic characterization of these elements. Instead, he proposes that a description in terms of features such as [+/- contrastive], [+/- old] is more adequate in this case.

The issue about the type of syntactic information that is prosodically encoded is more complex. Aside from the notions of topic and focus (or however we decide to define them), which proved to be necessary not only theoretically but also based on our experimental results, and the results obtained from numerous studies on focus constructions in Spanish, one can wonder about the relevance of syntactic information
like verb types. In English, for instance, there has been a long-standing claim about subjects of unaccusative and unergative verbs being prosodically distinct. On the other hand, in Spanish, there is no prosodic distinction between these two types of subjects, as Zubizarreta (1998) and others have proposed, and as our results confirm. The issue discussed here can be taken as a mere technicality. However, I believe that it can shed some light on the direction in which things operate. This is important because it prevents us from seeking or expecting a parallel prosodic behaviour in other types of syntactic environments that could be similar.

My hypothesis is that it is not the case that differences in verb types are relevant, in the sense that the distinct prosodic marking of the two structures contributes to a different interpretation of their subjects, but the way that the stress assigning mechanism operates has an effect on the prosodic realization of these two types of subjects. This is what I believe that Kahnemuyipour’s (2009) proposal of sentential stress assignment captures. By taking an approach like this, one can reanalyze the prosodic differences between subjects of unergative and unacusative verbs as a by-product of the stress-assigning mechanism.

Note that the prosodic prominence of unaccusative subjects in English is a problem because it cannot be explained based on differences in information structure, given that this prosodic behaviour is attested in broad focus constructions. As stated in Chapter 4, a relevant question about this differentiation is: why does the prosody signal this syntactic difference? Moreover, since we associate differences in intonation patterns to different meanings, what would the meaning of the distinction between sentences that include these two types of subjects? Assuming a hypothesis like the previous one, by which differences in stress patterns between unaccusative and unergative sentences arise as a result of the way in which the stress assignment mechanism operates, has the advantage of no meaning associated with the prosodic prominence of unaccusative subjects. Furthermore, if we analyze the prosodic prominence of unaccusative subjects as a result of the stress assigning mechanism along with the idiosyncratic use of the F0—especially in contexts in which no salient property of information structure of the given element is being signalled, i.e., cases of broad focus—the variability described in the prosodic realization of the unaccusative subjects in English simply follows. In opposition, a
context in which unaccusative subjects should never display variable behaviour in English is one in which they are marked as focus or contrastive topics. The hypothesis is that there should be no ambiguity in these cases.

A similar point can be made about linking phrasing patterns to syntactic structure in general, and specific phrasing patterns to particular syntactic structures. One of the problems in this case is that even if we were to find cases in which a one-to-one equivalence exists, such phrasing patterns may change once syllabic quantity is modified. Thus, theoretically, an important point to consider is the feasibility to propose a theory of phrasing linked to syntactic structure.

Take for instance the case of unaccusative and unergative verbs with long preverbal subjects in the data analyzed in this study, which were always phrased independently, along with the possibility of a different phrasing pattern when both the subject and the verb are short. In order to propose a theory that relates phrasing to syntax—specifically, syntactic constituents—it would be necessary to account for syllabic quantity. An ideal environment, depending on the specific licit prosodic combinations in each dialect/language, in which the predicted behaviour occurs, would have to be proposed first. Next, it should be noted that anything that deviates from this ideal environment is not expected to behave in the same way. This instance somehow resembles the case of secondary stress assignment in a sentence. Differently from primary stress, which can be derived from syntactic structure, secondary stress is subject to change based on the number of elements in the sentence, among other factors. This is the reason why some theories like Zubizarreta’s (1998) do not include a mechanism for secondary stress assignment. Interestingly, other models that include it do not mention that the results could vary, given that secondary stress does not only depend on the syntax, but is also subject to phonological constrains.

Ladd (2008) summarizes the main two problems about linking phrasing patterns to syntactic structure: “the most important—and most complex—conflict of criteria arises from the twin assumptions that (a) the division of sentences into ProsPs [prosodic phrases] in some way reflects syntactic, semantic or discourse constituency, but (b) prosodic structure is somehow fundamentally simpler than syntactic structure” (p. 288).
He further notes that the problem gets amplified when mapping theories such as Nespor and Vogel’s (1986) are taken as definitions and not as hypotheses. A clear example of how this type of approach fails is the conclusions that several studies about phrasing in Romance languages have proposed about the position that preverbal subjects occupy (see Chapter 4, section 4.5.3 for details).

Finally, our data also exemplifies how the idea of a one-to-one correspondence between syntactic structure and prosody fails. The specific instances in this case are structures with preverbal subjects of transitive verbs and structures with preverbal non-nominative subjects, in broad focus and topicalized contexts. Syntactically, these two types of subjects can be analyzed along the same lines, prosodically only preverbal subjects of transitive verbs display different phrasing patterns in the two conditions of information structure. A similar case is attested in the comparison of these two subjects and CLLDs. As noted, syntactically, these two types of subjects occupy the same position, which is different from the position that CLLDs occupy. Prosodically, only broad focus preverbal subjects of transitive verbs are distinct from topicalized CLLDs, i.e., broad focus non-nominative subjects and topicalized CLLDs behave alike with respect to phrasing. Crucially, there are no differences in word order between the constructions that include different types of non-nominative subjects and CLLDs, whereas, in the case of subjects of transitive verbs and CLLDs, there are differences in word order. Even though these sentences have different syntactic structures, phrasing differences only arise in cases in which the word order is different, i.e., strings that have the exact same elements cannot be teased apart by their phrasing patterns, despite the fact that their syntactic and information structures differ. Finally, the different behaviour that preverbal nominative and non-nominative subjects display with respect to phrasing in the two contexts of information structure—note that the strings are the same in both cases, and only nominative subjects show differences in the phrasing patterns in the two conditions—demonstrates that neither a one-to-one approach can be adopted, nor can we expect the same prosodic behaviour across instances whose syntactic structures may be equivalent. Thus, even though prosody can be informative in some cases, we should not expect to find a direct correspondence between syntactic structure and prosody.
Chapter 6
An alternative view on Spanish subjects

This chapter re-examines the syntactic position that preverbal and postverbal subjects occupy in Spanish. It reviews the results of the prosody experiment detailed in Chapter 5 in light of the existing syntactic analyses, and identifies and discusses some problems that arise from them. A new analysis that includes an event node (Event Phrase, EP), based on Borer (2005b), is proposed to account for the behaviour of Spanish subjects.

6.1. Data and current proposals

Based on the two competing theories about the position that preverbal subjects occupy in Spanish, the CP and the IP, the theory of the grammar, and the prosodic properties that focalized elements display with respect to broad focus elements in Spanish (see Chapter 4 for details), I carried out the experiment detailed in Chapter 5, which contrasted old information topics in preverbal position, such as CLLDs and topicalized subjects, with preverbal subjects in broad focus contexts. The intention was to obtain data that could help to disambiguate the syntactic status of preverbal subjects in Spanish.

The findings from this study show that the prosodic realization of preverbal phrases varies according to information structure. Specifically, old information preverbal phrases in topicalized constructions (CLLD datives and direct objects, and subjects) differ from subjects of transitive verbs in broad focus constructions. Old information elements display less duration than subjects in broad focus contexts. There are also differences with respect to the fundamental frequency between CLLD phrases, on the one hand, and preverbal subjects in broad focus contexts, on the other one. Preverbal subjects in broad focus contexts also differ from old information preverbal subjects with respect to fundamental frequency, which again points to a syntactic distinction between these two types of phrases and their structures. Differences in the fundamental frequency, however, are only relevant for some speakers of Havana. This behaviour is not unexpected given the idiosyncratic use of this parameter (see Chapter 5, sections 5.6 and 5.7, for details and discussion).
With respect to preverbal non-nominative subjects of psychological and other unaccusative predicates, which have been claimed to bear subject-like properties in Spanish (Masullo, 1992; Fernández-Soriano, 1999a, 1999b; Cuervo, 1999, 2003), our findings are similar to those of preverbal subjects of transitive verbs. Speakers systematically distinguish between elements that are old information, and as a result presumably move to the left periphery, such as CLLDs and topicalized non-nominative subjects, and elements that are new information, i.e., preverbal non-nominative subjects in broad focus contexts. Old information items display less duration than the same item in a broad focus context. However, differently from preverbal subjects of transitive verbs and CLLDs, there are no differences in the fundamental frequency between constructions with preverbal non-nominative subjects in broad focus contexts and CLLDs.

Finally, preverbal subjects of unaccusative and unergative verbs in broad focus contexts are not prosodically distinguishable. As both instances correspond to broad focus contexts, it seems logical that no prosodic difference should arise between them. The following scheme summarizes these findings. Underlined elements are old information.

```
SVO ≠ SVO  
SVO ≠ CLLDVO  
n-NomSVS ≠ n-NomSVS  
n-NomSVS ≠ CLLDVO  
S_{(unacc)} = S_{(unerg)}
```

---

77 Although there is movement to the preverbal field in the case of preverbal subjects of unaccusative constructions, this movement cannot be classified as a topic movement to the left periphery as old information topics do. Moreover, unaccusative structures with preverbal subjects are possible in broad focus contexts. Their interpretation, however, varies. Most proposals agree on the idea that unaccusative structures with preverbal subjects are interpreted about the subject, in opposition to a thetic statement in the case of postverbal subjects. Zubizarreta (1998).

78 See Chapter 5, section 5.7.3, for a hypothesis about the differences in the sentential stress patterns that these structures are claimed to display in English.
The distinct prosodic behaviour that topicalized and broad focus elements display suggests that these two types of elements do not occupy the same structural position. Although these findings do not entail that preverbal subjects in broad focus contexts do not appear in the left periphery, nor do they tell us the specific position preverbal subjects occupy, they are compatible with the syntactic analyses that propose that preverbal subjects in broad focus contexts can appear in the inflectional layer (Masullo, 1992; Goodall, 1999; Suñer, 2003; Ortega-Santos, 2008). Additionally, our results also support the proposals that argue that preverbal non-nominative subjects of psychological and other unaccusative predicates are not instances of left dislocations (Masullo, 1992; Fernández-Soriano, 1999a, 1999b; Cuervo, 1999, 2003; Ortega-Santos, 2008). Moreover, since the behaviour of preverbal subjects of unaccusative and unergative verbs is prosodically indistinguishable in the broad focus contexts analyzed in our study, an option may be that they also appear in the same syntactic position. Two issues arise when we consider these possibilities together. The first one is about the locus that these elements occupy: is it the same? (If this is so, where is it?) The second issue is about the trigger of the movement to that position.

6.1.1. Preverbal subjects' locus and their motivation for movement

With respect to the locus that broad focus preverbal subjects and subject-like elements occupy, the different IP analyses agree on the idea that they appear in the specifier of IP (Masullo, 1992; Goodall, 1999; Fernández-Soriano, 1999a, 1999b; Cuervo 1999, 2003; Suñer, 2003; Ortega-Santos, 2008). The reason for movement into this position is less clear, however. Some of these analyses do not deal with movement (Suñer, 2003). Others have suggested that the motivation for movement is an EPP requirement, likely analogous to the one found in English. Yet, differently from the English EPP, the element that fulfills this principle in Spanish does not have to be a nominative subject (Goodall, 1999; Fernández-Soriano, 1999b; Cuervo, 1999, 2003; Ortega-Santos, 2008). Locatives and

---

80 It is important to bear in mind that most of these analyses focus on specific constructions. For instance, Fernández-Soriano (1999a, 1999b) and Cuervo (2003, 2010) mainly work with dative experiencers and locatives. Goodall’s (1999) analysis, on the other hand, refers to nominative subjects, as does Suñer’s (2005). Masullo (1992) refers to nominative, dative and locative subjects.

81 There are some exceptions in English, but they are not relevant for the current discussion.
Datives seem to be possible in that position as well, as (1) and (2) illustrate. Additionally, some of these analyses propose that these non-subject elements move to the preverbal field because of a need for predication, which some of them extend to the context of nominative subjects of transitive and unergative verbs in order to account for their preverbal position as well.

(1) A Marcos le gusta la música coral.

Marcos\textsubscript{DAT} cl\textsubscript{DAT} like\textsubscript{3S} the music choral

‘Marcos likes choral music.’ (Masullo, 1992, p. 116, (1))

(2) En esta casa falta café.

in this house miss\textsubscript{3S} coffee

‘Coffee is missing in this house.’ (Fernández-Soriano, 1999a, p. 105, (4a))

(3) Me falta café.

cl\textsubscript{DAT} miss\textsubscript{3S} coffee

‘I’m missing coffee.’ (Fernández-Soriano, 1999a, p. 121, (32a))

Within this scenario, a question that still remains is what happens in cases such as (4)-(6)? If there is an EPP requirement in Spanish that calls for an element in the specifier of Inflection, how can we account for sentences in which no element appears in such position, as is the case with post-verbal subjects of unaccusatives (4), meteorological verbs (5), and null subjects (6)?

(4) Llegó Juan.

arrived\textsubscript{3S} Juan

82 These researchers have proposed that datives and locatives can fulfill the EPP in Spanish, based on the preverbal position that such elements occupy and the similar syntactic behaviour that they display with respect to nominative subjects. See Chapter 3 for details.
‘Juan arrived.’

(5) Llueve.

Rain$_{3S}$

‘It rains.’

(6) Comió pescado ayer.

ate$_{3S}$ fish yesterday

‘He/ she ate fish yesterday.’

To solve some of these issues, Ortega-Santos (2008) reformulates the EPP. He proposes that the EPP only applies in broad focus contexts in Spanish. He further argues, in line with the minimalist framework, that overt movement, i.e., internal merge, correlates with expressiveness in the language, which, according to his analysis, is supported by the fact that in Spanish sentences are interpreted with respect to the element that moves into that position.\(^{83}\) This idea is to some extent equivalent to the aboutness property that Masullo (1992) links to preverbal datives of psych predicates and some locatives (see Chapter 3).

Regarding the lack of an element in the preverbal field in unaccusative structures with postverbal subjects, Ortega-Santos proposes that they constitute instances of locative inversion. In the analysis that he adopts, some unaccusative verbs include a locative in their lexical entry. This locative occupies the specifier of Inflection, and is the subject of predication. As a result, sentences with these unaccusative verbs are interpreted with respect to the abstract locative, and a presentational/existential interpretation arises. When there is an overt locative, he suggests that it occupies the specifier of Inflection. He further claims that such instances are compatible with a broad focus interpretation.\(^{84}\)

---

\(^{83}\) He proposes the same for English.

\(^{84}\) It is not clear to what extent the possibility of interpreting this type of sentences as broad focus statements is determined by an interplay of different factors like frequency and/or type of locative expression, type of complement, etc. Deictic expressions such as aquí ‘here’, allí ‘there’ seem to work better than PPs. In addition, the complement of the verb also plays a role. This can be seen in the following
this case, locative movement is possible because no minimality constrains are violated (since the argument of the verb is generated in the complement position). The same, he concludes, cannot be proposed for transitive and unergative verbs, which do not include a locative in their lexical entry, and whose nominative subject is generated as an external argument. He further notes that the combination of these two properties makes it impossible for transitive and unergative verbs to be compatible with locative inversion structures. The nominative subject generated as an external argument is closer to T, and, therefore, it appears preverbally in broad focus structures. As a result, sentences with transitive and unergative verbs are interpreted with respect to their preverbal subject, according to his reformulation of the EPP.

Ortega-Santos (2008) extends the unaccusative analysis to psych predicates as instances of unaccusative structures. Although no derivation is provided for this type of sentence, in this case too, he proposes that the dative experiencer moves to the specifier of Inflection due to the EPP, and does not violate any minimality constrains. This entails that the dative should be generated higher than the nominative argument.

Thus far this proposal offers the most complete instantiation of the EPP in Spanish. It explores the possibility that the same principle that applies in some syntactic environments, such as transitive and unergative structures, can account for cases that

What happens/happened?

(a)   i. *A casa vino/ llegó Juan.
       home came/ arrived Juan

       ii. ?Aquí vino/ llegó Juan.
          here came/ arrived Juan

       iii. ?Aquí llegaron guaguas/insectos
            here came/ arrived buses/ insects

(b)   i. A casa/ Aquí llegan guaguas/insectos.
       Home/ here arrive buses/ insects

       ii. ??A casa/ aquí viene/ llega Juan.
          Home/ here come/ arrive Juan
have been previously considered exceptions, like unaccusative constructions with postverbal subjects, and psych predicates with preverbal dative experiencers. It also incorporates the general intuition that various analyses have suggested that movement to the preverbal field in Spanish occurs because of a need for predication, and as a result sentences are interpreted with respect to or about the element that occupies the preverbal position. However, some theoretical and empirical issues still call for our attention.

6.1.2. Against a predication requirement as movement trigger and a semantic reinterpretation of the EPP in Spanish

Theoretically, it is not clear how this predication requirement develops, nor is the exact nature of the notion of predication itself. If we take predication to be the relation that develops from the element that saturates a given predicate and the predicate itself, then, by entailment, there is a predication relation between the predicate and the element that saturates it. Once the element is merged, the predication relation is established. Thus, predication emerges as a result of a predicate being saturated. Although it can be argued that predicates need to be saturated, predication itself is not the goal, but a by-product. Moreover, the fact that in some instances a given predicative relation has no interpretative effects, e.g., the case of expletives, overt and null, suggests that predication, in the sense that these analyses claim, cannot be the trigger of the movement. Hence, the more general idea that the need to saturate a predicate is the trigger of the movement seems more likely. Finally, the extent to which these analyses propose that the existence of such predication relation, in cases in which it correlates with semantic effects, determines the interpretation of the sentence is not accurate either.

Let us assume that these analyses mean by predication something along the lines that sentences need an element in a specific position, let us say SpecIP, in order to be interpretable at the discourse level. Then given their general claim that sentences with preverbal nominative and non-nominative subjects are interpreted about such elements in

---

85 The reader should recall that this has been proposed specially for cases of non-nominative elements (Masullo, 1992; Fernández-Soriano, 1999a, 1999b; Cuervo, 1999, 2003).
86 Bošković (2007) raises the same point. He notes that deriving the EPP from a deep semantic requirement like predication is problematic when we consider the fact that the EPP can be satisfied by semantically null elements.
Spanish, they limit the interpretation of broad focus statements to being about the element that appears in specIP. This turns out to be too restrictive. A reformulation of the EPP in terms of predication is problematic for the same reason. Additionally, a semantic reformulation of the EPP would seem to entail that this type of predication should be found across-the-board. As in the previous case, this is not necessarily so. The analyses that propose that elements move because of the need for predication face a similar problem. Since there are contexts in which there is no predication, for instance, existential constructions and some contexts that include postverbal subjects, it is hard to determine the cases in which sentences are subject to this predication requirement, and the cases in which they are not.

In order to unfold these issues, I start by looking at the interpretation of broad focus sentences that relates to the event or state, along with examples in which there is no predication about any element at the sentence level. Next, I examine the evidence upon which the claim that predication is a sentence requirement is based.

6.1.2.1. When broad focus statements are not interpreted about the preverbal (overt or covert) element

It is a fact that any broad focus statement can potentially be interpreted about the event or state that it denotes. This is the simplest meaning that any broad focus sentence can convey; therefore one would expect it to be available. Furthermore, such meaning might plausibly arise from the structure. For instance, sentence (7), used in a context like the one that follows, is not interpreted about Luisa, the subject, but about the whole event that it denotes.

Context: Juan was supposed to go grocery shopping on Thursday, but Luisa, his roommate, did so the day before. When Javier, a common friend who knew about his plans, arrives on Thursday and sees Juan at home asks him: what are you doing here? What happened?

(7) Luisa hizo la compra ayer. (Entonces no tuve que ir hoy.)

Luisa did the shopping yesterday (‘So I didn’t have to go today.’)

‘Luisa did the shopping yesterday.’
The same point can be made about any broad focus statement independently of the verbal configuration that it includes, as the following examples show.

Context: María invited Juan to watch a soccer game on Friday night. On Saturday, María’s sister arrives and sees that the TV set is broken. She asks María, what happened?

(8) A Juan no le gustó el partido (‘y rompió la tele’).\(^{87}\)

\[\text{Juan}_{\text{DAT}} \text{ not } \text{cl}_{\text{DAT}} \text{ liked the game (‘and broke the TV set’)}\]

‘Juan didn’t like the game.’

Context: Luis and Ana share a small office. On Thursday, Ana left earlier. When Ana arrives on Friday she notices that there are new chairs, and asks Luis if he knows anything about it.

(9) a. El jefe vino ayer y las trajo.

\[\text{the boss } \text{came}_{1\text{S}} \text{ yesterday and } \text{cl}_{\text{ACC}} \text{ brought}_{1\text{S}}\]

‘The boss came yesterday and brought them’

Other possible answers in this case are (b) and (c).

b. Ayer vino el jefe y las trajo.

\[\text{yesterday } \text{came}_{1\text{S}} \text{ the boss and } \text{cl}_{\text{ACC}} \text{ brought}_{1\text{S}}\]

c. Vino el jefe ayer y las trajo.

\[\text{came}_{1\text{S}} \text{ the boss yesterday and } \text{cl}_{\text{ACC}} \text{ brought}_{1\text{S}}\]

If other principles of the grammar, such as this semantic reformulation of the EPP, prevent the previous interpretations from emerging, then such meanings should arise in some other way. One such way, as M. C. Cuervo (p.c.) suggests, could be some type of

\[^{87}\text{Sentence and context based on an example from M. C. Cuervo (p.c.).}\]
reanalysis, similar to the global reinterpretation that occurs with idioms, where *X kicked the bucket* is not interpreted as a transitive structure, based on the meanings of the verb *kick* and its object *bucket*, but as the intransitive *die*. While this could be an option, it seems like a last resort, as it would involve reanalyzing the structure. Moreover, since these are simple meanings associated with broad focus statements, the simplest configuration in terms of information structure, the need for a reanalysis seems counterintuitive.

In addition to the potential meaning about the event/state that any broad focus statement with a preverbal subject carries in Spanish, which entails that such sentences are not interpreted exclusively with respect to their subject despite the fact that there is a predication, the claim that predication is a sentence requirement is difficult to justify in the light of the evidence from existential constructions, for there is no predication in these cases, as in (10). The same can be said about other types of impersonal constructions, (11-13) in Spanish.

(10) Hay muchos animales.

There is many animals

‘There are many animals.’

(11) Se come bien en Francia.

SE eat_{3S} well in France

‘One eats well in France.’

(12) Se venden/ vende puertas.

SE sell_{3S}/ sell_{3P} doors

‘Doors are sold.’

(13) Tocan a la puerta.

knock_{3P} at the door
‘Someone is knocking at the door.’

Further evidence that predication is not a general sentence requirement comes from constructions that include unergative verbs with postverbal subjects and a preverbal locative such as (14) and (15). It is important to note that all these sentences are valid in broad focus contexts, which makes them the appropriate candidates for Ortega-Santos’ (2008) EPP reformulation to apply.

(14) Aquí duermen animales.

here sleep<sub>3p</sub> animals

‘Animals sleep here.’ (Torrego, 1989)

(15) En este parque juegan niños.

in this park play<sub>3p</sub> children

‘Children play in this park.’ (Example based on Suñer (1982))

(14) and (15) are not necessarily interpreted with respect to the preverbal locative, which supports the idea that predication, in the sense that these analyses use it, is not a necessary condition for the well-formedness of sentences in Spanish or their interpretation at the discourse level. As these examples show, unergative verbs can participate in what seems to be a similar configuration to that of unaccusative verbs. They are possible with postverbal nominative subjects. Moreover, these subjects can be bare plurals, which is impossible when they appear in the preverbal field (Suñer, 1982; Torrego, 1989; Zubizarreta, 1998; Casielles Suárez, 2004). Unlike unaccusative structures, however, unergative verbs with postverbal bare subjects are claimed to only be possible when there is an overt locative in preverbal position (Torrego, 1989).<sup>88</sup> The interpretation of the previous examples corresponds to presentational/ existential sentences. This is not a

---

<sup>88</sup> Torrego (1989) proposes that (14) constitutes an unaccusative structure, even though it includes a verb that is normally used as unergative. See footnote 99 in this chapter for Borer’s (2005b) arguments against Torrego’s analysis.
coincidence, as there is ample cross-linguistic evidence that locatives are linked to existential interpretations, even if the reason for it is still not clear (Borer, 2005b).

While analyzing these examples as exceptions to the predication requirement at the sentence level is an option, another approach would be to re-examine the properties of these structures, along with those that include preverbal subjects, in order to see if there is any common element that connects them. This is the path pursued in the next sections. The reason for pursuing it is that, despite their differences, all these structures correspond to broad focus contexts; thus there is a possibility that the same principle underlies all of them. Taking into account the interpretational and syntactic differences among them, predication seems not to be such a principle.

6.1.2.2. Analysis of the predication requirement

Generally speaking, the predication principle that these theories propose develops from an analysis that conflates movement and interpretation in some contexts, and extends the same analysis to other contexts in which it may not be relevant, or at least in which there is no enough evidence to propose it. It should be noted that even in the contexts in which there seems to be a link between movement and interpretation, this might be a correlation, while the reason for movement may be a different one than predication. In other words, while it is the case that there is predication in the cases that these analyses propose, suggesting that predication is the trigger of the movement may not be accurate, especially, if we take into account all previous syntactic and interpretational evidence.

Part of the problem comes from the nature of Spanish data. Spanish offers a far more complex picture with respect to subjects and subject realization. There are cases like unaccusative structures, where the possibility of postverbal (16) and preverbal (17) subject realization exists. These two options have been claimed to correlate with two different interpretations: thetic and about the subject, respectively. In this case, there is a difference in interpretation that arises from a contrast between two possible positions of
subject realization. Yet, it is worth noting that the two options are potentially valid in broad focus contexts.89, 90

(16) Vino Juan.
     came3S Juan
     ‘Juan came.’

(17) Juan vino.
     Juan came3S
     ‘Juan came.’

Psych predicates of the piacere class, on the other hand, display a different word order than regular transitive structures. In broad focus contexts, the nominative subject occupies a postverbal position and the dative experiencer appears in the preverbal field. Psych predicates are relevant because they represent a disassociation between movement to the preverbal field and the grammatical subject. It has been proposed that these constructions with a preverbal dative experiencer are interpreted with respect to the dative, which is considered the logical subject of the sentence (Masullo, 1992; Fernández-Soriano, 1999a; Cuervo, 1999, 2003; Ortega-Santos, 2008).

(18) A Juan no le gusta la sopa
     JuanDAT not clDAT like3S the soup
     ‘Juan does not like soup.’

89 It is important to note that analyses that suggest the aboutness interpretation of the sentence with respect to the preverbal element cannot rule out the possibility of an unaccusative structure with a preverbal subject in a broad focus context.
90 Preverbal subject realization with unaccusative verbs is very common when there is some modification. This is attested in sentences that include temporal adverbs (or expressions) like temprano ‘early’, tarde ‘late’ or a tiempo ‘on time’. The word order in broad focus contexts in this case is SVAdv. The same occurs when they include adjectives, as (a) shows.
(a) Juan llegó cansado.
     Juan came3S tired
Finally, in constructions with transitive and unergative verbs, subjects appear in preverbal position in broad focus contexts.

(19) Los niños ven la película.

the children watch3p the movie

‘The children are watching the movie.’

(20) Mi gata duerme mucho.

my cat sleeps3s a lot

‘My cat sleeps a lot.’

We can conclude that while it could be claimed that in unaccusative structures predication is linked to movement, the same conclusion does not necessarily hold in other verbal configurations. Moreover, even if we agree on the fact that unaccusative constructions with preverbal subjects are interpreted about their subjects, this may be only so because two options for subject realization are available. It is based on the contrast in the interpretation of these two options that a link between movement and interpretation seems likely, and, as a result, so does its extrapolation to other contexts.

Such a link is less clear once we add transitive constructions. Although transitive structures are not necessarily interpreted about their subject, as already shown, typical subject realization is preverbal. To some extent, the same can be said about psych predicates of the piacere class, which share properties of both transitive and unaccusative constructions. Similarly to transitive constructions, the word order of psych predicates seems restricted to one possibility in broad focus contexts. Differently from transitive constructions, this word order corresponds to a postverbal nominative subject and a preverbal dative experiencer. In these sentences there is predication with respect to the dative experincer and not the nominative subject.

While a possibility may be that a principle like this semantically triggered EPP is applying in the previous cases, and meanings related to the event/state emerge from a type of reanalysis, another possibility is that something else may be at play, and that the
correlation between the different syntactic behaviours in some of the previous structures, and different interpretations are masking it. In other words, even if some of these word orders happen to correlate with an interpretation linked to the element that occupies a specific position, it does not entail that there is a predication requirement that makes such elements move into that position.

Cardinaletti (2004) raises similar points with respect to subject movement and the EPP in Italian. She concludes, based on a similar type of interpretational evidence and theoretical problems as the ones discussed here, that predication should not be equated with the EPP. In order to maintain the EPP as a universal principle, she proposes two functional heads within IP: one that checks EPP, and another one related to predication. Broad focus statements with preverbal subjects would check EPP, but not predication. When the sentence is interpreted with respect to its subject, it checks both EPP and predication. Although Cardinaletti’s solution provides the accurate interpretational results, theoretically, the idea of maintaining the EPP as a universal principle when there is no enough evidence for it (Bošković (2007), or when its effects may be the result of a different syntactic requirement (Borer, 2005b) seems unnecessary at this point. Regarding predication, as we previously saw, it is a product of the structure, a relation between two elements that arises from a structural configuration; therefore there is no need for a functional head that checks it.

Returning to Spanish, I will argue that the structure, which imposes its own constraints, determines the different word orders, and that predication, in the sense that these analyses claim, is not a necessary condition for the well-formedness of sentences or their interpretation at the discourse level for that matter. I consider predication to be a by-product of a structural configuration and not the reason why certain elements move to a given position. Such configuration should include a c-command relationship between a given element and the predicate, and also adjacency between the two. Another relevant feature is that the element that c-commands the predicate has to be an argument of the predicate. Thus the predication property that these analyses suggest boils down to the

---

91 Similar conditions are proposed by Williams (1980), who labels them as c-command restriction and c-subjacency.
relation that emerges from an element that saturates a predicate and the predicate itself, when such a relation has interpretative effects due to the properties of the element that saturates the predicate.

This type of approach fares better when applied to the different structural configurations. Recall that the predication principle/requirement fails in the case of unaccusative structures with postverbal subjects, since there is no predication in such cases. Whether or not there is an abstract locative that occupies the specifier of Inflection (fulfilling the need for an element there), strictly speaking the presence of such an element does not have interpretative effects, i.e., there is no predication. Therefore, this type of construction would be an exception to the claim that there is a need at the discourse level for the sentence to be about something. The same logic can be applied to weather constructions and impersonal constructions. Even though it can be argued that there is a covert element that occupies the specifier of Inflection or Tense in these cases, predication does not arise.

Based on these facts, one should question if indeed this type of predication is a need at a sentence level or rather an effect that is visible in a great number of contexts. I believe the latter to be the case, and that the existence of existential sentences, along with the interpretation about the whole event that any sentence with a preverbal subject potentially carries in Spanish, can be taken as evidence that this type of predication is not a necessary condition of the sentence. This type of predication would arise in cases in which both the structural configuration and the element that occupies a given position make it possible. On the other hand, movement into a preverbal position in Spanish, which makes this c-command relationship available, may be the result of a different principle. This approach captures the fact that this type of predication is not found in all types of sentences.

Another welcome consequence is that while there could always be predication, assuming the right structural configuration and the presence of the appropriate element, predication itself is not the reason behind movement. This has the advantage of not liking the interpretation of broad focus sentences to a specific element. As a result, sentences that include a preverbal subject or any other preverbal subject-like element in broad focus contexts can be interpreted about the event/state, at the same time that there is
predication. Since the two meanings are derivable from the structure, the two meanings are available in every instance that a broad focus utterance with a preverbal subject is used.

To summarize, we can say that the different word orders associated with different verbal configurations, along with interpretational differences related with the different positions of some of their elements, and the lack of a constant, visible cue, such as an overt element in the specifier of IP in English, led to the identification of predication as a possible reason for movement to the preverbal field in Spanish. However, if we take into account the potential interpretation that any broad focus statement with a preverbal subject carries in Spanish, along with the existence of existential sentences, and other contexts in which this requirement does not hold, this type of predication seems to be neither a requirement for the well-formedness of the sentence, nor a requisite for their interpretation at the discourse level. Thus the idea that this predication principle develops from an analysis that conflates movement and interpretation in some contexts, and extends it to others seems a plausible one. As a result, it is possible to conclude that the correlation between movement and interpretation in unaccusative structures, along with the particular word order that psych predicates display are not enough to propose that there is a predication requirement at the sentence level that accounts for movement to the preverbal field in Spanish.

Assuming that the verb raises to T in Spanish, along with the preverbal position that these subjects and subject-like elements occupy, the next question that we have to answer is about the trigger of their movement to a position higher than the verb’s. The other side of the coin is the locus that they occupy. I will argue that the structure, and the constraints that it imposes can account for the movement of these elements. Specifically, building on Borer (2005b), I will propose that these preverbal elements move to the specifier of the Event Phrase (EP), which is located above IP, and that the trigger of the movement to this position is the need to license the event variable in EP (Borer, 2005b). As will be seen, the different word orders attested in broad focus contexts in Spanish can be accommodated under this analysis.
6.2. Borer’s (2005b) Event Phrase (EP)\(^{92}\)

In Borer’s proposal, the Event Phrase (EP) is an independent functional category that projects above TP (or without TP altogether). In order to justify the position of the EP node in the structure, Borer argues that an event interpretation can emerge in the absence of tense. (21) shows that it is possible to obtain such an interpretation despite the lack of tense. She further notes that the event interpretation does not depend on the nominal itself, but on whether or not the arguments (including the event argument) are retained. (22), which also includes a nominal that denotes an event, shows that an event interpretation does not arise in the case of result nominals. She concludes that the EP should not be associated with tense, as previous analyses did; thus TP cannot function as the existential binder of the event argument.

(21) the examination of the students by the faculty \(\text{(Borer, 2005b, p. 263, (20))}\)

(22) The examinations were thorough. \(\text{(Borer, 2005b, p. 263, (22))}\)

The idea of including the notion of event as part of the structure can be traced back to Davidson (1967), and his seminal proposal about the need for incorporating it in order to derive the logical form of action sentences. Davidson proposed that “verbs of action – verbs that say what someone did– should be construed as containing a place, for singular terms or variables, that they do not appear to” (p. 118). He argued that a sentence like *Shem kicked Shaun*, which was considered to be “two names and a two place predicate” (p. 118), should have the following logical form: \(\exists x (\text{Kicked(Shem, Shaun, } x))\), which reads “there is an event x, such as that x is the kicking of Shaun by Shem” (p. 118).\(^{93}\) In such representation the event variable is included as an extra argument.

\(^{92}\) This section summarizes Borer (2005b) proposal. There will be gaps in this summary, as she devotes a complete book to this matter. The reader is encouraged to review her work for a detailed account and the syntactic evidence that she examines in order to postulate the event node as part of the structure, and its specific localization.

\(^{93}\) The ontological status of events is not unanimously accepted, and even for those who recognize events as a distinct category, there is still disagreement about their exact nature. For instance, defining events as something that happens or takes place is problematic, if we want to include in the notion of event sentences that refer to states, for states are usually considered something that simply exists. In linguistics, the term has been variously used too. One thing is clear in this field though, events are considered “grammatically or linguistically represented objects, not (...) events in the world” (Tenny & Pustejovsky, 2000, p. 3).
Borer’s representation of the EP is closer to the original idea of Davidson. She points out that if the event argument is indeed an argument “it must emerge in the context of some functional structure” (2005b, p. 263). In her proposal, sentences have an extra argument, which is responsible for licensing the event. Specifically, the event argument is linked to a predicate that includes the verb, but is independent of the verb itself.

Borer makes a parallel between the structure of the DP (Borer, 2005a) and the EP, and proposes that an element should bind the variable linked to the head of each of the phrases. She suggests that like “<e>_d, the head of DP, where range assignment amounted to establishing a mapping from predicates to reference, so range assignment to the head of EP, <e>_E, is responsible for establishing a mapping from predicates to events” (p. 82). In both cases, range assignment to the variable “consists of binding it, in effect rendering it a logical variable” (p. 138).

In her analysis, range assignment to <e>_d e can be done by assigning a referential index to it, which can be done by merging an element in the specifier of DP. She notes that definite articles and demonstratives have a discourse antecedent from which they inherit such an index. By merging these elements in the specifier of DP, they bind the variable, and transmit the referential index to it. As a result they make it a logical variable. Similarly, in the case of EP, an element should appear in [Spec, EP]. Such element existentially binds the event variable, “provided that the element itself is existentially closed” (Borer, 2005b, p. 274).

Existentially closed elements are variables that refer to an object or set. This is possible because such variables are bound, i.e., they have received range. We already saw that <e>_d can be bound by a definite determiner or a demonstrative within the DP. When there is no determiner, i.e., when <e>_d is null, as in the case of bare nouns, for instance, Borer notes that <e>_d can be bound, from outside the DP, by the operation of Existential Closure (EC) (Heim, 1982), or by a generic operator. She assumes, in line with Diesing’s (1992) Mapping Hypothesis, that material within the c-command domain of the verb maps into the nuclear scope of the existential quantifier, and gets bound by it, i.e., it receives existential closure. When outside of the scope of EC, a generic operator may bind the variable. These two binding possibilities give rise to two different interpretations.
of bare nouns: existential and generic. Taking into account the type of elements that are licensed in the structure, she proposes subjects and expletives among possible range assigners to $\langle e \rangle_E$.

Borer suggests that EPP effects emerge from the need to assign range to $\langle e \rangle_E$. Taking this idea further, she notes that if indeed EPP effects emerge due to the need to assign range to $\langle e \rangle_E$, then other elements (aside from subjects and expletives) that are existentially closed may assign range to $\langle e \rangle_E$ as well. Her prediction is borne out by the fact that sentences with preverbal locatives and postverbal subjects exist in several languages. Based on examples from Catalan (23) and Spanish (24) that include a locative in the preverbal position and a postverbal subject, Borer concludes that locatives can assign range to $\langle e \rangle_E$. The same is attested in some contexts in English, as (25) and (26) illustrate.

(23) Hi canten nens.
   here sing$_{3P}$ children
   ‘Children sing here.’ (Example based on Borer (2005b), p. 276, (47a))

(24) Aquí duermen animales.
   here sleep$_{3P}$ animals
   ‘Animals sleep here.’ (Borer, 2005b, p. 276, (46a))

(25) Opposite the landing-place stood half-a-dozen donkeys with saddles on their back, etc.
   (Borer, 2005b, p. 287, (67a))

(26) On the third floor worked two young women called Maryanne Thomson and Ava Brent…
   (Borer, 2005b, p. 287, (67b))

---

94 In most analyses, the two operations take place at different levels: EC at the VP level, and binding by a generic operator at the IP level, based on Diesing (1992). However, Borer (2005b) argues that EC “while defined with respect to the c-command domain of the verb, cannot be associated with the verb as such, but rather must be associated with the functional head which we labelled E” (295). As will be seen, this accounts for the domain of EC that she proposes in her structures, which includes TP (IP) and not only VP.
The following structures, (27) and (28), depict these two possibilities of range assignment. She proposes that they are valid in Hebrew, English, Catalan, and Spanish, and suggests that they could probably be extended to all languages. In these cases, range assignment occurs indirectly by the specifier that assigns range to its head. However, range assignment to \(<e>_E\) can also be done directly, as we shall see.

(27)

(Borer, 2005b, p. 297, (80a))

(28)

(Borer, 2005b, p. 297, (80b))

(27) represents sentences that include any standard preverbal subject with a determiner, \(DP^{str}\) in her notation, such as (29) and (30), or with an expletive subject like (31).

(29) The students finished the exam.

(30) Los estudiantes terminaron el examen.

‘The students finished the exam.’
(31) It rains in the summer.

On the other hand, (28) represents sentences that have a locative in the preverbal field, and a postverbal subject, such as (23-26). Borer (2005b) modifies the structure in (28) for languages like Hebrew and Catalan in which the locative is a clitic of the verb. See (36) in this chapter.) She proposes that XP is a constituent with existential force. In both instances the existentially bound element, either a subject, expletive or locative, which occupies the specifier of EP, indirectly assigns range to the event variable, its head, and licenses the event, as the arrows indicate.

In order to complete the inventory of possible range assigners, Borer explores sentences that include bare subjects (which she labels as weak). She notes that most languages do not allow preverbal bare subjects; Spanish is one such language. In English, however, they are possible, and, in cases like (32-34), they can receive an existential interpretation. The first question that arises in this case is about the licensing of such subjects. Moreover, if \(<e>\) needs to be assigned range, what is possibly doing so in these cases?

(32) Farmers were sleeping/dancing.

(33) Butter was melting.

(34) Linguistic books are available.

(Borer, 2005b, p. 289, (70a-c))

Borer reformulates Dobrovie-Sorin and Laca’s (1996) (as in Borer (2005b)) proposal for these cases, and suggests that they are possible because they include a covert locative that binds not the weak preverbal subject—as in the Dobrovie-Sorin and Laca analysis—but the event variable, which in turn assigns range to the weak preverbal subject. As a result,

\[95\] Such interpretation is a problem because given that these are preverbal subjects, presumably in the IP layer, they should receive generic interpretation. Diesing’s (1992) analysis can only derive their existential interpretation by proposing some lowering to the VP domain at LF so these subjects can be bound by the operation of existential closure.
the existential interpretation emerges. This seems not to be the case in languages like Hebrew and Spanish, as this type of subject is not possible. The existence of a covert locative in English, but not in Spanish, and the possibility of indirect range assignment by a locative to the head \(<e^3>_E\), which licenses the weak preverbal subject, Borer concludes, can be taken as part of language variation.

Based on these differences she proposes two possible ways of range assignment to \(<e>_E\). The first one, (35), is for English, which allows weak preverbal subjects, while (36) depicts the structure for Hebrew and Catalan, where this type of subject is always postverbal and the verb appears in preverbal position linked to a locative clitic. Note that in (35) the locative, either covert or overt, is part of TP and indirectly assigns range to the event variable, which assigns range to the preverbal weak subject, as the arrows indicate in both cases. In (36), on the other hand, range assignment takes place from the locative clitic to the event variable, as the arrow signals.

(35)

(Borer, 2005b, p. 297, (81))

(36)

(Borer, 2005b, p. 297, (82))
Examples (37) and (38), in Catalan and Hebrew, respectively, correspond to the structure in (36).

(37) Hi canten molts nens.

\[ c_{\text{LOC}} \text{ sing}_{\text{3PL}} \text{ many children} \]

‘Many children sing here.’

(Based on Borer (2005b, p. 276, (47a))

(38) ‘\(\text{abad-}\)\(\text{šam} \text{ ganan} \)

\[ \text{worked-there gardener}^{96} \]

(Based on Borer (2005b, p. 284, (65a))

Since Spanish does not have locative clitics, we can conclude that the only possibility to license existential bare subjects in transitive and unergative structures is in the presence of an overt locative, as depicted in (28). In such cases the subject is postverbal, and its interpretation is also existential.

However, sentences with postverbal bare subjects without an overt locative are possible in Spanish as well as in Italian and Hebrew with a subset of unaccusative verbs, as sentences (39-41) show. Examples (40) and (41), in Italian and Hebrew are taken from Borer (2005b, p. 307, (4) and (5a), respectively). All these cases have an existential interpretation as well.

(39) Llegaron estudiantes.

\[ \text{arrived}_{\text{3P}} \text{ students} \]

‘Students arrived.’

(40) Sono arrivati studenti.

---

96 This example is reproduced as in the original text. Since Borer (2005b) does not include a translation into English, no translation is provided in this work either.
are$_{3P}$ arrived students

‘Students arrived.’

(41) parca mehuma (ha.boqer).

erupted$_{FS}$ riot$_{FS}$ (this.morning)

‘Riots erupted this morning.’

Borer argues that the previous cases are licit because they contain a locative that is part of the verb, and licenses the postverbal subject. In her analysis these types of verbs are instances of idioms. (42), adapted from Borer (2005b, p. 339, (64)), some details omitted, illustrates the structure of sentences that include this type of unaccusative verb.

(42)

Note that the verb, which includes the covert locative, moves to the specifier of EP and assigns range to the event variable, $<e>_E$, which licenses the event. The subject appears in the specifier of TP where it checks nominative case. It is also the verb (with the incorporated locative) that assigns range to the variable in D, $<e>_d$, within Asp$_Q$.\(^97\) This

\(^97\) Asp$_Q$, where Q stands for quantity, “is the syntactic projection responsible for telic interpretation.” (Borer, 2005b, p. 72).
gives rise to the existential interpretation of this type of structure. The reader should bear
in mind that locatives have existential force, and there is cross-linguistic evidence that
shows that they are linked to existential interpretations. It is also possible to have a strong
subject in these constructions. In such cases, the variable in D is not bound by the
locative that is part of the verb, but from within the DP, through the presence of a lexical
functional element such as a determiner.

6.3. EP and Spanish data

Assuming that there is an event variable that needs to be bound in order to do the
mapping from predicates to events, and different types of elements can bind such
variable, I will propose that the different word orders in Spanish in broad focus contexts
are the result of the alternative elements that can satisfy this variable, and the different
argumental frames in which verbs participate. The three main word order patterns found
in Spanish, which were introduced in the first chapter, and have occupied most of the
discussion of this work, can be accommodated under this proposal, as can transitive and
unergative structures with postverbal subjects.

Moreover, if EPP effects derive from the need to license the event variable, as Borer
(2005b) suggests, and different languages have different types of elements in their
inventories, including functional elements such as expletives, covert locatives, existential
pro, etc., which can bind this variable, then differences in the existence of this
requirement are expected across languages. Consequently, the fact that the EPP does not
hold in Spanish as originally formulated seems not to be a problem, but rather one of the
expected possible outcomes of variation.

6.3.1. SV(O)

Based on Borer (2005b), we can propose that preverbal subjects of transitive and
unergative structures in broad focus contexts, which correspond to SV(O) word order in
Spanish, occupy the specifier of EP. The reason for their movement to that position is the
need to bind the event variable (<e>_E) in order to license the event. (43) represents the
structure of the sentence: \textit{Los gatos comieron pescado} ‘The cats ate fish’. In this case, the
preverbal subject receives nominative case in the specifier of IP on its way to the
specifier of EP.
Since subject movement takes place because of the need to bind $<e>_E$, this proposal does not restrict the interpretation of a sentence to its subject, as previous analyses that linked subject movement to a predication requirement did.

Given the correspondence between the properties of the element that occupies the specifier of EP and preverbal subjects in Spanish, it seems likely that the latter occupy that position. If we go back to the properties of the event phrase, there is a restriction regarding the element that appears in its specifier: it has to be existentially closed; thus it can bind the event variable and license the event. As already noted, an element is existentially closed if it is a logical variable, i.e., if it has received range. Range assignment can occur within the DP, which is the case for strong DPs (DPs with determiners or demonstratives), and also proper names. Range assignment can also occur through a generic operator or the operation of EC. Since preverbal subjects in Spanish are strong DPs, they are good candidates to bind the event variable and license the event.

It should be noted that the fact that a position like spec EP requires an existentially bound element in order to bind the event variable and license the event does not explain why preverbal subjects in Spanish cannot be bare nouns. It simply offers a possible link between a syntactic property of preverbal subjects and a restriction of the position in which they presumably appear.

6.3.2. LocVS

The ungrammaticality of bare nouns as preverbal subjects in Spanish has been extensively noted. Suñer (1982) captures this syntactic behaviour in her Naked Noun Constraint.
An unmodified common noun in preverbal position cannot be the surface subject of a sentence under conditions of normal stress and intonation. (Suñer, 1982, p. 209)

This property has been used as a test for subject positions and subjecthood. For instance, the ungrammaticality that results from having a preverbal bare subject in broad focus contexts has been taken as evidence that preverbal subjects are not in the left periphery (Casielles-Suarez 2004; Goodall, 1999; Suñer, 2003). This is because while bare nouns can appear in preverbal position in contexts of topicalization or focalization, they are ungrammatical when they appear as preverbal subjects in broad focus contexts.

However, the restriction on the distribution of bare subjects in Spanish seems to be more complex. Not only preverbal subjects need a determiner in order to be grammatical, some postverbal subjects, such as those found in psych predicates, also do so, as (44) and (45) show.

(44) A María le gusta *(la) música.
\[ \text{Maria}_{\text{DAT}} \text{ cl}_{\text{DAT}} \text{ like}_{3S} *(\text{the}) \text{ music} \]
‘Maria likes music.’

(45) A María la preocupan *(las) guerras.
\[ \text{Maria}_{\text{DAT}} \text{ cl}_{\text{ACC}} \text{ worry}_{3PL} *(\text{the}) \text{ wars} \]
‘Maria worries about wars.’

Cuervo (2010), building upon Suñer (1982), reformulates The Naked Noun Constraint to include the behaviour of postverbal nominative subjects of psych predicates of the *piacere* class.\(^98\) She notes that the notion of subject in this case should be understood as a specifier of a verbal projection. This specification is relevant because in her analysis,

\(^98\) Cuevo’s reformulation of the Naked Constraint is also intended for postverbal subjects of inchoatives, unaccusatives with *se*, and objects of causatives. These other contexts, however, are not relevant to present discussion.
psych predicates are unaccusative structures that license their nominative subject in the specifier of $vP_{BE}$.

The Bare Noun Constraint, revised

An unmodified common noun cannot be the subject of a predicate under conditions of normal stress and intonation. (Cuervo, 2010)

Yet we face a problem with this new version of the constraint. It is the case that subjects of transitive and unergative verbs, presumably generated in the specifier of $vP$, or some equivalent projection, can be bare nouns when they appear in a postverbal position in cases that correspond to broad focus; thus these constructions are expected to bear normal stress and intonation, (46) and (47).99

(46) Aquí duermen animales.

here sleep$_{3PL}$ animals

‘Animals sleep here.’ (Torrego, 1989)

(47) En este parque juegan niños.

She notes that the structural requirement for $en$ cliticization to arise is that “$en$ must c-command its extraction site within the DP, but also (the highest occurrence of) the DP cannot c-command $en$” (p.279). Thus she concludes that any postverbal subject within the c-command domain of the verb is predicted to allow $en$ or $ne$ cliticization, which is not the case for a subject in [Spec, EP]. Moreover, Borer also points out that there is no evidence to propose that these structures can be analyzed as unaccusative because there are no changes in the event structure and “the postverbal argument retains its originator role” (p. 281).

99 The possibility of adopting an analysis such as the one proposed by Torrego (1989), in which these instances correspond to unaccusative structures, seems unlikely. Borer (2005b) argues that the $en$ cliticization evidence that Torrego uses to demonstrate the change in the paradigm in Catalan, sentences (a) and (b) (Borer, 2005b, p. 277), which parallels the $ne$ cliticization in Italian in the same environments, can be obtained by “any analysis that places the subject within the scope of the verb, at some relevant stage of the derivation, in (a) but not in (b)” (p. 279), without having to propose a change in the argument structure.
In this park play\textsubscript{3PL} children

‘Children play in this park.’ (Example based on Suñer (1982))

The same behaviour is attested in cases of transitive structures. Although these sentences need a larger context for the nominative subject to be interpreted as the subject and not as a direct object, once the context is available such interpretation is possible. The postverbal subject is also interpreted as existential in these cases.

(48) Por el reguero y las sillas pequeñas se sabe que aquí comen niños.

for the mess and the chairs small SE know\textsubscript{3S} that here eat children

‘By seeing all this mess and the small chairs, you can tell children eat here.’

The key here is how the word subject is understood, as the new version of the constraint is meant to include not only the traditional subject position (the highest specifier related to IP or TP, the canonical subject position, which corresponds to the preverbal subject once the structure is linearized), but also the position in which elements are merged. Specifically, it is intended to capture the distinction between elements merged in the complement position, which, for the most part, do not bear any grammaticality restriction linked to being bare, and those merged in a specifier position, which in the case of psych predicates of the piacere class need some determiner in order to be grammatical, whether they are realized in the preverbal field or not. Note the asymmetry with postverbal bare nouns, which are possible in a complement position of transitives and unaccusatives with an existential interpretation, (49) and (50), respectively.

(49) María oye música.

Maria listens\textsubscript{3S} music

‘Maria listens to music.’

\footnote{This is an oversimplification of the issue. See below for cases in which objects also need modification in order to be grammatical in similar contexts.}
Llegaron estudiantes.

came\textsubscript{3p} students

‘Students arrived.’

However, if we propose that the distinction between merging positions is also relevant to the distribution of bare nouns in Spanish, the appearance of postverbal bare subjects in unergative and transitive structures is still a puzzle. If, on the other hand, the behaviour of nominative subjects of psych predicates is due to properties of the structure in which they appear, namely the fact that they refer to states, we can start reanalyzing their syntactic behaviour based on the properties of such structures. I explore this option below. As far as this idea is on the right track, it allows for an analysis of unergative and transitive structures with postverbal subjects that is independent of the merging position of their subjects.

Torrego (1989) notes that unergative structures with postverbal subjects are only possible when there is a preverbal locative.\textsuperscript{101} According to Borer’s (2005b) analysis, the locative binds the event variable, which licenses the event, as depicted in (51) (adapted from Borer (2005b)). In this case, no grammaticality issues arise from the fact that the subject is a bare noun, independently of its merging position.

\textsuperscript{101} There seems to be some variation among the word order of these constructions if we take into account the different works that have used similar examples. It is not clear, however, to what extent such variation depends on the type of syntactic evidence that researchers are looking at. Torrego (1989) argues that the locative must occupy the preverbal position. In her analysis the locative is crucial for licensing this type of structure, which represents a change in the argument structure. See note 24 for criticisms of this proposal. Suñer (1982), who works with the distribution of bare nouns, and is looking at differences between presentational and declarative sentences, considers (a) and (b) grammatical. In both cases, however, the locative is present.

(a) Juegan niños en el parque.

play\textsubscript{3pl} children in the park (Suñer, 1982, p. 211)

(b) Durante la mañana juegan niños en el parque.

during the morning play\textsubscript{3pl} children in the park (Suñer 1982, p. 212)

Modifications of sentences (a) and (b) also appear in Zubizarreta (1998, p. 109) and Caso-Suárez (2004, p. 101), (c) and (d), respectively.

(c) A menudo juegan niños en este parque.

often play\textsubscript{3pl} children in this park

(d) Jugaban niños en la calle.

played\textsubscript{3pl} children on the street
This type of construction carries an existential/presentational interpretation, similar to that of unaccusative verbs with postverbal subjects. This meaning arises from the presence of a locative, which carries existential force. As previously noted, there is ample cross-linguistic evidence that locatives are linked to existential interpretations. Borer (2005b) points out that Dobrovie-Sorin and Laca (1996) arrive at similar conclusions: they argue that by localizing in space objects or eventualities we are asserting their existence. Thus we do not need to propose that there is a change in the argument structure of these constructions in order to account for the existential interpretation of the postverbal subject.

A problem emerges from (51) though: how does the postverbal subject DP receive nominative case? Taking into account that there are compelling theoretical and empirical reasons to assume that nominative and accusative case must be checked overtly in a spec-head configuration, and not through Agree or the operation of Move F (Bošković (2007), and references therein), sentences with postverbal subjects and a filled specifier in EP such as *Aquí duermen animales* ‘Here sleep animals’ represent a problem for case assignment theories, if we assume that the verb raises to T, and the subject stays lower in the structure.

A possible solution would be to treat LocVS structures in Spanish in a similar way to their counterparts in Catalan and Hebrew. In these languages such constructions include a locative clitic that attaches to the verb. Since Spanish does not have overt or covert locative clitics, and a locative element is necessary in these contexts, the need for an overt locative expression in Spanish simply follows. Although the locative in Spanish is, in principle, an independent element, a possibility may be that it functions more like a
functional element, similarly to locative clitics in Catalan and Hebrew. Assuming that this is the case, one way for capturing it is by merging the locative earlier; thus it fuses with the verb. A possibility that derives from this is that the same locative licenses the subject DP. Thus, that the postverbal DP must be weak also follows.\footnote{Borer (2005b) argues the same for Hebrew. She notes that since T is “in the upward path of the movement of V+loc to E, the DP in [Spec, TP] (…) agrees, perforce, with the copy of V+loc” (p. 300). From the existence of such an agreement relation, she proposes that the locative assigns range to $<e>$. She bases her analysis on the behaviour of perfective prefixes of Slavic languages, which assign range not only to $\{AspQ\}^{<e>\#}$, but also the DP in [Spec, AspQ]. She concludes that “a range assigner to an open functional head value must also assign range to some open value within that functional projection’s DP specifier, if one projects” (pp. 300-301).} If this is the case, the verb would be higher than T, and the postverbal subject could be in the specifier of T, where it receives nominative case. (52) represents this derivation.

\[(52) \quad [EP \quad \text{loc}^3 \cdot \text{V} \quad <e^3>_E [\text{IP} \quad \text{DP}_{\text{weak}}^3 \quad \text{loc}^3 \cdot \text{V} \quad […]]]\]

\[
[\text{EP} \quad \text{Aquí-duermen} \quad <e^3>_E [\text{IP animales} \quad \text{aquí-duermen} \quad […]]]
\]

It is important to note that the fact that these constructions are only possible when the DP is a bare subject supports the idea that it is the locative element the one that is licensing such DP. Although postverbal strong subjects are possible in the sequence LocVS, the information structure in such cases does not correspond to that of broad focus contexts. Finally, if this analysis is correct, it predicts that the so-called variation regarding the word order of this type of construction is not so (see note 26 in this chapter). In other words, the default word order of these constructions is LocVS, while the other possibilities correspond to different information structures.

6.3.3. VS

As previously argued, unaccusative constructions with postverbal subjects contain verbs that are like idioms (Borer, 2005b), in the sense that they that include a locative as part of their structure. As a result, these verbs, when they appear in preverbal position, are able to bind the event variable and license the event. The diagram in (53), based on (42), depicts the structure of unaccusative constructions in Spanish. Note that the verb is also able to bind the postverbal XP.
The verb that has a locative incorporated raises to T and then to the specifier of EP, which guarantees its first position in the string once the structure is linearized. In (52) the DP raises to the specifier of T to check nominative case. The interpretation of these constructions is always presentational/existential. Borer (2005b) points out that the postverbal subject can be a bare noun (this is the case in which the verb licenses not only the event variable, but also the variable in D), but it does not have to be. Strong DPs are possible in these constructions, unlike unergative and transitive verbs with postverbal subjects in which the postverbal DP is obligatorily weak, as seen in the previous section.\(^{103}\) When there is a strong DP, the verb does not take part in the licensing of the variable in D, which is licensed internally within the DP. There are interpretational differences linked to the two licensing possibilities. Bare postverbal NPs are interpreted as existential, while postverbal strong DPs are interpreted as specific. These constructions retain their presentational interpretation even when they include a postverbal strong DP.

This analysis with a locative incorporated in the verb is similar to Ortega-Santos’s (2008) proposal for unaccusative structures in Spanish. However, there are important differences.

---

\(^{103}\) The asymmetry between these two constructions could suggest that the operation of binding the variable in D does not take place in unergative and transitives until the verb with the locative occupies the specifier of EP.
between the two accounts. The main structural difference corresponds to the position that the locative or the element that contains the locative, i.e., the unaccusative verb, occupies: the specifier of IP, for Ortega-Santos, and the specifier of EP, for Borer. Further differences correspond to the nature of the locative itself.

In Ortega-Santos’s (2008) proposal, the locative is a covert element that is part of the verb’s lexical entry. However, the fact that it can move independently of the verb, as it does when it occupies spec IP, according to Ortega-Santos’s analysis, suggests that it is a functional element, a sort of null pronoun with locative meaning. If this is the case, one can wonder why such a locative is not available in other contexts in which postverbal bare subjects are possible in Spanish (also compatible with overt locatives in the preverbal position, such as unergative and transitive structures with postverbal subjects). In Borer’s (2005b) proposal, on the other hand, since the locative is incorporated into the verb, it follows that it is not available in any other context than those that include unaccusative verbs. In order to license postverbal bare subjects in structures that include unergative and transitive verbs there has to be an overt locative in Spanish, as seen in section 6.3.2.

6.3.4. DatVS

Psych predicates of the *piacere* class constitute a more interesting and puzzling instance as the nominative subject is realized postverbally, and the dative experiencer appears in the preverbal position. Yet these postverbal nominative subjects cannot be bare nouns; thus they seem to fall under the same restriction as preverbal subjects. As already noted, the behaviour of these subjects is part of the motivation for Cuervo’s (2010) reformulation of Suñer’s (1982) Naked Noun Constraint. Under the analysis adopted here, it would be tempting to propose that since the restriction that these subjects bear correlates with the properties of the element that occupies the specifier of EP, it may be the case that they may occupy the specifier of EP. A problem immediately emerges: if these subjects occupy that position, how can we account for their postverbal realization? Moreover, how can we accommodate the dative in the structure?

The structure of psych predicates of the *piacere* class has been a source of extensive debate (Belletti & Rizzi, 1988; Cuervo, 1999, 2003). The relevant structure for the
present argument is the one proposed by Cuervo (2010) depicted in (54). In her analysis, psych predicates are instances of unaccusative structures for they do not license an external argument (i.e., a DP licensed by Voice or vP) and accusative case is not available. However, unlike regular unaccusatives, they license their sole argument in the specifier of vP. Another relevant property is that the dative is licensed by an Applicative head as an extra external argument, and not as an argument of the verb. These two features differentiate Cuervo’s analysis from that of Belletti and Rizzi’s.

(54)

\[ \text{Cuervo's (2010) representation captures the interpretation of datives in this type of construction, in which a state is interpreted in relation to an individual. Datives, in this case, seem to be working as placeholders for these states to be realized in, a meaning close to that of locatives. It is possible for psych predicates to be realized with a locative element instead of a dative, as (55) and (56) show.}^{104} \text{ In such cases, their interpretation does not vary. When a locative appears instead of a dative, the state is interpreted with respect to an abstract set of individuals linked to the place that the locative denotes.} \]

(55) En Francia gusta el chocolate.

In France like$_{3S}$ the chocolate

‘People in France like chocolate.’

(56) En mi casa no gusta$_{3S}$ la música.

\[^{104} \text{The relation between locatives and datives has been discussed in Cuervo’s work, as well as in De Miguel (2015), among others.} \]
In my house not likes the music

‘No one likes music at my place.’

There are also cases in which the locative is not realized, (57-58). In such instances, the sentence is interpreted in reference to an abstract collective notion of individuals linked to a certain place, some sort of deictic abstract location which gets actualized depending on the context.

(57) El chocolate gusta.

the chocolate like$_3$S

(58) Gusta el chocolate.

like$_3$S the chocolate

The default word order for the previous sentences is not clear to me, nor is the communicative situation in which they are uttered. Both instances seem to be possible in broad focus contexts. This is especially true when the sentence is embedded in a subordinate, (59), which is probably close to what happens in regular acts of speech.

(59) a. Lo que pasa es que la música clásica no gusta, por eso la gente se ha ido.

b. Lo que pasa es que no gusta la música clásica, por eso la gente se ha ido.

‘What happens is that people don’t like classic music, that is why they have left.’

Since datives and locatives seem to be working as placeholders of the event in these constructions, an option to ponder is that they may occupy the specifier of EP and license the event. A potential problem with this hypothesis is that, assuming that the same behaviour is found in all verbal constructions, it predicts that postverbal nominative subjects of psych predicates would not display any restrictions on bare nouns. In other words, nominative subjects of psych predicates of the \textit{piacere} class would be expected to behave as cases of postverbal subjects of transitive and unergative structures with a preverbal locative, in which bare XPs are possible. As this is not the case, a second possibility is that locatives and datives in psych predicates do not occupy the same
position in the structure as locatives in unergative constructions with postverbal subjects.
Yet, there is a third possibility. It may be the case that datives and locatives are licensing
the event in psych predicates, and the behaviour of the postverbal nominative subject is
due to other restrictions linked to the specific properties of these constructions, namely
the fact that they refer to states, as discussed below in section 6.3.4.3.

The first scenario has to be discarded because the postverbal nominative subject cannot
be a bare noun. In the second scenario, preverbal dative experiencers of psych predicates
and preverbal locatives of unergative constructions do not occupy the same position.
Since [specEP] is not occupied by the dative element, it offers the possibility to
accommodate the postverbal nominative subject in it. This allows us to match the
restriction on nominative subjects of psych predicates—the impossibility of being bare
nouns—with the restriction of the element that occupies [spec EP], its need to be
licensed. A possible way of capturing this may be by merging the applicative above the
EP node. In this case, we would have to propose a in which the EP node projects above
vP_{BE}, then ApplP merges, then TP.

Even though this scenario allows matching the properties of the postverbal nominative
subject with the properties of the element that appears in the specifier of EP, a
fundamental question arises: is there any evidence for treating psych predicates
differently from previous cases regarding the localization of the event node?
Additionally, what are the consequences of proposing an intermediate position for the
event node if we assume something along the lines that syntactic representation is a
reflection of the event structure? By the same token, how is the dative incorporated to the
event? These questions are part of a more important one: whether or not it is possible to
extend to psych predicates, which are instances of states, the EP analysis under the same
conditions of eventive predicates (understood as dynamic).

6.3.4.1. A parenthesis on states

The idea of treating sentences that denote states along the same lines as those that denote
events (activities, accomplishments and achievements) was proposed by Parsons (1987,
1995), in order to derive their logical meaning. In the case of states, Parsons argues that
such meaning consists of an In-ness relation, which should be understood as being in the
state: “x is in s” (1995, p. 644) (where x is the individual/entity and s the state).

According to Parsons, the logical form of sentences that denote states such as *Mary is sick* is \((\exists s)[s \text{ is being-sick} & \text{In}(s, \text{Mary})]\). For *Mary believes what John said*, the logical form is \((\exists s)[s \text{ is a state of believing} & \text{In}(s, \text{Mary}) & \text{Theme}(s, \text{what John said})]\). He notes that “when the verb is one of psychology or perception, the In-ness relation coincides with (...) the Experiencer relation” (Parsons, 1995, p. 644, note 18). Although there is an overlapping between the two, he points out that the In-ness relation is more general, as sometimes being in a state does not constitute an experience.

One property of the logical form of state sentences is that the subject relates to the state by the In-ness relation that it establishes with it. Another characteristic that derives from the previous one is that it would seem necessary that the property to which the state refers be licensed for the subject to be able to be in an In-ness relation with it. If we assume that event structure corresponds to syntactic functional structure, then these properties may have an impact on the differentiation of states and events (understood as activities, accomplishments and achievements).

While Borer’s (2005b) work does not address states in detail, she suggests that they may contain some structure that is not present in eventive predicates. Her idea is that statives may include a projection, which she labels Stative Phrase (SP), responsible for licensing the subject-of-state, which in turn licenses the event. Although she does not offer an exact position for SP, she says that it should project above vP for two reasons. The first one is that SP could deverbalize vP, which would give rise to the possibility for statives to appear with APs as in *Mary is sick*. The second reason is that since SP introduces the subject-of-state, and given that statives with direct objects are possible, e.g., *Mary loves cats*, then SP should be higher than any projection in which objects receive/check (depending on the approach) case. (60) depicts her structure for stative events.

\[(60) \quad [_{\text{EP}} <_{\text{E}} [_{\text{TP}} [_{\text{SP}} [_{\text{VP/AP}} ]]]]] \quad \text{(Borer, 2005b, p. 265, (25a))}\]

If we apply the previous structure to psych predicates of the *piacere* class, the subject-of-state is the dative experiencer, since it is the element that corresponds to that interpretation. This is similar to Cuervo’s structure where the dative is merged as part of an Applicative Phrase. In this structure the dative is not an argument of the verb, but
relates to it because it is merged in the specifier of a verbal predicate, which makes it its argument. The idea of having a special projection like SP for stative sentences in which the subject-of-state is merged seems more desirable because it can accommodate not only dative experiencers of the *piacere* class, but also other experiencers that are realized as nominative subjects or direct objects, as will be seen. It also allows for distinguishing between these subjects (nominative and non-nominative) and subjects of eventive predicates. Additionally, other differences between stative and non-stative sentences can be accommodated in this projection as well.

In this structure, the localization of the event node is the same as in non-stative sentences. Thus it corresponds to the scenario where the dative experiencer licenses the event, which accounts for its preverbal position. A problem with this scenario is that it does not say anything about the postverbal nominative subject, which cannot be a bare noun. As previously suggested, a possibility may be that the restriction that nominative subjects of psych predicates bear arises from other properties of stative structures. Thus the next question that we have to answer is about what other properties, in addition to the Stative Phrase, characterize these constructions.

6.3.4.2. More on the properties of states

Generally speaking, stative sentences refer to states that hold over a stretch of time as in *Mary loves cats*. They have been characterized as homogeneous situations, “which do not change over time: they hold for an undifferentiated interval, with no structure: a state is the same at all times for which it holds” (Smith, 1999, p. 480). The same property is also found in sentences that refer to activities, such as *Mary eats apples*. Given their atelicity, there has been a long tradition in treating states and activities along the same lines.¹⁰⁵ There are, however, important differences between the two. Smith (1999) notes that while

---

¹⁰⁵ Note that several studies have challenged the idea that states are homogeneous; see Husband (2012), for instance. For the present argument, however, this is not a relevant point, as we are focusing on other differences between states and events.
states are static, activities are dynamic. One of the consequences of this is that “verbs expressing states do not have progressive forms” (Mourelatos, 1978, p. 416).

Another piece of evidence that points to the possibility of different syntactic structures for eventive and stative predicates—and this case is more relevant for nominative subjects of psych predicates—comes from the interpretation of bare nouns in object position in English in generic and/or habitual statements in predicates that correspond to events and states. Laca (1990) argues that the two possible interpretations of bare nouns in object position in English, traditionally labelled as generic and existential, (61) and (62) respectively, are linked to “the dimension of the verb’s being tied or not to actual occasions of performance (…) (what distinguishes *hate* from *beat*)” (p. 43). She also notes the relevance of agentivity in these cases.

(61) Mary hates coffee.

(62) Mary drinks coffee.

As Laca points out, although in English there seem to be no differences between these two types of objects, aside from their interpretations (generic and existential), in Spanish there is a difference regarding the use of the definite determiner. Transitive sentences that refer to events (63), understood as habitual activities, whose object is interpreted as existential (as in (62), do not require its direct object to have a determiner. In contrast, in sentences that refer to states (64) and whose object is understood as generic (as in English (61)), Spanish requires the direct object to include the definite determiner (Laca, 1990, p. 27).

(63) Los guamba-mamba comen salmón.

the Gwamba-Mamba eat3PL salmon

‘The Gwamba-Mamba eat salmon.’ (Laca, 1990, p. 27, (7a))

---

106 Some of these verbs can actually appear with progressive forms but their meaning changes from a state to an instantiation of that state, sometimes even interpreted as an event.
(64) Los guamba-mamba detestan *(el) salmón.

the Gwamba-Mamba hate_{3PL} *(the) salmon

‘The Gwamba-Mamba hate salmon.’ (Laca, 1990, p. 27, (7c))

Can the requirement for the definite determiner in nominative arguments of psych predicates respond to the same reason as the requirement in generic objects in stative contexts? This seems to be the case, as the syntactic behaviour and the interpretation of the nominative subjects in (65) and (66) illustrate.

(65) A María le gusta *(el) salmón.

Maria_{DAT} cl_{DAT} like_{3S} *(the) salmon

‘Maria likes salmon.’

(66) A María le encantan *(las) manzanas.

Maria_{DAT} cl_{DAT} like_{3P} *(the) apples

‘Maria loves apples.’

Taking this into account, a possible hypothesis is that the restriction on nominative subjects of psych predicates does not arise from such DPs being subjects, or being generated in a specifier position as opposed to a complement position, but from the fact that they are interpreted as generics in habitual contexts.\(^{107}\), \(^{108}\) Note that generic bare nouns are not possible in Spanish regardless of the position in which they appear (Longobardi, 1994).

\(^{107}\) This behaviour ultimately depends on the properties of Spanish as a language in the context of DPs with generic interpretation, as will be seen next. Also note that while it is true that non-specific indefinites also show restrictions in this type of predicates (cf. (a) ?A María le gusta un muchacho. ‘Maria likes a guy.’ vs. (b) A María le gusta un muchacho feo. ‘Maria likes an ugly guy.’), I am referring here to existential/habitual contexts, in which bare nouns are supposed to be acceptable, as in (64). I thank A.T. Pérez-Leroux for calling my attention on this issue. Interestingly, sentence (a) is acceptable in a context in which is not expected that María likes guys.

\(^{108}\) Another possibility is that the restriction arises from a combination of these two properties: structural position and being part of a stative predicate.
At this point it is worth pointing out that the restriction on preverbal subjects that the
Naked Noun Constraint (Suñer, 1982) and its reformulation (Cuervo, 2010), which
includes some postverbal subjects, aim to capture is relevant for contexts that refer to
either generic or existential interpretations of the XPs, since these are the only instances
in which bare nouns, i.e., non-referential XPs, are potentially possible. In contexts that
are not existential/habitual XPs are referential; thus they fall outside the jurisdiction of
these constraints.

The motivation for extending the Naked Noun Constraint (Suñer, 1982) to some
postverbal DPs comes from the similarity that these DPs—among them, nominative
subjects of psych predicates of the *piacere* class and nominative subjects of inchoatives
constructions—display with respect to preverbal subjects regarding their impossibility of
being bare nouns, (67) and (68). As already noted, in Cuervo’s analysis such DPs are
generated in a specifier position, as opposed to subjects of unaccusative verbs, which are
generated in a complement position. The fact that they pattern with preverbal subjects
with respect to modification is expected given their subject-like nature, in Cuervo’s
analysis.

(67) A todos les gustan *(los) gatos.

       everybodyDAT clDAT  like3PL *(the) cats

       ‘Everybody likes cats.’ (Cuervo, 2010)

(68) A Laura se le derritió *(el) chocolate.

       LauraDAT SE leDAT  melted3S *(the) chocolate

       ‘The/ some chocolate melted on Laura.’ (Cuervo, 2010)

Although not impossible, a problem immediately emerges when other constructions are
considered: the same restriction on bare nouns is found on objects of transitive stative
verbs, despite these being generated in a complement position. Additionally, while it is
ture that the ungrammaticality of the previous examples is due to the absence of
determiners, it may be the case different factors are playing a role in the need for
determiners in these two sentences. Specifically, one thing should be noted: while (67) refers to an unbounded state, (68) refers to a specific event that happened in the past. Also note that these differences have an impact on the interpretation of both DPs: in (67) the DP interpretation is generic, whereas the DP in (68) carries an interpretation close to that of an existential XP (but not exactly the same). Thus it is likely that there are different requirements for the need for a determiner in these two cases in Spanish.

For the DP in (67), given the generic meaning of the sentence, the need for a determiner may be due to the impossibility of licensing bare nouns with a generic interpretation in Spanish. This is found in all contexts in which the interpretation of the DP is generic, independently of the DP merging or final position, and independently of their case. (64), repeated below, shows that the same restriction applies on direct objects. In these cases, the presence of the determiner is problematic because determiners are supposed to encode specificity, and they do not do so in these contexts. Compare the Spanish version with the English equivalent in which no determiner appears.

(64) Los guamba-mamba detestan *(el) pescado.

The Gwamba-Mamba hate3PL *(the) fish

‘The Gwamba-Mamba hate fish.’ (Laca, 1990, p. 27, (7c))

On the other hand, (68) corresponds to some event that took place in which there was some chocolate that melted, and such an event happened to Laura. Although chocolate does not have to be specific in sentences like (68), (it may, as it is in this case), it does need some kind of anchoring in order to encode the meaning of a type of event, lets say different than an activity, which took place. Crucially, it is possible to license postverbal subjects of inchoative constructions even when they refer to events different than activities in the past if there is a locative present, as (69). Given the properties of locatives, a possibility may be that they are taking part in licensing such postverbal subjects.

(69) A Laura se le derritió chocolate en la bolsa.

LauraDAT SE clDAT melted3S chocolate in the bag
‘Chocolate/ some chocolate melted on Laura’s bag.’

While I will not attempt an analysis of the properties of inchoative constructions, a possible hypothesis is that their restriction on bare nouns develops from a combination of some of their properties, both structural and semantic. What is clear, however, is that there is a difference between inchoative constructions and psych predicates regarding the interpretation of their nominative DPs and the elements that seem to be taking part in their licensing proposes. Moreover, psych predicates also differ from unergative and transitive non-stative predicates regarding the possibility to license postverbal bare subjects, and the elements that participate in the licensing process.

At this point, we can assume that in (69) the postverbal bare XP is licensed by the operation of existential closure or by a locative element. We already saw that locatives can license postverbal bare subjects in Spanish. As shown before, postverbal bare subjects of unergative verbs are possible when there is a preverbal locative element. In this case the XP receives an existential interpretation. The same is attested with subjects of transitive verbs, which are possible in a postverbal position, where they can also receive an existential interpretation. These sentences need a larger context for the nominative subject to be interpreted as the subject and not as a direct object. Nonetheless, their existence demonstrates that it is possible to license existential postverbal subjects of transitive verbs. (48), repeated here, shows that a locative is also present.

(48) Por el reguero y las sillas pequeñas se sabe que aquí comen niños.

‘By seeing all this mess and the small chairs, you can tell children eat here.’

On the other hand, in psych predicates, or stative predicates in general, when they refer to unbound states, and appear either in present tense or some atelic past like imperfect in Spanish, the interpretation of the DP is always generic. Thus we must assume that the operation of existential closure, for some reason, seems not to be available in stative contexts, or that it gets blocked due to some other operation that takes place before it. Along the same lines, the presence of locative elements, which can license or take part in the licensing process of events and also seem to be able to license postverbal bare XPs
with an existential interpretation in most constructions in Spanish, do not exhibit the same effects for states. Unlike all previous cases that refer to activities, and even though locatives are compatible with psych predicates, locatives cannot license bare XPs in stative constructions. If this reasoning is on the right track, the behaviour of nominative subjects of psych predicates can be reduced to the combination of their semantic properties, i.e., generic interpretation, which is part of the structure in which they appear (unbound states), and the restrictions that Spanish has regarding licensing nouns in generic contexts.

A similar problem, minus the determiner issue, emerges in English in cases of stative contexts with their object. Although in English bare nouns with generic interpretations are possible, the only context in which they appear postverbally is in stative predicates. In eventive predicates, postverbal bare nouns are always interpreted as existential. In order to capture these differences, Borer (2005b) suggests that the Stative Phrase (SP), the projection in which the subject-of-state is merged, could also license generic direct objects. The motivation for including the generic feature in this projection is that, potentially, such objects in English may be licensed by the operation of existential closure since they are within the VP domain. By incorporating a generic feature in SP, this proposal does not exclude the possibility of existential closure at the VP level, but such an operation does not take place because the generic operator binds the variable in D associated with these objects.

A tentative implementation, based on the previous interpretational requirements and Borer’s insights can be found in (70) for Spanish. For the DP structure, I adopt Borer’s representation, which I discuss below in relation to the use of the definite determiner in generic contexts in Spanish. Although not represented in this structure because it is not relevant, I assume, also based on Borer (2005a), that since GEN, the generic operator, assigns range not only to \(<e>_d\), the variable linked to D, but also to \(<e>_#\), the one related to quantity, it forces the projection of \(<e>_#\) in the DP.
(70) corresponds to stative contexts as it includes a SP (Stative Phrase). This projection is responsible for licensing the subject-of-state. Given that the only requirement for an element to appear in this position is its interpretation with respect to the state, this position can accommodate any DP that corresponds to such an interpretation. This has the advantage of not linking the specifier of SP to any element with specific case, in other words, the case with which the DP that merges in this position surfaces is not a problem. However, as will be seen next, it does have an impact on the case that is assigned to the other DP of the structure. This structure also includes a functional projection (FP) related to generic interpretation. This FP is the host of the generic operator, which binds the DP that moves into it. Additionally, this FP is also the locus where accusative case is assigned when this operation takes place.

For instance, in psych predicates of the *piacere* class, which include a dative experiencer DP, as (71), the dative DP is merged in the specifier of SP, as its interpretation corresponds to that of the subject of state. The theme DP, which is realized as the nominative subject, is merged lower in the structure within the VP (or vP).

(71) A Juan le gustan los gatos.

\[ \text{Juan}_{DAT} \text{ cl}_{DAT} \text{ like}_{3PL} \text{ the cats} \]

‘Juan likes cats.’
I assume the dative experiencer to bear inherent case. Since an element with dative case is merged in the specifier of SP, the only possibility for the other DP is to receive nominative case, presumably in spec IP (given that accusative case cannot surface in the absence of nominative). Note that, in principle, this structure is compatible with both, an analysis that merges the theme DP in the complement position, as Belletti and Rizzi (1988) do, or in a specifier position, as Cuervo (2010) does. For the present discussion, and based on the evidence covered, I am more inclined towards the first option.

In transitive stative constructions such as (72), which include an experiencer nominative subject and a direct object, the experiencer is also merged in the spec SP, but it is a caseless DP. This DP raises to spec IP to receive nominative case. The theme DP, which is merged in the complement position in VP, receives accusative in the specifier of FP, where it also gets bound by the generic operator.

(72) Juan adora los gatos.

Juan love3S the cats

‘Juan loves cats.’

Finally, psych predicates with accusative experiencers like (73) represent the same scenario described for psych predicates with dative experiencers, the only difference being that the experiencer is merged with inherent accusative case.

(73) A Juan lo asustan las guerras.

JuanACC clACC scare3PL the wars

‘Wars scare Juan.’

As seen, the structure in (70) has the advantage that it can accommodate any of the three psychological predicate types in Spanish, independently of the specific cases that the

109 Note that several issues have to be dealt with in order to derive the correct linear word order. I leave this matter for further research.
experiencer DP bears: dative, accusative or nominative. Additionally, it also captures the restriction related to the generic interpretation of the theme DP, whether it bears nominative or accusative case. This is possible because the generic operator associated with the stative projection is responsible for binding either the variable in D or some other element, such as a definite determiner in the case of Spanish, that binds the variable in D. As a result, the generic interpretation arises. The generic operator is necessary at this level because otherwise these elements would be bound by the operation of existential closure, and their interpretation would be different.

6.3.5. Some thoughts on generic and existential XPs in Spanish, and their distribution

Given that DPs with generic interpretations in Spanish always appear with a definite determiner, and assuming that language variation within the functional domain is due to the way in which open values associated with functional categories receive range (Borer 2005b), a possibility may be that the operation of range assignment by the generic operator, GEN, to the variable in D, <e>d, occurs indirectly in Spanish. This seems to be the case if we compare Spanish, in which generic meaning always emerges in the context of strong DPs, to English, which allows bare nouns with generic interpretations. To illustrate the Spanish case, I examine the way in which range assignment to <e>d takes place in English. For the DP structure, I adopt Borer’s (2005a) representation, depicted in (74).

(74)

In (74), the head of the DP is the variable <e>d, which needs to receive range in order to carry out the mapping from predicates to objects. One way for <e>d to receive range is by merging a definite determiner in the specifier of DP. In this case, the determiner assigns range to <e>d and binds it, as indicated by k and i, respectively in (75). Borer notes that
definite determiners have an index that they inherit from the discourse. This property, in a sense, makes definite determiners similar to anaphors.

(75)

Another way of assigning range to \(<e>_d\) is through a generic operator, GEN, which binds \(<e>_d\). The grammaticality of (76) and the interpretation of the DP *cats* suggest that this is the case for English.

(76)  Cats eat mice.  (Borer, 2005a, p. 138)

Based on this possibility we can assume that, in English, the generic operator assigns range directly to the variable and binds it. (77) shows the representation of this operation for the DP *cats* in the previous sentence.

(77)  \(\text{GEN}^i [\langle e \rangle_d \langle \#P \langle e \rangle_h \langle \# \text{cats} \rangle] \ldots \langle \text{eat mice} \rangle\) (Borer, 2005a, p. 138, (8))

In Spanish, on the other hand, direct range assignment by the operator to the variable seems not to be available, as all sentences that include DPs with generic interpretations suggest. Thus a possibility may be that range assignment occurs indirectly through a specifier that assigns range to its head. If this is the case, the generic operator assigns range to the definite determiner, which in turn assigns range to the variable, which ultimately does the mapping from predicates to objects.

One of the properties of definite determiners is that they inherit an index from the discourse, which makes them like anaphors. Crucially, in generic contexts, determiners do not bear this anaphoric property. This derives from the fact that they do not inherit any index from the discourse, i.e., they have no antecedent. In a sense, they simply offer the possibility of binding \(<e>_d\). As a result, the generic interpretation emerges in these cases.
because the definite determiner lacks such an anaphoric property. To some extent the interpretation of the definite determiner in generic contexts in Spanish is similar to that of a quantifier. However, differently from quantifiers which point to the whole class or to some members of it, definite determiners just point to the class as an entity without restricting their reference to some, all or none of its members. For this reason, they seem to be the least marked among functional elements, which makes them a good choice for generic contexts in the cases where the generic operator cannot bind the variable directly.

On the other hand, existential bare XPs are possible in Spanish (in most contexts except for states) as both objects and subjects, but they are restricted to the postverbal position. In the case of subjects, such properties have been taken as an indication that they remain lower in the structure. Casielles-Suárez (2004) argues that bare subjects are possible in Spanish only when they stay in the VP, where they are mapped into the nuclear scope of the clause and receive existential closure. This proposal, however, represents a problem for theories of case assignment. As pointed out before, there is cross-linguistic evidence and theoretical arguments that support the idea that case unlike agreement is checked in a spec-head configuration (Bošković, 2007, and references therein). Assuming that this is so, the fact that postverbal bare subjects in broad focus statements in Spanish do not leave the VP entails that they must receive case via Agree. While this could be a possibility, another one, based on the analysis adopted in this work, is that bare postverbal subjects are licensed by other elements in the structure, such as locatives; thus, despite their postverbal position, they may be higher than the vP, and could receive nominative case in a spec-head configuration. Note that the results in terms of the interpretation are the same if these subjects are licensed by the operation of existential closure or by a locative element.

Borer (2005b) argues that the domain of existential closure should be extended to the EP node. This, however, does not entail that XPs that are higher than the VP are licensed by the operation of existential closure, but that in the presence of other elements, such as locatives, it is possible to license XPs with existential interpretations higher than the VP domain. Since sentences that allow postverbal bare subjects in Spanish always include either an overt or covert locative element, a possibility may be that these locative elements license the postverbal subjects. If this idea is on the right track, such subjects
could actually be higher in the structure, let’s say in spec IP, where they receive nominative case. (See sections 5.3.2 and 5.3.3 for an implementation of this idea in cases of unergative/ transitive and unaccusative constructions with postverbal subjects, LocVS and VS, respectively.)

Some evidence for this analysis comes from the asymmetry between objects and postverbal subjects with existential interpretations. While bare objects with existential interpretations can be licensed without any locative in the structure, likely though the operation of existential closure, postverbal bare subjects with existential interpretation always appear in structures that include a locative element (either overt or covert). This difference suggests that these two types of XPs are licensed through different means, even though they both bear an existential interpretation. This proposal has the advantage that it allows for the checking of nominative case in a spec-head configuration in all verbal configurations in Spanish.

6.4. Summary

The analysis developed in this chapter is compatible with the experimental findings of Chapter 5. The reader should recall that, based on the data examined, preverbal subjects of transitive verbs and preverbal dative experiencers of psych predicates of the *piacere* class behave alike in broad focus contexts. Additionally, they differ from elements that appear in contexts of topicalization. Thus it was suggested that these findings were compatible with the analyses that propose that preverbal nominative subjects and dative experiencers remain within the inflectional domain in Spanish (Goodall, 1999; Suñer, 2003; Ortega-Santos, 2008; Masullo, 1992; Cuervo, 1999, 2003; Fernández-Soriano, 1999a). Although the IP proposal can also accommodate the prosodic data, I noted that there were two problems: 1) there was no agreement among the different analyses about the exact position that these preverbal elements occupy, 2) the motivation for movement to the preverbal field that some of the analyses have suggested was problematic.

With respect to movement to the preverbal field, I argued that neither EPP, nor predication could be postulated as movement triggers in Spanish, as they prove to be insufficient or too restrictive. Thus, taking into account the interpretational properties of broad focus structures and differences in subjects’ distribution in the same contexts, I
proposed a new analysis, based upon Borer’s (2005b) Event Phrase (EP). In this analysis, the specifier of EP is the locus of preverbal subject realization, and the need to license the event by binding the event variable is the trigger of the movement to this preverbal position. Since specEP can also be filled by non-nominative phrases, this analysis has the advantage of accommodating the different word orders attested in broad focus contexts in Spanish. Another advantage of this analysis is that it provides room for two DPs to be outside of vP, such that some postverbal nominative subjects can be in the “regular” subject position (specIP), while a locative or dative DP occupies specEP, allowing for nominative case checking to take place in a spec-head configuration. This proposal also offers a structure for stative events, based upon Borer’s (2005b) insights, that accommodates the three types of psychological predicates in Spanish regardless of the case of their arguments. Finally, the prosodic behaviour of preverbal nominative subjects of transitive verbs and dative experiencers of psych predicates in broad focus contexts can be captured in this analysis, as they both occupy specEP, a position lower than the CP layer.
Chapter 7
On verbal agreement: subject positions and variable agreement. In Spanish?

I have argued in Chapter 6 that a postverbal subject in Spanish does not necessarily stay within the extended verbal projection. Instead, postverbal nominative subjects in constructions that include preverbal dative experiencers and other unaccusative constructions occupy the specifier of IP, allowing for nominative case checking to take place in a spec-head configuration. Data from variable agreement with postverbal subjects, as opposed to obligatory full agreement with preverbal DPs, presents a challenge to this approach. Thus in this chapter I examine three syntactic contexts where verbal agreement patterns different than the expected ones seem to surface in Spanish. The specific instances are: partial agreement between the verb and a postverbal nominative subject, as (1), lack of agreement in SE-passive constructions with postverbal subjects, as (2), and mismatches between the number feature on the verb and the subject’s, when the latter is a collective noun with singular morphology, as (3).

(1) Vino mi hermana y mi cuñado.
   came$_3S$ my sister and my brother-in-law
   ‘My sister and my brother-in-law came.’

(2) Ahora ya no se suele construir esos edificios.
   now anymore not SE use$_3S$ to build those buildings
   ‘Nowadays, buildings like those are not built anymore.’

(3) El jurado, presionados unos y apocados los más,
   the jury$_S$ under pressure some and fainthearted the majority
   declararon inocente al terrorista.
declared\textsubscript{3P} innocent \textsubscript{DOM}\textsubscript{the terrorist}.

‘The jury, under pressure some and fainthearted the majority, declared the terrorist innocent.’ (Martínez, 1999, p. 2767, (188a))

The position that the subject occupies, either postverbal or separated from verb, as in S-XP-V, has been noted to be relevant for the possibility of an agreement pattern different than the expected one to emerge. Specifically, the behaviour that these cases display has been contrasted to that of preverbal subjects (presumably in specIP), which show regular agreement, to argue (implicitly or explicitly) in favour of differences in the operation of agreement in Spanish, whether it takes place in a spec-head configuration or long distance.

This analysis, however, is problematic for two reasons. First, within the current version of the minimalist framework, the operation of agreement does not need to take place in a spec-head configuration, as previously proposed, and has been replaced by the more general version: Agree (Chomsky, 2000). There is ample empirical evidence that supports the idea that adopting Agree, and as a result, only one version of the agreement operation, is the right choice.\textsuperscript{110} Moreover, as Soltan (2006) points out, “on a conceptual level, it is not clear if there is anything natural about agreement being a reflex of a phrase structure relationship between a head and its specifier” (p. 246), as the same could be expected in an equally primitive relationship between the head and its complement.

But even if this turns out not to be the case, i.e., supposing that the way in which the agreement operation takes place is relevant for its outcome, if we take into account the analysis proposed in Chapter 6, an agreement pattern different than the expected one should not be possible regardless of the surface subject position in cases like (1).

According to the analysis in Chapter 6, postverbal subjects of unaccusative verbs in broad focus structures are hosted in the specifier of IP, consequently at some point of the derivation they are in a spec-head relation with the verb before this one raises to the specifier of EP. A similar point can be made about left dislocated subjects.

\textsuperscript{110} See Boeckx (2008) for arguments in favour of this idea, and ample discussion of data.
Thus, the goal of this chapter is to examine these and the other cases of unexpected agreement patterns, and provide an analysis of them. Although the analysis developed in this chapter is compatible with the proposal of Chapter 6, it does not depend on it, meaning that the fact that some of these elements are in a spec-head configuration is irrelevant. Hence, this analysis suggests that at least in these contexts, there is no evidence that supports the idea that spec-head agreement versus long distance agreement may result in different outcomes in Spanish. Theoretically, my hope is that the data examined here will contribute to the discussion of the properties of agreement as a syntactic operation in Spanish, and also in the language system in general.

7.1. Agreement data: preliminarily interpretation and general discussion

There are cases in Spanish where, at first sight, agreement appears to show sensitivity to the position in which the subject is realized. Examples (4) and (5) contain subjects that constitute a conjunct. These subjects have been argued to be able to trigger partial agreement when they occupy a postverbal position (Camacho, 1999; Fernández-Soriano, 1999a). Note that the possibility of full agreement still holds in these cases, examples (a). On the other hand, if they appear in the preverbal field, the only option is full agreement, compare (b) and (c). Partial agreement refers to the instance in which the verb only agrees with the first element of the conjunct, thus the phi features on the verb are those of the first element: third person singular. Full agreement refers to the case in which the verb agrees with both elements of the conjunct; hence a plural phi feature on the verb is realized.

(4) a. Vino/ vinieron mi hermana y mi cuñado.  

\[\text{vino}_3S/ \text{vinieron}_3P \quad \text{mi hermana y mi cuñado}\]

‘My sister and my brother-in-law came.’

b. *Mi hermana y mi cuñado vino.  

\[\text{mi hermana y mi cuñado vino}_3S\]

c. Mi hermana y mi cuñado vinieron.  

\[\text{mi hermana y mi cuñado vinieron}_3P\]
my sister and my brother-in-law came\textsubscript{3p}

\begin{itemize}
\item a. Me gusta/ gustan el chocolate y el café.
\begin{itemize}
\item the chocolate and the coffee me\textsubscript{DAT} like\textsubscript{3S}/ like\textsubscript{3P}
\end{itemize}
\end{itemize}
‘I like chocolate and coffee.’

\begin{itemize}
\item b. *El chocolate y el café me gusta.
\begin{itemize}
\item the chocolate and the coffee me\textsubscript{DAT} like\textsubscript{3S}
\end{itemize}
\end{itemize}

\begin{itemize}
\item c. El chocolate y el café me gustan.
\begin{itemize}
\item the chocolate and the coffee me\textsubscript{DAT} like\textsubscript{3P}
\end{itemize}
\end{itemize}

A similar pattern linked to DP position (movement) and variable agreement seems to hold in (6) as well, in SE-passive constructions. In this instance, although the DP is plural, when it appears postverbally, it has been claimed that it can trigger either singular or plural agreement (Martínez, 1999). If the DP appears in the preverbal field, plural agreement is the only option.

\begin{itemize}
\item a. Ahora ya no se suele/ suelen construir esos edificios.
\begin{itemize}
\item buildings like those are not built anymore nowadays.
\end{itemize}
\end{itemize}

\begin{itemize}
\item b. *Esos edificios ya no se suele construir ahora.
\begin{itemize}
\item those buildings anymore not SE use\textsubscript{3S}/ use\textsubscript{3P} to build now
\end{itemize}
\end{itemize}

\begin{itemize}
\item c. Esos edificios ya no se suelen construir ahora.
\begin{itemize}
\item those buildings anymore not SE use\textsubscript{3P} to build now
\end{itemize}
\end{itemize}

The last example of an unexpected agreement pattern comes from sentences whose subject is a collective noun with singular morphology, and the verb bears plural morphology (\textit{ad sensum} agreement), as (7a) shows. In this case, adjacency between the
verb and the subject has been suggested to play a role (Suñer, 2003), compare examples (7a) and (7b).

(7)  

a. El jurado, presionados unos y apocados los más,  

the juryS under pressure some and fainthearted the majority  

declararon inocente al terrorista.  

\[\text{declared}_{3P} \text{ innocent } \text{DOM.} \text{the terrorist} \]

‘The jury, under pressure some and fainthearted the majority declared the terrorist innocent.’


the juryS declared\[3P \text{ innocent } \text{DOM.} \text{the terrorist} \]

(Martínez, 1999, p. 2767, (188a))

All these cases have been noted in both traditional and descriptive grammars (Camacho, 1999; Martínez, 1999), and some of them have been used in generative studies to argue in favour of the existence of different subject positions (Suñer, 2003), or different possibilities of agreement in Spanish (Fernández-Soriano, 1999a; Suñer, 2003).

Based on the previous data, one could assume, as previous works have done, that the agreement operation is sensitive to the position in which the subject is realized in Spanish. A hypothesis within this scenario is that an agreement pattern different than the expected one can only arise in cases in which this operation takes place via long distance, given that all previous examples, in which a different agreement pattern is attested, include either a postverbal or left dislocated subject. On the other hand, cases that arguably involve a spec-head relation between the verb and the subject can only display regular agreement.

Note that this hypothesis is not easy to implement. An analysis that postulates that depending on how the agreement operation takes place its results would be different predicts that full agreement should not obtain in configurations that do not include a spec-
head relationship. Contrary to this, full agreement obtains in the case of postverbal conjoined subjects (cases that, according to these works, do not represent a spec-head relation). To account for this possibility, we need to take an alternative route to allow full agreement via long distance. This theoretical move generates a separate problem: we would also need a mechanism to determine how the two agreement possibilities emerge, i.e., under what circumstances each option takes place. I reject this possibility for two reasons. First, under the analysis adopted in this work, both preverbal and postverbal subjects of broad focus sentences in Spanish stop at or stay in the specifier of IP, which entails that at some point of the derivation they are in a spec-head relation with the verb. If this analysis is correct, and assuming that a spec-head relation is relevant to the operation of agreement (note that this claim has been largely challenged), it predicts that there should not be differences in agreement patterns independently of the surface subject position in some of these constructions. Second, there are empirical and theoretical reasons that should make us suspect that differences in agreement patterns in these constructions are the result of differences in other features than the agreement operation itself. Let us review the interesting case of the different analyses of the subject-verb agreement asymmetry in Standard Arabic.

Standard Arabic has been classified as a VSO language; however, SVO is a possible word order too. Asymmetries in subject-verb agreement arise in the context of these two word orders. When the subject is postverbal, the verb agrees in person and gender, but not in number with it. On the other hand, when the subject is preverbal, the verb agrees in all phi features (person, gender and number), as (8) and (9) show.

(8) a. ʔal-ʔawlaad-u qaraʔ-uu d-dars-a
    the-boys_{NOM} read_{3PM} the-lesson_{ACC}

    b. qaraʔa l-ʔawlaad-u d-dars-a

Interestingly, most of the analyses of Standard Arabic do not make any distinction regarding word order differences in VSO and SVO sentences (similarly to what was done in Spanish, see Chapter 3 for discussion). Crucially, as in Spanish, differences in subject realization in these two cases correlate with different interpretations of the subject, as Soltan (2006) points out. Thus the conclusions and proposals that derive from such analyses are limited by this theoretical glitch.
Traditional analyses suggest that, although agreement takes place in a Spec-head configuration in both word orders, different elements participate, and/or different processes happen in each case. For instance, Mohammad (2000) proposes that in SVO sentences, due to the spec-head relationship between the subject and the verb, full agreement surfaces. On the other hand, in the case of postverbal subjects, a null expletive sits in specIP, and partial agreement arises. Aoun, Benmamoun and Sportiche (1994) suggest that some type of agreement loss takes place due to verb movement to a higher projection in the case of postverbal subjects. In SVO sentences, on the other hand, full agreement is preserved (as the verb does not move any higher). As Soltan (2006) notes, both options are problematic within the minimalist framework and, more importantly, they rely on stipulations: Mohammad’s null expletive is not independently motivated, and Aoun, Benmamoun and Sportiche’s agreement loss is not attested in any other constructions of the same language.

In a different take, Soltan analyzes agreement asymmetries in these two contexts in Standard Arabic as a consequence of two different base-generated structural representations. He takes evidence from the behaviour of null and overt subject pronouns, which always trigger full agreement on the verb (irrespective of their position), and
argues that full agreement is obtained in the case of preverbal subjects, because there is a *pro* in the structure that triggers full agreement, and not because the subject and the verb are in a spec-head configuration. In Soltan’s account, the preverbal subject DP appears in the CP layer. He points out that this analysis can not only explain agreement differences between the two constructions, but also account for interpretational, syntactic and semantic differences between them, among them the interpretation of preverbal subjects. He notes that the preverbal subject DPs in Standard Arabic are categorically interpreted as the topic of a topic-comment structure. On the other hand, VS(O) bears a “(default/unmarked) thetic interpretation, whereby an event is neutrally reported with the participants involved” (Soltan, 2006, p. 248).

Contexts that display lack of agreement such as constructions that include quirky subjects in Icelandic are another interesting instance. In this case a dative element appears preverbally, in the canonical subject position, and a nominative element appears postverbally. Although the nominative element has object-like properties, it triggers agreement on the verb. However, as Boeckx (2008) notes, “this agreement is partial and optional, and even barred in some cases” (p. 21). Taking into account the subject behaviour, he proposes a new analysis by which agreement obtains between the quirky element and the verb, “even though it fails to show up morphologically on the verb for non-syntactic reasons” (p. 21). In other words, although there is no overt manifestation on the verb, agreement takes place.\(^{112}\)

While Spanish data is different, the point is that differences in agreement patterns do not necessarily have to be the result from differences in the agreement operation, and the previous two cases point to that direction. Thus I will pursue the simpler hypothesis that differences in agreement patterns in Spanish can be explained by different reasons—depending on the specific syntactic structure of each type of construction—all of them unrelated to the operation of agreement itself. The specific idea that I want to put forward is that the two agreement possibilities in sentences (4-6) version (a), and (7a) and (7b)

\(^{112}\) See Boeckx (2008), Chapter 1, for details.
correspond to different underlying structures. I will argue there is no dual possibility of agreement patterns in sentences with postverbal and left dislocated subjects.

7.2. Analysis of the different agreement patterns

Given the heterogeneity of the previous examples, I grouped them in order to carry out their analysis. These seeming cases of variable agreement have as common feature sensitivity to the position in which the nominative subject is realized, but may be result from different syntactic phenomena.

7.2.1. Agreement and postverbal subjects

The first group, sentences (4) and (5), corresponds to partial agreement in descriptive grammars (Camacho, 1999) (and first conjunct agreement in generative grammar).

(4) a. Vino/ vinieron mi hermana y mi cuñado.
   came\textsubscript{3S}/ came\textsubscript{3PL} my sister and my brother-in-law
   ‘My sister and my brother-in-law came.’

(5) a. Me gusta/ gustan el chocolate y el café.
   me like\textsubscript{3S}/ like\textsubscript{3P} the chocolate and the coffee
   ‘I like chocolate and coffee.’

Two factors have been identified in Spanish for the possibility of partial agreement to arise: the postverbal realization of the nominative subject (which should include two elements joined by the coordinating conjunction y ‘and’), and the type of referent expressed by the subject. It has been noted that nouns that refer to abstract entities favour partial agreement, such as (10) (Camacho, 1999).

(10) Pero el viejo rey estaba acusado de anglofilo
    but the old king was accused of anglophile
    y en la corte existía el favoritismo y la corrupción.
and in the court existed the favouritism and the corruption

‘But the old king was accused of being an anglophile, and there were favouritism and corruption in his court.’ (Camacho, 1999, p. 2650)

However, cases in which the referent is not abstract, partial agreement takes place up to 10%, according to a corpus analysis cited by Camacho (1999). This instance corresponds to sentences (4a) and (5a) with singular agreement. Generally, this type of agreement is found in broad focus structures that include unaccusative verbs and psych predicates of the *piacere* class, because these are some of the syntactic environments that include postverbal nominative subjects in Spanish.

If we adopt an analysis that matches the differences in subject positions (postverbal versus preverbal), to phi feature differences on the verb, and focuses on these two properties only, it is expected that such analysis would conclude that there are differences in the agreement mechanism that depend on the position that the subject occupies. As a result, it would seem plausible to suggest that partial agreement arises in one context, (4a) and (5a), but not in the other, (4b) and (5b).

(4)  
   b. *Mi hermana y mi cuñado vino.*
   
   my sister and my brother-in-law came

(5)  
   b. *El chocolate y el café me gusta.*
   
   the chocolate and the coffee me like

As noted, this type of analysis is silent about the possibility of full agreement with a conjoined postverbal subject. Even if it were able to include it, it would run into problems when implementing the mechanism in each case, as pointed out before. Moreover, if we take subject positions to be the only relevant feature from which agreement differences arise, we would be tacitly assuming that these clauses are equivalent in terms of their structure. Crucially, there is syntactic evidence that suggests that this is not the case.

I would like to propose that the agreement variability in (4a) and (5a) is just apparent, and what we are actually seeing is two sentences that correspond to two different syntactic
structures, whose linear order is the same. The first structure corresponds to a clause that includes a plural subject (formed by two singular elements, \([el\ \text{chocolate}]\) and \([el\ \text{café}]\) in (5a)). In this case the verb agrees with the plural subject. Such sentence would be equivalent to the version with a preverbal subject and plural agreement, (c), in terms of the elements that they comprise. The second structure would be one that includes two clauses with a VP ellipsis in the second clause. (This corresponds to a coordination extraction context, in generative terms.) In this case, there is no counterpart with a preverbal subject because the same structure cannot be obtained in that configuration.

Martínez (1999), also notes this possibility. He proposes that cases of so-called partial agreement are possible because the two postverbal elements correspond to different clauses. He argues that the second clause includes a verb ellipsis, and offers the following example where a modifier is added to demonstrate that this is the case.

(11) \(Llegó\ \text{el\ gobernador\ y\ luego}\) [llegó] \(el\ \text{alcalde}\).

\[\text{arrived}_3\text{s the governor and then [arrived}_3\text{s the major}\]

‘The governor arrived and then did the mayor.’ (Martínez, 1999, p. 2762)

Although Martínez’s analysis and the suggestion above are equivalent, the previous example only proves that the two subjects may belong to different clauses. The following examples that include a temporal adverb at the edge of the clause, which forces the two elements to be treated as a syntactic unit, demonstrate that it is impossible to consider them as part of the same clause. This behaviour goes against an analysis that argues that these cases are instances of partial agreement. The ungrammaticality of (12) confirms that indeed these two elements are not part of the same clause; otherwise the grammaticality of (13) would be hard to explain.

(12) \(*Vino\ \text{mi\ hermana\ y\ mi\ cuñado\ ayer}.*\)

\[\text{came}_3\text{s my sister and my brother-in-law yesterday}\]

\[\text{\textsuperscript{113}}\]

\[\text{I thank Ana Teresa Pérez-Leroux for the idea of using this type of evidence, especially this particular example.}\]
Vinieron mi hermana y mi cuñado ayer.

came$_{3PL}$ my sister and my brother-in-law yesterday

‘My sister and brother-in-law came yesterday.’

Evidence that supports this analysis comes from the behaviour of first conjunct agreement in languages in which this phenomenon is attested. One such language is Standard Arabic. Although, at first sight, Standard Arabic data seems similar to Spanish’s, first conjunct agreement constructions in Standard Arabic display different syntactic properties than Spanish constructions.\(^{114}\)

As noted, a feature of Standard Arabic is that the verb agrees with postverbal subjects in person and gender, but not in number, as (14) and (15) show.\(^ {115}\)

(14) qadim-a (/\*qadim-uu) al- ?awlaadu.

came$_{3MS}$ came$_{3MPL}$ the-boys$_{3MPL}$

\(^{114}\) The syntactic phenomenon of first conjunct agreement in Arabic is rather complex, as there are differences among dialects. Some dialects, such as Lebanese and Moroccan Arabic, seem to allow partial and full agreement with postverbal conjoined subjects. Standard Arabic, on the other hand, only allows partial agreement with postverbal conjoined subjects. Given this behaviour, first conjunct agreement in Arabic has received at least three different analyses.

Aoun, Benmamoun and Sportiche (1994, 1999), who work with Lebanese and Moroccan Arabic, argue in favour of a biclausal analysis, similar to the one proposed here for cases that display singular agreement on the verb. They conclude that first conjunct agreement does not exist in Lebanese and Moroccan Arabic. Munn (1999), on the other hand, analyzes cases of first conjunct agreement as monoclausal, but distinguishes between semantic and syntactic plurality, and proposes that coordinated subjects in Arabic are semantically plural and syntactically singular. He analyzes coordination as an instance of adjunction. In his proposal there is also a difference between the operation of agreement depending on the position of the subject: agreement with postverbal subjects takes place through government, while in the case of preverbal subjects, it takes place in a spec-head configuration. The latter accounts for plural agreement.

Larson (2013), whose goal is to develop an analysis of first conjunct agreement that comprises cross-dialectal differences, claims that none of these two analyses can account for the full paradigm: Aoun, Benmamoun and Sportiche (1994, 1999) cannot account for the lack of variability with preverbal subjects, while Munn (1999) cannot predict variability with postverbal subjects. Larson’s analysis solves these issues by decomposing the Merge operation in two parts: Concatenate and Label (based on Hornstein 2009). Since coordination is also a case of adjunction in his analysis, and “adjuncts not being necessary to the derivation, do not have to necessarily have to undergo Label” (618), when two DPs are coordinated they can undergo concatenate and not label, or concatenate and label. The first scenario accounts for agreement with the first element only, while the second case accounts for plural agreement. He further claims that only subjects that have undergone Label can move to SpecTP.

\(^{115}\) Note that this statement is only true for subjects that are DPs. When the postverbal subject is a pronoun, the verb agrees not only in person and gender, but also in number with the subject in Standard Arabic.
‘The boys came.’  
(Harbert & Maher, 2002, p. 45, (1a))

(15) qadim-at (/\*qadim-ataa\/) al-bint-aani  
came\textsubscript{3FS} came\textsubscript{3MD} the-girl\textsubscript{3D}

‘The two girls came.’  
(Harbert & Maher, 2002, p. 45, (2a))

In cases in which the postverbal subject is a conjunct, i.e., two coordinated DPs, as in (16), the verb only agrees with the first DP.

(16) a. ʒaaʔa Zayd-un wa Hind-u  
came\textsubscript{3SM} Zayd\textsubscript{NOM} and Hind\textsubscript{NOM}

‘Zayd and Hind came.’

b. ʒaaʔa-t Hind-u wa Zayd-un  
came\textsubscript{3SF} Hind\textsubscript{NOM} and Zayd\textsubscript{NOM}

‘Hind and Zayd came.’

c. *ʒaaʔaa Zayd-un wa Hind-u  
came\textsubscript{3DM} Zayd\textsubscript{NOM} and Hind\textsubscript{NOM}

(Soltan, 2006, p. 242, (7a-c))

Evidence that (16a-b) are instances of first conjunct agreement and not VP ellipsis in the second conjunct (or coordination extraction) comes from the fact that these structures are compatible with elements like ‘together’, ‘each other’ and ‘both’, as (17a-b) shows.

(17) a. ʒaaʔa-t Hind-u wa ʃamr-u maʃa  
came\textsubscript{3SF} Hind\textsubscript{NOM} and Amr\textsubscript{NOM} together

‘Hind and Amr came.’

b. tuḥibbu Hind-u wa ʔixwat-u-haa baʔD-a-hum əl-baʔD
love$_{3SF}$ Hind$_{NOM}$ and brothers$_{NOM}$-her some$_{ACC}$-them the-some

‘Hind and her brothers love each other.’

(Soltan, 2006, p. 244, (13a-b))

According to Habert and Maher (2002), and Soltan (2006), this syntactic behaviour is also attested in other languages that allow first conjunct agreement like Welsh and Irish.

Although Spanish data is more restricted given that postverbal subjects in broad focus contexts are only found with a subset of verbs, those that can appear in such configuration reject any plurality element compatible with such contexts.

(18) *Vino mi hermana y mi cuñado juntos/ los dos.

came$_{3S}$ my sister and my brother-in-law together/ both

(19) Vinieron mi hermana y mi cuñado juntos/ los dos.

came$_{3P}$ my sister and my brother-in-law together/ both

‘My sister and brother in law (both) came (together).’

Based on the previous examples, the ungrammaticality of (4b), repeated below, cannot be taken as an indication that there are differences in the agreement mechanism linked to the position in which the subject is realized in this context.

(4) b. *Mi hermana y mi cuñado vino.

my sister and my brother-in-law came$_{3S}$

The same analysis can be extended to psych predicates of the piacere class, as these also include postverbal subjects. In (20a), the version that shows singular agreement on the verb does not accept any element that denotes plurality. Unlike the previous examples of Standard Arabic, (20a) does not constitute an instance of first conjunct agreement, but two clauses. Once the subject is plural, as (20b), these plural elements are possible.

(20) a. *Me gusta el chocolate y el café por igual/ los dos/ ambos.
me like\textsubscript{3S} the chocolate and the coffee equally/ the two/ both

b. Me gustan el chocolate y el café por igual/ los dos/ ambos.

me like\textsubscript{3P} the chocolate and the coffee equally/ the two/ both

Interestingly, among these constructions, there is another group that contains an infinitive instead of a DP. These cases, unlike other unaccusative constructions, never allow plural agreement, examples (15) and (16).\textsuperscript{116}

(21) Me gusta cantar y bailar.

me like\textsubscript{3S} singing and dancing

‘I like singing and dancing.’

(22) *Me gustan cantar y bailar.

me like\textsubscript{3P} singing and dancing

Finally, if Spanish data were real instances of first conjunct agreement, then we would expect other combinations to also be possible. As the following examples illustrate this is not the case. Even if we maintain the third person feature, once we have a plural DP as the second member of the conjoined subject the result is deviant if not ungrammatical, as (23) and (24) show. Note that this is not a problem in Standard Arabic in which the second member of the conjunct can be plural, example (17b), repeated for convenience.

(23) */Vino Pedro y los inspectores a revisar la casa.

came\textsubscript{3S} Pedro and the inspectors to check the house

(24) */Me gusta el pan y las frutas por igual.

me like\textsubscript{3S} the bread and the fruits equally

\textsuperscript{116}The preverbal realization of the infinitives in this case is quite uncommon.
Based on the previous discussion and syntactic evidence, sentences that include two postverbal subjects and singular agreement—as a result of a VP ellipsis in the second sentence or coordination extraction—share the same string with cases that include a postverbal subject that constitutes a conjunct: [X and Y]. As a result, the apparent possibility of two agreement patterns arises when such sentences are compared to those that display full agreement with a preverbal subject. However, only the postverbal version that shows full (plural) agreement can be compared to the sentence with a preverbal subject. Since the postverbal version with partial agreement and the preverbal version with regular agreement correspond to different structures, differences in their agreement patterns do not constitute evidence that the agreement mechanism is sensitive to the position in which the subject is realized in this context.

### 7.2.2. Ad sensum agreement

The second group of unexpected agreement patterns comprises sentences that display *ad sensum* agreement. This type of agreement describes the relation between a noun with singular morphology, whose semantic meaning corresponds to a group, i.e., a collective noun, and a verb that shows a plural phi feature. As (7a) shows, although the DP *el jurado* ‘the jury’ bears singular morphology, the verb shows third person plural agreement. Note that the same type of agreement is not possible in (7b).

(7) a. El jurado, presionados unos y apocados los más, declararon

    the jury$_S$ under pressure some and fainthearted the majority declared$_3P$

    inocente al terrorista.

    innocent DOM the terrorist

    ‘The jury under pressure some and fainthearted the majority declared the terrorist innocent.’

The jury declared innocent the terrorist

(Martínez, 1999, p. 2767, (188a))

The necessary condition for *ad sensum* agreement is, what has been called in descriptive grammars, some distance between the subject and the verb. Within the generative framework, such distance has been taken as an indication that these subjects occupy different structural positions in the two sentences. This is precisely the type of analysis that Suñer (2003) proposes, based on Bosque (1999) who “suggests that subject verb agreement with collective nouns shows that preverbal subjects are not in Topic or LD position” (Suñer, 2003, p. 350). Suñer follows Bosque’s proposal and argues that left dislocated subjects “those clearly in an A’-position (…) admit [*ad sensum* agreement]” (p. 350). On the other hand, she states that preverbal subjects, presumably in the IP layer, such as the one in (7b), “reject *ad sensum* even in the conversational register” (p. 350). Her examples follow.


The jury were pressed

b. *La familia se pusieron de acuerdo.

The family were in agreement

c. *La escuadra atacaron a los enemigos.

The squad attacked the enemy

(Suñer, 2003, p. 350, (25))

(26) a. El jurado, Mara nos aseguró que estaban presionados.

The jury Mara us assured that were pressed

‘The jury, Mara assured us that they felt pressured.’
b. La familia, Mara nos aseguró que finalmente se pusieron de acuerdo.

The family Mara us assured that finally were\textsubscript{3P} in agreement

‘The family Mara assured us that they reached an agreement.’

c. La escuadra, el general explicó que atacaron a los enemigos.

The squad the general explained that attacked\textsubscript{3P DOM} the enemy

‘The squad, the general explained that they attacked the enemy.’

(Suñer, 2003, pp. 350-1, (26))

While Suñer’s claim that there may be different subject positions (and indeed these subjects can potentially be an example of it, although not necessarily so) is right, her analysis implicitly assumes that depending on how the agreement operation takes place, in a spec-head configuration or via long distance, there may be differences in its results. By the same token, it assumes that the operations themselves may be different as well. More important, in her analysis, it seems that the two versions are equivalent in terms of the elements that participate in the syntactic operation of agreement, while this needs not be the case. The key here is the use of the terms accept and reject linked to the subject DP, which seems to imply that the same element participates in the operation of agreement in both cases.

The first two assumptions are problematic when we consider that theoretically the operation of agreement should be the same independently on whether it takes place via long distance or in a spec-head configuration. Empirical evidence against the claim that the two versions are equivalent with respect to the elements that take part in the agreement operation comes from the fact that ad sensum agreement is not the only possibility in cases that include a subject separated from the verb (presumably a left dislocated subject), as (27) and (28) show.

(27) El jurado, presionados unos y apocados los más, declaró

the jury\textsubscript{S} under pressure some and fainthearted the majority declared\textsubscript{3S}
inocente al terrorista.

innocent DOM-the terrorist

‘The jury under pressure some and fainthearted the majority declared the terrorist innocent.’

(28) El jurado, según María, declaró inocente al terrorista.

the jury according Maria declaredS innocent DOM-the terrorist

‘According to Maria, the jury declared the terrorist innocent.’

The existence of examples such as (27) and (28) suggest that structures that display ad sensum agreement and those that show regular agreement include different elements. Assuming that in (7a) the subject is generated in the specifier of vP (or some equivalent projection) and then raises to the CP domain, establishing agreement with the verb at some point, differences in agreement patterns between (7a) and (7b) should not arise, independently of the existence of intervening elements in (7a). If, on the other hand, in (7a), the DP el jurado ‘the jury’ is merged in the CP domain directly, there may be a pro lower in the structure, which would be the real argument of the verb. The phi features of pro would be the ones realized on the verb. Another possibility is that the DP el jurado ‘the jury’ in (7a) is an extra-clausal constituent. In this scenario there is also a pro.117

Evidence for these two possibilities comes from the behaviour of wh-questions. It is widely accepted that in wh-questions that include simple wh-elements in Spanish, there is fronting of the wh-element and verb preposing in most dialects. Compare the word orders in (29a) and (29b).

(29) a. Pedro dijo eso.

117 Note that even in the scenarios that include pro regular agreement can be found. If the phi features of pro are the same of those of the DP merged in the CP domain, the sentence would display the same agreement pattern that a sentence that includes the same DP in the specIP, or a left dislocated subject merged lower in the structure, and later raised to the CP domain. If, on the other hand, pro bears different phi features from the DP in the CP layer, the so-called possibility of ad sensum agreement arises.
Pedro said that

‘Pedro said so.’

b. ¿Qué dijo Pedro?

what said Pedro

‘What did Pedro say?’

The following versions of examples (7a) and (7b) that correspond to questions about the object show that there may be differences between the structures of these two sentences, based on the dissimilar results in both cases. While (31) is licit, (30) is ungrammatical.

The ungrammaticality in this case may arise because wh-extraction through a constituent in the left periphery is not possible.

(30)  *¿A quién declararon inocente el jurado, presionados unos y

\[\text{DOM who declared}_{3P} \text{ innocent the jury}_{S} \text{ under pressure some and}
\]

\[\text{apocados los más?}
\]

\[\text{fainthearted the majority}
\]

(31)  ¿A quién declaró inocente el jurado?

\[\text{DOM who declared}_{3S} \text{ innocent the jury}_{S}
\]

‘Who did the jury declare innocent?’

The behaviour of (30) is compatible with a scenario in which the DP el jurado ‘the jury’ in (7a) appears in the left periphery, thus a possible hypothesis is that it is a topic. Interestingly, as Goodall (1999) notes, topics may be followed by a wh-phrase in Spanish. He argues that this is possible because the left dislocated constituent occupies the Topic projection, which is higher than the Focus projection that hosts the wh-phrase, according to Rizzi’s (1997) system. (32) illustrates this behaviour.

(32)  Ese libro, ¿cuándo lo compraste?
that book_{MS} when \overline{cl}_{ACC/MS} bought

‘That book, when did you buy it?’ (Goodall, 1999, p. 99, (10))

If the DP *el jurado* ‘the jury’ in (7a) is a topic and occupies the topic phrase in the CP layer, we would expect that it could appear to the left of the wh-phrase. Although (33), below, is better than (30), still it is not fully acceptable.

(33) ¿*El jurado, presionados unos y apocados los más, ¿a quién declararon inocente?*

the jury_{S} under pressure some and fainthearted the majority _DOM_ who declared_{PL} innocent

I will argue that the reason for this is that the DP *el jurado* ‘the jury’ is not actually a topicalized subject, but a topicalized element or an extra-Clausal constituent. Note that in (32), the topicalized phrase *ese libro* ‘that book’ is reproduced in the question by the resumptive clitic *lo*. The same pattern is found in (34), when the goal or recipient is left dislocated.

(34) A María, ¿quién le dio *el libro?*

to María_{i} who \overline{cl}_{i} gave the book

‘Who gave María the book?’

Interestingly, when the topicalized element is a subject, a resumptive nominative pronoun usually appears in the question, as (35) shows.

(35) LUISA, ¿qué trajo ella a la fiesta?

Luisa_{i} what brought she_{i} to the party

‘LUISA, what did she bring to the party?’
If we try to replicate the behaviour of (35) in (33), and add a resumptive pronoun to the question with the phi features of the left dislocated subject, a mismatch between the verb and the pronoun would arise. If the resumptive pronoun maintains the verb’s phi features, the mismatch arises between the DP in the left periphery and the resumptive pronoun. In either case, the sentence cannot be saved.

This case, however, is not straightforward. There seem to be restrictions (and also variation) on the substitution of collective nouns by pronouns that match their phi features. In Spanish, for instance, after the introduction of a noun like jurado ‘jury’ in the discourse, the next time that it is mentioned, in most contexts, it is substituted by the overt form ellos/ellas, 3PM/F, ‘they’ or by a covert pro with these phi features. (Note that this could be the case of Suñer’s (2003) examples in (26a-c), as they correspond to different clauses.) When this is not an option, the noun is repeated. A comparable behaviour is attested in tag questions in English.

One of the properties of tag questions in English is that the nominative subject is repeated in the form of a pronoun, as (36) shows.

(36) John wrote a book last year, didn’t he?

Crucially, when the subject is a collective noun, there is variation. Note that when the context activates the plural interpretation of a collective noun like jury, native speakers tend to substitute the DP the jury by they, as in (37). On the other hand, when the noun is seen as a homogenous entity, it is replaced by it, example (38). Also note that while (37) may also appear with it, (38) can never appear with they.\textsuperscript{118}

(37) The jury found Jane guilty, didn’t they?

(38) The jury contained only white males, didn’t it?

\textsuperscript{118} I thank Matt Patience and Olivia Marasco for their help with judgments and intuitions about these contexts in English.
Although this part of the evidence does not prove that DP *el jurado* ‘the jury’ in (7a) is not the subject, the ambivalent behaviour that collective nouns show in the previous contexts may be contributing to the acceptability of the agreement mismatch.

Moreover, as M.C. Cuervo (p.c) suggests, the acceptability of plural agreement on the verb in (7a) may also be related to the fact that the AP *presionados unos y apocados los más* ‘under pressure some and fainthearted the majority’ that modifies the DP *el jurado* ‘the jury’ is plural. Evidence for this possibility can be found in (39) which, despite having an intervening element between the DP *el jurado* ‘the jury’ and the verb, is not fully grammatical.

(39) ??El jurado, después de tres días, declararon inocente al terrorista.

Finally, the intonational properties of (7a) in general, and in particular those of the DP and the AP may be playing an important role in the acceptability of the agreement mismatch in (7a) too. Since these features do not carry over to (33), which imposes its conditions (there is a wh-question after the DP), the sentence remains odd.

As with partial agreement, examples of *ad sensum* agreement do not constitute evidence that in Spanish there are differences in the agreement mechanism that depend on subject positions. It is not the case that a left-dislocated subject accepts *ad sensum* agreement, while a subject in the specifier of inflection rejects it, but that there are different elements that take part in the agreement operation in the two cases. In this instance, as well, the conclusion is that there is an agreement alternation is incorrect, as the two alternatives correspond to different underlying structures.

7.2.3. SE constructions

The last group of unexpected agreement patterns comprises sentences in (6), which correspond to what has been labelled as *SE* passive constructions. According to Martínez (1999), in these cases, the verb alternates between the two possibilities of agreement, singular and plural, when the plural DP is realized in postverbal position, as (5a). When
the DP moves to the preverbal position, he notes that the only possibility is plural agreement, (6c), as the ungrammaticality of (6b) demonstrates.

(6)  

a. Ahora ya no se suele/ suelen construir esos edificios.

   now anymore not SE use3S/use3P to build those buildings

b. *Esos edificios ya no se suele construir ahora.

   those buildings anymore notSE use3S to build now

c. Esos edificios ya no se suelen construir ahora.

   those buildings anymore not SE use3P to build now

   ‘Buildings like those are not built anymore nowadays.’

Martínez analyzes the previous sentences as a type of passive construction that uses SE as a marker. As in regular passive constructions, such as (40), in passive constructions with SE, agreement is expected. However, as (6a) shows, an alternation seems to be possible in this case.119

(40)  

Esos edificios fueron construidos el siglo pasado.

those buildings were built the century last

   ‘Those buildings were built last century.’

Similarly to cases that include two postverbal singular DPs and so-called partial agreement, the string in (6a), SE-Verb-DP, is ambiguous between a passive construction with SE, and an impersonal construction. The hallmarks of impersonal constructions in Spanish are the use of SE, and default third person singular agreement. Taking this into account, one possibility may be that sentences like (6a) do not represent an agreement alternation, but two different structures with the same linear order (the only difference

---

119 Martínez (1999) proposes that the alternation in (5a) arises due to a reanalysis by which the postverbal DP is reinterpreted as the direct object of the verb.
being the phi features that verb expresses, given the differences of the structures themselves).

The possibility that these sentences may correspond to different structures can be seen more clearly when we have an animate object, which requires the differential object marker (DOM) a in the case of impersonal constructions. It is important to note that in impersonal constructions the DP is never realized preverbally, i.e., they only allow one word order, as opposed to passive constructions with SE that allow both postverbal and preverbal realization of the DP.

(41) Aquí se recoge a los niños. IMPERSONAL

here SE pick up3S DOM the kids

‘Kids are picked up here.’

(42) Aquí se recogen los niños. PASSIVE with SE

here SE pick up3P the kids

‘Kids are picked up here.’

Supposing that the alternation is actually found in passive constructions with SE, we would expect that when these sentences include a plural animate DP, without DOM, they would allow the possibility of singular agreement. This, however, seems not to be the case. The following examples, which depict this possibility, are not accepted among native speakers.

(43) /*Aquí se recoge los niños.

here SE pick3S up the kids

(44) /*Ya no se llama los pasajeros antes de los vuelos.

anymore not SE call3S the passengers before of the flights

(45) /*Ahora no se educa los niños como antes.
now not SE teach the kids like before

Based on the behaviour of the previous examples, it is possible to argue that what we are seeing is an overlap among cases of SE passive constructions with postverbal subjects, and SE impersonal sentences that share the same string, and different phi features on the verb, as opposed to an alternation in agreement in passive constructions with SE.

Interestingly, impersonal constructions with SE do exhibit agreement alternations in Spanish. It has been claimed that agreement should never appear in impersonal constructions with direct objects that require the DOM a (Rigau 2008, p.801; Torrego 1998; Mendikoetxea, 1999, pp. 1637-38). Contrary to this, such sentences exist in Spanish. The following examples are taken from Martínez (1999), but similar cases can be found in regular speech, newspaper headlines, and written and spoken language in general.120, 121

(46) Se contratan a eventuales.

SE imp Olson hire DOM temporary workers

‘Temporary workers are hired (here).’ (Martínez, 1999, p. 2772, (201))

(47) Con estas obras se molestan a los peatones.

120 The following examples illustrate that this type of agreement is widespread in the language. (a) is a newspaper headline, and (b) is taken from a book.

(a) Esto es lo que se sabe del edificio colapsado en México
This is what SE know of the building collapsed in Mexico

‘This is what is known about the building that collapsed in Mexico in which 37 people are still missing’


(b) Cada año se contratan a miles de trabajadores jóvenes poco calificados.

‘Each year SE hire thousands of young under qualified workers are hired’

https://books.google.ca/books?id=UkWaAyHmBswC&pg=PA53&lpg=PA53&dq=%22se+contratan+a%22&source=bl&ots=YwOqP7GJWZ&sig=llzYMexYCqJtvA_OuAXsmuzZLk&hl=en&sa=X&ved=0ahUKEwiKwzXAhVM9JMKhfgeCLEQ6AEIMzAC#v=onepage&q=%22se%20contratan%20a%22&f=false

121 Ordóñez and Treviño (2011) also note that agreement in impersonal SE constructions between the verb and the object, even when this one is marked by the DOM a, is widespread in all Spanish dialects.
with these works $SE_{IMP} bother_{3P} \text{ DOM the pedestrians}$

‘Pedestrians are disturbed with these works.’/ ‘These works disturb pedestrians.’

(Martínez, 1999, p. 2772, (201))

(48) Hasta el momento se carecen de noticias.

until the moment $SE_{IMP} \text{ lack}_{3P} \text{ of news}$

‘There is no news so far.’/ ‘We have no news so far.’

(Martínez, 1999, p. 2772, (201))

A possible way to account for this alternation is to propose that impersonal constructions with SE have a null pronoun that lacks phi features, similar to the English $there$. In this scenario, such pronoun takes the phi features of the associate, and agreement between the verb and the non-subject DP takes place. Since both agreement with the associate and the default third person singular agreement are possible in these constructions, we need to postulate two types of null pronouns: one with phi features, the default third person, singular, and another that lacks phi features. If this analysis is correct, unlike to what has been claimed, the possibility of agreement alternation is not found in passive sentences with SE, but in impersonal sentences with SE and plural DPs.

Catalan presents us with a similar scenario. According to Rigau (2008), in Central Catalan, existential constructions show agreement between the verb and what she labels as its object, whereas in Ribagorçan, another Catalan dialect, the verb shows default third person singular agreement.

(49) a. Hi ha una cadira.  (Central Catalan)

cl. has a chair

‘There is a chair.’

b. Hi han tres cadires.
cl. have three chairs

‘There are three chairs.’

(Rigau, 2008, pp. 280-281, (11a-b))

(50)  a. Hi ha una cadira. (Ribagorçan Catalan)

cl. has a chair

‘There is a chair.’

b. Hi ha tres cadires.

cl. has three chairs

‘There are three chairs.’

(Rigau, 2008, p. 281, (12a-b))

Rigau analyzes existential impersonal constructions as light verb structures with a covert preposition of central coincidence, which is incorporated into the verb. This interrelational preposition has two arguments: a complement, the NP, and a subject, the locative clitic. Rigau proposes that this locative clitic works as an impersonalizer, and she labels it as a “quirky clitic subject”.

In her proposal, Romance quirky clitic subjects are overt or covert determiners with interpretable features and the locative property L, which have oblique clitic φ-features: $L_\{\text{person, number, . . .}\}$. Since they have this L property, they behave as locatives; they are “impersonal”, according to Rigau. She further proposes that “T can manifest the L property in its person feature or in both person and number features” (786). When the T matches the oblique D in both person and number features, the default third person singular agreement arises. On the other hand, if the number feature in T is not L-number feature, but α-number, where α-number has singular or plural value, number agreement with the complement arises.
Rigau extends this analysis to impersonal sentences with SE in Catalan, which display the same agreement patterns in these two dialects. In this case, it is SE the element that bears L φ-features. Since the alternations in Spanish are not found in two different dialects, but across all dialects, it is more accurate to propose that the two options exist in all dialects. Further differences between Rigau’s (2008) analysis and the one proposed in this work derive from the elements that participate in these constructions. However both proposals account for agreement variation without proposing differences in the agreement mechanisms, re-analyses of the elements that participate in the agreement operation, and differences that derive from the preverbal or postverbal positions that the elements occupy, which ultimately results in a better account of the syntax of these structures.

Other cases of variable agreement in Spanish also support the hypothesis of the existence of these two types of null pronouns. Existential constructions with haber (51), impersonal weather constructions with hacer (52), and other impersonal constructions such as (53) show the possibility of variable agreement.

(51) Habían/ había muchos animales.
    there.were/ there.was many animals
    ‘There were many animals.’

(52) Aquí hacen/ hace unos días preciosos en verano.
    here make3p/ make3p some days beautiful in summer
    ‘Days are beautiful in the summer here.’

(53) Van/ va ya dos horas que vino.
    go3p/go3S already two hours that came3s
    ‘It’s been two hours since he/she came.’

It is also possible to find existential constructions with haber that show person agreement different than the default third person. This syntactic behaviour is highly sanctioned by traditional normative grammars. Interestingly, some speakers claim that they are only
able to match number features, and not person features. To account for these cases, we may need to propose a fine-grained division of the phi features in the null pronoun, similar to the model proposed by Rigau (2008), and not a binary analysis such as the one suggested before.

(54) Habíamos cinco en la habitación.

there.were\textsubscript{1P} five in the room

‘There were five of us in the room.’

(55) Habemos cuatro esperando.

there.are\textsubscript{1P} four waiting

‘There are four of us waiting.’

(56) ¿Cuántos habíais en la fiesta?

how many there.were\textsubscript{2P} in the party

‘How many of you were at the party?’

Less frequently, it is also possible to find the polysynthetic form hay, ha+y(loc), pluralized.

(57) En el centro también hay cafés.

in the center also there.are cafes

‘There are also cafes in downtown.’

In all previous cases, it has been suggested that there is a reanalysis, and, as a result, the postverbal DP is interpreted as the grammatical subject. If we adopt an analysis that incorporates a null pronoun that lacks phi features (or some other option along these lines), the previous alternations in agreement patterns simply follow. This analysis is preferable since it does not rely on a reanalysis of the elements, which is a stipulative solution.
7.3. Final comments

This chapter examined different contexts that have been described as cases of unexpected agreement patterns or variable agreement in Spanish. A more systematic analysis revealed that such variability is only apparent. More precisely, the idea that variability in agreement patterns exists in Spanish develops from the comparison of sentences that share the same string, except for verbs’ phi features, but whose syntactic structures are different. As we saw, in all cases, the so-called variability in agreement patterns can be explained by independent reasons, all of them unrelated to the operation of agreement (Agree) itself.

This is a welcome result because it corroborates the theoretical assumption that Agree is a universal operation whose results depend on the elements that take part in it. Additionally, this analysis supports the hypothesis that language variation in the functional domain depends on both the way in which open values associated with functional categories receive range (Borer, 2005b), and the inventory of functional elements that languages have and their specific properties. The latter accounts for cases of agreement in impersonal constructions with the non-subject DP in Spanish. By postulating the existence of two versions of a null pronoun, one with the default third person singular phi features, and another one that lacks phi features all together (or something similar to this), we can account for the so-called variability in agreement patterns in all impersonal constructions in Spanish without the need to call for a reanalysis of these structures.
Chapter 8
Conclusion

This study re-examined the distribution of preverbal and postverbal subjects in Spanish, and proposed a new analysis, built upon Borer’s (2005) Event Phrase proposal, that allows for a unified treatment of broad focus structures in this language. Additionally, this work provides new prosodic data on preverbal left-dislocated old information topics that adds to the description and characterization of informationally marked elements in Spanish.

In this work, I employed a combination of theoretical and experimental approaches as a means to look for empirical evidence to disambiguate the syntactic status of preverbal subjects. Taking into account the two contrasting analyses about the position that preverbal subjects occupy in Spanish (the left periphery and the inflectional layer), as well as the relationship between syntax and phonology, and the prosodic properties of Spanish, I hypothesized that preverbal subjects should resemble the prosodic realization of left dislocated elements, if they appear in the left periphery. If, on the other hand, preverbal subjects appear in the inflectional layer, they should not pattern with left dislocated elements. These two possibilities of subject realization were assessed in two contexts of information structure, broad focus and topicalization, with respect to the prosodic features of duration, fundamental frequency, and intensity. The prosodic analysis employs data collected using a controlled elicitation task from 10 female speakers from the Spanish varieties spoken in Havana and Madrid.

The results showed that there are prosodic differences in duration and, for some speakers, in fundamental frequency, between elements that are old information and appear in the left periphery, and elements that are new information and appear in broad focus structures, as the following pairs illustrate. Old information is underlined.

\[
\begin{align*}
\text{SVO} & \neq \text{SVO} \\
\text{SVO} & \neq \text{CLLDVO}
\end{align*}
\]
Thus, I suggested that the distinct prosodic behaviour that subjects in the two conditions of information structure display is compatible with the syntactic analysis in which preverbal subjects in broad focus structures appear in the inflectional layer in Spanish.

The prosodic behaviour of preverbal subjects of unergative and unaccusative verbs in broad focus contexts was also examined. In this case, no prosodic differences were found between the two subject types. This result corroborates Zubizarreta’s (1998) claim regarding the prosodic properties of these constructions in Spanish, and potentially constitutes a problem for mechanisms that generate different sentential stress patterns for the two constructions.

Although the IP proposal can accommodate the experimental findings, I noted that there were some problems with the arguments about movement to the preverbal field that some of the analyses had suggested. Specifically, I argued that the idea that predication was the trigger of the movement to the preverbal field (Masullo, 1992; Cuervo, 1999, 2003, 2010; Ortega-Santos, 2008) could not be maintained for two reasons. First, it restricts the interpretation of broad focus statements to categorical judgements about the element that moves to the preverbal position, which leaves out the event-related interpretation that any broad focus sentence carries. Second, such a predication requirement is not attested in all broad focus contexts. For instance, various types of impersonal constructions would be an exception to this requirement, as there is no predication in such cases.

Taking into account the interpretational properties of broad focus structures and differences in subject distribution in the same contexts, I proposed a new analysis, based upon Borer’s (2005b) Event Phrase (EP). In this analysis, the specifier of EP is the locus of preverbal subject realization, and the need to license the event by binding the event variable is the trigger of the movement to this preverbal position. This proposal is compatible with the experimental findings, as the EP is located below CP.

Since specEP can be filled by nominative and non-nominative phrases, this analysis has the advantage of accommodating a wider range of word orders attested in broad focus contexts in Spanish. Other properties of Spanish constructions like the distribution and
interpretation of strong DPs and locatives also show a correlation with the features of EP and the element that moves into its specifier.

Additionally, this study offers a structure for stative events, based upon Borer’s (2005b) insights, that accommodates the three types of psychological predicates in Spanish regardless of the case of their arguments. The rationale for proposing a special structure for statives is based on differences in the interpretation of DPs and NPs, and the restrictions on the distribution of the latter, which have been correlated in Spanish with structural positions (Suñer, 1982; Torrego, 1989; Masullo, 1992; Cuervo, 2011). I argued that only by taking into account DP/ NP interpretation, distinctive patterns of stative constructions emerge, which ultimately makes a unified treatment of these structures possible.

This study also examined the operations of nominative case checking/ assignment and agreement since these are linked to the subject. With respect to nominative case, the derivations suggested in this work are compatible with the theoretical assumption that the case-checking/ assignment operation, unlike agreement, takes place in a spec-head configuration (see Bošković (2007), and references therein), as the specifier of IP is available to host nominative subjects. Regarding agreement, I argued, through the examination of so-called variable agreement contexts, that this variability in agreement is only apparent, as each instance corresponds to a different underlying structure. I concluded that Spanish data does not support the hypothesis that there may be differences in the output of agreement that depend on whether it takes place in a spec-head configuration or via long distance, contrary to what has been explicitly or implicitly assumed. Thus, Spanish data indirectly bolsters the minimalist hypothesis that Agree should be the single option of this operation in the grammar (Chomsky, 2000; Boeckx, 2008; among others).

The analysis developed in this work departs from previous studies in two crucial aspects. First, it adopts a composite understanding of subjecthood (Harley, 1995; McCloskey, 1997). Second, it incorporates the notion of argument structure, i.e., the different word orders that derive from the different argumental frames are considered equally possible, as long as they correspond to a broad focus interpretation. The result of the combination
of these two features is a less rigid approach to the notion of default word order, i.e., SVO is not assumed as the default option against which the rest of the constructions must be compared. By adopting a more general approach with respect to word order differences one can take a step back and look for the common element that underlies broad focus constructions. In doing so, I focused on the event-related interpretation that these constructions carry, and proposed an analysis that not only accommodates a larger number of cases, but also provides a unified treatment of broad focus constructions.

However, even though the current proposal differs from earlier studies about Spanish subjects in the aforementioned ways, it includes many of their insights. In a sense, the analysis developed here represents the integration of the previous research on Spanish subjects and their properties into one proposal. The current analysis retains the idea that some non-nominative phrases, such as dative experiencers and locatives, have subject-like properties and behave as nominative subjects in many syntactic contexts, which has been a recurrent argument of several studies (Masullo, 1992; Cuervo, 1999, 2003; Fernández-Soriano, 1999a, 1999b). It also preserves the idea that in some structures the presence of locatives in the preverbal position is crucial to the possibility of licensing such constructions (Torrego, 1989). Additionally, it maintains that covert locatives are part of some constructions that include unccusative verbs (Torrego, 1989; Ortega-Santos, 2008). Lastly, generalizations about the distribution and interpretation of NPs and DPs, which have been correlated to structural positions in Spanish (Suñer, 1982; Torrego, 1989; Casielles-Suárez, 2004; Cuervo, 2010), are a central part of the arguments used to develop and propose the structure of psychological predicates as stative events (as opposed to eventive events).

Finally, in a different vein, but equally important, the adoption of two approaches—experimental and theoretical—in this work is another significant element, as it expands the empirical base of our understanding of subject properties in the language. Furthermore, the examination of the data provided within the syntax-phonology interface directly contributes to the theoretical discussion about topics related to the relationship between the two modules. Specifically, the experimental findings of this work raise two important issues: 1) the kind of information that should be available at the syntax-
phonology interface, and 2) the type of syntactic information that is prosodically represented.

With respect to the first question it was suggested that the notions of topic and focus may not be appropriate to characterize informationally marked elements, and distinctions like old versus new, [+- contrastive] may be more adequate, in line with López’s (2009) proposal. Empirical support for the idea that either the notions of topic and focus need further characterization or they should be abandoned comes from the fact that attempting to generate the stress pattern of a sentence with a left dislocated contrastive topic would be problematic given that, unlike focus, topicalization is not included in the mechanisms that generate sentential stress.

An additional consideration is that sentential stress is not the only relevant prosodic property that sentences bear. Specifically, the findings of this work show that preverbal elements show prosodic differences in terms of their duration when they bear different informational properties, e.g., old versus new. Thus, even though old information may not influence the computation of sentential stress, elements that bear this property are prosodically distinct from elements that are new information.

To solve this issue we could entertain the possibility that features like [old] need to be visible at the interface, even if they are not related to the computation of sentential stress. However, things are more complex, as depending on the position (prefocal versus postfocal) that elements occupy, differences like old versus new may not be prosodically realized. Specifically, elements that are old information and occupy a postfocal position have been found to be prosodically indistinguishable from the same element in the same position of the string in a broad focus context (Féry & Kügler, 2008; Kügler, 2008). It remains to be tested if this is also the case in Spanish, as the previous studies refer to German. In any case, we are faced with the problem of what information should be included at the syntax-phonology interface, and in what cases such information is prosodically realized.

Lastly, there is a generalized assumption in the syntax that old information is unstressed; however, it is not clear how this unstressed property manifests itself or should be understood. As stated, old information topics differ from new information elements
mainly in their duration. Thus, we need a better characterization or a more appropriate description, as the notion of unstressed seems to refer to differences in fundamental frequency, such as differences in pitch, specific pitch accents or lack of pitch accents, none of them suitable to describe the data analyzed in this work.

The issue about what type of syntactic information—aside from information structure—is prosodically encoded is equally complex. For instance, differences in stress patterns between English and Spanish in the contexts of unaccusative and unergative constructions have been a problem for theories of sentential stress, because these constructions display different stress patterns in the two languages (Zubizarreta, 1998). The data analyzed in this work confirm Zubizarreta’s claims, as no prosodic differences between subjects of unaccusative and unergative verbs were found. A question that still remains is what prosody is possibly marking in English by distinguishing between these two types of subjects? A hypothesis suggested in this work is that prosodic differences between these subjects in English are the result of the stress assigning mechanism. By taking this approach, one can reanalyze such prosodic differences as a by-product. Taking this into account, one has to be careful when linking prosodic features to syntactic structures/features, as some instances may represent just correlations or a by-product of a different phenomenon. Finally, our data also exemplifies how a one-to-one correspondence between syntactic structure and prosody fails. Given the differences in phrasing patterns that similar constructions display in similar environments, it was concluded that we should not expect the same prosodic behaviour across instances whose syntactic structures may be comparable. Thus, our data supports the idea that a direct correspondence between syntactic structure and prosody seems unattainable. One of the possible reasons is that prosodic structure is generally simpler than syntactic structure (Ladd, 2008). Although more research needs to be done regarding these topics, it is my hope that this study about Spanish subjects contributes to the dialogue about these and other issues, and opens new research avenues.
References


Kügler, F. (2008). The role of duration as a phonetic correlate of focus. Handout of presentation at Speech Prosody, May 6-9, Campinas, Brazil.


Appendix A

Stimuli

1) Preverbal non-nominative subjects

A Juan le gusta el dinero.
A Juan le gusta el dinero.
A Juan le gasta el dinero.

Al hombre le encantó el cuadro.
Al hombre le encantó el cuadro.
Al hombre le encontró el cuadro.

A Luis le importa el amor.
A Luis le importa el amor.
A Luis le inculca el amor.

A Manuel le molestan los perros.
A Manuel le molestan los perros.
A Rubén le amaestran el perro.

Al niño le duelen las muelas.
Al niño le duelen las muelas.
Al niño le miran las muelas.

A Julia le crece el cabello.
A Julia le crece el cabello.
A Julia le ofrece un camello.

A Juan le falta el llavero.
A Juan le falta el llavero.
A Juan le compra el llavero.

A la mesa le sobra una pata.
A la mesa le sobra una pata.
A la mesa le suena una pata.

Al árbol le salen las hojas
Al árbol le salen las hojas.
Al árbol le suenan las hojas.

2) Preverbal subjects of transitive verbs

Un hombre revende dulces.
El hombre revende dulces.
Al hombre le vende dulces.

El empleado midió la tela.
El empleado midió la tela.
Al empleado le dio la tela

Amanda pintó la casa.
Amanda pintó la casa.
A Juan le pintó la casa.
Graciela viró las maletas.
Graciela viró las maletas.
A Julia le abrió las maletas.

Laurita remodela la casa.
Laurita remodela la casa.
A Rita le cuidaba la casa.

El chico bebió ayer.
El chico bebió ayer.
El coche lo vio ayer.

La niña repasa mucho.
La niña repasa mucho.
A Dora la pesa mucho.

Los barcos entrenan temprano.
Los barcos entrenan temprano.
Los barcos los guardan temprano.

Marina bailó ayer.
Marina bailó ayer.
A Lina la vio ayer.

Javier saludó también.
Javier saludó también.
A Juan lo llamó también.

Manuela corre mucho.
Manuela corre mucho.
A Mou lo quiere mucho.

La niña estudia mucho.
La niña estudia mucho.
A Mou lo aprecia mucho.

El nieto de Manuela estudia mucho.
El nieto de Manuela estudia mucho.
Al nieto de René lo llama mucho.

El hijo de Manuela besa mucho.
El hijo de Manuela besa mucho.
Al hijo de René lo besa mucho.

3) Preverbal subjects of unaccusative and unergative verbs

Ana María llegó.
Ana María bailó.

Los estudiantes de Malta vinieron.
Los estudiantes de Marta leyeron.

Iván Bustamante salió.
Iván Bustamante comió.
Un barco grande entró.
Un barco grande ancló.

Juan Marcel creció.
Juan Marcel comió.

María Luisa apareció.
María Luisa escribió.

Los estudiantes nadaron.
Los estudiantes vinieron.

La prima de riesgo cayó.
La prima de Luisa lloró.

El niño de María creció mucho.
El niño de María copió mucho.