The Morphosyntax of Wh-questions: Evidence from Spanish-English Code-switching

BY

SHANE EBERT
B.A., Beloit College, 2002
M.A., University of Illinois at Chicago, Chicago, 2008

THESIS

Submitted as partial fulfillment of the requirements for the degree of Doctor of Philosophy in Hispanic Studies in the Graduate College of the University of Illinois at Chicago, 2014

Chicago, Illinois

Defense Committee:

Kay González-Vilbazo, Chair and Advisor
Richard Cameron
Luis López
Susanne Rott
Almeida Jacqueline Toribio, The University of Texas, Austin
ACKNOWLEDGEMENTS

It is difficult to trace the many sources of inspiration and support that have contributed to this dissertation, but I would like to acknowledge a few of them.

First and foremost, I am very grateful to my advisor, Kay González-Vilbazo, not only for his support and guidance throughout the writing of my dissertation but also for the many ways he has helped me to grow as a scholar.

I am also grateful to the other members of my dissertation committee: Richard Cameron, Luis López, Susanne Rott, and Almeida Jacqueline Toribio. Their advice and feedback were crucial in shaping my dissertation.

I would also like to thank the current and former members of the Bilingualism Research Laboratory (BRL), both for all the hours of discussion on various aspects of my dissertation throughout the design, data collection and analysis stages as well as the equally important experience gained from being a member of the lab. Together we have collaborated on many experiments and discussed many aspects of experimental research, bilingualism and code-switching, and theoretical linguistics.

With respect to data collection, I would like to thank all the BRL research assistants who helped with the various stages of data collection and the UIC Spanish for Bilinguals program for help in participant recruitment. I am particularly grateful to Bryan Koronkiewicz for all his help while collaborating on our many rounds of data collection, from small practical questions to important methodological ones.

Finally, I want to thank my friends and family for their love and support throughout my whole graduate career.
# Table of Contents

1 **Introduction** .......................... 1  
  1.1 Research Questions ............................ 4  
  1.2 Dissertation Overview ............................ 6  

2 **Assumptions and Theoretical Framework** .......................... 9  
  2.1 Object of Study ............................. 9  
  2.2 Structural Assumptions ............................ 10  
  2.3 Code-Switching Framework ........................... 13  
  2.4 Role of Monolingual Judgments ............................ 15  
  2.5 Premises for Answering Research Questions ............................ 17  
  2.6 Dialect, Context and Linguistic Knowledge ............................ 19

3 **Literature Review** .......................... 23  
  3.1 Wh-Questions ............................... 25  
    3.1.1 Descriptive Generalizations ............................ 25  
    3.1.2 Inversion in Spanish Wh-Questions ............................ 27  
      3.1.2.1 T-to-C Movement in Spanish? ............................ 28  
      3.1.2.2 Preverbal Subjects in Spanish ............................ 29  
      3.1.2.3 The Trigger for Inversion ............................ 32  
    3.1.3 Inversion in English Wh-Questions ............................ 33  
      3.1.3.1 Rizzi (1996) ............................ 33  
      3.1.3.2 Pesetksy and Torrego (2001 et seq.) ............................ 35  
      3.1.3.3 Conclusions ............................ 41  
    3.1.4 The Role of the Wh-Phrase ............................ 42  
      3.1.4.1 Syntactic Function of the Wh-Phrase ............................ 42  
      3.1.4.2 D-Linked and Non-D-Linked Wh-Phrases ............................ 43  
      3.1.4.3 Simple and Complex Wh-Phrases ............................ 45  
    3.1.5 Conclusions ............................ 49  
  3.2 Code-Switching .............................. 49  
    3.2.1 Early Formal Approaches to Code-Switching ............................ 50  
      3.2.1.1 Poplack (1980) ............................ 50  
      3.2.1.2 Joshi (1985) ............................ 52  
    3.2.2 Generative Approaches to Code-Switching ............................ 53  
      3.2.2.1 Woolford (1983, 1984) ............................ 53  
      3.2.2.2 Di Sciullo, Muysken, and Singh (1986) ............................ 55  
      3.2.2.3 Belazi, Rubin and Toribio (1994) ............................ 56  
      3.2.2.4 Mahootian and Santorini (1996) ............................ 59  
    3.2.3 Minimalist Approaches to Code-Switching ............................ 60  
      3.2.3.1 MacSwan (1999) ............................ 60  
      3.2.3.2 Toribio and González-Vilbazo (Forthcoming) ............................ 61  
      3.2.3.3 González-Vilbazo and López (2011, 2012) ............................ 63  
    3.2.4 Conclusions ............................ 64  
  3.3 Hypotheses ............................... 65
4 Methodology

4.1 Pilot Experiments ........................................ 70
4.2 Stimuli Design ........................................... 72
  4.2.1 General Considerations ............................. 72
    4.2.1.1 Limiting Stimuli Number ......................... 72
    4.2.1.2 Design Factors .................................. 73
    4.2.1.3 Lexicalizations .................................. 76
  4.2.2 Code-switching Stimuli Design ...................... 78
  4.2.3 Monolingual Stimuli Design ......................... 81
4.3 Participant Selection ..................................... 83
4.4 Materials .................................................. 85
  4.4.1 Consent and Background Questionnaire .............. 86
  4.4.2 Acceptability Judgment Training .................... 87
  4.4.3 Distractor Tasks .................................... 87
  4.4.4 Code-switching Stimuli ............................ 88
  4.4.5 Monolingual Stimuli ................................ 89
  4.4.6 Proficiency Tasks .................................. 89
4.5 Procedure ................................................ 90
  4.5.1 General Considerations ............................. 90
    4.5.1.1 Type of Experimental Task ....................... 90
    4.5.1.2 Types Acceptability Judgment Tasks ............. 92
    4.5.1.3 Confounds in Collecting Code-Switching Data .... 94
  4.5.2 Participant Recruitment ............................. 95
  4.5.3 Experimental Procedure ............................ 95
4.6 Methods of Analysis ..................................... 96
  4.6.1 Final Participant Selection ......................... 96
    4.6.1.1 Background and Proficiency ...................... 96
    4.6.1.2 Data Characteristics ............................ 97
    4.6.1.3 Participant Characteristics ..................... 98
    4.6.1.4 Participant Groups .............................. 99
  4.6.2 Subject Language ................................... 103
  4.6.3 Lexicalization Analysis ............................ 104
  4.6.4 Items Analysis ..................................... 105
  4.6.5 Interpreting Standard Deviations ................... 105

5 Discussion ............................................... 107

5.1 Factors Contributing to Word Order ..................... 109
  5.1.1 Subject ............................................. 110
  5.1.2 Wh-Phrase ......................................... 114
  5.1.3 Tense Head ......................................... 116
  5.1.4 Complementizer Head ............................... 120
    5.1.4.1 C as a Determinant Head ......................... 121
    5.1.4.2 Grammatical Subject Positions and the Language of C 122
    5.1.4.3 Additional Evidence for C as the Determinant Head 124
  5.1.5 Conclusions ......................................... 127
5.2 Generalizations Based on the Language of C ............ 128
  5.2.1 The Language of C .................................. 129
5.2.2 Code-Switching Between C and T .................................................. 137
5.2.3 Code-Switching Between C and the Wh-Phrase .................................. 138
5.3 Empirical Generalizations .............................................................. 143
5.4 Implications .............................................................................. 145
  5.4.1 Wh-Questions in Spanish and English ........................................ 145
    5.4.1.1 Wh-Questions in English .................................................. 145
    5.4.1.2 T-Based Accounts of Spanish Inversion ............................. 147
    5.4.1.3 Other Accounts of Spanish Wh-questions ........................... 151
  5.4.2 Wh-Questions More Generally .................................................. 153
  5.4.3 Wh-Questions in Code-Switching .............................................. 154
  5.4.4 Code-Switching More Generally .................................................. 156
    5.4.4.1 Code-Switching as a Tool of Linguistic Analysis ................. 157
    5.4.4.2 Theoretical Considerations .............................................. 157
    5.4.4.3 Empirical Tests for the Code-switching Framework ............ 158
  5.5 Summary and Conclusions ......................................................... 163

CITED LITERATURE .................................................................................. 166

APPENDIX ............................................................................................. 172

VITA .................................................................................................... 199
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>PARTICIPANT BACKGROUND OVERVIEW</td>
<td>99</td>
</tr>
<tr>
<td>II</td>
<td>THE EFFECT OF RATINGS ON STANDARD DEVIATION</td>
<td>106</td>
</tr>
<tr>
<td>III</td>
<td>LANGUAGE OF C: OVERALL GENERALIZATIONS</td>
<td>130</td>
</tr>
<tr>
<td>IV</td>
<td>C≠T WITH CLEAR RATINGS (EMBEDDED SIMPLE WH-PHRASES)</td>
<td>131</td>
</tr>
<tr>
<td>V</td>
<td>C≠T WITH MIDDLE RATINGS (EMBEDDED SIMPLE WH-PHRASES)</td>
<td>132</td>
</tr>
<tr>
<td>VI</td>
<td>C=T AND C≠T BOTH GRAMMATICAL (EMBEDDED SIMPLE WH-PHRASES)</td>
<td>133</td>
</tr>
<tr>
<td>VII</td>
<td>EXCEPTIONS TO OVERALL GENERALIZATIONS (GROUP 3)</td>
<td>134</td>
</tr>
<tr>
<td>VIII</td>
<td>EXCEPTIONS TO OVERALL GENERALIZATIONS (GROUP 5)</td>
<td>136</td>
</tr>
<tr>
<td>IX</td>
<td>LICIT CODE-SWITCHES BETWEEN C AND T</td>
<td>144</td>
</tr>
<tr>
<td>X</td>
<td>LICIT CODE-SWITCHES BETWEEN C AND THE WH-PHRASE</td>
<td>145</td>
</tr>
<tr>
<td>XI</td>
<td>CODE-SWITCHING RATINGS PARALLEL MONOLINGUAL RATINGS (GROUP 1)</td>
<td>161</td>
</tr>
<tr>
<td>XII</td>
<td>SUBSET OF POSSIBLE SUBJECT POSITIONS IS GRAMMATICAL (ENGLISH T)</td>
<td>163</td>
</tr>
<tr>
<td>XIII</td>
<td>MONOLINGUAL BASE SENTENCES</td>
<td>186</td>
</tr>
<tr>
<td>XIV</td>
<td>CODE-SWITCHING LEXICAL UNIT 2 SENTENCES</td>
<td>187</td>
</tr>
<tr>
<td>XV</td>
<td>MONOLINGUAL LEXICAL UNIT 2 SENTENCES</td>
<td>188</td>
</tr>
<tr>
<td>XVI</td>
<td>ADDITIONAL MONOLINGUAL SPANISH STIMULI</td>
<td>189</td>
</tr>
<tr>
<td>XVII</td>
<td>ADDITIONAL MONOLINGUAL ENGLISH STIMULI</td>
<td>190</td>
</tr>
<tr>
<td>XVIII</td>
<td>GROUP 1 ACCEPTABILITY JUDGMENTS</td>
<td>192</td>
</tr>
<tr>
<td>XIX</td>
<td>GROUP 2 ACCEPTABILITY JUDGMENTS</td>
<td>193</td>
</tr>
<tr>
<td>XX</td>
<td>GROUP 3 ACCEPTABILITY JUDGMENTS</td>
<td>194</td>
</tr>
<tr>
<td>XXI</td>
<td>GROUP 4 ACCEPTABILITY JUDGMENTS</td>
<td>195</td>
</tr>
<tr>
<td>XXII</td>
<td>GROUP 5 ACCEPTABILITY JUDGMENTS</td>
<td>196</td>
</tr>
<tr>
<td>XXIII</td>
<td>GROUP 6 ACCEPTABILITY JUDGMENTS</td>
<td>197</td>
</tr>
<tr>
<td>XXIV</td>
<td>GROUP 7 ACCEPTABILITY JUDGMENTS</td>
<td>198</td>
</tr>
</tbody>
</table>
List of Figures

1. Simple wh-phrase ............................................. 47
2. Complex wh-phrase ........................................... 47
3. Swiping example ............................................. 48
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>complementizer</td>
</tr>
<tr>
<td>CS</td>
<td>code-switching</td>
</tr>
<tr>
<td>D-linked</td>
<td>discourse-linked</td>
</tr>
<tr>
<td>DELE</td>
<td>Diploma of Spanish as a Foreign Language (from the Spanish abbreviation)</td>
</tr>
<tr>
<td>DP</td>
<td>determiner phrase</td>
</tr>
<tr>
<td>EL</td>
<td>Embedded Language</td>
</tr>
<tr>
<td>ERP</td>
<td>event-related potentials</td>
</tr>
<tr>
<td>FHC</td>
<td>Functional Head Constraint</td>
</tr>
<tr>
<td>FL</td>
<td>faculty of language</td>
</tr>
<tr>
<td>fMRI</td>
<td>functional magnetic resonance imaging</td>
</tr>
<tr>
<td>LF</td>
<td>logical form</td>
</tr>
<tr>
<td>ML</td>
<td>Matrix Language</td>
</tr>
<tr>
<td>MLF</td>
<td>Matrix Language-Frame</td>
</tr>
<tr>
<td>MP</td>
<td>Minimalist Program</td>
</tr>
<tr>
<td>NP</td>
<td>noun phrase</td>
</tr>
<tr>
<td>PF</td>
<td>Phonetic Form</td>
</tr>
<tr>
<td>PHH</td>
<td>Phase Head Hypothesis</td>
</tr>
<tr>
<td>PLD</td>
<td>primary linguistic data</td>
</tr>
<tr>
<td>SpecCP</td>
<td>specifier of the complementizer phrase</td>
</tr>
<tr>
<td>SpecTP</td>
<td>specifier of the tense phrase</td>
</tr>
<tr>
<td>SpecvP</td>
<td>specifier of the little v phrase</td>
</tr>
<tr>
<td>SV</td>
<td>subject-verb</td>
</tr>
<tr>
<td>T</td>
<td>tense</td>
</tr>
<tr>
<td>TP</td>
<td>tense phrase</td>
</tr>
<tr>
<td>UG</td>
<td>Universal Grammar</td>
</tr>
<tr>
<td>V</td>
<td>verb</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------</td>
</tr>
<tr>
<td>V2</td>
<td>verb-second</td>
</tr>
<tr>
<td>VP</td>
<td>verb phrase</td>
</tr>
<tr>
<td>vP</td>
<td>little v phrase</td>
</tr>
<tr>
<td>VS</td>
<td>verb-subject</td>
</tr>
</tbody>
</table>
SUMMARY

A growing body of research stretching back to the 1970’s has shown there to be systematic restrictions on when bilinguals consider it permissible to switch from one language to another within the same sentence, in a process known as code-switching (CS). Formal research on these restrictions has largely focused on accounting for these restrictions by drawing on current linguistic theory. Turning this process on its head, this dissertation instead draws on the systematic nature of these restrictions to provide evidence concerning the properties of the individual languages as well as human language more generally. CS allows one to independently control the language of individual words, and their associated syntactic properties, and observe the effect that those properties have on the acceptability of a given sentence. In this way, CS becomes a tool of linguistic analysis. This makes CS particularly well-suited to disentangling language-specific properties from more general, cross-linguistic properties of syntactic phenomena. Wh-questions are just such a phenomenon. There are a variety of parallels in the formation of wh-questions across languages, yet their specific syntactic structure varies greatly by language.

This dissertation draws its data from Spanish-English CS. Both Spanish and English exhibit a phenomenon known as subject-version, in which the subject is sometimes required to appear after the verb or auxiliary, but the structural requirements vary by language and include a number of different factors. This raises two important questions that are the focus of this dissertation: (i) what head or phrase ultimately determines the word order (i.e. the grammatical subject positions) in a given wh-question? and (ii) what are the restrictions on code-switching between the complementizer head, the tense head, and the wh-phrase?

To answer these questions, a controlled experiment was conducted in which participants provided written acceptability judgments for both CS and equivalent monolingual sentences.

In response to the first question, the results of this dissertation provide evidence that the complementizer is ultimately responsible for determining the properties of inversion for a given wh-question. Looking at wh-questions in Spanish in particular, these results provide evidence against the hypothesis that properties of the Spanish T head are fully responsible for the inversion requirements in Spanish (Toribio 1993, Barbosa 2001, Goodall 2001, Gutiérrez-Bravo 2005, Zubizarreta 2012 inter alia).
SUMMARY (Continued)

Regarding the second question, the results also provides evidence for restrictions on code-switching between C and T and C and the wh-phrase. In particular, it provides evidence that a code-switch between an English simple wh-phrase and a Spanish C is ungrammatical in embedded questions with a Spanish T. This suggests a previously unknown syntactic relationship between embedded Spanish complementizers and simple wh-phrases.

Turning to more general implications for code-switching, this dissertation supports a broader hypothesis according to which phase heads, including the complementizer, determine the grammatical properties of their complement, as proposed in González-Vilbazo and López (2012). In addition, both the results and the approach to CS from this dissertation extend a line of research (e.g. Belazi et al. 1994; Mahootian and Santorini 1996; González-Vilbazo 2005; González-Vilbazo and López 2011, 2012) in which CS can tease apart the role of individual constituents and their associated properties in a range of syntactic phenomena, thereby providing a unique source of evidence concerning fundamental properties of the human language faculty. Finally, this dissertation serves as one example of an experimental approach to investigating code-switching, particularly with respect to syntactic structure. There are a variety of different concerns that need to be addressed in conducting such research, and this dissertation offers one perspective.

In conclusion, the combination of controlled experimental methods and the unique analytical potential of CS to tease apart otherwise opaque syntactic relationships makes for a valuable tool for addressing a wide range of questions in theoretical linguistics.
1 Introduction

This dissertation takes as its starting point the assumption that linguistic structure is the result of one overarching set of principles applied across languages. This assumption has profound implications when two languages are combined in the same sentence, both analytical and empirical consequences. First, this assumption implies that data from combining two languages, known as code-switching, bears on those overarching principles, providing an additional type of empirical data concerning the structure of language that goes beyond the more commonly-used monolingual data. Importantly, most work on code-switching (CS) that is interested in those overarching principles characterizes CS data in terms of these principles rather than using the CS data to inform our understanding of the principles. At the analytical level, CS allows one to independently control the language of individual words, and their associated syntactic properties, and observe the effect that those properties have on the acceptability of a given sentence. In this way, CS becomes a tool of linguistic analysis.

These unique properties of CS are particularly well-suited to disentangling language-specific properties from more general, cross-linguistic properties of syntactic phenomena. Wh-questions are just such a phenomenon. There are a variety of parallels in the formation of wh-questions across languages, yet their specific syntactic structure varies greatly by language, suggesting some underlying, universal mechanisms interacting with language-specific properties, i.e. features.

To see the sorts of structural relations and features that make wh-questions a fertile ground for analysis, consider some properties of English wh-questions. For example, in English, in addition to the movement of the wh-phrase to the front of the clause, the tense head, T, and the auxiliary that it hosts raise above the subject to the complementizer head, C, creating an inversion between the subject and the auxiliary. This is illustrated in the following example in (1a) and the simplified representation of that sentence in (1b).

(1) a. What has John bought this week?          auxiliary-subject-participle
b. \([\text{CP What } \text{has John} \text{ bought this week}]\)?

We see, then, that the relationship between C and T is an important factor in the formation of wh-questions in English. As with any syntactic process, one can ask what drives the process of inversion. Given the central role that features play in the approach to syntax outlined above, we would expect that some feature causes T to raise to C, producing inversion. At first glance, that would suggest that this feature is enough to explain the presence of inversion in English, but it turns out that inversion only occurs in some structures, indicating that there may be more than this feature involved. Consider the following pair of sentences:

(2)  a. *I don’t know [what \text{has John bought this week}] auxiliary-subject

b. I don’t know [what \text{John has bought this week}]. subject-auxiliary

Both of these sentences have an embedded clause which is a wh-question, indicated with square brackets. In contrast to the matrix wh-question in (1a), the embedded wh-question in (2a) is ungrammatical with inversion but perfectly grammatical without it, as in (2b). In other words, there is an asymmetry in English between matrix and embedded wh-questions. The process of inversion, then, must be mediated by more than just a single, simple interrogative C.

Spanish also exhibits a type of inversion, though we will see that determining the underlying process that derives Spanish inversion is, in some ways, more complicated than in English. For example, unlike in English inversion, an inverted subject in Spanish appears after the entire verb complex, whether that consists of a simple finite verb, as in (3a), or an auxiliary and a participle as in (4a).

(3)  a. ¿Qué compró Juan ayer?
    \(\text{what bought Juan yesterday}
    \text{} \quad \text{verb-subject}
    \text{‘What did Juan buy yesterday?’}

b. *¿Qué Juan compró ayer?
    \(\text{subject-verb}

(4)  a. ¿Qué has comprado Juan esta semana?
    \(\text{what has bought Juan this week}
    \text{} \quad \text{auxiliary-participle-subject}
    \text{‘What has Juan bought this week?’}

1Note that this does vary by dialect and by context. See §2.6 for additional discussion.
As will be discussed in §3.1.2, there is evidence that inversion in Spanish also stems from properties of C and of T, but in the case of Spanish, inversion occurs because properties of C and T force the subject to remain in situ rather than requiring T to raise to C. This once again highlights the importance of C and T as factors in the formation of wh-questions. Spanish also offers evidence that another factor, the syntactic function of the wh-phrase, plays a role in the formation of wh-questions. Consider the contrast between a question where the wh-phrase is an argument, such as ‘what’ in the previous examples, and a parallel question with a wh-phrase that is an adjunct, as in (5a) and (5b).

(5) a. ¿Por qué compró María ese libro ayer?  
   ‘Why did Mary buy that book yesterday?’

b. ¿Por qué María compró ese libro ayer?  
   ‘Why did Mary buy that book yesterday?’

When the wh-phrase has the syntactic function of an adjunct, inversion becomes optional, so both the typical verb-subject word order in (5a) and the subject-verb word order in (5b) are grammatical. Essentially, there is an asymmetry between wh-arguments and wh-adjuncts in Spanish. Thus the syntactic function of the wh-phrase is another piece of the puzzle for determining when inversion is available in a given wh-question.

We have seen, then, that the syntactic structure of wh-questions depends on the complex interaction of a number of different factors, including a few more factors that will be discussed in §3.1. And in the framework adopted here, each of these factors is, at least in part, a consequence of the syntactic features involved. This raises two questions. First, what is it about these factors and the syntax of wh-questions that causes them to interact? Second, do these complex interactions have something larger to say about wh-questions more generally? It may turn out that these apparent interactions are an idiosyncratic result of some language-specific properties of the languages in question, but if they are a reflection of

---

\[\text{2} \text{Some researchers argue that it is not the argument status but rather other properties of the wh-phrase that determine when inversion becomes optional, including whether the wh-phrase is complex (Ordóñez 1997) and whether it is D-linked (Buesa-García 2008).} \]

\[\text{3} \text{Extra-syntactic properties of these factors may also play a role, such as the pragmatics of D-linked elements (cf. Pesetsky 1987), but the syntactic or formal effects that these factors have, which go beyond wh-questions, suggest that syntactic properties are also crucial.} \]
the formation of wh-questions more generally, then they can provide valuable insight into not only the morphosyntax of wh-questions but also the computational system more broadly.

1.1 Research Questions

The overarching goal of this study is to develop an understanding of the complex interaction of factors that determines the availability of inversion and their implications for the morphosyntax of wh-questions. In order to address this goal, I propose the following research question:

RQ1: What factors determine the word order in wh-questions in Spanish-English code-switching?

The most visible reflection of the influence of particular factors will be the availability of inversion in a given structure, and this is the word word effect we will be focusing on. Note, however, that the process of wh-fronting is also important, even if the ultimate result, i.e. the first constituent of a wh-question being a wh-phrase, is superficially the same in Spanish and English. For example, the wh-phrase may stop off in one or more intermediate landing sites before landing in its final position. If so, the presence of the wh-phrase in those landing sites, even if transitory, may have secondary effects that become apparent in the data.

Based on the previous discussion of some of the factors that are relevant for the formation of wh-questions in Spanish and English, we can break this overarching research question into four sub-questions, starting with research sub-question 1A:

RQ1-A: Do the syntactic properties of the subject determine which subject positions are grammatical?

Since we will be focusing on the availability of inversion in answering RQ1, we are essentially interested in the grammatical subject positions of a given sentence. It is important to consider, then, whether the syntactic properties of the subject itself determine the grammatical subject positions.

Given the argument-adjunct asymmetry in Spanish, where inversion becomes optional with wh-adjuncts, it appears that properties of the wh-phrase also play a role in inversion and therefore in the formation of wh-questions, suggesting the following research sub-question:
**RQ1-B:** Do the syntactic properties of the wh-phrase determine which subject positions are grammatical?

While this is a Spanish-specific phenomenon, we will also see that additional properties of the wh-phrase, including its role in the discourse, also have syntactic effects in wh-questions in both Spanish and English. How will these properties of the wh-phrase interact with the properties of C and T in the formation of wh-questions?

Considering again the contrast between inversion in Spanish and English, I have suggested that Spanish inversion is a consequence of leaving the subject in situ rather than raising it to SpecTP. If inversion in Spanish has to do with whether the subject raises to SpecTP, then it appears that T plays an important role in Spanish inversion, suggesting the following question:

**RQ1-C:** Do the syntactic properties of the tense head determine which subject positions are grammatical?

Given the contrast with English, where T raises to C to form inversion, what role does T play in wh-questions in Spanish-English code-switching? What does this role in code-switched wh-questions tell us about wh-questions in each language? And across languages?

Finally, we can consider the role of the complementizer head.

**RQ1-D:** Do the syntactic properties of the complementizer head determine which subject positions are grammatical?

There are various reasons to believe that C plays some role, but without further investigation it is not clear how far that role goes. The fact that the wh-phrase appears to front to SpecCP in both Spanish and English wh-questions suggests that C plays at least some role in the formation of wh-questions in these languages. Still stronger evidence comes from the fact that T also incorporates into C in English, yet it appears that the same thing cannot be said of Spanish, where inversion is not due to T-to-C movement. Does this mean that C plays a lesser role in wh-question formation in Spanish, or is C ultimately responsible for inversion in Spanish for different but equally important reasons?

So far we have focused on the factors determining the grammatical subject positions that we observe, whether that be the syntactic properties of the subject, the wh-phrase, the tense head, or the
complementizer, without particular regard for the language involved. In other words, with RQ1, we are more interested in any cross-linguistic generalizations that we can uncover. It would also be informative to consider the possibility of language-specific restrictions, particularly with respect to specific code-switches, or more concretely:

**RQ2:** Are there restrictions on code-switching between particular constituents in wh-questions in Spanish-English code-switching?

If a constituent has syntactic requirements that are specific to a particular language and require a syntactic relationship with another constituent, then CS between those constituents should render the sentence ungrammatical. It is generally functional heads that have these sorts of syntactic requirements, which among the constituents we are focusing on here consist of the complementizer head and the tense head. Following results from a series of pilot experiments\(^4\) we do not expect code-switches between T and the wh-phrase nor between full DP subjects and other constituents to restrict grammaticality, so we will not consider these factors. This leaves code-switches between C and T and between C and the wh-phrase, or more concretely:

**RQ2-A:** Are there restrictions on code-switching between the tense head and the complementizer head?

**RQ2-B:** Are there restrictions on code-switching between the wh-phrase and the complementizer head?

We have discussed a variety of factors that play a role in the formation of wh-questions in Spanish and English, including C, T, the type of wh-phrase, and the clause of the wh-question. This led us to the question of exactly what role each of these factors plays in code-switched Spanish-English wh-questions.

### 1.2 Dissertation Overview

Before reviewing the structure and content of this dissertation, I would like to make two comments concerning reading and navigating this document. First, in addition to the full table of contents near the

\(^4\) See §4.1 for some discussion of the pilot experiments.
beginning of this document, each chapter contains its own table of contents to facilitate navigation. Like all references to sections, examples or citations, these tables of contents contain links to the referenced sections or subsections. Second, for a quick overview of this dissertation, I recommend focusing on the following sections first: (1) the framework adopted for code-switching in §2.3, (2) the role of monolingual judgments in §2.4, (3) the hypotheses being tested in §3.3, (4) the introduction to the discussion in §5, and (5) the final summary and conclusions for the dissertation in §5.5.

Turning to the dissertation content, Chapter 2 lays out the assumptions and theoretical framework on which this proposal is based. These are important not only for analyzing the results of any data collected but also for designing the experiment so that the results are meaningful in the chosen framework. The chapter begins with a discussion of the ultimate object of study for this dissertation, the faculty of language (Chomsky, 1995, 2000), and the implications this has for syntactic structure. Following that, I discuss the ways in which these assumptions affect how we analyze and interpret code-switching. In particular, these assumptions lead us to consider code-switching as just another expression of I-language, one which follows the same principles as monolingual language. I then lay out two premises that will guide my discussion of the research questions and the implications that follow. Finally, I consider the role of dialect given the framework laid out in this chapter.

Chapter 3 reviews previous work in two key areas: wh-questions and code-switching. The discussion on wh-questions focuses on the factors, including those outlined above, that affect the formation of wh-questions in Spanish and English. We will see that, while the relevance of each factor to wh-questions is clear, the exact role these factors play in the morphosyntax of wh-questions remains unclear. Following that is a discussion of code-switching, with a particular focus on how previous CS research addresses wh-questions formation, directly or indirectly. We will see that, while much fruitful and insightful work has been done, little of it has focused on wh-questions. This section concludes by presenting concrete hypotheses to address the research sub-questions introduced above.

In Chapter 4 I review the methodology that was used in administering the dissertation experiment, along with a variety of considerations that went into arriving at the final design and the steps taken to prepare the resulting data for analysis. This includes discussion of a series of pilot experiments, the experimental stimuli, criteria and materials for selecting participants, the experiment materials, how the
experimental task was chosen and conducted, and the processing for selecting the final set of data and participants.

Chapter 5 is dedicated to answering the research questions and considering the implications for these answers both for Spanish-English code-switching specifically as well as Spanish, English, and the structure of wh-questions in general. For Research Question 1, repeated below, we will see evidence that the complementizer is ultimately responsible for determining the position of the subject in wh-questions, although other constituents also play an important role.

**RQ1:** What factors determine the word order in wh-questions in Spanish-English code-switching?

**RQ2:** Are there restrictions on code-switching between particular constituents in wh-questions in Spanish-English code-switching?

Turning to Research Question 2, repeated above, we will see that there are important restrictions on code-switching between a simple English wh-phrase and a Spanish complementizer, and this has implications for the relationship between these two constituents.

At a language-specific level, this dissertation provides counter evidence against the hypothesis that inversion in Spanish is triggered by properties of the tensed head (Toribio 1993; Barbosa 2001; Goodall 2001; Gutiérrez-Bravo 2005; Zubizarreta 2012, inter alia). At a cross-linguistic level, the results support a broader proposal that phase heads, including the complementizer, determine the grammatical properties of their complement, as proposed in González-Vilbazo and López (2012).

Finally, this study makes contributions to both theoretical and experimental aspects of research in code-switching. In terms of the relationship between code-switching and theoretical linguistics, this dissertation extends a line of research in which code-switching can serve as a unique tool of linguistic analysis, allowing the researcher to vary the language of the words included in a sentence and observe the effect the properties associated with those words have on the grammaticality of the sentence. This study also serves as one example of an experimental approach to code-switching research, with a particular focus on syntactic structure.
2 Assumptions and Theoretical Framework

2.1 Object of Study

In this chapter, I begin with a discussion of the ultimate object of study for this dissertation, the faculty of language, and the implications this has for syntactic structure. Following that, I discuss the ways in which these assumptions affect how we analyze and interpret code-switching. I then lay out two premises that will guide my discussion of the research questions and the implications that follow. Finally, I consider the role of dialect given the framework laid out in this chapter.

2.1 Object of Study

I begin with a brief discussion of the object of study for this dissertation. While at first glance it might appear to take us too far afield, it will be important for grounding our approach to understanding and interpreting code-switching.

In this dissertation, I seek to contribute to theoretical linguistics. Ultimately, then, I am interested not in language per se but in the linguistic knowledge that underlies language. There are many interesting and important descriptive observations that can be made about individual languages and differences between languages, but in the end, theoretical linguistics is concerned with the underlying processes and properties that produce the phenomena we observe. Note that these underlying properties and principles of interest are by no means limited to syntax. In fact, they encompass essentially every major field of linguistics, including language acquisition, sociolinguistics, phonology, morphology, semantics, and pragmatics.
Returning to the object of study, I am interested in the linguistic knowledge that speakers have concerning which sentences are and are not part of the languages they speak. This knowledge, known as linguistic competence, is the result of the interaction of an innate language module in the brain, the faculty of language (FL), with exposure to language, the primary linguistic data (PLD). Underlying principles of language structure and language acquisition, known as Universal Grammar (UG), determine how the PLD are interpreted and integrated into speakers' linguistic competence, building up the properties of the language(s) being acquired. The final product of this process is an I-language: internal, intensional and individual linguistic knowledge. I-language is internal in the sense that it is internal to the speaker, and it is individual in the sense that every speaker has their own individual I-language. Everyone is exposed to different linguistic data, and these linguistic data are sometimes interpreted differently during the process of building linguistic competence, producing a unique I-language. Finally, I-language is intensional, rather than extensional, because it is not a list of all possible sentences but an application of general principles to some language-specific properties.

In this dissertation, then, the ultimate object of study is the faculty of language, particularly the principles driving its relationship with language, Universal Grammar. Unfortunately, we do not have direct access to FL, but we can observe the product of UG and primary linguistic data: I-language. The object of study, then, is not the sentences that speakers utter, or even might utter, but the linguistic competence underlying those sentences, i.e. I-language. By exploring the properties of individual I-languages and the range of possible I-languages, we can infer more general properties and mechanisms generating those I-languages, those of UG. Ultimately, this leads us to FL itself.

2.2 Structural Assumptions

Having established the overarching object of study, I would like to layout my basic assumptions concerning the architecture of FL, focusing specifically on syntax. Following current implementations of the Minimalist Program (MP) (Chomsky 1995, et seq.), I take syntax to be a computational system, one that takes as input groups of syntactic properties, essentially words, and produces as output hierarchically structured syntactic properties, essentially sentences. Crucially, there is just one computational system across languages, based on the same structure-building mechanisms. Each individual word in a
sentence contributes to the overall syntactic properties of a given sentence through the syntactic properties that it is composed of. Among other properties, individual words are made up of abstract syntactic properties known as features. The sentence building mechanisms act on these features to generate the final output of the computational system. This output is then passed on to the interpretation and articulation systems, to be interpreted and pronounced, respectively.

With respect to the key structural components of a sentence, I assume the following structure, represented in bracketed form in (6a) and in tree form in (6b):

(6) a. \[CP \[TP \[T \[\text{vP} \[\text{vP} \[\text{VP} \[\text{DP} \[\text{NP} \[\text{N}]\]]]\]]]\]

b. 

The structural elements are as follows, from bottom to top. N is a noun. D is a determiner and takes a noun phrase (NP) as a complement. V is the verb, and it often takes a determiner phrase (DP) complement, though this will depend on the verb. \(v\) or ‘little \(v\)’ is a functional head that, among other things, introduces the external argument (essentially the subject in transitive predicates). \(v\) takes a verb phrase (VP) as its complement. T, the tense head, is the locus of tense and aspect properties of the clause. T takes a little v phrase (vP) as its complement. C, the complementizer head, is a functional element that heads up a clause and determines certain discourse properties of that clause, including whether it is declarative or interrogative. C takes a tense phrase (TP) as its complement. In some
theories, some of these heads are actually split into a more articulated hierarchy of functional heads, but I will focus on the structure outlined above for the sake of simplicity and to ease exposition.\footnote{For a more cartographic approach, with a more finely articulated CP, see Rizzi (1997, 2004).}

The properties reviewed above are sufficient to answer the research questions and to consider various implications of those answers, but for the sake of concreteness, I will formulate the assumptions in more explicit terms, following ideas common to most current implementations of the Minimalist Program. Though I will be using this terminology at times for ease of exposition, none of the core argumentation hinges on these details.\footnote{For a more complete overview of current implementations of the Minimalist Program, see the introductory chapter of Boeckx (2011).}

There are two core mechanisms for building syntactic structures: Merge and Agree. Merge takes two syntactic objects and combines them into one larger syntactic object. For example, a noun and a determiner may be combined to form a determiner phrase, or DP, as in (7a) and (7b).

\[\text{(7) a. } \text{DP D N} \]
\[\text{b. DP} \]
\[\text{D N} \]

The syntactic objects may be the result of previous operations of merge, in which case the internal structure of the syntactic objects remains intact. For example, a verb might merge with the DP above to form a VP, yielding the embedded structure in (8).

\[\text{(8) a. } \text{VP } \text{[DP D N]} \]
\[\text{b. VP} \]
\[\text{V DP} \]
\[\text{D N} \]

Some syntactic features on a given head need to form a relationship or dependency with a syntactic feature on another head. This relationship is established by means of the Agree operation, where agreement occurs between the two closest heads with matching features. The particular implementation of
this agreement processing, and what counts as the closest head, varies by the specific theory. Sometimes agreement not only establishes a relationship between two heads but also requires the lower head to move or re-merge above the higher head to a position known as the specifier. For example, in English, T has features which require a relationship with a DP, so agreement takes place and the DP moves to the specifier of T (SpecTP). In this way, the subject, as the closest DP, moves from the specifier of v (SpecvP), where it first merges, as seen in (9a), to SpecTP, as seen in (9b). For the sake of exposition, a trace (t) is often left in the original location of the moved element to indicate the original merge position. A more visual version can be seen in (9c).

(9) a. \[TP \ [\ell_v \ DP \ v \ VP]]\]

b. \[TP \ DP \ T \ [\ell_v \ t \ v \ VP]]\]

c. TP
   \[\text{DP} \quad T' \quad vP \]
   \[T \quad vP \]
   \[t \quad v' \quad \text{VP} \]

2.3 Code-Switching Framework

The approach to syntax outlined above has important implications for the study of code-switching. If every language uses a subset of universal syntactic features and these are manipulated by the same computational system to build sentences, then code-switching is just another example of a set of features being manipulated by the computational system in order to build syntactic structures (cf. MacSwan 1999 for the first proposal along these lines and González-Vilbazo and López 2011, 2012 for further discussion). Put another way, just as with linguistic knowledge of one language, code-switching is also a reflection of a speaker’s linguistic competence, or I-language. Given that code-switched sentences are constructed with the same computational system as monolingual sentences, I also assume that there is no third grammar specific to CS, following Woolford (1983), Di Sciullo et al. (1986), Belazi et al. (1994),
Mahootian and Santorini (1996), MacSwan (1999), González-Vilbazo (2005) and González-Vilbazo and López (2011, 2012). In other words, there are no specific rules or operations whose sole purpose is to govern code-switching.

As with monolingual sentences, code-switching is underlyingly a hierarchical syntactic structure, so there are no code-switching points per se.

(10) a. \[ CP [TP [vP vL_B L_A ...] T_{LA}] ]

b. \[
\text{CP} \\
\text{C}_{LA} \\
\text{TP} \\
\text{vP} \\
\text{T}_{LA} \\
\text{vL_B} L_B ... \]

For example, in (10a), we can see that linearly there is a code-switch between C, in language A (L_A), and v, which is in language B (L_B). Hierarchically, however, C and v have no direct relationship, as seen in (10b). Instead, the first two heads in the tree from top to bottom, C and T, are both in the same language.

Still, there are two important differences from the building of monolingual sentence structure. First, the combination of features from more than one language often means that features can be combined that are never part of any single language’s set of features. For example, there are features in Spanish that give rise to subject-verb inversion (resulting in VS word order), and there are features in English that give rise to subject-auxiliary inversion (resulting in V2 word order). Spanish-English wh-questions can have both kinds of inversion, while, to my knowledge, there is no single language that has both kinds of inversion as part of the process of wh-question formation. Importantly, these unique combinations of features can provide additional insight into the computational system through exploring and analyzing their interaction in the computational system. The second difference between the building of CS and monolingual sentences is that we are able to manipulate the language of the lexical items used in a given CS sentence. When we combine words from different languages, we are also combining the features associated with those words. By carefully controlling the language of the words that we combine, we can
tease apart the role that those associated features play in the syntactic derivation, making CS a unique tool for linguistic analysis.

Broadly speaking, analysis of CS data can provide two kinds of insight. On the one hand, an exploration of the ways in which features from two different languages interact can lead to insights into general properties of the computational system. In fact, CS raises questions not just about the computational system but also other interesting and important questions about the faculty of language and language more generally, including issues concerning the application of phonological processes from multiple languages in one utterance and the online processes involved in making particular code-switches in a particular discourse. I will not be addressing such questions in the present study. On the other hand, using CS for linguistic analysis can also provide insight into the individual languages that are being code-switched. With a careful design and analysis, it becomes possible to disentangle the role that specific features play in the syntax of one or both of the individual languages. In particular, we can manipulate the language of different parts of code-switched sentences and observe the effect of this manipulation on the structure and grammaticality of those sentences. Such an analysis must, of course, be grounded in previous syntactic work on each of the individual languages.

2.4 Role of Monolingual Judgments

Given the primacy of I-language in the framework adopted here, monolingual acceptability judgments must necessarily play a central role in organizing and interpreting any CS acceptability judgments. As discussed in §2.1, we are using I-language as a window into the properties of the faculty of language (FL). In the context of syntax, I-language consists of an available set of syntactic features interacting based on Universal Grammar (UG) principles. The specific set of syntactic features will vary based on the particular speaker’s I-language.

CS can provide insight into particular I-languages, and therefore into FL itself, through the following process. First, we gather acceptability judgments on CS sentences to determine which sentences are acceptable for a specific speaker or group of speakers, and then we work to infer what combination of syntactic features and syntactic principles could generate those acceptable sentences. The proposed

---

3The exception to this would be cases in which the characteristics of the participants have already previously been established.
features and principles would be part of UG and hence FL. In order to do that, we need to know as much as we can about the available syntactic features. The monolingual acceptability judgments are a reflection of the available syntactic features in each language. Together, the syntactic features of both languages reflect the entire set of syntactic features available for the computational system to use in building the CS sentences. In summary, in order for CS data to inform our understanding of FL, we need the context of monolingual acceptability judgments.

So far we have focused on the I-language of individual speakers, but in the experimental approach of the current study, we are interested in aggregating data from groups of participants. Because the I-language of each participant will by definition be different to some degree, these groups will necessarily be heterogenous in some sense. Scientific data generally require some degree of leveling or idealization, so this is not problematic in and of itself. Still, the I-languages of the individual participants should overlap as much as possible to mitigate any effects those differences might have.

Since it is not possible to test every facet of a participant’s knowledge of each language, researchers should focus on those features and those structures that are relevant to the phenomenon or phenomena being investigated. In the context of the present work, we are interested in the syntactic properties of Spanish and English wh-questions. Note that this encompasses not only issues of word order, such as the position of the subject in a given wh-question, but also related syntactic properties. For example, as we will see in §3.1.2.1, it is possible that the difference between inversion in English, as in (11a), and inversion in Spanish, as in (11b), comes down to whether only the auxiliary ‘has’ moves above the subject or whether the auxiliary and participle ‘has bought’ both move above the subject.

(11) a. What has John bought today?
   b. ¿Qué ha comprado Juan hoy?

‘What has Juan bought today?’

Alternatively, the subject may remain in situ below the verb and auxiliary. The word order in questions such as (11) is not sufficient for distinguishing between these accounts, but tests based on the position

---

4 Strictly speaking, this is not necessarily the case. For example, González-Vilbazo and López (2011) argue that a light verb is available in a variety of Spanish-German code-switching that is not available in either of the input grammar (cf. González-Vilbazo and López, 2011, p. 846, for additional discussion). Still, it is likely the case with most phenomena, and I will assume that this simpler assumption holds for the phenomena under study here, unless the data force a reconsideration.
of adverbs, as proposed by Goodall (1992) and Suñer (1994), can provide evidence. Thus, additional structures such as those used in the adverb tests must also be included in the stimuli used for the monolingual acceptability judgments.

Importantly, the monolingual acceptability judgments do not necessarily have to overlap with monolingual speakers, though there are important benefits to such an overlap. The greater the degree of overlap with monolingual speakers, the greater the degree to which the analysis can draw from previous syntactic work on the languages. Still, any reasonably well-studied phenomenon will have plenty of cross-linguistic research from which to draw, at which point the interest is much more on the properties of the particular I-languages under study, irrespective of the language they are drawn from. More generally, any I-language can provide a window into properties of FL, regardless of the degree to which it resembles particular well-studied languages.

2.5 Premises for Answering Research Questions

Pulling together the assumptions laid out in this chapter, I would like to propose two premises concerning syntactic structure and code-switching. These premises will be useful for answering each of the research questions as well as discussing the implications that follow from those answers.

First, all things being equal, it would be preferable to derive a given phenomenon from the properties of one specific syntactic head, in the spirit of a common scientific heuristic, Occam’s razor. In other words, I would like to assume that there is a specific head which plays a central role in explaining a given phenomenon. This is as opposed to deriving the phenomenon from a constellation of heads. In the latter case, it is not clear how to generate hypotheses concerning any exceptions to the general properties of the phenomenon, nor is it clear what predictions it would make if properties of just one head were to change, such as in a CS sentence. This is problematic, given the central role of predictions in evaluating scientific proposals. Applying this reasoning to wh-questions, I will assume the following premise and see how far it can take us in answering our research questions:

**Premise 1:** There is a head which has a specific set of syntactic features that ultimately determine the grammatical subject positions in a given wh-question. Call this a determinant head.
The premise is couched in terms of syntactic features following assumptions in §2.2 regarding how particular syntactic structures or word orders are derived. The focus is on grammatical subject positions because they are the most salient word order difference between Spanish and English wh-questions. This premise was already implicit in the sub-questions proposed for the first research question, repeated below.

**RQ1**: What factors determine the word order in wh-questions in Spanish-English code-switching?

- **RQ1-A**: Do the syntactic properties of the subject determine which subject positions are grammatical?
- **RQ1-B**: Do the syntactic properties of the wh-phrase determine which subject positions are grammatical?
- **RQ1-C**: Do the syntactic properties of the tense head determine which subject positions are grammatical?
- **RQ1-D**: Do the syntactic properties of the complementizer head determine which subject positions are grammatical?

If the phenomenon of interest, word order in wh-questions, is derived from the properties of a specific syntactic head, the determinant head, and there are four plausible heads or phrases containing that head that could be the determinant head, then there are four research questions to consider, hence RQ1-A, RQ1-B, RQ1-C, and RQ1-D.

The second premise largely follows from previously discussed assumptions concerning the nature of code-switching.

**Premise 2**: A head in a given language and in a parallel structure has the same set of syntactic features whether in a monolingual or CS sentence.

The syntactic structure of a given sentence is derived from the syntactic features present in that sentence, and syntactic features come from the words that make up that sentence. In CS, speakers have the

---

[^5]: At first glance, it might appear that this premise is already implied by the first one, since the same grammatical subject positions in both CS and monolingual sentences will necessarily involve the same set of syntactic features. It turns out, though, that this premise will be necessary for ascertaining which head is the determinant head.
ability to combine words from the lexicon of either of the two languages that they speak, but the words themselves still have the syntactic features of the languages that they are drawn from.

2.6 Dialect, Context and Linguistic Knowledge

There is one remaining issue to consider, the issue of dialect. Broadly speaking, the goal of this dissertation is to use the linguistic knowledge of Spanish-English bilinguals to inform our understanding of wh-questions, as well as syntax more generally. To do this, participants are asked to complete an acceptability judgment task, and then the resulting data is analyzed by contrasting what is rated more acceptable with what is rated less acceptable. Because of this, it is important to carefully consider the connection between acceptability judgments and linguistic knowledge in order to ensure that any conclusions that we draw are valid and generalizable. This is particularly true in the case of inversion in embedded wh-questions in English, where some speakers accept both (12a), which has no inversion, and (12b), which has inversion, while others appear to only accept sentences like (12a), judging sentences like (12b) to be unacceptable.

(12) a. I wonder what he is like.
   b. I wonder what is he like.

In traditional syntactic analysis, acceptability judgments of a particular set of structures are taken to fully reflect the degree to which those structures are possible for a given speaker, with the possible exception of a missing or incompatibly discourse context. While context is taken into account when it is considered important, it is generally limited to issues of information structure or pragmatics, including what is given, focused, contrasted, or a topic. For example, many speakers of English, particularly non-linguists unfamiliar with the structure, would find the sentence in (13a) highly odd, if not unacceptable. By adding some context, as in (13b), the judgment becomes much clearer.

(13) a. Oranges, John eats.
   b. LEMONS, John would never eat, but ORANGES, John eats.

What is not often considered is the social context. As has been known at least as far back as Labov’s seminal work in sociolinguistics, when participants tell a researcher that they cannot not say
something, it does not necessarily mean that they cannot say it. They may be consciously aware that a particular form is stigmatized or of a prescriptive rule concerning that form, and this may influence how they complete any tasks involving this form. In addition, participants may subconsciously restrict their responses in a task based on their choice of dialect, style, or register. This is especially important in the case of acceptability judgments because these are administered in the formal context of an experiment, not to mention the fairly unnatural character of an acceptability judgment task. Any structures that participants only use in casual or informal settings might be rated as unacceptable in spite of the fact there are contexts in which the structures are acceptable.

On the one hand, any sentences that are judged as acceptable clearly constitute sentences that the participant’s linguistic competence allows him or her to generate (i.e. that are part of his or her dialect). On the other hand, sentences that are judged unacceptable may be truly impossible for a participant to say (i.e. not part of their linguistic competence), or the sentences may simply not be part of the dialect, style, or register they are currently using. In the latter case, an account of the participant’s linguistic competence that categorically forbade such sentences would be incorrect, even though the acceptability judgments specifically suggest that any account should predict these sentences to be ungrammatical.

While in theory the potential mismatch between the provided acceptability judgments and actual judgments could occur with any structure, this is likely to be an issue only in the case of structures or properties for which there is evidence of variation across dialects, styles, or registers. For example, except for a few very specific structures, one wh-phrase always appears in the left periphery of the clause in English information-seeking questions, with any additional ones remaining in situ. In this case, judgments related to this phenomenon can likely be taken at face value. Now consider the previously mentioned case of inversion in embedded questions in English, as in (12) above. It turns out that there is good evidence that structures such as (12b), with inversion, are acceptable for many more dialects than is generally accepted (cf. McCloskey 2006 for discussion). This is a case where the effect of dialect, style, or register is significant and therefore must be taken into account.

There are a number of different ways of taking into account such effects, and the best option will vary based on the situation. Notice that the we are specifically interested in the use of acceptability judgments...
for informing syntactic theory. There are many interesting questions at the intersection of syntax, variation and dialect, but they are beyond the scope of this dissertation. In terms of solutions, it may be helpful to pair the acceptability judgment task with a sociolinguistic interview in order to attempt to elicit nature examples of the structures of interest, particularly those that previous research suggests are affected by dialect, style, or register. While one can never guarantee that a speaker will use the particular register you are interested in, this is more likely in the context of an appropriately administered interview than during an acceptability judgment task. Still, such an approach may not always be practical, given the population or the context in which the study is being conducted. Unfortunately, this is the case with the data collected for this dissertation.

Another solution is to modify the acceptability judgment task. For example, participants can be trained to keep in mind that some sentences may only sound possible or acceptable when imagining them as part of a casual conversation with friends. An additional solution for studies using constructed sentences, such as is the case with this dissertation, is to also test an additional set of stimuli that is based on natural examples from corpora or published texts.

I would like to briefly touch on how the specific issue of inversion in embedded questions in English affects my data and analysis. While the monolingual stimuli do include more natural examples of inversion in embedded questions, in addition to the constructed examples, the current CS data does not include inversion in the embedded stimuli. Because of that, the issue of dialect, style, or register effects does not directly affect the analysis in this case. Still, the data on the acceptability of inversion in embedded questions does play an important indirect role. As discussed in §4.6.1.4, the participants for this study were grouped according to their monolingual judgments in both Spanish and English due to the high degree of variation in their monolingual judgments. Among the factors used for comparing participants were the acceptability of both the constructed and more natural embedded questions with inversion. Since some participants judged these two sets of stimuli differently, it was important to include each set as individual factors.

\footnote{It would also be possible to only use the natural examples, depending on the design of the study. For the present study, it was important to have sets of stimuli based on the same translation-equivalent vocabulary across structures for both monolingual and CS stimuli. See section §4.2.1.3 for further discussion of the use of lexicalizations in the stimuli design.}
Moving beyond inversion in embedded English questions, the same issue potentially holds for all the structures in question, particularly because of the variety in the requirements for inversion in Spanish across dialects. Beyond the aforementioned separation of participants according to their monolingual judgments, which ensures that we are not assuming particular features of wh-questions for all participants, the present study does not have other ways of controlling for dialectal, stylistic, or register differences. Because of this, conclusions drawn from the current data do not necessarily apply to all aspects of the participants’ linguistic competence in syntax but only to those of the dialect, style, or register used during the study. Still, the data do offer insights into wh-questions, as well as more general properties of syntax. Finally, the issues discussed in this section are also something that must be taken in to account for future CS work, not only when including English embedded V2 word order but also more generally.
# Literature Review

## 3.1 Wh-Questions

- **3.1.1 Descriptive Generalizations** .................................................. 25
- **3.1.2 Inversion in Spanish Wh-Questions** ........................................ 27
  - **3.1.2.1 T-to-C Movement in Spanish?** ............................................. 28
  - **3.1.2.2 Preverbal Subjects in Spanish** .......................................... 29
  - **3.1.2.3 The Trigger for Inversion** ................................................. 32
- **3.1.3 Inversion in English Wh-Questions** ........................................ 33
  - **3.1.3.1 Rizzi (1996)** ................................................................. 33
  - **3.1.3.2 Pesetsky and Torrego (2001 et seq.)** .................................... 35
  - **3.1.3.3 Conclusions** ................................................................. 41
- **3.1.4 The Role of the Wh-Phrase** .................................................... 42
  - **3.1.4.1 Syntactic Function of the Wh-Phrase** ................................... 42
  - **3.1.4.2 D-Linked and Non-D-Linked Wh-Phrases** ............................... 43
  - **3.1.4.3 Simple and Complex Wh-Phrases** ...................................... 45
- **3.1.5 Conclusions** ........................................................................ 49

## 3.2 Code-Switching

- **3.2.1 Early Formal Approaches to Code-Switching** ............................. 50
  - **3.2.1.1 Poplack (1980)** .............................................................. 50
  - **3.2.1.2 Joshi (1985)** ................................................................. 52
- **3.2.2 Generative Approaches to Code-Switching** ............................... 53
  - **3.2.2.1 Woolford (1983, 1984)** .................................................. 53
  - **3.2.2.2 Di Sciullo, Muysken, and Singh (1986)** .............................. 55
  - **3.2.2.3 Belazi, Rubin and Toribio (1994)** ....................................... 56
  - **3.2.2.4 Mahootian and Santorini (1996)** ........................................ 59
- **3.2.3 Minimalist Approaches to Code-Switching** ............................... 60
  - **3.2.3.1 MacSwan (1999)** .......................................................... 60
In order to situate this project in the broader field of theoretical linguistics, it is necessary to review pertinent research in two areas of inquiry. First, in §3.1, I will review some relevant background information and key work on wh-questions, with a particular focus on the language-specific properties of Spanish and English. This information will serve two purposes. First, it will make clear the factors that are most relevant to understanding inversion and wh-movement in the two languages. Second, it will demonstrate the many open questions that remain with respect to inversion.

In the second section of the literature review, §3.2, I will survey previous work in code-switching, with an emphasis on the ability of that work to make concrete predictions for wh-questions in code-switched sentences. Ultimately, we will see that, while much fruitful work has been done in code-switching research, there is ample opportunity to bring this unique type of evidence to bear on a number of issues in the area of wh-questions, particularly the phenomenon of inversion, where many open questions remain.

Finally, having reviewed pertinent research on wh-questions and on code-switching, we are in a position to consider the most appropriate hypotheses for each of the sub-questions for Research Question 1 and Research Question 2, repeated here:

**RQ1:** What factors determine the word order in wh-questions in Spanish-English code-switching?

**RQ2:** Are there restrictions on code-switching between particular constituents in wh-questions in Spanish-English code-switching?

Each sub-question is considered in turn in §3.3.
3.1 **Wh-Questions**

In this section we begin by reviewing the key generalizations concerning wh-questions in Spanish and English. With the background these generalizations provide in place, we turn to the intricate properties of inversion in Spanish wh-questions. We will see that the interaction between the position of the preverbal subject and the landing site(s) of the wh-phrase is crucial to understanding inversion. Next we turn to inversion in English wh-questions. Here we consider the interaction of the properties of C and T as motivated by either semantic requirements, the focus in [Rizzi 1996](#), or more formal features and operations, the primary focus of [Pesetsky and Torrego 2001](#) et seq.). Each section will help in the goal of understanding to what degree there is consensus on the basic syntactic processes involved in wh-movement and inversion.

### 3.1.1 Descriptive Generalizations

There are two syntactic processes that are characteristic of wh-movement in English. First, the wh-phrase fronts to what is generally considered to be SpecCP ([Chomsky 1986](#)). Second, there is a process of inversion whereby the auxiliary verb moves above the subject, generally analyzed as the T head (occupied by the auxiliary) moving to the C head ([Den Besten 1983](#)). This results in the following word order: Wh-phrase + Auxiliary + Subject + Participle, as illustrated by the following simple matrix question in English:

(14) What has John bought? \[Wh, Aux, S, Part\]

When T-to-C movement occurs and there is otherwise no overt auxiliary, T is expressed as the semantically vacuous auxiliary ‘do’ in a process known as ‘do-support’. An example like (14) without a lexical auxiliary turns into (15).

(15) What did John buy? \[Wh, Aux, S\]

While inversion is obligatory with a wh-object in a question such as (14), if the wh-phrase is a subject, T-to-C movement is no longer grammatical, as seen in example (16).

(16) What has John eaten? \[Wh, Aux, S\]

---

1 For a more cartographic approach, with a more finely articulated CP, see [Rizzi 1997, 2004](#).
(16)  a. Who bought the book?  
b. *Who did buy the book? (non-contrastive reading)

If the wh-question is in an embedded clause, such as in the following sentence, inversion, or V2 word order, is also ungrammatical:

(17)  a. I don’t know [what book John has bought].  
b. *I don’t know [what book has John bought].

Thus there is an asymmetry between the availability of inversion in matrix and embedded clauses in English.

Spanish wh-questions also involve the fronting of the wh-phrase and a form of inversion. Importantly, this inversion does not produce the subject-auxiliary inversion of English but rather a simple subject-verb inversion. In fact, the subject cannot intervene between an auxiliary and a participle in Spanish wh-questions, as illustrated by the contrast between (18b) and (18c):

(18)  a. ¿Qué compró Juan?  
verb-subject inversion (VS)
    what bought Juan  
    ‘What did Juan buy?’

b. ¿Qué ha comprado Juan?  
verb-subject inversion (VS)
    what has bought Juan  
    ‘What has Juan bought?’

c. *¿Qué ha Juan comprado?  
auxiliary-subject inversion (V2)
    what has Juan bought  
    ‘What has Juan bought?’

Also in contrast to wh-questions in English, there is no matrix-embedded asymmetry in Spanish, as illustrated by the grammaticality of inversion in both example (18a) above and example (19) below.

2This nomenclature, derived from the verb-second phenomenon of many Germanic languages, is not intended to represent a specific syntactic analysis. Rather, it serves to distinguish this word order from the VS word order of ‘auxiliary participle subject’. See Rizzi (1996) for arguments that this phenomenon in English is a residual V2 phenomenon.

3In some dialects of English, there are contexts in which inversion is also possible in embedded questions (McCloskey 2006). See §2.6 for some discussion.

4Strictly speaking, the matrix clause encompasses the entire sentence, unless two sentences are conjoined. Here, the term is used to refer to the part of the sentence that does not included the embedded clause, essentially the matrix subject, verb and complementizer, if present. I will use the term ‘matrix clause’ to refer to this part of the sentence throughout the dissertation, for the sake of simplicity.

5Note that this study focuses on non-Caribbean varieties of Spanish. In Caribbean varieties of Spanish, inversion is optional in structures such as (18a), as long as the subject is pronominal (cf. Ordóñez and Olarrea 2006).
Instead, there is an argument-adjunct asymmetry where wh-phrases that are arguments (e.g. ‘who’ and ‘what’) require inversion, but inversion in wh-questions with wh-adjuncts is optional, as exemplified by (20a) and (20b). 

\begin{align*}
(19) & \text{No sé [qué compró Juan],} \\
& \text{not know-I what bought Juan} \\
& \text{‘I don’t know what Juan bought.’}
\end{align*}

\begin{align*}
(20) & \text{a. ¿Por qué trabaja tanto Juan?} \\
& \text{for what works so much Juan} \\
& \text{‘Why does Juan work so much?’} \quad \text{VS}
\end{align*}

\begin{align*}
& \text{b. ¿Por qué Juan trabaja tanto?} \\
& \text{for what Juan works so much} \\
& \text{‘Why does Juan work so much?’} \quad \text{SV}
\end{align*}

In summary, inversion is optional in questions with wh-adjuncts but required elsewhere.

### 3.1.2 Inversion in Spanish Wh-Questions

This section reviews in more detail the key factors that play a role in inversion in Spanish wh-questions. I begin by considering whether inversion in Spanish is due to T-to-C movement as it is in English, ultimately concluding that it cannot be. If that is the case, then what position is the subject in when there is inversion? I consider the possibility that preverbal subjects are in SpecTP and postverbal subjects are in situ, noting along the way that this question remains controversial. Following that, I turn to the connection between movement of the wh-phrase and ungrammaticality of preverbal subjects. A variety of accounts (e.g. [Toribio] 1993, [Goodall] 2001) posit a conflict between preverbal subjects and wh-movement by having the wh-phrase move through the position that a preverbal subject normally occupies, SpecTP, before finally landing in SpecCP. The next section on the properties of wh-phrases will discuss this dual movement approach to inversion in light of the sensitivity to the type of wh-phrase that Spanish inversion evidences.

Some researchers argue that it is not the argument status but rather other properties of the wh-phrase that determine when inversion becomes optional, including whether the wh-phrase is complex ([Ordoñez] 1997) and whether it is D-linked ([Buesa-García] 2008).
3.1.2.1 T-to-C Movement in Spanish?

While on the surface the process of inversion in Spanish appears similar to inversion in English, which is considered to be a result of T-to-C movement, there is strong evidence that this is not the case. One compelling piece of evidence comes from the position of adverbs and negation in wh-questions, as argued by Suñer (1994). She begins by pointing out a number of adverbs that may appear before the verb:

(21) a. La viejita apenas puede leer los periódicos.
   the old woman barely can to read the newspapers
   ‘The old woman can barely read the newspaper.’

   b. Los estudiantes ya casi/jamás terminaron el examen.
   the students already almost/never finished the test
   ‘The students almost/never finished the test.’

   c. Bri todavía/aún estudia historia del arte.
   Bri still/still studies history of the art
   ‘Bri still studies art history.’

Because verbs in Spanish raise to T (Suñer, 1994), these adverbs must occupy a position higher than T. When these adverbs are used in questions, they continue to precede the verb and therefore remain higher than T:

(22) a. ¿A quién jamás ofenderías tú con tus acciones?
   to who never would offend you with your actions
   ‘Who would you never offend with your actions?’

   b. ¿Qué idioma todavía estudia Pepita en su tiempo libre?
   what language still study Pepita in his time free
   ‘What language does Pepita still study in his free time?’

If T had raised to C, then these adverbs would follow T, as they clearly do in the English translations. Additional evidence comes from negation. Both the previously mentioned adverbs and negation itself precede the verb:

(23) a. Drea todavía/aún no sabe el resultado.
   Drea still/still not knows the result
   ‘Drea still does not know the result.’

   b. El otro ya no decía nada.
   the other already not said anything
   ‘I don’t know what Hans has sold.’
When negation is used in a wh-question, it continues to precede T:

(24) a. ¿A quién ya casi no le escribes tú cartas?
    ‘Whom do you almost never write letters to anymore?’

b. ¿A cuántos todavía no les has mandado (tú) tarjetas para las fiestas?
    ‘How many have you still not sent cards to for the holidays?’

Once again, the adverbs and negation precede the verb, in spite of the inversion. Note that in the declarative sentences, both the adverbs and negation are after the subject, which is in SpecTP. That being the case, neither the adverbs nor negation can be between SpecCP and C in the wh-question example. That only leaves a position below C, so T (and hence the verb) cannot have raised to C.

It is important to note that there are also other kinds of evidence against the T-to-C movement account of inversion, including evidence based on sentence negation (Suñer 1994), sluicing, (Suñer 1994), exhortative subjunctive (Villa-García 2013) and the interpretation of adverbs (Buesa-García 2008). In addition, there are a variety of potential problems with such an account based on the fact that T-to-C movement is essentially a V2 phenomenon and there is no historical reason why Spanish would have any V2 phenomena (Goodall 1993).

3.1.2.2 Preverbal Subjects in Spanish

If inversion in Spanish is not due to T-to-C movement, then the question becomes how the verb comes to precede the subject. This issue has important implications for wh-question formation in Spanish, particularly for the underlying source of inversion. A variety of theories (e.g. Toribio 1993, Goodall 2001) posit that inversion is obligatory with wh-objects because a wh-phrase prevents a subject from raising to SpecTP. If preverbal subjects do not raise to SpecTP but some other position, then inversion in wh-questions must arise not because the subject is forced to remain in situ but for other reasons.

By examining the evidence concerning the position of preverbal subjects in Spanish, we can determine which theories on inversion in Spanish are worth pursuing.
Consider the possibility that the subject has remained in situ. Before T merges with little v, we have the following structure, in which the subject is in Spec,v, the verb has risen from the VP and incorporated into little v, and the wh-object is the complement of the verb:

\[(25) \begin{array}{c}
vP \\
Juan \\
\end{array} + comprar_i \begin{array}{c}
T \end{array} \begin{array}{c}
Qué \\
\end{array} \]

‘What did Juan buy?’

When T merges with vP, then v+comprar raises and incorporates into T, as in both of the examples below. Now there are two possibilities for the subject. The subject may move to SpecTP, placing it above the verb, which has incorporated into T, and yielding the subject-verb (SV) word order in (26a). Alternatively, if the subject is not required to, or is forbidden from, raising to SpecTP, it will remain in situ. Notice that, because the verb has raised to little v and then to T, an in situ subject is actually below the verb, as seen in (26b). In either case, C will then merge with the structure and force the wh-phrase to raise to SpecCP and form a complete question, as in (26c).

\[(26) \begin{array}{c}
TP \\
Juan_j T+v+comprar_i vP \begin{array}{c}
t_j t_i \end{array} \begin{array}{c}
Qué \\
\end{array} \]

\[SV\]

\[(26) \begin{array}{c}
TP \\
T+v+comprar_i \begin{array}{c}
vP \end{array} \begin{array}{c}
Juan \end{array} t_i \begin{array}{c}
vP t_i \end{array} \begin{array}{c}
Qué \\
\end{array} \]

\[VS\]

While it is largely agreed that a postverbal subject (cf. Suñer (1994) for arguments and Zubizarreta (1999) for another perspective) is in situ, much more controversial is the position of the preverbal subject. I will outline some of the numerous proposals on Spanish preverbal subjects and conclude with some evidence in favor of an analysis in which preverbal subjects raise to SpecTP (though they may later be dislocated).

Under some accounts, including Contreras (1991), Ordóñez (1997), Alexiadou and Anagnostopoulou (1998) and Ordóñez and Treviño (1999), a preverbal subject in Spanish is always left dislocated. Evidence for this perspective is usually based on the fact that Spanish preverbal subjects seem to have

\[\text{I do not consider here the distinction between VSO and VOS postverbal subjects. See Ordóñez (1998) and López (2009) for two perspectives on this issue.}\]
some properties of constituents that have undergone A-movement. For example, consider the following sentences from [Jaeggli 1987].

(27) a. ¿Qué dijiste que compró quién el otro día?
   what you say that bought who the other day
   ‘What did you say that who bought the other day?’

   b. *¿Qué dijiste que quién compró el otro día?
   what you say that who bought the other day
   ‘What did you say that who bought the other day?’

Both examples are multiple wh-questions, which in Spanish means that only one wh-phrase needs to front. These data show that if a wh-subject does not front, then it must remain in a postverbal position, as in (27a), or the sentence is ungrammatical, as in seen in (27b). [Ordóñez 1997] argues that a preverbal subject is ungrammatical in such a sentence because the wh-subject would be required to undergo A-movement twice, once overtly to a preverbal position and a second time covertly to take scope over the clause. Covert movement is posited based on the idea that all wh-phrases must front at LF in order to provide the correct scopal relations for operator-variable binding (cf. [Huang 1982], i.e. for semantic interpretation. This double movement violates the following constraint, proposed by [Lasnik and Uriagereka 1988] (see also [Epstein 1992], [Kayne 1998]):

(28) Overt A-movement may not be followed by covert A-movement of the same item.

As [Ordóñez 1997] argues, since the postverbal wh-subject only undergoes A-movement once, moving covertly to take scope, an in situ postverbal wh-subject is perfectly grammatical.

Others, including [Goodall 2001], [Ortega-Santos 2005], and [López 2009], have argued that preverbal subjects occupy SpecTP, though they may later raise to an A-position. Perhaps the simplest argument against a requirement that all preverbal subjects be dislocated comes from the discourse properties of dislocated constituents. [López 2009] observes that in an all-focus context, where all information is new (e.g. answering the question ‘What happened?’), dislocation is not felicitous. This stems from the fact that a dislocated constituent must be given information, roughly. If a preverbal subject is dislocated, then it would be expected to be infelicitous in an all-focus context. Instead, the most natural position
for the subject is postverbal. It turns out that only a preverbal subject is felicitous\footnote{A number sign (#) is used to indicate that a sentence is infelicitous, i.e. that it does not fit into the given discourse.} in an all-focus context:

(29)  
[Context: What happened?]

\begin{itemize}
\item[a.] Juan compró tres libros  
\hspace{2cm} SV

\textquote{Juan bought three books.}'

\item[b.] #Compró Juan tres libros.
\hspace{2cm} VS

\textquote{Juan bought three books.}'
\end{itemize}

The fact that VS word order is not grammatical in such a context is strong evidence that preverbal subjects do not have to be dislocated (cf. \cite{Goodall2001, Ortega-Santos2005, López2009} for additional arguments). Though the topic remains controversial in the literature, I will conclude that preverbal subjects in Spanish occupy (or pass through) \textit{SpecTP}, following the above argument and cited references.

### 3.1.2.3 The Trigger for Inversion

If inversion in Spanish is not due to T-to-C movement and postverbal or inverted subjects are in situ, then the question becomes what prevents the subject from raising to \textit{SpecTP}, the preverbal position, in some types of \textit{wh}-questions. The most common approaches involve moving the \textit{wh}-phrase to \textit{SpecTP}, and the presence of the \textit{wh}-phrase either blocks the subject from raising or satisfies the requirements of T, obviating the need to raise the subject.

As an example of the blocking approach, consider the proposal by \cite{Toribio1993}. In order to account for the argument-adjunct asymmetry discussed in \S3.1.1, Toribio begins by assuming that T (INFL in her terms) is specified as [\textit{wh}] through feature-checking with the \textit{wh}-argument. The \textit{wh}-phrase then raises to \textit{SpecTP} to satisfy the [\textit{wh}] feature of T. In matrix clauses, this is the final landing site of the \textit{wh}-phrase. In embedded clauses, on the other hand, the matrix verb selects a C specified for [\textit{wh}], so the \textit{wh}-phrase moves again to satisfy the [\textit{wh}] feature of C, this time to \textit{SpecCP}. Whether with matrix or embedded questions, the \textit{wh}-phrase passes through or lands in \textit{SpecTP}, so the subject cannot raise to this position. For \textit{wh}-adjuncts, Toribio proposes that they move directly to \textit{SpecCP}, without passing
through SpecTP like wh-arguments do. Since SpecTP is never occupied by the wh-phrase, the subject is free to raise to SpecTP, though it is not forced to do so. This accounts for the optional availability of preverbal subjects in questions with wh-adjuncts.

Rather than blocking the raising of subject as in Toribio (1993), Goodall (2001) provides an account of inversion in which the wh-phrase satisfies the EPP of T, eliminating the need for the subject to raise to SpecTP. To do this, he proposes that both C and T have a [Q] feature that requires a wh-phrase in their specifier. Once the wh-phrase raises to SpecTP, it satisfies the T’s EPP feature, thus leaving the subject in situ. In order to derive the argument-adjunct asymmetry, Goodall assumes that T does not have a [Q] feature when the wh-phrase is an adjunct because wh-adjuncts are not quantifier-variable binders, as noted by Ordóñez (1997). C still has a [Q] feature, so the wh-phrase moves directly to SpecCP. This permits the subject to optionally fulfill the EPP requirement of T by raising to SpecTP, allowing for an optional preverbal subject.

3.1.3 Inversion in English Wh-Questions

English contrasts with Spanish in that there is broad consensus about the basic mechanisms behind wh-fronting and inversion. Still, as we will see, there are many details and complexities underlying these mechanisms that remained unexplained. We begin with Rizzi (1996)’s account for inversion based on a well-formedness condition known as the Wh-Criterion. According to this account, wh-movement and inversion are driven by semantic requirements on the structure of wh-questions. Following that, we discuss a different approach by Pesetsky and Torrego (2001 et seq.) that brings together a number of interrelated phenomena concerning C and T in both declarative and interrogative clauses. Some examples of these phenomena, and the account they offer, will serve to illustrate the complex ways in which the properties of C and T interact in declarative and interrogative clauses.

3.1.3.1 Rizzi (1996)

Rizzi (1996) proposes the following principle of well-formedness for wh-questions:

\[(\text{30}) \quad \text{The Wh-Criterion} \]

\[a. \quad \text{A wh-operator must be in a Spec-head configuration with } X^0 \text{.} \]
b. A \(X^0 \ [+\text{wh}]\) must be in a Spec-head configuration with a wh-operator.

This criterion applies universally at logical form (LF), and for some languages, it also applies at S-structure (i.e. after movement), among them English and the Romance languages.

Notice that the Wh-Criterion refers to a ‘wh-operator’ and not simply a wh-phrase. This is crucial for accounting for multiple wh-questions such as the following, where one wh-phrase remains in situ:

(31) What did you give to whom?

Rizzi (1996) defines a wh-operator as follows:

(32) wh-operator = a wh-phrase in a scope position (i.e. in a left-peripheral \(\bar{A}\)-position)

In example (31), then, the wh-phrase ‘(to) whom’ does not violate Clause A of the Wh-Criterion, in spite of not being in Spec-head configuration with a head that is \([+\text{wh}]\), because it is not a wh-operator.

And where does the \([+\text{wh}]\) feature come from? There are four possibilities. In English embedded clauses, we can assume that the matrix verb selects a C that is \([+\text{wh}]\). In order to satisfy the Wh-Criterion, a wh-phrase must raise to SpecCP, as in (33a). If the wh-phrase remains in situ, as in (33b), both clauses of the Wh-Criterion are violated.

(33) a. I wonder [who \(i\) [Mary has seen \(t_i]\]]

b. I wonder [[Mary has seen who]]

In English matrix clauses, Rizzi (1996) assumes that \(\text{T}^\text{IP}\) is independently assigned a \([+\text{wh}]\) feature and that it moves to C in order to establish a configuration that can satisfy the Wh-Criterion. This is the source of T-to-C movement. Without it, a wh-operator in SpecCP would not be in a Spec-head relation with a \([+\text{wh}]\) head.

In English, wh-subjects do not occur with inversion, even in matrix clauses. Rizzi (1996) accounts for this by positing that the Wh-Criterion is satisfied as long as the chain of the relevant head has \([+\text{wh}]\).

Consider example (39), a wh-question with a subject wh-phrase.

(34) \([CP \ [\text{Who} \ C_i \ [TP \ t_i \ T_i \ [+\text{wh}] \ [vP \ t_i \ bought \ the \ book]]]]\)

\(^4\)Rizzi (1996) uses I and IP, but I will be substituting T and TP, respectively, throughout this dissertation.
The subject wh-phrase ‘who’ begins in Spec\(v\)P, and then raises to SpecTP, as indicated by the trace, ‘\(t\)’. Notice that the trace has the same index ‘\(i\)’ as the wh-phrase. When the wh-phrase raises to SpecTP, it also becomes co-indexed with T through agreement, giving T an index ‘\(i\)’ as well. Following that, the wh-phrase moves from SpecTP to SpecCP, indicated once again by a ‘\(t\)’. The wh-phrase also undergoes agreement with C, becoming co-indexed with the wh-phrase in the process. Since the wh-subject is co-indexed with C and with T, by the law of transitivity, C is also co-indexed with T. Since T has the [+wh] feature and is co-indexed with C, the Wh-Criterion is met at S-structure.

In summary, to account for the formation of wh-questions, and in particularly wh-movement and inversion, Rizzi (1996) begins with the premise that wh-phrases front in order to satisfy a semantic requirement on the structure of wh-questions. In some cases, this same requirement also forces T to move to C in order to produce the correct structural relation. This contrasts with proposals in Pesetsky and Torrego (2001, et seq.) that are generally focused on the formal properties of wh-questions, among other structures, and the syntactic operations that derive them.

### 3.1.3.2 Pesetksy and Torrego (2001 et seq.)

Pesetksy and Torrego (2001, et seq.) (henceforth P&T) propose a complex system of features and principles to explain the intricacies of wh-questions and related phenomena. As they develop their framework throughout the series of articles, they make a variety of changes to both the proposed features involved and the mechanisms behind how they agree. I will use the most up-to-date proposal, involving feature sharing, from Pesetsky and Torrego (2004).¹⁰

Before illustrating P&T’s framework with a derivation, I will first sketch their approach to agreement, known as feature sharing and defined as follows:

\begin{align}
\text{(35) Agree (Feature sharing version)} \\
\quad a. \text{An unvalued feature } F (\text{a probe}) \text{ on a head } H \text{ at syntactic location } \alpha (F_\alpha) \text{ scans its c-command domain for another instance of } F (\text{a goal}) \text{ at location } \beta (F_\beta) \text{ with which to agree.} \\
\quad b. \text{Replace } F_\alpha \text{ with } F_\beta, \text{ so that the same feature is present in both locations.}
\end{align}

¹⁰There is a more recent paper in this series of articles, but it focuses on issues of complementation rather than the sorts of properties that we are interested in here.
Simplifying a little bit, once features agree, they form a co-indexed chain of now identical features. As long as a feature remains unvalued, it may continue to probe, agreeing with the nearest feature that it hasn’t agreed with, until it agrees with a valued feature. At that point, all of the features in the chain become valued, since they are identical and co-indexed.

One key innovation in their framework is the proposal that any combination of valued/unvalued and interpretable/uninterpretable features is possible.

(36) **Types of features**

\[
\begin{align*}
&uF [\text{val}] \text{ uninterpretable, valued} & iF [\text{val}] \text{ interpretable, valued} \\
&uF [\text{ }] \text{ uninterpretable, unvalued} & iF [\text{ }] \text{ interpretable, unvalued}
\end{align*}
\]

The features in bold are additional combinations not proposed in Chomsky (2000) and Chomsky (2001). This allows interpretable features to be unvalued. Let’s look at why this would be useful. In order for a head \( \alpha \) with feature \( F \) to probe and agree with a head \( \beta \) with feature \( F \), that head must have an unvalued feature. This is what ultimately motives agreement. What happens if the feature on head is naturally a reflection of the semantic content of that feature, i.e. interpretable? In Chomsky (2000, 2001)’s framework, such an interpretable feature would also have to be valued, as those characteristics go hand in hand. What happens if there are reasons to believe that this same head is also a probe, in spite of being valued? \( T \) is a perfect example of this. A tense feature, \( Tns \), on \( T \) would naturally be interpretable, but \( T \) c-commands the head that ultimately receives tense, little \( v \), so \( T \) must be the probe. How can \( T \) be a probe when the \( Tns \) feature is valued? In P&T’s framework, this can be accounted for very naturally by assuming that \( T \) has an unvalued but interpretable feature \( Tns \).

One of the primary goals of the P&T series of articles is to account for a variety of seemingly disparate but ultimately interrelated phenomena concerning the properties of \( C \), \( T \) and case, making every phenomenon a piece of a very interconnected puzzle. Here I will focus on wh-movement and T-to-C movement in order to illustrate the complex relationship between \( C \) and \( T \) and how this relationship

---

11. When representing the features of a head, the initial italicized letter indicates whether the feature is interpretable ‘i’ or uninterpretable ‘u’, and the brackets represent whether or not the feature is valued. Empty brackets indicate that the feature is not valued.

12. To avoid confusion, I will refer to the tense feature in P&T’s work as ‘Tns’ so as to distinguish it from the tense head, \( T \). Note that this is exactly the reverse of the nomenclature P&T use, but it allows me to consistently use ‘\( T \)’ for the tense head throughout the dissertation.
can be accounted for. I’ll begin with a simple matrix wh-question in which the wh-phrase has raised to SpecCP and T has moved to C:

(37)  What did Mary buy?

We will focus here on the fronting of the wh-phrase. P&T propose that C has an interpretable, unvalued $iQ$ feature, indicated by empty brackets, and the wh-phrase has a valued, uninterpretable $uQ[interrogative]$ feature. When C merges with the TP, this yields the (simplified) structure in (38).

(38)  $\langle CP \ C_iQ[\, +EPP\] [TP Mary buy what$_{uQ[interrogative]}]\rangle$

Since C has an unvalued feature, it probes for a valued feature with which it can Agree. C finds the valued $uQ[interrogative]$ feature on the wh-phrase, so they Agree, valuing C’s $iQ$ feature. To motivate movement, P&T use the traditional EPP feature, but with an important change. Rather than being a feature of the head, EPP is a property of a feature, and so any given feature may be [+EPP] or [-EPP]. In this case, C’s $iQ$ feature is [+EPP], so the wh-phrase fronts to SpecCP, as shown in (39).

(39)  $\langle CP \ what$_{uQ[interrogative]} C_iQ[interrogative, +EPP] [TP Mary T [vP v+buy t]]\rangle$

Following that, T is incorporated into C to satisfy a $uTns$ feature on C, proceeding along similar lines. This produces the final structure in (40). Further details of T-to-C movement and its implementation will be discussed below.

(40)  $\langle CP \ what$_{uQ[interrogative]} C_uQ[interrogative] T+i \ [TP Mary tT [vP v+buy twhat]]\rangle$

What did Mary buy?

It is also important to understand other aspects of the relationship between C and T so that any proposal that aims to understand inversion in English is also consonant with these related syntactic phenomena. One such phenomenon is known as the ‘that-trace effect’. Normally, an embedded declarative clause in English may optionally begin with ‘that’, whether the matrix clause is declarative, as in (41a) and (41b), or interrogative, as in (41c) and (41d):

---

13Note that it is C that has the interpretable feature because the clause type, whether interrogative or declarative, is interpreted on C. The fact that wh-phrases are inherently interrogative (in English) is reflected in the fact that their $uQ$ feature is valued.

14Notice that the verb has not incorporated into T here. Like in Spanish, as discussed in §3.1.2.2, verbs in English incorporate into $v$, however, unlike in Spanish, verbs do not raise to T in English.
(41) a. John said [that Mary bought three books].
    b. John said [Mary bought three books].
    c. What did John say [that Mary bought]?
    d. What did John say [Mary bought]?

Notice that in the second pair of sentences, the wh-phrase is extracted from the embedded clause and fronted to the matrix clause. Interestingly, if the wh-phrase that is extracted is a wh-subject, then the presence of ‘that’ becomes ungrammatical:

(42) a. *Who did John say that saw Mary?
    b. Who did John say saw Mary?

Traditionally, ‘that’ in such sentences is considered to represent the complementizer of the embedded clause, meaning that the complementizer is optionally covert in English in such structures. P&T instead analyze the word ‘that’ in such sentences as an instance of T-to-C movement where C+T is pronounced as ‘that’. Notice that we have some kind of interaction between wh-movement of a subject and the expression of C. Let’s see how P&T account for this interaction by contrasting sentence (41c) with a wh-object and sentence (42a) with a wh-subject.

First, we should review the features that the key lexical items have. P&T propose that T has an unvalued interpretable Tns feature \[\text{\textsc{ttns}} \text{[+\text{epp}]},\] both C and any DPs have an unvalued uninterpretable Tns feature \[\text{\textsc{u_tns}}[]\], and little v has a valued uninterpretable Tns feature \[\text{\textsc{u_tns}}[\text{\text{present}}]\]. In addition, the wh-phrase has a \[\text{\textsc{u_q}}[\text{interrogative}]\] feature, as discussed above. Importantly, the C of the embedded clause has an uninterpretable, rather than interpretable, \[\text{\textsc{u_q}}[\text{,+epp}]\] feature. As a consequence, C will still need to agree with the wh-phrase and value its \[\text{\textsc{u_q}}\] feature, but, because the feature is uninterpretable, that [interrogative] value will not be interpreted at LF, making the clause declarative.

Now we are ready to contrast the derivation of (41c) and (42a), repeated as (43a) and (43b) below.

(43) a. What did John say [that Mary bought]?
    b. *Who did John say [that saw Mary]?
The derivation for both sentences begins in the same way, starting with the embedded clause. When T merges with vP, it probes for a Tns feature to agree with. It finds the subject DP, which raises to SpecTP and agrees with T. This is the case whether the subject is a wh-phrase, as in (44a), or a simple DP subject, as in (44b).

\[(44)\]

a. \[TP \, who^uTns[\, ,uQ[interrogative]] \, T^uTns[\, ,+EPP] \, [vP \, t \, v^uTns[past] + saw \, Mary]\]

b. \[TP \, Mary^uTns[\, T^uTns[\, ,+EPP] \, [vP \, t \, v^uTns[past] + bought \, what^uQ[interrogative]]]\]

Because the DP is \(uTns[\, \), Agree forms a chain of co-indexed features, but the chain remains unvalued. Since the probe, T, continues to be unvalued, it probes again and finds finite little v, which is valued. The chain has been valued, and probing stops, leaving T, the DP, and little v with the same valued feature, as seen by the \([past]\) values in bold. This is the case whether the subject is a wh-phrase, as in (45a), or a DP, as in (45b).

\[(45)\]

a. \[TP \, who^uTns[past],uQ[interrogative] \, T^uTns[past, ,+EPP] \, [vP \, t \, v^uTns[past] + saw \, Mary]\]

b. \[TP \, Mary^uTns[past] \, T^uTns[past, ,+EPP] \, [vP \, t \, v^uTns[past] + bought \, what^uQ[interrogative]]\]

Now C merges with TP. C has both a iQ[ ,+EPP] and a \(uTns[ ,+EPP]\) feature, but we will begin with \(uTns[\, ,+EPP]\). Since the \(uTns[\, \) feature is unvalued, C probes. Because both the subject and T now contain an identical, valued Tns feature, either of them has the appropriate features to agree with C. Importantly, P&T consider both the subject DP, in SpecTP, and the T head to be equidistant (cf. Pesetsky and Torrego, 2001 for discussion). Agreement occurs with the closest possible goal, and since the subject and T are equally close, C can agree with either the subject or T. Let’s see what happens in each case.

First we consider what happens if C agrees with the subject DP. Whether or not it is a wh-phrase, the subject raises to SpecCP and forms an Agree relation. Because the subject is in a feature-sharing relationship with T and little v, its \(uTns\) feature is already valued. When C enters in an Agree relation with the feature chain containing the subject, C also becomes valued, leaving C, T, the DP, and little v all with the same value, as seen here:

\[(46)\]

a. \[CP \, who^uTns[past],uQ[interrogative] \, C^uQ[ ,+EPP],uTns[past, ,+EPP] \, [TP \, T^uTns[past, ,+EPP] \, [vP \, t \, v^uTns[past] + saw \, Mary]]\]
b. \[[\text{CP } \text{Mary}_{\text{Tns[past]}}] \text{ C } [uQ[+EPP], uTns[past,+EPP]] \text{ TP } \text{T}_{uTns[past,+EPP]} [\text{VP } t \text{ v}_{uTns[past]} \text{ bought} \text{ what}_{uQ[interrogative]}]||\]

C also has an $uQ[]$ feature, so if the subject is a wh-phrase, C can immediately agree with the wh-phrase in its specifier and value its $uQ[]$ feature. This means that the movement of the wh-phrase to SpecCP has satisfied both the $uQ[]$ and $uTns[]$ features of C:

\[(47) \text{ [CP who}_{uTns[past],iQ[interrogative]} \text{ C } [uQ[interrogative,+EPP], uTns[past,+EPP]] \text{ TP } T_{uTns[past,+EPP]} [\text{VP } t \text{ v}_{uTns[past]} \text{ saw Mary}]]||\]

Turning to the second option for satisfying $uTns[]$ on C, if C agrees with T rather than the subject, then T raises and incorporates into C, valuing the $uTns[]$ of C and satisfying its EPP property. In this type of structure, a T incorporated into C is spelled out as ‘that’, and the verbal material is spelled out in situ (cf. Pesetsky and Torrego, 2001 for discussion). Recall that this contrasts with the standard assumption that the word ‘that’ is a realization of C. When the wh-phrase is an object, this produces the optional structure with an overt ‘that’ as in (41c), repeated below as (48). The derivation of the matrix clause proceeds along the lines described for (37), repeated below as (49a), and this yields the final (simplified) structure in (49b).

(48) What did John say [that Mary bought]?

(49) a. What did Mary buy?

b. \[[\text{CP What did John say } [\text{CP C } [uQ[,+EPP], uTns[past,+EPP]] \text{ TP } T_{uTns[past,+EPP]} [\text{VP } t \text{ Mary}_{uTns[past]} \text{ bought}]]]||\]

Now consider what happens if the wh-phrase is a subject and T incorporates into C. The wh-subject still needs to front to SpecCP in order to value the $uQ[]$ feature on C and satisfy its $[+EPP]$ property. Notice that if the wh-subject fronts to satisfy $uQ[]$ and T incorporates into C to satisfy $uTns[]$, then two separate operations are required in order to satisfy the requirements of $uQ[]$ and $uTns[]$. P&T
propose that structures such as (42a) are ungrammatical because they violate the following Economy Condition:\footnote{The Economy Principle was originally based on the idea that uninterpretable features trigger probing. Here it is updated to refer to unvalued features, the proposed trigger in the newest framework, \textcite{pesetsky_torrego:04}.}

(50) A head H triggers the minimum number of operations necessary to satisfy the properties (including EPP) of its unvalued features.

Since fronting of the wh-subject can satisfy both \(uQ\) and \(iTns\) on C with one movement operation while incorporating T into C forces a second movement operation to front the wh-subject, incorporating T into C violates the Economy Condition and hence is ungrammatical. In this way, P&T derive the that-trace effect.

To summarize, in most embedded declarative sentences in English there are two options to satisfy the \(uT\) feature of C. Either the T incorporates into C and is spelled out as ‘that’, or the subject raises to SpecCP and there is no ‘that’ pronounced. When a wh-object is extracted from such a clause, there continue to be two options, and so ‘that’ may or may not be pronounced. On the other hand, if the wh-phrase is a subject, then the wh-subject not only is able to but must satisfy both the \(uQ\) and \(iTns\) features of C with one movement operation, due to the Economy Condition. Because of this, ‘that’ is ungrammatical with an extracted wh-subject. P&T’s account of the that-trace effect illustrates that a seemingly simply phenomenon, optional ‘that’, can be the result of the complex interaction of various syntactic properties of a language, including extraction of a wh-subject. Importantly, these syntactic properties include properties of C and T in both declarative and interrogative contexts.

3.1.3.3 Conclusions

Though not mutually exclusive, we have seen one approach to inversion in English that is ultimately based on the semantic requirements of wh-questions and another approach that focuses on the formal properties of wh-questions and how these can derive phenomena such as optional ‘that’ and the that-trace effect. Any analysis of wh-movement will need to be able to account for these two distinct sides of wh-questions.
3.1.4 The Role of the Wh-Phrase

Up to this point, we have reviewed the general properties of the syntactic process of wh-question formation, with a particular focus on inversion in Spanish and English, but we must also consider the role that the specific syntactic and semantic properties of a given wh-phrase play in inversion. We will begin by looking at the distinction between simple and complex wh-phrases and some important syntactic differences between these two types of wh-phrases. Following that, we will examine the effect that linking a wh-phrase to previous discourse can have on inversion. This discussion will make it clear that the properties of inversion depend not only on the relationship between C and T but also on the syntactic and pragmatic properties of the wh-phrase.

3.1.4.1 Syntactic Function of the Wh-Phrase

In §3.1.1, we saw that the syntactic function of the wh-phrase affected the availability of inversion in Spanish such that inversion was required with wh-arguments but optional with wh-adjuncts. We also saw in §3.1.2 two different proposals for deriving Spanish inversion, Toribio (1993) and Goodall (2001), and the way that each proposal accounts for the argument-adjunct asymmetry of Spanish wh-questions. Baković (1998) discusses the fact that the availability of inversion can depend not only on whether the wh-phrase is an adjunct but also what kind of wh-adjunct. He distinguishes between three kinds of wh-adjuncts: reason wh-phrases (e.g. why), location wh-phrases (e.g. when, where) and manner wh-phrases (e.g. how). For example, in some dialects of Spanish, inversion is obligatory in matrix wh-questions with wh-arguments and location wh-phrases but not manner and reason wh-phrases. In other dialects, inversion is required in matrix questions with wh-arguments, location wh-phrases, and manner wh-phrase but not with reason wh-phrases. Interestingly, a sort of matrix-embedded asymmetry also emerges when looking at wh-questions with wh-adjuncts. In some dialects, embedded wh-questions are more permissive with inversion in that inversion becomes optional with more types of wh-adjuncts in embedded clauses than when those same wh-adjuncts are in matrix questions. For example, a dialect may permit SV word order with reason wh-phrases in matrix clauses but both reason and manner wh-phrases in subordinate clauses.

\[^{16}\] With a very limited number of exceptions, Baković (1998) does not make clear whether the dialects that he discusses are from specific cities or regions or whether they are individual idiolects of those who participated in his research.
Returning to the two previously discussed accounts of inversion in Spanish, one disadvantage of the Toribio (1993) and Goodall (2001) accounts is that they are unable to explain more fine-grained distinctions between different kinds of wh-adjuncts and the effects that these distinctions have on the availability of inversion. Toribio (1993) proposes that the [wh] feature of T, which forces the wh-phrase to move to SpecTP and the subject to remain postverbal, comes from a wh-argument. With this particular approach, there is no straightforward way to differentiate between the various types of wh-adjuncts without stipulating the differences. In the approach pursued by Goodall (2001), wh-arguments and wh-adjuncts are distinguished by whether they are quantifier-variable binders. It is unclear how to allow different kinds of wh-adjuncts to be quantifier-variable binders based on the dialect, as would be required for this account. Note, however, that I am not aware of any approach for capturing such a difference except those that use Optimality Theory (cf. Prince and Smolensky 1993; Grimshaw 1997) such as Baković (1998)’s account.

3.1.4.2 D-Linked and Non-D-Linked Wh-Phrases

Another property of wh-phrases is their discourse status. More concretely, some wh-phrases are considered to be D(iscourse)-linked because they refer to a previously established set of referents from the discourse. For example, in the following sentence, the answer to the wh-question must be a book from among a set of books that has been established, explicitly or implicitly, in the discourse.

(51) Which book are you going to buy this week?

This is in contrast to a non-D-linked wh-phrase such as in (52) where the answer could be any object that one could logically buy at the store.

(52) What are you going to buy at the store this week?

D-linked wh-phrases have several properties that distinguish them from non-D-linked wh-phrases. For example, Pesetsky (1987) observes that non-D-linked wh-phrases show Superiority Effects, but D-linked wh-phrases do not. We can see this in the contrast between (53a), where the higher wh-phrase has fronted and the sentence is grammatical, and (53b), where the lower wh-phrase has fronted and the sentence is ungrammatical or strongly dispreferred.
(53)  a. Who did you persuade to read what?
       b. ?? What did you persuade who(m) to read?

When the wh-phrases are D-linked, on the other hand, no such contrast occurs, as evidenced by the following sentences:

(54)  a. Which man did you persuade to read which book?
       b. Which book did you persuade which man to read?

Pesetsky (1987) attributes the contrast between (53b) and (54b) to a condition on movement which prohibits two chains of movement from crossing over each other. If we assume that all wh-phrases must move to SpecCP at LF for interpretive reasons, then both (53) and (53b) have two instances of movement. In (55), repeated from above, the wh-subject is already above the wh-object, so when the wh-object moves to SpecCP, it simply passes over the movement chain of the wh-subject. In (56), in contrast, the wh-object first moves above the wh-subject, and then the wh-subject moves above the wh-object at LF, crossing over the movement of the wh-phrase and violating Pesetsky’s condition on movement.

(55)  a. Who, did you persuade \( t_i \) to read what?
       b. Who what \( j \) did you persuade \( t_i \) to read \( t_j \)?

(56)  a. ?? What \( i \) did you persuade who(m) to read \( t_i \)?
       b. ?? who(m) \( j \) What did you persuade \( t_j \) to read \( t_i \)?

So what allows a D-linked wh-object to move in (54b)? Pesetsky (1987) takes the grammaticality of (54b) to be evidence that D-linked wh-phrases do not move at LF. Since there is no second instance of movement, there is no second movement to violate the condition on movement. This is one example of a syntactic difference between D-linked and non-D-linked wh-phrases.

Another difference between D-linked and non-D-linked wh-phrases can be seen in their effect on the availability of inversion in Spanish wh-questions. In contrast to a question with a non-D-linked wh-object such as (57a), a D-linked wh-object allows for SV word order, such as in (57b).

(57)  a. ?? What \( i \) did you persuade who(m) to read what?
       b. ?? who(m) \( j \) What did you persuade t\( j \) to read t\( i \)?
3.1.4.3 Simple and Complex Wh-Phrases

At first glance, it might appear that the difference between simple, one-word wh-phrases such as ‘who’ or ‘what’ and longer, more complex wh-phrases like ‘which book’ lies in their discourse status. This would suggest that the fact that some wh-phrases are complex is epiphenomenal and not meaningful in and of itself. However, consider a wh-phrase such as ‘how many times’ that is clearly complex in the sense that it consists of a wh-word and another constituent. Note that this wh-phrase does not require or suggest any previously established set of referents, meaning that it is not D-linked, in spite of being a complex wh-phrase. The following examples of wh-movement in Spanish-English code-switching also suggest that the distinction between simple and complex wh-phrases is important, even when those wh-phrases are not D-linked.

In (58a), an embedded wh-question with a Spanish verb and VS word order is perfectly grammatical, while in (58b), the same structure and word order is ungrammatical. The key difference is that in

(57)  a. *¿A quién María no quiere ni ver?  
      to who María not want either to see  
      ‘Who does María just not want to see?’

      b. ¿A cuál de estos chicos María no quiere ni ver?  
      to which of these boys María not want either to see  
      ‘Which of these boys does María just not want to see?’

(58)  a. No recuerdo how much money han robado esos delincuentes in the last decade.  VS
      Not I remember has stolen those criminals

      ‘I don’t remember how much money those criminals have stolen in the last decade.’

      b. *No recuerdo what han robado esos banqueros since the last economic crisis.  VS
      Not I remember has stolen those bankers

      ‘I don’t remember what those bankers have stolen since the last economic crisis.’

17 Note that I am only claiming that such phrases are not D-linked by default. Even simple wh-phrases can be D-linked in the right context. See [Pesetsky] for discussion.

18 Unless otherwise indicated, the Spanish-English CS examples cited in this chapter come from the final data set for this dissertation. They do not necessarily represent the judgments of all participant groups, but they are generally representative. For discussion of the participant groups, see [4.6.1.4] and for an example of how a pattern emerges from the variation across groups, see [4.5.2.1].
the wh-phrase is a non-D-linked complex wh-phrase, and in (58b) the wh-phrase is a non-D-linked simple wh-phrase.

Now let’s consider one possible analysis to account for the syntactic differences between simple and complex wh-phrases. Van Craenenbroeck (2010) proposes that the difference lies in the fact that simple wh-phrases move through two landing sites in the left periphery while complex wh-phrases only land in one site. To support this account, he reviews a variety of evidence form different Germanic languages that can be accounted for by assuming that simple, but not complex, wh-phrase move through two specifiers in the left periphery. Consider the case of a special kind of ellipsis known as ‘swiping’ (Merchant 2002).

In the example (59a), we have an embedded wh-question where everything following the wh-phrase may optionally be elided, as indicated by the brackets. The elided constituent is generally considered to be the TP (cf. Merchant 2002). This yields the sentence in (59b).

(59)  a. Mary gave a speech, but I don’t know about what <Mary gave a speech>.

b. Mary gave a speech, but I don’t know about what.

This process is known as ‘sluicing’. Importantly, it is also possible for the preposition and the wh-phrase to swap positions, as seen in (60) below, producing what is known as ‘swiping’.

(60) Mary gave a speech, but I don’t know what about.

This particular structure adds two complications that any account of sluicing needs to address.

First, this structure looks similar to preposition stranding, but notice that if we pronounce the elided part of the sentence, then the elided constituent appears between two pronounced elements, as in example (61).

(61) Mary gave a speech, but I don’t know what [TP <Mary gave a speech> about].

This contradicts the hypothesis that it is the TP that is elided in sluicing, as the stranded preposition is also in the TP. An additional complication comes from introducing a complex wh-phrase, such as in the following example.

(62) * Mary gave a speech, but I don’t know which topic about.
When the simple wh-phrase of previous examples is replaced with an equivalent complex wh-phrase, the sentence becomes ungrammatical. In order to account for these two complications, Van Craenenbroeck (2010) proposes that simple wh-phrases are operators and must move through the specifier of a C[Op] head before moving to a C[wh] head. Complex wh-phrases, which are not operators, either move directly to the C[wh] head or are base-generated there. Instead of being satisfied by the movement of a wh-phrase, the [Op] feature of the C[Op] head is satisfied by the movement of a null operator. The contrast between the derivation of these two types of wh-phrases is illustrated below.

Figure 1: Simple wh-phrase

Figure 2: Complex wh-phrase

In Figure 1, the simple wh-phrase has a [Q] and a [Op] feature, so it raises to the specific of C[Op] to check its [Op] feature and then to the specific of C[wh] to check its [wh] feature. On the other hand, in Figure 2, the complex wh-phrase only has a [wh] feature, and therefore it does not raise to the specifier of C[Op]. In order to account for some of the wh-phenomena that he reviews, Van Craenenbroeck (2010) 19

---

19See below for further discussion.
argues that the complex wh-phrase is based-generated in the specifier of C_{wh}, rather than raising there from an A-position in TP. In order to maintain the sensitivity to islands that complex wh-phrases exhibit, which suggests that some kind of movement is involved, he proposes that a null operator is raised to C_{Op}, as seen in Figure 2. This also checks the [Op] feature on C_{Op}. In order to account for some reconstruction effects, he also considers the possibility that it is instead the complex wh-phrase that moves and the null operator that is base-generated. We will not be concerned with which of these two proposals is correct. The key insight is the featural differences between simple and complex wh-phrases.

With the proposal for simple and complex wh-phrases outlined above, we are now in a position to consider how Van Craenenbroeck (2010) accounts for swiping. Recall that the issue was how the preposition could be stranded and yet not be elided with the rest of the TP. Van Craenenbroeck (2010) proposes that simple wh-phrases pied-pipe the preposition when they raise to SpecC_{Op} and then strand the preposition when they raise to SpecC_{wh}, as illustrated in Figure 3, focused just on the embedded CP. The complete sentence can be seen in (63).

Figure 3: Swiping example

```
[wh]_i
[+Q, +Op]

C_1'

C_1
[+Q]

[PP about what]

CP_2

C_2
[+Op]

TP

Mary gave a speech [PP about what]
```
(63) Mary gave a speech, but I don’t know what about <Mary gave a speech>.

Because complex wh-phrases do not move through SpecC_{Op}, whether they are base-generated in or move to SpecC_{wh}, there is no way for them to strand a preposition in SpecC_{Op} and hence no way for them to derive swiping. In this way, Van Craenenbroeck (2010) can straightforwardly account for swiping by appealing to just two potential C features. He also reviews a number of other phenomena in a couple of other Germanic languages, and in each case, his proposal for the features of simple and complex wh-phrases allows him to straightforwardly account for the phenomenon. This, in turn, serves as strong support for his proposal on the syntactic differences between simple and complex wh-phrases.

3.1.5 Conclusions

In this section we have reviewed a variety of key properties of C, T, and wh-phrases and their relationship to the formation of wh-questions in Spanish and English. We have seen that inversion in Spanish can be accounted for by assuming that preverbal subjects and wh-objects occupy the same position, SpecTP. According to Toribio (1993), a wh-object in SpecTP blocks the raising of the subject to SpecTP, while Goodall (2001) proposes that the wh-object in SpecTP satisfies the EPP of T, leaving no trigger for movement of the subject. In either case, the subject remains in situ, deriving inversion. In our discussion of inversion in English wh-questions, we saw that inversion can be analyzed by focusing on semantic requirements, as in Rizzi (1996) or syntactic requirements, as in P&T, though both types of requirements must ultimately be met. In addition, in our discussion of P&T, we also saw the way that one phenomenon, optional ‘that’ in English, can interact with various properties of C and T in both interrogative and declarative clauses. Finally, we reviewed the role that various properties of wh-phrases play in the formation of wh-questions, including the syntactic function, whether they are D-linked, and whether they are simple or complex.

3.2 Code-Switching

Any theory purporting to account for wh-questions must account for two central features of wh-questions. First, it must have an analysis for movement, specifically A-movement. Second, it must account for syntactic phenomena that co-occur with wh-movement in some languages, such as inversion. As I review
the literature on code-switching, I will consider whether each theory of code-switching has something to say about these two key aspects of wh-questions.

I will begin my review of code-switching theories by discussing two early, influential approaches to analyzing code-switching, [Poplack (1980)] and [Joshi (1985)]. While neither of their approaches is generative, the framework that my own work is based on, both approaches are formal in nature. In other words, they are interested in grammatical or structural constraints on possible code-switched sentences. Although it will turn out that neither proposal has much to say about wh-movement in general, as they do not address movement of any sort, they both make specific predictions that my own data are able to address.

Following that, I will review a number of the most influential generative approaches to code-switching, including [Woolford (1983, 1984), Di Sciullo et al. (1986), Belazi et al. (1994), and Mahootian and Santorini (1996)]. Many, though not all, of these proposals will also have little to say about wh-movement in general, again because they do not focus on movement. Nevertheless, all of them have at least some data from my dissertation experiment which they are unable to account for.

Finally, I will turn to [MacSwan (1999)]'s proposal for a minimalist approach to code-switching and subsequent work in that vein. In some ways, these proposals will turn out to have the opposite problem of previous generative work. On the one hand, they will have general predictions about wh-movement in code-switching, as they offer the tools to deal with movement. On the other hand, in some cases, they will be unable to make specific predictions concerning my data because they focus on phenomena not related to wh-movement.

3.2.1 Early Formal Approaches to Code-Switching

3.2.1.1 Poplack (1980)

[Poplack (1980)] made the first influential proposal of a formal constraint on code-switching. Of interest here is the Equivalence Constraint:

(64) Code-switches will tend to occur at points in discourse where juxtaposition of L1 and L2 elements does not violate a syntactic rule of either language, i.e. at points around which the surface structures of the two languages map onto each other.
Poplack (1980) illustrates her constraint by lining up two monolingual sentences with a Spanish-English code-switched sentence, as in (65).  

(65) I told him that so that he would bring it fast.

(Yo) le dije eso pa’ que (él) la trajera ligero

I told him that pa’ que la trajera ligero.

The first line is the English translation of the sentence, the second line is the Spanish translation of the sentence, and the third line is the actual CS sentence. The order of the constituents in the code-switched sentence matches the order of the same constituents in the monolingual equivalent sentence. In other words, the code-switched and monolingual constituents are equivalent in meaning and, in her example, the syntactic rules for English have applied to the English words and the syntactic rules for Spanish have applied to the Spanish words to produce their respective monolingual word orders.

Because the Equivalence Constraint is ultimately about linear word order, there are limitations to the types of predications that it can make about wh-movement. For example, it would have nothing to say about the intermediate positions of successive cyclic movement. Still, the Equivalence Constraint does make two predictions for wh-movement in Spanish-English code-switching. First, if the wh-phrase is fronted and there is inversion when appropriate, the sentence should be grammatical, as all constituents would be equivalent in meaning and correctly placed according to their respective syntactic rules. The following sentence, originally from Woolford (1984), shows that this prediction does not hold.

(66) *When vino él?

came he

‘When did he come?’

The wh-phrase is fronted, which is grammatical in both English and Spanish, and the Spanish verb and subject have been inverted as required in a Spanish wh-question. Yet, the sentence is ungrammatical, contra the Equivalence Constraint.

The second prediction is that sentences in which syntactic requirements on inversion in one language are applied to words from another language will be ungrammatical. For example, in (67), the subject

---

20 Following a convention for reporting code-switching data, one of the two languages is always in italics, to highlight where the code-switches take place. Spanish will always be the italicized language throughout the dissertation.

21 While inversion is not required with wh-adjuncts, it is still an option, so application of this optional rule should still render this sentence grammatical.
and the verb of the embedded question are in Spanish, but they have not been inverted as required in Spanish wh-questions.

(67) I don’t remember what tus colegas han comprado this week.
     your colleagues have bought
     ‘I doesn’t remember what your colleagues have bought this week.’

In spite of the lack of inversion, this sentence was still judged grammatical by my participants, once again violating the Equivalence Constraint. Although Poplack’s constraint does make reference to syntactic rules and therefore potentially structure, notice that it is ultimately a linear constraint.

3.2.1.2 Joshi (1985)

Joshi (1985) proposes a structural approach to constraining code-switching based on the following four assumptions:

(68) a. In any example of CS, there is one language which is perceived as the language the sentence is "coming from", called the matrix language (ML), and another language, called the embedded language (EL)

b. You can only switch from the ML to the EL.

c. Closed-class items such as determiners, prepositions, and tense cannot be switched.

d. It is possible to switch at the level of maximal projections rather than just heads.

The last assumption implies that a determiner, for example, can still be in the EL if the whole determiner phrase is switched, as long as no element in the DP is later switched back to ML. This would violate the assumption that it is only possible to switch from the ML to the EL.

Because Joshi (1985) is interested in processing and not competence, his theory has nothing specific to say about movement of any kind, including wh-movement.22 Even so, his proposal does make predictions about some wh-questions. One prediction is that tense cannot be switched to the embedded language. Let’s see why. Tense is a closed class item, so while the verb itself can be switched to the embedded language if tense is separate (i.e. on a light verb), a tensed verb cannot be switched to the

22However, Goodall (2004, 2010) has investigated the effect that processing can have on the availability of inversion in Spanish wh-questions.
embedded language, as that entails a switch of tense itself and closed-class items cannot be switched. Unfortunately, this prediction does not hold for the data I have collected. Since it is not clear what might be the ‘matrix’ language in Spanish-English CS, I offer counterexamples switching from a Spanish matrix verb to an English embedded verb, (69a), and from an English matrix verb to Spanish embedded verb, (69b):

(69) a. *No recuerdo qué* those rabbits have eaten *desde que planté el jardín.*
    not I remember what since that I planted the garden
    ‘I don’t remember what those rabbits have eaten since I planted the garden.’

    b. I don’t remember what *tus colegas han comprado* this week.
       your colleagues have bought
    ‘I doesn’t remember what your colleagues have bought this week.’

Whatever the ‘matrix’ language in these sentences is, the sentences are grammatical in spite of the fact that the matrix and embedded verbs, which are closed-class items, are in different languages.

Even though Joshi (1985)’s proposal does not make general predictions about how wh-questions can be derived in code-switching, it offers a very precise proposal for structural constraints on code-switching.

### 3.2.2 Generative Approaches to Code-Switching

#### 3.2.2.1 Woolford (1983, 1984)

Woolford (1983) proposes a model of code-switching in which code-switched sentences are generated in the same way that monolingual sentences are, the only difference being that the grammars of two languages are available to generate any given part of the sentence. In other words, there are no code-switching-specific rules but only general syntactic processes for building sentences. The work in Woolford (1983) characterizes the grammars of the languages in terms of phrase structure rules, but it has little to say about movement, i.e. transformations in the then current work in generative linguistics.

Woolford (1984) further explores the model proposed in Woolford (1983) by investigating wh-movement in Spanish-English code-switching. Like phrase structure rules, she proposes that transformations that form part of each language’s grammar apply in the same way that they do in monolingual sentences. This means that they are triggered in the same way and have the same restrictions. Woolford (1984) focuses on one specific transformation to test her model: subject-verb inversion in Spanish wh-
questions. In particular, she looks at embedded wh-questions where there is an inversion transformation in Spanish but no such transformation in English. Following Torrego (1984), Woolford (1984) assumes that it is a wh-word in SpecCP (my terminology) that triggers the inversion. This assumption leads to the prediction that embedded questions with a wh-word in Spanish will permit or require inversion, depending on the syntactic function of the wh-phrase, and embedded questions with an English wh-word will not allow inversion.

This prediction does not hold for a number of sentences from my data. In (70), the wh-phrase in the embedded clause is in Spanish, but there is no inversion in the embedded clause. In spite of the fact that the required inversion transformation has not taken place, the sentence is still grammatical.

(70) No recuerdo cuántas manzanas tus hijos han comido desde el domingo.
not I remember how many apples your children have eaten since the Sunday
‘I don’t remember how many apples your children have eaten since Sunday.”

If we extend Woolford (1984)’s predictions to matrix wh-questions, more problems with the analysis emerge. First, we will quickly review some key facts concerning matrix wh-questions in Spanish and English. Both languages require inversion in most matrix questions, but the type of inversion is different. In Spanish, the verb and auxiliary, when present, both precede the subject. In English, on the other hand, only the auxiliary and subject invert, leaving the subject between the auxiliary and the participle. If the Spanish wh-phrase triggers inversion, then it would be expected to trigger a Spanish-style inversion of the entire verb complex. Likewise, it would be predicted that an English wh-phrase would trigger subject-auxiliary inversion. Once again, the predictions do not hold. When the wh-phrase is in English, the sentence is grammatical with Spanish-style inversion, as in (71a), but ungrammatical with English-style inversion, as in (71b). This is the reverse of what was predicted.

(71) a. How many blouses han comprado tus sobrinas this fall?
   ‘How many blouses have your cousins bought this fall?”

   b. *How many blouses han tus sobrinas comprado this fall?
      have your cousins bought
   ‘How many blouses have your cousins bought this fall?”

---

23 See §3.1.1 for additional discussion.
We see, then, that the language of the wh-phrase does not always predict position of the subject in Spanish-English wh-questions.

While using her model with the specific assumptions that Woolford (1984) makes leads to incorrect predictions, the insight that code-switching can be analyzed using only the rules of the monolingual grammars involved is a significant one.

3.2.2.2 Di Sciullo, Muysken, and Singh (1986)

Subsequent research into code-switching using a generative perspective began to focus on specific properties or relations within the computational system. Di Sciullo et al. (1986) make use of the structure relation of government and the idea that lexical items in a sentence are indexed to indicate which lexicon they came from. Specifically, they propose that a governed category must be in the same language as its governor.

In order to make their proposal more precise, they have to introduce a number of constructs and formalisms. First, their analysis relies on the notation of a language index in which every word drawn from the lexicon is marked according to the language it comes from. Second, under their definition of government, heads govern maximal categories, so they develop a mechanism for determining the language index of a particular maximal projection. Essentially, the language index of a maximal projection is determined by the language of the highest lexical element of that projection, and this lexical element is called an $L_q$ carrier. More formally, then, the government constraint can be stated as follows:

(72) Government Constraint: the $L_q$ carrier of a governed category must have the same $L_q$ index as its governor.

Di Sciullo et al. (1986) state that this constraint applies at S-structure, the equivalent of saying that it applies after movement.

In the case of wh-movement, this means that the $L_q$ carrier of any particular maximal projection will be determined after the wh-phrase has fronted and any required inversion has occurred. Since $V$ is a governing category, this would predict that the language index of a matrix verb would have to

---

2Di Sciullo et al. (1986) define government in the following manner: “$X$ governs $Y$ if the first node dominating $X$ also dominates $Y$, where $X$ is a major category N, V, A, P and no maximal boundary intervenes between $X$ and $Y$.”
match the language of the fronted wh-phrase in an embedded question, as the wh-phrase would be the highest lexical element of the CP projection. In the following example, the language of the matrix verb is Spanish, but the language of the wh-phrase in the embedded question is English. In spite of this, the code-switched sentence is grammatical, according to my data.

(73)  
No recuerdo how much money esos delincuentes han robado in the last decade. 
not I remember those criminals have stolen 
‘I don’t remember how much money those criminals have stolen in the last decade.’

While Di Sciullo et al. (1986)’s proposal does make concrete predictions about wh-movement, it does not have anything to say about wh-questions more generally. For example, it makes no predictions about the availability of inversion in a given code-switched sentence based on the language of the verb, in spite of the fact that we might expect the language of the verb to be relevant. In addition, from a Minimalist perspective, in which a sentence is built up in various stages, it seems unlikely that the only relevant point in the derivation would be the final stage. Consider the case of postverbal subjects in wh-questions in Spanish. If position of the subject is due to a wh-phrase moving through SpecTP before landing in SpecCP, then that intermediate stage of wh-movement is crucial to deriving inversion. Di Sciullo et al. (1986) could not predict the outcome of having an English wh-phrase or subject in such a structure.

Although Di Sciullo et al. (1986) cannot fully account for wh-questions in code-switching, it does offer a crucial insight: the fundamental properties of the computational system help determine possible code-switches. We will see that one such property, phases, can play an important role in the structure of code-switched sentences, as argued by González-Vilbazo and López (2011, 2012).

3.2.2.3 Belazi, Rubin and Toribio (1994)

Belazi et al. (1994) also assume that lexical items are marked for the language they come from, but they propose a much simpler mechanism to constrain code-switching: selection. Specifically, they use Abney (1987)’s idea of f-selection: a functional head selects the features of its complement. Combining this with a language index (which they call a language feature), they propose the Functional Head Constraint (FHC):
(74) The language feature of the complement f-selected by a functional head, like all other relevant features, must match the corresponding feature of that functional head.

Because the Functional Head Constraint says nothing about code-switches involving only lexical heads, Belazi et al. (1994) also propose the Word-Grammar Integrity Corollary:

(75) A word of language X, with grammar $G_x$, must obey grammar $G_x$.

In order to evaluate the predictions that Belazi et al. (1994) make with respect to wh-questions, we need to look at what their proposal has to say about C and T, since both of these heads are involved in the formation of wh-questions (depending on the analysis). Belazi et al. (1994) argue that T, as a null functional head, is unmarked with respect to the language feature when it enters the derivation, so its language feature does not (and cannot) conflict with its complement. Consequently, the FHC is not violated up to this point in the derivation. Once the subject is in SpecTP, however, it undergoes spec-head agreement with T, so T acquires the language feature of the subject. Essentially, this means that TP is in the language of the subject DP. By the FHC, C must be in the same language as TP and therefore in the same language as the subject in SpecTP. Consider example (76a). The embedded subject is in Spanish, so the embedded TP should also be in Spanish. Notice, though, that C, which is realized as ‘that’, is in English. According to the FHC, (76a) should be ungrammatical because C and its complement, TP, are in different languages, but this sentence was found to be grammatical in data from Hoot (2011).

(76) a. What did John say that el jefe compró?
   the boss bought
   ‘What did John say that the boss bought?’

   b. John said that el jefe compró el libro.
      the boss bought the book
      ‘John said that the book bought the book.’

González-Vilbazo (2007) has pointed out that similar sentences such as (76b) may be two separate sentences (or predicates, at least), in which case there really is no switch between C and its complement, since they are not in the same sentence. Notice, however, that example (76a) cannot be analyzed in the same way, as the wh-phrase has moved out of the lower CP, a type of syntactic dependency that does not occur across sentences.
Returning to the Word-Grammar Integrity Corollary, let’s consider what predictions it would make for the position of the subject in a Spanish-English wh-question. To do that, we first need to ascertain what determines the position of the subject, i.e. whether it appears preverbally or postverbally. Looking at Spanish, if we assume an analysis under which a subject is forced to be postverbal when a wh-object passes through SpecTP (cf. §3.1.2.3), then the position of the subject is really due to two factors. First, T must have a [wh] feature which attracts the wh-object. Following Toribio (1993), this feature comes from T undergoing feature checking with the wh-object, something that is specific to the grammar of Spanish, as evidenced by the fact that Spanish, but not English, has an argument-adjunct asymmetry. The second factor is that T in Spanish is able to assign case to an in situ subject, something that is not the case for English, for example. We see, then, that the particular properties of Spanish T require a postverbal subject in questions with a wh-argument. Importantly, if we just consider the subject itself and whether it obeys the grammar of its language of origin, then the position of the subject does not affect grammaticality of the sentence because that position is due to external factors, namely the properties or position (in English inversion) of T. In order to test the predictions of the Word-Grammar Integrity Corollary for inversion, then, we need to look at the language of T and the subject position that the grammar of that language requires. For example, this constraint would predict that a Spanish subject in an embedded question with a wh-object would be postverbal, since in a monolingual Spanish sentence such subjects are always postverbal due to subject-verb inversion. Now consider the following sentence.

(77) I don’t remember what tus colegas han comprado this week.
     your colleagues have bought
     ‘I doesn’t remember what your colleagues have bought this week.’

Notice that the subject in the embedded clause is in Spanish and in preverbal position and the wh-phrase is an argument, and yet the sentence remains grammatical. Therefore, the Word-Grammar Integrity Corollary does not fully predict the grammaticality of code-switched sentences.

Like Woolford (1984), Belazi et al. (1994)’s proposal cannot account for the Spanish-English wh-questions in my data based on the specific assumptions of their original proposal. On the other hand, the general approach offers an important insight, specifically, the intuition that selection plays a key role in the derivation of code-switched sentences.
3.2.2.4 Mahootian and Santorini (1996)

In Belazi et al. (1994), the relationship between a head and its complement is important to the extent that their language features must match. Mahootian and Santorini (1996) broaden this relationship to include other properties of the complement. In particular, they propose that “[h]eads determine the syntactic properties of their complements in code-switching and monolingual contexts alike.” While Mahootian and Santorini (1996) focus on word order as determined by the head, they do point out that heads also determine other properties of their complements, such as syntactic category and finiteness. Still, the proposal offers no criterion to determine exactly what properties the head can determine nor how far into the complement’s structure the influence of the head can go.

Under one interpretation, the matrix verb of an embedded question, being a head, should determine the formation of the embedded wh-question, being its complement. If that is true, then any grammatical sentence where the position of the subject does not match that predicted by the language of the matrix verb will serve as a counterexample. Take, for example, the following code-switched sentence from my data:

(78) No recuerdo cuántas manzanas your children have eaten desde el domingo.
    I don’t remember how many apples since the Sunday

In this sentence, the matrix verb is in Spanish, so the embedded wh-question should have the properties of a Spanish embedded wh-question. That is to say, the wh-phrase should be fronted, and there should be subject-verb inversion. While the wh-phrase has fronted in (78), there is no inversion, which is not what Mahootian and Santorini (1996) would predict.

Once again, while specifics of this proposal fail on empirical grounds, Mahootian and Santorini (1996)’s proposal still offers a significant insight: the potential importance of the relationship between a head and its complement. While this relationship cannot be the only factor which determines the grammaticality of code-switched sentences, as illustrated by (78), it may still play a role in understanding how code-switched wh-questions are formed. For example, in the following sentence, the matrix verb is in English, and there is no inversion in the embedded wh-question, exactly as Mahootian and Santorini
This is in spite of the fact that the wh-phrase and the verb complex in the embedded wh-question are both in Spanish, a language that does require inversion in embedded wh-argument questions.

(79) I don’t remember cuántas invitaciones my parents han enviado in the last week.
     how many invitations have sent
     ‘I don’t remember how many invitations my parents have sent this week.’

3.2.3 Minimalist Approaches to Code-Switching

3.2.3.1 MacSwan (1999)

MacSwan (1999) proposes a research program for code-switching based on the following thesis: “Nothing constrains code switching apart from the requirements of the mixed grammars.” This sounds surprisingly similar to Woolford (1984): “Nothing more than the rules of the two monolingual grammars can be used to explain constraints on code-switching”, but with a very important difference. In Woolford (1984)’s approach, there really are two distinct, language-specific grammars, and code-switching is constrained by the ways in which the rules of those grammars can interact. MacSwan (1999), on the other hand, proposes a minimalist approach to code-switching in which syntactic derivations are built by one computational system, rather than individual, language-specific grammars.

MacSwan (1999)’s approach to CS is different from previous approaches in two important ways. First, it adopts the newest approach to generative linguistics, specifically Minimalism. In doing so, it automatically incorporates both the conceptual and the empirical advantages of the framework. Second, it turns the focus of grammaticality to the features and their interaction in the computational system ($C_{HL}$). While the use of features (cf. Belazi et al., 1994) and the reliance on mechanisms of the computational system is not new (cf. Woolford, 1983; Mahootian and Santorini, 1996, for discussion), the shift to a focus on features is an important one. The degree to which any two languages can be code-switched is not based on a specific feature or small set of features but rather all of the features of both languages. Put another way, it is not the languages as a whole, or their respective ‘grammars’, that interact in the derivation of code-switched sentences but rather all of the individual features of those languages.

\(^{25}\)Exceptionally, this example is from a series of pilot experiments conducted before undertaking the final dissertation experiment (cf. \cite{11} for discussion). None of the stimuli from the final experiment had both the wh-phrase and the verb in the same language.
Being more of a programmatic choice rather than a specific set of constraints, it can be difficult to evaluate predictions of a particular Minimalist approach to code-switching with respect to a specific syntactic phenomenon. For example, MacSwan (2000) makes a specific proposal using a Minimalist approach, but the proposal is concerned with the Phonetic Form (PF) component of human language. Since my work focuses on a largely syntactic phenomenon, wh-movement, it is hard to develop and test predictions from this specific proposal. In particular, MacSwan (2000)'s is interested in switches below the word level, something that no analysis of wh-movement is likely to touch on.

3.2.3.2 Toribio and González-Vilbazo (Forthcoming)

Toribio and González-Vilbazo (forthcoming) observe that the subject-verb/auxiliary inversion that occurs in wh-questions in Spanish and English\(^{27}\) also occurs in negative fronting such as in (80a) and (80b).

\[
(80) \begin{align*}
\text{a. Ni } & \text{ un alma vio Juan en el cementerio.} \\
& \text{not even a soul saw Juan en the cemetery} \\
& \text{‘Not a soul did John see at the cemetery.’}
\end{align*}
\]

\[
\text{b. Not a soul did John see at the cemetery.}
\]

Given this general pattern of fronting which sometimes requires inversion, they propose a more general type of movement called operator movement, and they layout an analysis for how this movement is derived based on feature checking and specific checking domains.

While the article begins with a detailed account of wh-question formation in a Government and Binding framework, the details of exactly how the derivation of a wh-question proceeds become less clear as the analysis is adapted for a Minimalist framework. The important insight from Toribio and González-Vilbazo (forthcoming) is that feature checking combined with language-specific feature specifications has the potential to account for the grammaticality of operator movement constructions in monolingual and code-switched sentences alike.

While an approach based on feature checking can clearly be fruitful, there is also evidence suggesting that the proposed combination of features and mechanisms for feature checking are not sufficient to fully

\(^{26}\)For an exceptional case in which the possibility of a switch below the word level may matter, see discussion in §5.2.3.

\(^{27}\)They also look at German and German-Spanish code-switching, but I will not be discussing those data or analyses here.
account for all code-switched sentences. For example, Toribio and González-Vilbazo (forthcoming) state that the features of C and T+V of the clause from which the wh-phrase originates determine the pattern of subject-verb inversion. Consider example (70), repeated here as (81a), which contains an embedded wh-question with a Spanish wh-phrase and an English verb complex. The subject of the embedded question is not inverted, exactly as the language of the verb, English, would predict. Notice that this is not what would be predicted by the language of the wh-phrase, which is in Spanish, but this is precisely what Toribio and González-Vilbazo (forthcoming) would predict. It is the features of C and T+V and not the features of the wh-phrase that determine the availability of inversion.

(81) a. No recuerdo cuántas manzanas tus hijos han comido desde el domingo.  
not I remember how many apples your children have eaten since Sunday  
‘I don’t remember how many apples your children have eaten since Sunday.’

b. No recuerdo cuántas invitaciones mis padres han enviado en los últimos días.  
not I remember how many invitations my parents have sent in the last week.  
‘I don’t remember how many invitations my parents have sent this week.’

In example (79), repeated here as (81b), on the other hand, the lack of inversion is unexpected. Let’s look at the problem in detail. In order to derive the argument-adjunct asymmetry of Spanish, Toribio and González-Vilbazo (forthcoming) propose, following Toribio (1993), that wh-arguments agree with T (my terminology), giving T a [wh] feature. The wh-phrase then moves to Spec,T because of the [wh] feature on T, preventing the subject from raising to SpecTP. In this way, the subject remains post-verbal, deriving subject-verb inversion. In (81b), the auxiliary is in Spanish, strongly suggesting that T is also in Spanish, due to incorporation. Notice that the wh-phrase, which is an argument, is also in Spanish. Since they are both in Spanish, it seems safe to assume that they have the features of Spanish, among others a [wh] feature on the wh-phrase that agrees with the T. For the reasons described above, this should cause there to be inversion in (81b), and yet no inversion has occurred. More importantly, in spite of the lack of the apparently required inversion, the sentence is still grammatical. This suggests that there are more features involved in the derivation of inversion than previously thought, although it may also be that the features are correct but the way they undergo agreement is not. I believe that the approach is on track but that there are still many details to discover and refinements to make.
3.2.3.3 González-Vilbazo and López (2011, 2012)

In their papers on little v in code-switching, González-Vilbazo and López (2011, 2012) develop an account for some puzzling data from light verb constructions in German-Spanish code-switching using feature checking, among other syntactic tools. For example, consider the contrast between (82a) and (82b). Both examples use a light verb construction ‘do sow’ to express the verb ‘sow’, but only the example with the light verb in Spanish (and the main verb in German) is a grammatical code-switch.

(82) a. *Juan **hace** nähen das Hemd.  
     Juan does.3rd sew.inf the shirt  
     ‘Juan sews the shirt.’  
     
     b. *Juan **tut** coser una camisa.  
     Juan did.3rd sew.inf a shirt  
     ‘Juan sewed the shirt.’  

Much of the papers focus on understanding the origins of the light verb constructions and puzzling properties such as asymmetry, an absence of these light verb constructions in the input grammar, word order, prosody, and focus/background expression. While these phenomena are interesting and, importantly, strongly support a Minimalist approach to code-switching, they do not directly relate to the formation of wh-questions.

Still, the syntactic tools that they use to account for their data do have important implications for any analysis of wh-movement in Spanish-English code-switching. Like Toribio and González-Vilbazo (forthcoming), features play an important role in their analysis of a complex syntactic phenomenon, in this case light verbs in code-switching. This suggests that features and feature checking may be a fruitful way to analyze wh-questions in code-switching, another complex syntactic phenomenon, as well.

While feature checking itself is too broad to make specific predictions about wh-movement in code-switching, a more concrete proposal concerning the role of phase heads does make a few predictions. González-Vilbazo and López (2011, 2012) base their proposal on the following assumption:

**Phase Head Hypothesis (PHH)**

The phase head determines grammatical properties of its complement.
This has important implications for phases headed by little v, which they explore in depth, as well as phases headed by C. I will briefly examine each phase head in turn in terms of its potential effects on wh-question formation.

Since both wh-fronting and inversion occur in the C-T phase, it would seem that little v would have no direct influence on wh-question formation. Still, the language of little v may have a more indirect effect on later phases if head-to-head movement or incorporation requires all the heads to be in the same language. If that were the case, this would limit the ways in which wh-questions could be derived. For example, in Spanish, where V raises to little v and then little v raises to T, all three heads would have to be in the same language.

In terms of the C phase, feature spreading would predict that a wh-question should behave like wh-questions behave in the language of C. For example, sentences in which an English C is the head of an embedded CP, would be expected to lack inversion, as in English embedded questions. If it is possible to code-switch between C and T in some circumstances, as example (76a), repeated below as (83), suggests, this would make an interesting prediction.

(83) What did John say that el jefe compró? 
the boss bought

‘What did John say that the boss bought?’

Essentially, inversion should occur according to the language of C, no matter what language the rest of the sentence is in. In addition, if the language of C determines not only whether or not there is inversion but also exactly how that inversion takes place (e.g. as T-to-C movement or a post-verbal subject), then it should be possible to have Spanish verbs undergoing T-to-C movement. The circumstances under which that prediction holds true, or if the prediction does not hold true under any circumstances, might have consequences for how inversion precedes in the two languages, particularly for Spanish where the exact mechanism that produces inversion is much more controversial. We will discuss these possibilities in more detail in the analysis section.

3.2.4 Conclusions

Much progress has been made since Poplack (1980) in refining the tools used to account for the grammaticality of code-switched sentences, but significantly less progress has been made on tools that lend
themselves to an analysis of wh-movement in particular. Fortunately, with the emergence of Minimalist approaches to code-switching, this is significantly less problematic. Because there is only one computational system driven principally by featural content, the task, while still as daunting as any other syntactic problem, is at least no more daunting than the parallel problem in monolingual contexts.

3.3 Hypotheses

We have seen that many questions remain concerning the nature of wh-questions in both Spanish and English. In particularly, it is clear that inversion is the result of a process involving the complex interaction of many different factors, but the details of those interactions remain unsettled. A better understanding of these issues has the potential to provide insight into properties of each individual language as well as the morphosyntax of wh-questions across languages.

We are now at a point where we can return to the research questions laid out in the introduction and consider some possible hypotheses. The first research question is the following:

RQ1: What factors determine the word order in wh-questions in Spanish-English code-switching?

We will review each research sub-question in turn and consider a hypothesis aimed at answering that question. We begin with the first sub-question for RQ1, focused on the subject, and propose the following hypothesis:

RQ1-A: Do the syntactic properties of the subject determine which subject positions are grammatical?

H1-A: No, the syntactic properties of the subject will not ultimately determine which subject positions are grammatical.

No theory that I am aware of gives the subject a determinative role in the position of the subject, so we will adopt the null hypothesis and assume that the subject does not ultimately determine the grammatical subject positions. Still, it is worth testing this possibility because it is a reasonable logical
possibility. We are talking about the position of the subject, so it is possible that the subject plays a central role in determining where and when it moves.

Next, consider the second sub-question that we proposed concerning the properties of the wh-phrase:

**RQ1-B:** Do the syntactic properties of the wh-phrase determine which subject positions are grammatical?

We have seen evidence that a variety of properties of the wh-phrase are important in the formation of wh-questions in both Spanish and English. Still, no theory that I am aware of gives the wh-phrase a determinative role in the position of the subject. Nonetheless, it is worth testing this possibility because the wh-phrase clearly plays an important role in the formation of wh-questions. This leads me to propose the following hypothesis:

**H1-B:** No, the syntactic properties of the wh-phrase will not ultimately determine which subject positions are grammatical.

The third sub-question centers on the T head, and I propose a hypothesis based on the relationship between T and inversion:

**RQ1-C:** Do the syntactic properties of the tense head determine which subject positions are grammatical?

**H1-C:** Yes, the syntactic properties of the tense head will ultimately determine which subject positions are grammatical.

Given the central role that T plays in some accounts of inversion in Spanish (e.g. Toribio 1993, Goodall 2001) and the role that T plays in T-to-C movement, it is entirely plausible that T determines the grammatical subject positions in at least one of the two languages.

Finally, the fourth sub-question for RQ1, RQ1-D, focuses on the complementizer head. I base my hypothesis for RQ1-D on the following proposal by González-Vilbazo and López (2012), known as ‘Phase Head Hypothesis’:
Phase Head Hypothesis (PHH)

The phase head determines grammatical properties of its complement.

Thus, in answer to RQ1-D, I propose the following hypothesis

RQ1-D: Do the syntactic properties of the complementizer head determine which subject positions are grammatical?

H1-D: Yes, the syntactic properties of the complementizer head will ultimately determine which subject positions are grammatical.

Since C is a phase head, this hypothesis would predict that C would determine the grammatical properties of TP. Let us see why this might apply to the phenomenon of inversion. In English, inversion occurs when T raises to C, so there is a straightforward link between C and T in English wh-questions and therefore between C and its complement. In Spanish, inversion occurs when the subject remains in situ, so the subject actually remains in the little v phase. Still, when there is no inversion in Spanish, the subject has raised to SpecTP, part of the C phrase. If the ability of the subject to raise to SpecTP has to do with the grammatical properties of TP, then by the Phase Head Hypothesis, C would determine the availability of inversion for a given wh-question by determining whether TP has the properties necessary to move the subject to SpecTP. In this way, C can determine the availability of inversion in both Spanish and English wh-questions.

Taken together, these four hypotheses make general predictions about the role that each factor that we have introduced plays in the formation of wh-questions in Spanish-English code-switching. Recall that we are also interested in any restrictions on code-switching between specific heads or constituents, as expressed in the second research questions:

RQ2: Are there restrictions on code-switching between particular constituents in wh-questions in Spanish-English code-switching?

The first sub-question for RQ2 focuses on the relationship between the tense head and the complementizer head. Following previous work on the relationship between C and T and restrictions on
code-switching between these two functional heads (Belazi et al. 1994; González-Vilbazo 2005), I propose the following hypothesis:

**RQ2-A:** Are there restrictions on code-switching between the tense head and the complementizer head?

**H2-A:** Yes, there will be restrictions on code-switching between the tense head and the complementizer head.

Finally, we turn to RQ2-B, focused on the relationship between C and the wh-phrase. Given previous research suggesting that CS with simple wh-phrases in matrix clauses is ungrammatical (Woolford 1983, 1984; D’Introno 1996), I propose the following hypothesis:

**RQ2-B:** Are there restrictions on code-switching between the wh-phrase and the complementizer head?

**H2-B:** Yes, there will be restrictions on code-switching between the wh-phrase and the complementizer head.

Having laid out the hypotheses for our research questions, we now turn to the methodology that this experiment is based on.
4 Methodology

4.1 Pilot Experiments .......................................................... 70

4.2 Stimuli Design .................................................................. 72

4.2.1 General Considerations .................................................. 72

4.2.1.1 Limiting Stimuli Number .............................................. 72

4.2.1.2 Design Factors .......................................................... 73

4.2.1.3 Lexicalizations .......................................................... 76

4.2.2 Code-switching Stimuli Design ....................................... 78

4.2.3 Monolingual Stimuli Design ........................................... 81

4.3 Participant Selection ........................................................ 83

4.4 Materials ........................................................................ 85

4.4.1 Consent and Background Questionnaire .......................... 86

4.4.2 Acceptability Judgment Training .................................... 87

4.4.3 Distractor Tasks ........................................................... 87

4.4.4 Code-switching Stimuli ................................................ 88

4.4.5 Monolingual Stimuli ...................................................... 89

4.4.6 Proficiency Tasks ........................................................ 89

4.5 Procedure ................................................................. 90

4.5.1 General Considerations ................................................ 90

4.5.1.1 Type of Experimental Task ........................................ 90

4.5.1.2 Types Acceptability Judgment Tasks ............................ 92

4.5.1.3 Confounds in Collecting Code-Switching Data ............ 94

4.5.2 Participant Recruitment ................................................ 95

4.5.3 Experimental Procedure ............................................. 95

4.6 Methods of Analysis ...................................................... 96

4.6.1 Final Participant Selection ............................................. 96

4.6.1.1 Background and Proficiency ................................. 96
In this chapter I review the methodology that was used in administering the dissertation experiment, along with a variety of considerations that went into arriving at the final design and the steps taken to prepare the resulting data for analysis.

I begin in §4.1 with a discussion of a number of pilot experiments that were conducted prior to the final dissertation experiment and the influence these experiments had on the final design choices. Following that, I review the design of the experimental stimuli in §4.2. In §4.3 I turn to the criteria and materials involved in selecting participants for the study. Next, §4.4 outlines in detail all of the materials that formed part of the experiment. In the following section, §4.5 I discuss in detail how I chose the final experimental task, some related matters concerning the structure of the experiment, and the step-by-step procedure of the experiment itself. Finally, in §4.6 I review the process by which I selected the final set of participants included in the data analysis and a few additional steps I took to prepare the data for analysis.

4.1 Pilot Experiments

Prior to the design and administration of the current study, I conducted a series of pilot experiments. The purpose of these experiments was to twofold. On the one hand, I was interested in gauging the relevance of the many different factors involved in wh-questions, in the context of code-switching. Some factors would be expected to produce more informative contrasts in acceptability judgments than others. Across the four experiments, I ultimately tested all of the following factors, though never within a single experiment:
(84) a. syntactic function of the wh-phrase (object or adjunct)
b. clause of the wh-question (matrix or embedded)
c. position of the subject (SV, V2, VS)
d. complexity of the wh-phrase (simple or complex)
e. language of the question verb
f. language of the wh-phrase
g. language of the matrix clause

On the other hand, these experiments also served to develop and refine methods for collecting experimental CS data. For example, the acceptability judgment training was expanded and refined throughout the course of the experiments.

In the context of the current study, the pilot experiments also served two purposes. First, they informed a number of decisions with regard to the factors that were ultimately included as independent variables. As discussed in detail in §4.2.1.1 I ultimately decided to limit the number of stimuli such that all stimuli could be included in a single experiment administered to all participants. Results from the pilot experiments suggested that three factors in particular did not affect the grammaticality of Spanish-English wh-questions in code-switching: the language of the matrix clause in sentences with embedded wh-questions, the language of adverbs and adverbial phrases, and a code-switch between the wh-phrase and the tense head. For that reason, these factors were not included as independent variables in the final experiment. Additional details concerning these factors can be found in §4.2.1.2 and §4.2.2.

The pilot experiments also played a role in the final experimental methodology. They did not directly determine specific methodological aspects of the study, as the final choices were mostly grounded in methodological standards and previous CS research, but they did inform what approaches were considered. For example, only the last pilot experiment included a proficiency measure and monolingual acceptability judgments. The effect these materials had on participant selection and data analysis compared to the previous experiments helped to underscore the value of those materials.
4.2 Stimuli Design

I will begin my discussion of the stimuli design by focusing on a number of general considerations, including the issue of limiting the total number of stimuli in the experiment, the factors that play a role in the final design, and how the lexicalizations are structured. Following that, I review the final stimuli design for the code-switching and monolingual stimuli.

4.2.1 General Considerations

4.2.1.1 Limiting Stimuli Number

Let’s begin with the issue of limiting the total number of stimuli presented to participants. As discussed in the literature review, there are numerous factors involved in the structure of wh-questions in Spanish and English, including various properties of the wh-phrase, the position of the subject, and the clause that contains the wh-question. Combining these factors with the possibility of two languages for every constituent of a given sentence has the potential to create a very large quantity of stimuli. Because making acceptability judgments does involve a level of performance (cf. Schütze 1996; Toribio 2001; Birdsong 1989 for discussion), it is important to limit the number of stimuli that any one participant judges. Otherwise, the judgments can become too unreliable to be meaningful. One way to limit the total number of stimuli that an individual participant must judge is to divide the stimuli into blocks containing a subset of the full set of stimuli. Each participant only judges one block, and judgments on the full set of stimuli are obtained by having each block of stimuli judged by a different set of participants. The disadvantage of this approach is that it requires a larger number of participants. The alternative is to restrict the factors under study to reduce the total number of stimuli.

While both approaches have their advantages and disadvantages, the approach under which participants judge all stimuli is more clearly preferred once we take into account the role of the monolingual stimuli. Essentially, participants may differ in their monolingual judgments based on factors that are important to interpreting the data, and this is not necessarily known ahead of time, at least not without
a separate study just of the monolingual judgments of the population being worked with. If there are differences between participants in their monolingual judgments but there is not a full set of stimuli for all participants, it becomes impossible to ensure that comparable participants are being grouped for analysis. The properties of the monolingual sentences are a reflection of the available syntactic features in each language. This in turn affects the syntactic features available in the CS sentences. The CS results can only be interpreted in light of these available syntactic features, so it is crucial to know what these features are. Without knowing the monolingual judgments, this would not be possible. For this reason, I ultimately decided to reduce the number of factors under study in order to be able to have all participants judge the full set of both CS and monolingual stimuli.

4.2.1.2 Design Factors

In this section, I discuss the major factors or variables that go into the design of the dissertation experiment and the reasons for including these specific factors. The role that these factors play in the design of the CS and monolingual stimuli will be discussed in §4.2.2 and §4.2.3 respectively.

Overall, the choice of factors is based on two criteria. First, structural factors, such as whether or not the question is a matrix question, are included whenever these factors are relevant in at least one of the individual languages. This makes it possible to observe the effect that other factors have on the grammaticality of these structures. Second, factors based on language, such as the language of the wh-phrase, are chosen based on whether or not a particular constituent may plausibly influence the overall structure of the sentence. If changing the language of a given constituent, while holding everything else constant, changes the grammaticality of the structure, then this suggests that the features associated with that constituent in that particular language are necessary for derivation of that sentence. By carefully choosing the structures to investigate and the constitutes to code-switch, it becomes possible to investigate the interaction between these structures and the language-specific properties of the corresponding constituents.

1In principle, this could be solved by having all participants judge all of the relevant monolingual sentences but only a subset of the CS stimuli so as to reduce the total number of stimuli. However, this has several practical issues, including the fact that the full set of monolingual stimuli will likely be as large as the CS stimuli, if not larger, and would therefore still constitute a potential burden on the participants.
The principle factors determining the design of the stimuli included in the dissertation experiment are listed in (85). I will discuss each factor and how it came to form part of the final experiment in turn.

(85) **Principle Design Factors**

a. Clause of the wh-question  
b. Complexity of the wh-phrase  
c. Position of the subject  
d. Language of the verb  
e. Language of the wh-phrase  
f. Language of the subject  

First, my dissertation experiment includes wh-questions in both matrix and embedded clauses. Since English has a matrix-embedded asymmetry where inversion only occurs in matrix questions, it is important to include both types of wh-questions. In other words, the stimuli include both questions that only consist of one clause, such as (86a), and questions that consist of two clauses where the wh-phrase originates in and remains in the lower clause, as in (86b).

(86) a. What has John bought this week?  
b. I don’t know what John has bought this week.

While a variety of different characteristics of the wh-phrase play a role in wh-question formation, as outlined in §3.1.4 my dissertation experiment focuses only on the contrast between simple and complex wh-phrases. As discussed in §4.2.1.1 I have chosen to limit the total number of factors wherever possible while still being able to make crucial comparisons between different types of wh-phrases. To that end, only object wh-phrases are included in the stimuli, leaving non-argumental and subject wh-phrases for future research. Among other complications, adjunct or non-argumental wh-phrases evidence a greater degree of variation with respect to the requirements for inversion than do wh-objects, as discussed in

---

2 See previous discussion in §2.6 concerning dialectal and register variation with respect to inversion in English embedded clauses.
3 I refer to the latter type as embedded questions, though the sentence as a whole is declarative. These are sometimes referred to as indirect questions or semi-questions. These terms are not necessarily synonymous, but I will not make a distinction between them here. See [Suñer 1991] for some discussion concerning such questions in Spanish.
Subject wh-phrases, on the other hand, necessarily have no inversion at all, removing an important property of wh-questions that meaningfully differs between Spanish and English. Furthermore, the D-linked status of the wh-phrase is not included as an additional factor. Instead of using clearly D-linked complex wh-phrases such as (87a) where the question most naturally refers to a specific set of books given in the discourse, all of the stimuli use complex wh-phrases as in (87b). While any wh-phrase may be D-linked, this type of wh-phrase may naturally be used in a discourse in which the speaker has no particular set of books in mind, such as following the statement “It sounds like you read a lot of books.”. This contrasts with a question in (87c), where ‘books’ refers to the assigned books.

(87) a. Which books have you read so far?
    b. How many books have you read this week?
    c. I’ve read two of the three assigned books so far. How many books have you read?

Since the position of the subject offers the only contrast in word order in most wh-questions when comparing Spanish and English, it is crucial to include the subject position as a factor. As discussed in §3.1.2 and §3.1.3 there are three possible subject positions in the present experiment: (1) preverbal, abbreviated SV, (2) postverbal, abbreviated VS, and (3) between an auxiliary and a participle, abbreviated V2.

For example, the following trio of stimuli sentences only differs by the position of the subject, where the subject is preverbal in (88a), after the auxiliary in (88b), and postverbal in (88c).

(88) a. What esos trabajadores han comido in the last four days? SV
   ‘What have those workers eaten in the last four days?’
   those workers have eaten

b. What han esos trabajadores comido in the last four days? V2
   ‘What have those workers eaten in the last four days?’
   have those workers eaten

c. What han comido esos trabajadores in the last four days? VS
   ‘What have those workers eaten in the last four days?’
   have eaten those workers

Because all participants see all the stimuli for a given experiment, it is important to include the minimum number of stimuli necessary. To that end, sentences with V2 word order are not tested with all possible combinations of the remaining factors. The particular combination of factors that are tested is discussed below in §4.2.2.
The three remaining factors are related specifically to the CS stimuli in the experiment. Since we are interested in using CS as a tool of linguistic analysis, we want to code-switch those constituents that may plausibly play a role in the phenomenon of interest, in this case the formation of wh-questions. These constituents are (1) the wh-phrase, (2) the tense head, (3) the subject and (4) the complementizer, as discussed in some detail in §1.1. Since the last constituent, the complementizer, is covert in the structures under study, it is not possible to directly manipulate the language of this constituent. This leaves the wh-phrase, the tense head and the subject. In order to limit the total number of stimuli, the language of the wh-phrase and the language of the tense head are controlled such that the wh-phrase and the tense head are always in opposite languages. For example, in (89a), the wh-phrase ‘what’ is in English, so the main verb and auxiliary, and therefore the tense head, are in Spanish. In (89b), the wh-phrase is in Spanish, and therefore the main verb and auxiliary are in English.

(89)  a. **What esos trabajadores han comido** in the last four days? 
     those workers have eaten 
     ‘What have those workers eaten in the last four days?’

   b. **Qué esos trabajadores** have eaten in the last four days? 
     what those workers 
     ‘What have those workers eaten in the last four days?’

Preliminary data from the pilot experiments suggests that there are no constraints on code-switching between a wh-phrase and a tense head, so linking the language of these two constituents is not expected to constrain the data analysis.

### 4.2.1.3 Lexicalizations

While lexicalizations can help to separate out the influence of particular lexical choices from the acceptability of the underlying structure, the standard use of lexicalization offers limited options for dealing with lexicalizations that have very different ratings. When one particular lexicalization varies notably from the other three lexicalizations, that suggests that there is some issue with the lexicalization and therefore that the rating does not reflect that participant’s intuitions. It may be, however, that the apparent outlier is in fact the most representative. This becomes particularly clear if two of the four

---

5 This is discussed in some detail in §2.3.
6 See §5.1.3 for argumentation.
sentences are rated high and the other two are rated low. In that case, it is even less clearly which lexicalizations are representative of the participant’s intuitions of the underlying structure. To mitigate the issue, I systematically generate lexicalizations such that the effect of particular lexical choices can be separated from the underlying structure being tested.

To do that, the lexicalizations are structured as follows. First, I identify factors that could be combined to create major categories of stimuli for a given language pair. In particular, I use the following two factors: the clause type (matrix or embedded) and the complexity of the wh-phrase (simple or complex). These factors generate four groups or categories of stimuli: matrix questions with simple wh-phrases, matrix questions with complex wh-phrases, embedded questions with simple wh-phrases, and embedded questions with complex wh-phrases. For each category, I then create a series of pairs of monolingual sentences, one in Spanish and one in English, such as in (90). Each monolingual sentence is a translation equivalent of the other monolingual sentence it is paired with.

(90) **A lexical unit**

a. Qué han comido esos trabajadores en los últimos cuatro días?
   what have eaten those workers in the last four days
   ‘What have those workers eaten in the last four days?’

b. What have those workers eaten in the last four days?

Together, these base sentences make what I call a *lexical unit*. These lexical units are then used to create the actual CS and monolingual sentences. For example, continuing with matrix questions with simple wh-phrases, pieces of the base sentences are combined and ordered based on the remaining factors in the experiment. With Spanish-English, these would be the following (with the levels of the variable in parentheses): the language of T (2, with the wh-phrase in the opposite language), the language of the subject (2, Spanish or English), and the position of the subject (3 for matrix: SV, V2, and VS). Multiplying these factors yields the following 12 combinations:

(91) a. What esos trabajadores han comido en los últimos cuatro días?

   What those workers have eaten in the last four days?

c. Qué esos trabajadores have eaten in the last four days?

7 The full set of factors included in the experimental design are detailed in §4.2.1.2.
d. Qué those workers have eaten en los últimos cuatro días?

e. What han esos trabajadores comido in the last four days?

f. What han those workers comido en los últimos cuatro días?

g. Qué have esos trabajadores eaten in the last four days?

h. Qué have those workers eaten en los últimos cuatro días?

i. What han comido esos trabajadores in the last four days?

j. What han comido those workers en los últimos cuatro días?

k. Qué have eaten esos trabajadores in the last four days?

l. Qué have eaten those workers en los últimos cuatro días?

Because these sentences cover all of the factors under study for each major category, it becomes possible to tease apart the role of the lexical items in the ratings of the sentences. If there is something odd about the word choice in the base sentences that lowers the ratings, it will do so for all of the sentences based on those lexical items. Even if the overall ratings for the group of sentences based on a lexical unit is lower, the effect that particular structures have on the acceptability ratings will still be clear because the ratings will still be comparatively high, or low.

There are four lexical units for each of the four previously mentioned types of wh-questions, which produces four lexicalizations of each final structure, whether that is a CS or a monolingual stimulus. For example, the structures of the 12 sample sentences above would be repeated three more times using three more lexical units (i.e. pairs of monolingual base sentences) to produce the 4 total lexicalizations for each structure or type of stimulus.

4.2.2 Code-switching Stimuli Design

There are two main aspects to consider in the design of the CS stimuli: the independent variables and the controlled variables. The independent variables are the structural aspects, or factors, that are manipulated in order to test their effect on the acceptability of a given sentence (i.e. the dependent variable). The controlled variables are characteristics of a given stimulus that need to be controlled so as to avoid confounding variables. This minimizes the effect of these variables to ensure that it is
the independent variables, or the factors under study, that affect the dependent variable. The factor of language has a unique status in that, to some extent, it is both an independent variable and a controlled variable. On the one hand, the language of some constituents is expected to play a role in the grammaticality of a given structure, making them independent variables that are manipulated to see their effect on the dependent variable. This is the case for the language of the wh-phrase and the language of the verb complex (i.e. the finite verb or the auxiliary and particle). On the other hand, the language of the remaining constituents of each sentence needs to be controlled so that it is consistent and predictable across stimuli. I will discuss the independent and controlled variables in turn.

The independent variables are based on the overall design factors discussed in §4.2.1.2. This gives the following six independent variables, with the possible values of each variable also listed:

(92) **CS independent variables and their values**

a. Clause of the wh-question: matrix or embedded
b. Complexity of the wh-phrase: simple or complex
c. Position of the subject: SV, V2 or VS
d. Language of the verb: Spanish or English
e. Language of the wh-phrase: opposite language of the verb
f. Language of the subject: Spanish or English

While the language of the wh-phrase and the verb are not fully independent, they both serve as independent variables, as both constituents are always tested in both languages with each of the remaining independent variables. As with any study making use of an acceptability judgment task, the dependent variable for this experiment is the acceptability of the sentences, as this is expected to vary depending on the values of the variables in (92). With regard to the three possible subject positions in (92c), V2 word order is only tested with the matrix questions in order to limit the total number of stimuli, since this word order was not expected to be acceptable in embedded questions for monolingual Spanish or for monolingual English.

The total number of stimuli, then, is as follows. First, we calculate the number of types of stimuli, before adding in the lexicalizations. Since matrix and embedded questions have a different number of
subject positions (three and two, respectively), these are calculated separately, as seen in (93). Because the wh-phrase and the verb are always in opposite languages, this produces two pairings or combinations: a Spanish wh-phrase with an English verb and an English wh-phrase with a Spanish verb.

(93) Calculating CS stimuli types

a. Matrix question types: 2 levels of wh-complexity × 3 subject positions × 2 language pairings for the verb and wh-phrase × 2 languages for the subject = 24 types

b. Embedded question types: 2 levels of wh-complexity × 2 subject positions × 2 language pairings for the verb and wh-phrase × 2 languages for the subject = 16 types

There are four lexicalizations of each type, yielding a total of 160 CS stimuli.

The controlled variables can be divided into two types, those whose value is held constant and those whose values are varied but balanced across all stimuli. Beginning with the former, various characteristics of the tense head, the wh-phrase, and the subject are held constant. In order to be able to test V2 word order, where only a tensed modal or auxiliary appears above the subject and the verb remains in situ, all of the questions are in the present perfect tense. While it would also be possible to have V2 word order by allowing do-support (e.g. ‘What did John buy?’), this would also introduce a confounding variable. There is no parallel process of do-support in Spanish, so this particular type of inversion may make structures that are otherwise grammatical ungrammatical. The present perfect, on the other hand, is more parallel in structure between the two languages. Furthermore, all verbs are in plural form because of the ambiguity between the English third-person singular present perfect has and the Spanish second-person singular present perfect ha. As previously discussed, all wh-phrases are objects in order to limit the overall number of stimuli, and complex wh-phrases are chosen so as to be non-D-linked in a neutral context. Finally, subjects are always full DPs, rather than proper names, in order to be able to control the language of the subject. In addition, the determiner is either a demonstrative or a possessive pronoun, based on the intuition that these determiners are more natural without context than definite articles, as seen in the contrast between (94a) and (94b).

(94) a. What have those kids bought at the mall so far?

8See §4.2.1.2 for discussion.
b. What have the kids bought at the mall so far?

The possessive pronouns also enabled the use of subjects that participants would encounter in daily life, including relatives (e.g. ‘sisters’ and ‘cousins’) and other people (e.g. ‘classmates’ and ‘neighbors’). The type of determiner is balanced so that half the subjects include a demonstrative and half the subjects include a possessive pronoun.

Two additional characteristics of the stimuli are balanced. First, in sentences with an embedded wh-question, the language of the matrix clause is balanced so that it is in Spanish for half of the stimuli and English for the other half of the stimuli. This was done based on data from the pilot experiments suggesting that the language of the matrix clause does not affect the grammaticality of a given structure. The language is varied, rather than being held constant, so as to reduce repetition. Second, an adverb or adverbial phrase is included with each sentence in order to add a little bit of context to the sentences. The adverb or adverbial phrase always explicitly or implicitly demarcates the relevant time for the question (e.g. ‘since last week’ or ‘today’) in order to make the use of present perfect more natural. Compare the question in (95a) without an adverb and the question in (95b) with an adverbial phrase.

(95)  a. What have you eaten?

b. What have you eaten since this morning?

Based on pilot data, the language of the adverb or adverbial phrase also did not affect the grammaticality of a given structure, but once again the language is varied so as to reduce repetition. In this case, the adverb is always in the opposite language as the subject so that, in the end, half the sentences have an adverb or adverbial phrase in each language.

4.2.3 Monolingual Stimuli Design

The monolingual stimuli are largely designed around the need to test the factors under study, as outlined in the previous section on the code-switching stimuli, but there are a few additional stimuli that are included in order to test more general properties of wh-questions in each of the languages.

Continuing with the careful use of lexical units (cf. §4.2.1.3), the majority of the monolingual stimuli are based off of these lexical units. However, in order to limit the overall number of stimuli, not all of
the monolingual base sentences, and their corresponding lexical units, are varied according to all of the factors included in the study. For example, not all subject positions are tested for all lexical units. Still, all lexical units, and therefore the lexical items underlying all CS stimuli, are tested minimally with one grammatical and one ungrammatical word order in order to separate the effects of lexical choice from the underlying intuitions concerning a given structure.

Of the 6 design factors outlined in §4.2.1.2, three of them involve the language of specific constituents and hence are not relevant in the context of monolingual sentences. This leaves three factors to act as independent variables for the monolingual stimuli, as outlined in (96).

(96) **Monolingual independent variables and their values**

a. Clause of the wh-question: matrix or embedded

b. Complexity of the wh-phrase: simple or complex

c. Position of the subject: SV, V2 or VS

As with the CS stimuli, the number of subject positions differs between stimuli with a matrix wh-question and stimuli with an embedded wh-question. The total number of principle monolingual stimuli also differs between languages, so we will consider each language in turn. First, the monolingual Spanish stimuli parallel the CS stimuli in that all three subject positions are tested with matrix questions, but only SV and VS word orders are tested with embedded questions, yielding a total of 10 types:

(97) a. Matrix: 2 levels of wh-complexity × 3 subject positions = 6 types

b. Embedded: 2 levels of wh-complexity × 2 subject positions = 4 types

The English monolingual stimuli include not only the parallel word orders tested in the CS stimuli but also an additional word order for English. Specifically, V2 word order is included with the embedded questions because some research has shown that these word orders may be grammatical for some speakers, as discussed in McCloskey (2006). For additional discussion see §2.6. This produces a total of 12 stimuli types for English, as seen in (98).

(98) a. Matrix: 2 levels of wh-complexity × 3 subject positions = 6 types

b. Embedded: 2 levels of wh-complexity × 3 subject positions = 6 types
As with the CS stimuli, there are four lexicalizations per type, for a total of 88 monolingual stimuli.

Finally, a number of additionally monolingual stimuli are included. First, a series of wh-questions involving superiority effects are included for potential future research\(^9\). These were included in both the monolingual Spanish and the monolingual English stimuli. Second, a set of sentences are included with the Spanish stimuli based on the adverb tests for T-to-C movement in Spanish proposed in Suñer (1994)\(^10\). These stimuli make it possible to distinguish between participants who use similar types of inversion for Spanish and English (i.e. those who also evidence T-to-C movement in Spanish) and those who have two distinct processes of inversion for Spanish and English, as in the case for monolingual speakers.

Thirdly, a number of more natural examples of embedded questions are included with the monolingual English stimuli to ensure that judgments on inversion in embedded questions are not overly affected by the artificial or constructed nature of the main stimuli. These are taken from McCloskey (2006). In particular, participants judge sentences where inversion is expected to be ungrammatical across dialects, as in (99a), sentences without inversion, as in (99b), and parallel sentences with inversion, as in (99c).

\[(99)\]
\[
a. \quad *I \text{ found out how } \text{did they } \text{get into the building.}
\]
\[
b. \quad I \text{ wonder what } \text{he is } \text{like.}
\]
\[
c. \quad I \text{ wonder what } \text{is he } \text{like.}
\]

### 4.3 Participant Selection

Because the object of study for this investigation is I-language, there are a number of important characteristics for participant selection that go beyond having experience with Spanish and English. First, we are interested individuals who are native speakers of both languages, in the sense of having grown up with the two languages. While second language learners also have linguistic competence in each language that they have experience with, most syntactic theory is developed based on studying the knowledge of native speakers, and I will follow in that tradition.

\(^9\)See §3.1.4.2 for some discussion of superiority effects. For the full list of these stimuli, see the appendix.
\(^10\)See §3.1.2.1 for discussion.
The most basic criterion for establishing native competence is the age at which participants began learning each language. For this study, only participants that began learning both languages by the age of 6 are included. This particular age was chosen based on the characteristics of the population under study. These are individuals who grow up in Spanish-speaking households and often do not receive significant exposure to English until pre-kindergarten or kindergarten, usually by the age of 6. Importantly, it’s entirely possible that participants learned both languages natively but have since stopped using one of the languages. For that reason, only those participants that report using both languages in various contexts on a daily or semi-daily basis are included. A variety of additional information about the participants is also collected on the background questionnaire, including basic language information on their parents, what languages they spoke with whom while growing up, and how they rate their language proficiency in a variety of domains.

Beyond background information, there are two additional types of data that further narrow the pool of participants to those most appropriate for studying the I-language of Spanish-English codeswitchers: language proficiency and monolingual acceptability judgments. Language proficiency serves as an additional signal of the degree to which participants have maintained their linguistic competence in both languages. As discussed earlier in §4.2.1.1 as well as in §2.3, monolingual stimuli play an important role in the analysis and interpretation of data for this experiment.

The ultimate goal of this project is to gain insight into the formation of wh-questions by looking at the way that the features of Spanish and English interact in the computational system when forming code-switched wh-questions. To do this, it is crucial to know what features Spanish and English are contributing individually in order to be able to infer what the result of combining those features means. To that end, participants are asked to judge monolingual sentences in order to gain insight into the features of the individual languages on a speaker-by-speaker basis. The acceptability judgments for these sentences play a role in the selection of participants to the extent that individual participants differ in their monolingual judgments. If most participants agree on the acceptability of the monolingual stimuli, those participants that do not agree could be excluded from the analysis so as to maintain one larger group of participants. On the other hand, if there is more variety in the monolingual judgments, it would be preferable to form smaller groups based on these judgments. In either case, the CS and
monolingual judgments of participants are analyzed together with those of participant with comparable
monolingual judgments in order to be able to make valid inferences from the data.

Participants’ Spanish proficiency is evaluated using a modified version of the Diploma of Spanish as
a Foreign Language (DELE, from the Spanish abbreviation), similar in form to a cloze test[^11]. While the
DELE has its limitations as a measure of proficiency, research by [Montrul](2002, 2004) suggests that it
is a reliable measure of proficiency for bilinguals, and it is commonly used in studies of Spanish-English
bilinguals and second language learners of Spanish to assess proficiency in Spanish. Because it focuses
on more academic Spanish, it is entirely possible that participants who are proficient in Spanish but
only learned the language in informal contexts might score poorly. Importantly, the reverse is not true.
It is highly unlikely that participants who are not proficient in Spanish will still score well on the DELE.
Thus, excluding participants with low DELE scores should narrow the pool of participants to those
that are reasonably proficient, even if it may also occasionally exclude other participants who are also
proficient. The modified DELE consists of two parts, a thirty-question multiple-choice section and a
twenty-question multiple-choice cloze test.

The English proficiency measure was a forty-question multiple-choice cloze test that has been used
in a variety of previous studies (cf. [Ionin and Montrul](2009, 2010) [Ionin et al.](2013) [Montrul](2001)).

4.4 Materials

The experiment was completed entirely online using open source software known as “Ibex” (“Internet
Based EXperiments”), which is designed for collecting self-paced reading and acceptability judgments
online. It is similar in some ways to presentation software such as E-Prime. The experiment was run using
version 0.3.2 of the software[^12] using a free website for hosting Ibex experiments[^13]. Participants simply
visited a webpage accessed through the provided link, and then they completed the entire experiment in
their web browser. The format of the experiment can be likened to a series of slides with small amounts
of text. Participants read the information on the current screen or slide, and then they click a link or
button to advance to the next screen. The experiment does not allow going back to previous screens.

[^11]: See the appendix for full text of the proficiency measures.
[^12]: The latest version can be found on the following website: [https://code.google.com/p/webspr/](https://code.google.com/p/webspr/)
[^13]: Found here: [http://spellout.net/ibexfarm/](http://spellout.net/ibexfarm/)
Additional details concerning how particular types of information (e.g. a background questionnaire or an acceptability judgment) are presented are included with the discussion of that information.

In addition to short transition screens between tasks, participants complete the following materials throughout the course of the experiment\footnote{With the exception of the consent form and the distractor tasks, all of these materials can be found in the appendix.}

\begin{enumerate}
\item Overview of experiment materials
\begin{enumerate}
\item Consent form
\item Background questionnaire
\item Training on performing acceptability judgments
\item Short distractor tasks
\begin{itemize}
\item Two memory games
\item Two puzzles
\end{itemize}
\item Spanish proficiency task (DELE)
\item English proficiency task (English cloze test)
\item CS stimuli
\item Monolingual stimuli
\end{enumerate}
\end{enumerate}

Each of the materials is reviewed in turn in the following sections.

\subsection{Consent and Background Questionnaire}

The consent form reviews the standard information necessary for participants to provide informed consent. The background questionnaire collects general demographic information, information on participants language experience both growing up and currently, and participants’ self-ratings of their language proficiency in a variety of domains. For additional details, see §4.3 and for the full background questionnaire see the appendix.
4.4.2 Acceptability Judgment Training

The training on performing acceptability judgments consists of written, code-switched instructions describing what it means to rate a sentence. The term grammatical is not used because of the prescriptivist approach to grammar that it elicits in many individuals. The instructions also discuss the distinction between sentences which are semantically odd but perfectly grammatical and sentences which are quite plausible or comprehensible but still ungrammatical. The instructions use a combination of code-switched, monolingual Spanish and monolingual English examples to illustrate the points being made.

4.4.3 Distractor Tasks

The experiment includes two types of distractor tasks. These short tasks are included in the experiment in order to give participants a break from performing acceptability judgments while also having them do a different sort of thinking. These tasks are used to break both the CS and monolingual stimuli into smaller blocks of stimuli. In addition, these tasks help to misdirect the participants with respect to the goals of the experiment.

There are two types of distractor tasks. First, there is a simple memory game. Participants are presented with a $5 \times 4$ grid of tiles. When they click on a pair of tiles, they are briefly shown the images beneath those tiles. If they are the same images, the tiles remain turned over. Participants are asked to click on tiles to reveal pairs of images, and as they click on tiles, the web page automatically shows the number of moves up to that point. Once they have matched all pairs, they are asked to enter the total number of moves they took to complete the task. Participants take approximately 4 minutes to complete this task. There were two different versions of this task, each with different images in a different order. The first version is presented between blocks of CS stimuli and therefore has CS instructions. The second version is presented between blocks of monolingual English stimuli and therefore has English instructions.

The second distractor task is a puzzle. Participants are presented with a $4 \times 3$ blank grid, and next to that grid is a jumbled set of squares that together form a complete image. Participants are asked to drag the puzzle pieces on to the grid and arrange them so that they form the complete image.
Participants take approximately 4 minutes to complete this task. There were two different versions of this task, each with a different image to complete. The first version was presented between blocks of CS stimuli and so had CS instructions. The second version was presented between blocks of monolingual Spanish stimuli and so had Spanish instructions.

4.4.4 Code-switching Stimuli

The code-switched stimuli are presented to participants as follows. For those parts of the experiment involving acceptability judgments, including the CS stimuli, participants are presented with a simple progress bar near the top of the browser window showing them how far along they are in the experiment. Each sentence is displayed individually and centered below the progress bar. Below the stimulus is a row of seven small boxes inside of which are one of the numbers for the 1–7 Likert scale, from smallest to largest going left to right. These boxes function as buttons that participants can click on to register their judgment. To the left of the boxes is the text “(completely bad)”, and to the right is the text “(completely good)”, to remind participants of how to use the scale. Below the Likert scale are the following instructions “Haga clic en los cuadritos para contestar” (“Click on the boxes to answer”). In this way, the instructions during the CS part of the experiment contain both Spanish and English. After providing a judgment for a given sentence, the software records the time in milliseconds that the participant took to judge the sentence, the screen is replaced with “...” for 500 milliseconds, and then the next sentence is displayed. This short pause between stimuli gives the participants a short recovery time and helps ensure that they do not rush through the experiment too quickly.

The CS stimuli consist of 160 target items and 170 filler items. In addition, 10 transitional stimuli were added to the beginning of the first block of CS stimuli. The transitional stimuli served to give the participants time to transition from the format of the training examples to the format of the experimental stimuli. The design of the target items is detailed in §4.2. The filler items consist of declarative, monoclusal sentences with code-switches between a finite verb and one of the following:

(101)  a. Subject lexical DPs

b. Object lexical DPs

c. Subject pronouns
d. Object Pronouns

e. Object Clitics

The DPs were always third person and varied in number and gender. The pronouns varied in person, number and gender. The target and filler items were independently pseudo-randomized so as to prevent the same type of stimuli from being presented too close together. Following that, the two types of items were combined such that no more than three of one type of item appears in succession.

Lastly, two versions of the final set of stimuli were created in order to reduce ordering effects (i.e. differences in acceptability judgments based solely on the order in which the stimuli were presented). To do this, the stimuli were first divided into three blocks, as required for the overall experimental procedure. Version 1 used the original order of the stimuli. The second version swapped the first and second halves of each block. For example, the first CS block had 110 stimuli, so the Version 2 started with stimuli 56 through 110, and then it finished with stimuli 1 through 55.

4.4.5 Monolingual Stimuli

The monolingual stimuli consist of 72 English target items, 96 English filler items, 80 Spanish target items, and 84 Spanish filler items. The design of the target items is detailed in §4.2. The monolingual filler items have the same structure as the CS filler items. The monolingual stimuli were presented in the same way as the CS stimuli except that the onscreen instructions were only in the language currently being judged. The monolingual target and filler items were pseudo-randomized and combined using the same procedure as the CS stimuli.

4.4.6 Proficiency Tasks

As discussed in §4.3, participants completed a written Spanish proficiency test, a modified version of the Diploma of Spanish as a Foreign Language (DELE, from the Spanish abbreviation). The format of the online version of the DELE used for the experiment is similar to an online or web-based form. Each blank where participants would normally write a letter corresponding to the correct answer instead contains a dropdown menu with the possible answers. By default, a blank answer is highlighted, rather
than any of the possible answers, to prevent biasing the responses. The actual order of the possible answers is identical to the paper version of the DELE.

Participants also completed a written English proficiency task in the form of an English cloze test. The test was formatted along the same lines as the DELE, similar to an online or web-based form.

4.5 Procedure

Before reviewing the procedure that participants followed in completing the experiment, it is important to consider a variety of methodological choices that are implicit in that procedure.

4.5.1 General Considerations

4.5.1.1 Type of Experimental Task

It is important to consider the type of task that participants are asked to perform and whether the kinds of data this task provides bear on the questions of interest for the present study. Much of the syntactic work in generative linguistics has focused on acceptability judgment tasks,

but there are a variety of other tasks that may offer insight into the linguistic competence of a given set of speakers. First, there are a number of production-based tasks, varying from less controlled examples such as corpora and interviews to more controlled types like elicited imitation tasks. While production data can provide important insights into various properties of language, they are less well-suited for phenomena that require a very precise combination of factors, including wh-questions. In other words, even the more controlled examples of this task make it difficult to gather data on all of the combinations of factors that are of interest for this study. In addition, data from production tasks, particularly when they are less controlled, preclude or limit the amount of negative evidence. That is to say, they may allow researchers to observe the kinds of structures that speakers can produce, and therefore those that are part of their linguistic competence, but they cannot tell researchers what speakers cannot say. I may never hear certain obscure syntactic structures outside of conversations with linguists, but that does not mean that

---

15 These tasks are often referred to as grammaticality judgment tasks, since they seek to determine the grammaticality of a given set of stimuli, but it is more accurate to refer to them as acceptability judgment tasks. We ask participants to judge the acceptability of a given sentence, and then we infer the grammatical of the underlying structure from the acceptability judgment.

16 Even this is not necessarily the case. Some production data represents errors and therefore does not reflect competence.
they do not exist. Given the importance of negative evidence in syntactic theory, these types of tasks are not suitable for the object of study.

There are also a number of online tasks where participants’ reactions to stimuli are recorded as they see or hear them. These may involve reading times, as in self-paced reading or eye-tracking tasks, or neuro-cognitive measurements such as event-related potentials (ERPs) or functional magnetic resonance imaging (fMRIs). This type of task is particularly good at reducing conscious biases that participants may introduce when responding to stimuli, such as metalinguistic knowledge, since the measurements are made before participants have time to apply such biases. They are also well-suited for research on processing, but this is of less interest here because this study focuses on competence rather than performance. Notice that these tasks by their very nature are focused on the first reaction or early reactions that participants have to a given stimulus, but we are interested in whether a given sentence is possible at all. In other words, we want participants to have the time to consider different ways of constructing the sentences as well as various possible contexts\(^17\). Sentences that are syntactically well-formed but difficult to process or less natural sounding than intended may have the same longer reading times as sentences that are truly not syntactically well-formed, for example. For this reason, online tasks are not particularly well-suited for the focus of this study, the syntactic structure of wh-questions and what this can tell us about the human language faculty more generally.

Instead, I will use the traditional source of data for syntactic theory, acceptability judgments\(^18\). There is still an important choice to be made concerning the modality used for presenting the stimuli, i.e. whether the stimuli should be presented in written or aural form. CS is primarily a spoken phenomenon used in informal settings (Grosjean, 1982; Montes-Alcalá, 2001; Mahootian, 2005), while acceptability judgment tasks are generally administered in a laboratory setting in written form. While the experimental task must necessarily be formal in its nature, given the priority placed on administering a controlled experiment, presenting the stimuli in aural form would have the advantage of more closely approximating the CS that participants are most familiar with. However, aural stimuli are not without their disadvantages. Beyond the additional level of labor and logistics that is involved in creating and

---

\(^{17}\)This process is, of course, largely if not entirely subconscious, particularly for a study such as this in which participants judge a fair amount of stimuli.

\(^{18}\)For some additional discussion on these issues and the benefits of acceptability judgments in the context of syntactic theory and code-switching, see MacSwan and McAlister (2010) and González-Vilbazo et al. (2013).
presenting the stimuli, there is also a potential disadvantage based on fixing the prosody of the stimuli. With aural stimuli, the researcher must ultimately pick a specific prosody to apply to a given stimulus, but this may not be the prosody that participants expect, thus affecting the acceptability judgments that they provide. With written stimuli, on the other hand, participants can apply whichever prosody they prefer for a given stimulus. In this way, stimuli with low acceptability ratings are truly unacceptable, regardless of the prosody. Interestingly, Gianaspro (2013) found evidence that Spanish-English heritage speakers rated otherwise grammatical stimuli lower due to their sensitivity to the prosodic form of stimuli that were presented aurally.

With regard to the issue of the formality of the experiment, the online nature of the task may have an advantage. Since the present study is completed by participants online, they are more likely to complete the task in a more familiar and comfortable setting. In addition, this may help mitigate aspect of what are known as demand characteristics (Orne, 1962; Rosnow, 2002), where participants consciously or unconsciously develop an interpretation of the purpose of an experiment and modify their behavior based on that interpretation. Specifically, it may be beneficial not having a researcher directly interacting with the participant during the experiment, as there are various subtle ways in which this can influence how participants complete the study.

4.5.1.2 Types Acceptability Judgment Tasks

There is one final issue to consider with respect to the design of the acceptability judgment task: the way in which participants are asked to judge sentences. The four most common types of judgment tasks are a forced-choice design, yes-no ratings, magnitude estimation, and a Likert scale. The first two are more qualitative in nature, being limited to two options, and the latter two are more quantitative in nature.

Beginning with the quantitative tasks, in a forced-choice design, participants are generally presented with pairs of sentences and asked to choose which sentence is the most (or least) acceptable. This type of task is not well-suited to experiments such as the present study with a large number of overlapping factors, due to the quantity of pairings that would be necessary to make all of the comparisons of interest. In a yes/no task, participants are presented with individual sentences and asked to rate them as either

\[19\] Of course the researcher does not know for sure how participants assign prosody to the stimuli that they are judging, but the same can be said of aural stimuli. Perhaps participants are ignoring the given prosody and instead applying their own.
acceptable or unacceptable. Because of the stigmatized nature of CS for some speakers (Montes-Alcalá 2001; Poplack 1980; Toribio 2001), this type of task is potentially problematic. Speakers may be more likely to rate sentences as unacceptable simply because they are code-switched, masking any true differences in acceptability between sentences.

Turning to the more quantitative tasks, in a Likert scale task, participants rate individual sentences using a numerical scale with the end points indicating a rating of fully acceptable and fully unacceptable. The scale generally contains an odd number of points in order to have an exact midpoint, commonly either 1–5 or 1–7. The main disadvantage of this type of task has to do with the distance between the numbers on the scale. While in principle the distance between each point on the scale should be uniform, it is not possible to know whether participants treat the difference between 2 and 3 the same as the different between 5 and 6, for example.

Finally, in a magnitude estimation task, participants rate sentences according to a reference sentence. For example, if a reference sentence is assigned a rating of 100 and another sentence is half as acceptable, it would receive a rating of 50. Conversely, a sentence that is twice as good as the reference sentence would receive a 200. While in the abstract such a task should allow participants to make more fine-grained distinctions than a Likert scale, there is some evidence that most participants perform the task in a manner equivalent to a Likert scale (Sprouse 2011). In addition, research also shows that the pattern of results is comparable (Bader and Häussler 2010; Weskott and Fanselow 2011), i.e. that those sentences that are found to be acceptable with one task are also found to be acceptable with the other. This being the case, the simpler task is to be preferred.

While each type of judgment task has its benefits and drawbacks, I believe that the Likert scale task is best suited for this experiment, given in particular the disadvantages reviewed above for the remaining tasks, particularly in the context of this experiment. I chose to use a 1–7 Likert scale in order to give participants a range of choices without overburdening them (e.g. with a 10-point scale). For example, on a 5-point scale, it is hard to know if a sentence that is rated 4 is really halfway between the middle (i.e. 3) and fully acceptable (i.e. 5), whereas participants using a 7-point scale can indicate whether a less-than-perfect sentence is closer to fully acceptable (i.e. with a 6) or closer to the middle rating of 4 (i.e. with a 5).
I would like to briefly touch on two potential confounds when collecting CS data. First, the sometimes stigmatized nature of CS (Montes-Alcalá 2001; Poplack 1980; Toribio 2001) may reduce the degree to which participants are willing to judge CS sentences as acceptable. This, in turn, may mask meaningful differences between stimuli. This issue is addressed in two ways, principally. First, the controlled design of the stimuli allows for systematic comparisons that compensate for any issues of scale compression. If the number of participants is sufficient, this may be done by statistical means. If not, comparisons across structures as well as between target and filler items make it possible to observe any persistent effects across stimuli. In addition, participants are trained using CS instructions and CS examples. This has a number of benefits. First, the CS instructions help emphasize that CS is perfectly valid even in the formal context of instructions for an experimental task. Second, the CS examples include both fully acceptable and fully unacceptable sentences and are indicated as such. In this way, participants are shown that some CS sentences can be rated as perfectly acceptable.

Also important is the degree to which each of a participant’s languages is activated, or where they are on the bilingual continuum (Grosjean 1985, et seq.). If participants have one language much more strongly activated than the other, this may affect how they rate CS sentences. Likewise, if both languages are strongly activated but they are rating monolingual sentences, this may also affect how participants judge those sentences.

To help mitigate this issue, each segment of the experiment begins with a task in the mode corresponding to the subsequent acceptability judgment task. The CS stimuli are preceded by the training for making acceptability judgments, and this training is written in CS and contains example CS sentences. Each set of monolingual stimuli are preceded by a cloze test or proficiency measure in the respective language. I believe that the training and proficiency measures are sufficiently involved to move participants in the necessary direction along the bilingual continuum.

There is an alternative approach to this issue where participants complete each task separately, to avoid one task influencing another task. While this largely prevents each set of acceptability judgments from influencing the other ones, it does not necessarily control the degree to which each language is activated.

For further discussion and a number of other concerns when collecting CS data, see González-Vilbazo et al. (2013).
activated. Since the task is performed online, we do not know the previous language context before participants begin the study. Moreover, there is an important benefit to having participants complete the entire experiment in one setting: less risk of incomplete data. Every study has some level of attrition, and it is particularly important to have a full set of data for this experiment, as discussed in §4.2.1.1.

4.5.2 Participant Recruitment

Participants were recruited through one of two means. Some participants were recruited at the University of Illinois at Chicago through classes designed for heritage speakers of Spanish. Those students who were interested in participating were emailed general information about the experiment as well as a link to the experiment itself. This constituted the majority of participants. Additionally, a number of participants were recruited through a Craigslist advertisement. These participants were also emailed general information about the experiment, as well as a link to the experiment itself. Half of the participants were sent the link for Version 1 of the experiment, and half of the participants were sent the link for Version 2 of the experiment.

4.5.3 Experimental Procedure

Once participants open the link to start the study, the experiment begins with some general information about how to complete the study, including the importance of working in a quiet location, removing all distractions and completing the entire study in one sitting. After the welcome screen and general information, participants are presented with a standard consent form allowing them to consent to participate in the experiment. Following that, they fill out the background questionnaire, and then they complete the training for performing acceptability judgments.

After that, participants are ready to start the main acceptability judgment task, beginning with the CS stimuli. For these stimuli, participants judge one third of the stimuli, and then they complete the first memory game. Following that, they complete another third of the stimuli and the first puzzle, and then they complete the last third of the CS stimuli.

Next, they fill in the English cloze test and complete the first half of the monolingual English stimuli. Then they complete the second memory game followed by the second half of the monolingual

\[21\] See the end of §4.4.4 for details on the difference between Version 1 and Version 2.
English stimuli. After that, they fill in the two parts of the DELE before completing the first half of the monolingual Spanish stimuli. This is followed by the second puzzle and then the second half of the monolingual Spanish stimuli. Once the last monolingual stimuli are complete, the participants are instructed to wait while the results of the study are sent to the server, and then they see a brief message thanking them and letting them know they can close their browser window.

In summary, the experiment proceeded as follows:

1. General experiment information
2. Consent form
3. Background questionnaire
4. Acceptability judgment training
5. CS stimuli Block 1
6. Memory Game 1
7. CS stimuli Block 2
8. Puzzle 1
9. CS stimuli Block 3
10. English cloze test
11. English stimuli Block 1
12. Memory Game 2
13. English stimuli Block 2
14. DELE
15. Spanish stimuli Block 1
16. Puzzle 2
17. Spanish stimuli Block 2

4.6 Methods of Analysis

There are a number of issues to consider before arriving at the final set of data that will be used for analysis. These include the final participant selection, an issue with the language of the subject, and potential issues with particular lexicalizations and particular items.

4.6.1 Final Participant Selection

4.6.1.1 Background and Proficiency

As outlined in §4.3, participant selection for this experiment begins by looking at two key pieces of background information. First, participants needed to have started to learn both English and Spanish by the age of 6. In addition, participants needed to report some daily use of both language. Based on these two criteria, of the 47 participants who completed the experiment, 38 participants qualified for the data analysis.
Proficiency was also an important criterion for this experiment. For the Spanish proficiency measure, the DELE, participants needed to score at the Advanced or Near Native levels to be included, based on the lenient evaluation scale standardly included with this modified DELE. I elected to include participants at the Advance level of proficiency because of the more formal register and academic nature of the test, as compared to the experience that most participants have with Spanish. For the English cloze test, I included participants with a score of at least 35 out of 40. This score is higher, by percent, than the DELE score without being overly strict. Based on these two criteria, 13 of the remaining 38 participants qualified for inclusion in the data analysis. However, this is not the final number of participants included for data analysis.

4.6.1.2 Data Characteristics

A number of additional participants needed to be excluded from the data analysis due to various characteristics of their data and the way that these characteristics would influence the process of data analysis. Since only 13 participants qualify to be included in the data analysis up to this point, any one participant can greatly affect the results.

First, to ensure that participants were consistent in their ratings, I considered the degree to which participants clearly rated a given stimulus type as acceptable or unacceptable. While it is entirely possible that individual structures are neither fully acceptable or fully unacceptable, a large number of unclear ratings suggests some difficulty in performing the task. If many participants rated a good number of structures to be neither fully acceptable or unacceptable, this would suggest an underlying pattern, but, as we will see, that was not the case. More concretely, I counted the number of stimulus types, for both monolingual and CS stimuli, that received an average rating greater than 3 but less than 5, near the midpoint of 4 on the 7-point Likert scale used to judge the sentences. I call these middle ratings. Two participants stood out when looking at the middle ratings. One participant rated 60% of the stimulus types with middle ratings, so I excluded this participant from the analysis. The second participant gave 43% of the stimulus types a middle rating, while the next highest percentage of middle ratings of participants not otherwise excluded was 15%. For this reason, this participant was also excluded.

22Not coincidentally, this participant also rated 60% of the stimuli with a 4.
Next, given the importance of the monolingual judgments for interpreting the CS data, I specifically considered the types of judgments that participants were making for these stimuli. I was interested in whether participants were making distinctions between acceptable and unacceptable word orders, without regard to which specific structures were unacceptable or unacceptable. For one participant, all but one structure for monolingual English and two structure for monolingual Spanish were rated unacceptable or in the middle, compared to about half and half for other participants. Because of the effect that so few contrasts in acceptability would have on interpreting this participant’s CS judgments, this participant was also excluded from the analysis. Finally, whether for performance reasons or otherwise, two participants had an unusually large number of middle ratings for the monolingual stimuli as compared to other participants (more than one standard deviation greater than the other participants). Given that the structures under study are expected to be clearly acceptable or clearly unacceptable in the monolingual context, these two participants were also excluded from the final analysis.

In total, then, five additional participants had to be excluded from the data analysis, leave a set of 8 participants.

4.6.1.3 Participant Characteristics

The final set of 8 participants have the following characteristics:

The participants are well balanced with respect to gender as well as the version of the experiment that they completed. Though they began learning Spanish earlier, their characteristics suggest that they are stronger in English than in Spanish, considering both their proficiency scores and their self-rated proficiencies. Importantly, they do continue to use both Spanish and English in their daily lives. While the parents of most of the final set of participants are from Mexico, a couple of participants have parents from other countries. Because participants were grouped according to linguistic competence, as discussed in the following section, any differences in dialect will be accounted for.

\[^{23}\text{Mean daily use for both languages was calculated by averaging participant’s reported daily use in the following three contexts: (1) at home, (2) with friends and (3) at home or at work. This encompasses the three contexts where people spend most of their time speaking in a given language or set of languages.}\]
TABLE I: PARTICIPANT BACKGROUND OVERVIEW

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Age range</td>
<td>19–42</td>
</tr>
<tr>
<td>Mean age</td>
<td>23.6</td>
</tr>
<tr>
<td>Gender (# male / # female)</td>
<td>4M/4F</td>
</tr>
<tr>
<td>Mean age of acquisition of Spanish</td>
<td>0.25</td>
</tr>
<tr>
<td>Mean age of acquisition of English</td>
<td>3.50</td>
</tr>
<tr>
<td>Mean Spanish proficiency test score (out of 50)</td>
<td>40.9</td>
</tr>
<tr>
<td>Mean English proficiency test score (out of 40)</td>
<td>36.4</td>
</tr>
<tr>
<td>Mean self-reported Spanish proficiency (1–5)</td>
<td>3.63</td>
</tr>
<tr>
<td>Mean self-reported English proficiency (1–5)</td>
<td>4.75</td>
</tr>
<tr>
<td>Mean daily Spanish use</td>
<td>34.2%</td>
</tr>
<tr>
<td>Mean daily English use</td>
<td>65.8%</td>
</tr>
<tr>
<td>Participants Completing Version A</td>
<td>4</td>
</tr>
<tr>
<td>Participants Completing Version B</td>
<td>4</td>
</tr>
<tr>
<td>Participants’ Parents’ Country of Origin</td>
<td></td>
</tr>
<tr>
<td>Mexico</td>
<td>6</td>
</tr>
<tr>
<td>Guatemala</td>
<td>1</td>
</tr>
<tr>
<td>Ecuador</td>
<td>1</td>
</tr>
</tbody>
</table>

4.6.1.4 Participant Groups

The final stage of participant selection involves organizing the participants according to their monolingual acceptability judgments. As discussed in §2.3, the monolingual acceptability judgments indicate the properties of each individual language for a given participant and the corresponding syntactic features that go along with those properties. Those two sets of the syntactic features, in turn, compromise the full set of syntactic features available to the speaker for forming CS sentences. Similar CS data may have different underlying causes, so the acceptability judgments of the CS sentences can only be understood

\[\text{\footnotesize It is, in fact, possible for there to be combinations of features that only occur in code-switching, for example when a light verb is inserted as a last resort, but I will not consider this possibility here, as there is no evidence for it. See González-Vilbazo and López (2011) for some addition discussion.}\]
in the context of the monolingual acceptability judgments. Thus, in order to be able to interpret the CS data for this study, it is important that participants have the same, or substantially similar, judgments for the monolingual data. In other words, they should have comparable I-languages, within the scope of the structures of interest.

For the present study, there are 27 total structures across the two languages that are relevant in comparing the monolingual judgments of the participants, which will be reviewed shortly. While aggregate data from Likert scale acceptability judgments is by its nature continuous, for the purposes of comparing participants, I elected to categorize each participant’s judgments for each of the 27 types according to three levels: grammatical, middle, or ungrammatical. Ratings greater than or equal to 5 were considered grammatical, ratings between 3 and 5 were considered middle ratings, and ratings less than or equal to 3 were considered ungrammatical. For the purposes of comparison, the middle ratings were considered neutral such that a middle rating and a grammatical rating would be considered matching judgments, and a middle rating and an ungrammatical rating would also be considering to be matching. In this way, any arbitrariness in the division between grammatical and middle or ungrammatical and middle ratings is prevented from arbitrarily categorizing some numerically close ratings as a match and others as not matching. Only ratings that are both on different ends from the midpoint and far enough from the midpoint are considered to not be a match.

I began by comparing structures that were more clearly relevant for distinguishing differences in wh-question formation. For example, V2 word order is characteristic of English wh-questions, and it is not expected to be grammatical in monolingual Spanish wh-questions. Even with this properties, there were already differences across participants, as some participants did accept V2 word order in Spanish for specific types of wh-questions. Continuing with another relevant structure, there are some dialectal differences in the grammatical subject positions with matrix wh-questions in Spanish when the wh-phrase is complex. Some speakers accept both SV word order, as in (102a), and VS word order, as in (102b), while other speaker only accept VS word order.

(102) a. ¿Cuántos libros María ha comprado esta semana?
   How many books María has bought this week
   ‘How many books has María bought this week?’
Once again, there was variation across participants with respect to this structure. Continuing this sort of analysis, there are enough significant structures for comparing participants, each of which showed non-overlapping differences between participants, that it became clear that comparing individual structures was not sufficient to form larger groups of participants. Had there been enough overlap between the most relevant structures, it may have been worthwhile to overlook other differences, but this was not the case. For this reason, the final process for grouping was based on comparing across all 27 structures and only grouping those participants who completely matched.

I will review the 27 structures used for comparison and then the process by which participants were compared. The structures come from crossing the main factors for each of the two monolingual sets of stimuli, along with a few extra sets of monolingual structures. For English, the following three factors were relevant, with the corresponding levels in parenthesis: clause (matrix or embedded), complexity (simple or complex) and subject position (SV, V2, VS). Crossing these factors produces 12 structures. For Spanish, the same factors and levels were relevant, except that there was no V2 in the two embedded structures, leaving 10 total structures.

For the additional monolingual stimuli, I focused on one set of related structures in each language. For English, I included the structures based on natural examples of embedded wh-questions. These consisted of a declarative sentence where the embedded clause cannot function as a question, such as (103a), an embedded question with SV word order, as in (103b), and an embedded question with V2 word order, as in (103c).

(103)  

a. *I found out how did they get into the building.

b. You’d be better off asking why he married me.  

c. You’d be better off asking why did he marry me.

These offer an important comparison with the parallel structures in the main set of monolingual English stimuli where the sentences are artificially constructed. For Spanish, I included the two structures based on the position of adverbs and negation that serve as evidence against T-to-C movement in Spanish.
a declarative sentence with an adverb and negation above the verb, as in (104a), and the equivalent interrogative version of the sentence, where the adverb and negation remain above the verb, as in (104b).

\[(104)\]
\[
a. \text{Yo sencillamente no aguanto a ese mocoso.} \\
   \text{I simply cannot stand that brat} \\
   \text{‘I simply cannot stand that brat.’}
\]
\[
b. \text{¿A cuál de esos mocosos sencillamente no aguantas?} \\
   \text{to which of those brats simply not you stand} \\
   \text{‘Which of those brats can you simply not stand?’}
\]

The judgments on these sentences can serve to distinguish between V2 word order that is only superficially similar to English and V2 word order that is a result of T-to-C movement\(^{25}\).

The final groups of participants were determined with the following process. First, all the possible pairings of the final set of 8 participants were compared according to their acceptability judgments across the 27 structures. Again, the comparisons were made using the three categories outlined above: grammatical, middle, or ungrammatical. Participants who categorized all 27 structures the same were placed in the same group. Because of the ambiguity caused by middle ratings, some participants were placed in more than one group. In those cases, a larger group could not be formed because the remaining participants in those individual groups had different overall ratings on one or more structures. This process produced 7 final groups for analysis, where the participant numbers are listed in parentheses:

\[(105)\]  
**Final Participant Groups**

a. Group 1 (108, 116)
b. Group 2 (112)
c. Group 3 (105)
d. Group 4 (101, 113)
e. Group 5 (106)
f. Group 6 (115)

\(^{25}\) Rather than T-to-C movement, V2 word order may also be the result of allowing the auxiliary and participle to be separated, something that can in fact occur in Spanish Suñer (1987). In that case, the subject still remains in situ, but the participle does as well. The auxiliary raises to T, as all finite verbs in Spanish do, moving the auxiliary above the subject. Since the subject remains in Spec\(vP\), it ends up between the auxiliary in T and the participle in V.
g. Group 7 (113, 116)

Again, given the general lack of overlap across the most relevant structures and the importance of grouping participants with comparable I-languages, it was preferable to have smaller but matching groups, even if some groups only contain one individual.

4.6.2 Subject Language

As we saw in §4.2.1.2, the language of the subject was included as an independent variable, and thus every type of wh-question was test with both an English and a Spanish subject. Unfortunately, careful analysis of the effect of the language of the subject on the grammaticality of a given sentence revealed a large amount of variation. Whether comparing across groups or across structures, the effect of the subject language was not clear. One straightforward way of making comparisons is to look at whether the subject and the tense head are in the same language. Since the subject raises to SpecTP in a variety of structures in both languages, it may make a difference whether or not T and the subject are in the same language. For example, a code-switch between the subject and T may make sentences ungrammatical, or it may generally lower the acceptability. Because T and the subject are in the same language in monolingual sentences, the reverse is not expected to be the case. Surprisingly, comparisons across groups and across structures revealed that in some cases switching between T and the subject made otherwise grammatical sentences ungrammatical, while in other cases, not switching between T and the subject made a given structure ungrammatical. Moreover, which switches affected which structures variety by group, or even by sets of group.

Given the large amount of variation and the lack of a discernible pattern, I elected to only analyze data where the subject and the tense head were in the same language. This was based on the aforementioned assumption that keeping two constituents in the same language should not make an otherwise grammatical sentences ungrammatical. Those cases where code-switching between the subject and the tense head appeared to improve grammaticality may have to do with additional properties of subjects or with contrasts between Spanish and English subjects that are beyond the scope of this dissertation.
4.6.3 Lexicalization Analysis

Before analyzing the data, it was also necessary to check whether any lexicalizations had a particularly strong effect on the acceptability judgments of the participants. Recall that the final stimuli were constructed from base monolingual sentences. Each sentence had a translation equivalent sentence in the other language, and the words from these sentence were reordered and combined according to the various structures that make up the final stimuli. Each Spanish-English pair of base sentences form what I call a lexical unit. If the particular vocabulary choices of these lexical units, i.e. the specific lexicalization, has an effect on the acceptability of the stimuli, this effect should be apparent across all structures that use this lexical unit. Otherwise, any difference are likely simple variation.

I chose to operationalize a lexical unit’s effect on acceptability as the degree to which a particular lexical unit moves acceptability ratings toward the midpoint on the rating scale. In other words, if the lexical unit makes the ratings less clear for participants, then it is negatively affecting the acceptability judgments and hence may need to be excluded from the data.

The degree to which a lexical unit moved ratings toward the middle was calculated by comparing the average for a given stimulus type with and without the lexical unit. For example, one stimulus type had an average of 6.04 with all lexical units and an average of 6.33 without lexical unit 7, so this lexical unit moved the rating 0.29 closer to the middle, negatively affecting the clarity of the rating. Using this calculation, positive numbers move the ratings closer to the middle, and negative numbers move the rating further from the middle. These numbers were calculated for all lexical units for all types for CS, monolingual Spanish and monolingual English stimuli, and then the number was summed across all types. The greater the degree to which ratings were moved closer to the middle, the larger the final net number.

To compare the lexical units and determine whether any lexical units stood out, I then calculated the standard deviation for all of the lexical units based on the net calculated numbers described above. Finally, I calculated how many standard deviations each lexical unit was from the average across all lexical units. It turns out that no lexical unit was more than one standard deviation from the average for all lexical units, so I concluded that no lexical unit had a larger enough of a differential effect on the acceptability judgments to warrant removing the lexical unit from the analysis.
4.6.4 Items Analysis

The final step in preparing the data for analysis was to exclude any items, or stimuli, that were markedly different from other items of the same type. Similar to the lexicalizations, this was operationalized as the distance the average acceptability judgment for a given item was from the overall average for that stimulus type. For example, for one stimulus type, the overall average was 5.69, which is 1.69 from the midpoint, while one item was rated 4.31, only 0.31 from the midpoint. In cases like this, where the difference between the distance for the average (here 1.69) and the distance for the individual item (here 0.31) was greater than one (here 1.38), the item was excluded from the final data analysis. A difference of 1 was considered large enough because it is a substantial fraction of the range of possible acceptability judgments (6 total, from 1 to 7). Seven items were ultimately excluded based on this process.

4.6.5 Interpreting Standard Deviations

As a final note on preparing the data for analysis, I would like to contextualize the relatively high standard deviations that appear in some of the results to follow. Essentially, since there are only four ratings per type (i.e. per structure) for some groups, the standard deviation can quickly become large when compared to the maximum rating of 7 (or compared to the average for that structure).

Consider the example ratings in Table II. Each row has a general description of the kind of ratings, the four ratings for the example, the average ($\mu$) and the standard deviation ($\sigma$). In the first three rows, we see various combinations of relatively low ratings and how one higher rating begins to increase the standard deviation. At this point, the standard deviation is already a sizable fraction of the average itself. Notice that if the ratings consistent only of 1s and 7s, the standard deviation increases sharply. If three of the four ratings are very low but one is very high, as in row 4, the standard deviation jumps to 3.00. If half the ratings are 1s and half the ratings are 7s, as in row 5, the standard deviation increases even further.

Notice that the sets of ratings in row 4 and row 5 actually have similarly high standard deviations. If sets of ratings like 4 are just as common as sets of ratings like 5, it becomes hard to tell if there is

---

Note that, in general, participants used the entire Likert scale across all stimuli. Only one participant, participant 115, limited him or herself to the extremes of the Likert scale, using only 1 or 7. Importantly, this participant did not have more middle ratings (those not clearly acceptable or unacceptable, i.e. between 3 and 5) than other participants.
TABLE II: THE EFFECT OF RATINGS ON STANDARD DEVIATION

<table>
<thead>
<tr>
<th>Rating 1</th>
<th>Rating 2</th>
<th>Rating 3</th>
<th>Rating 4</th>
<th>μ</th>
<th>σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Mix of low ratings</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>1.75</td>
</tr>
<tr>
<td>2. Low or middle ratings</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>2.50</td>
</tr>
<tr>
<td>3. Low and one higher rating</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>2.50</td>
</tr>
<tr>
<td>4. One opposite rating</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>2.50</td>
</tr>
<tr>
<td>5. Contradictory ratings</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>7</td>
<td>4.00</td>
</tr>
</tbody>
</table>

an important difference between the two. It may be that the participant is sometimes choosing between 7 and 1 more or less by chance, in which case the difference between row 4 and row 5 is just flipping the rating. Importantly, this does not appear to be the case. With 20 code-switching structures and 22 monolingual structures for each of the 7 groups, there are 294 total structures for which a standard deviation can be calculated and, only 12.9% of those structures have a standard deviation that is greater than or equal to 3.00. In addition, not all of those structures are like rows 4 and 5. That only a small fraction of the ratings have such a high standard deviation suggests that chance flipping of individual ratings from 1 to 7, or vice versa, is not appreciably affecting the results.
5 Discussion

5.1 Factors Contributing to Word Order
  5.1.1 Subject
  5.1.2 Wh-Phrase
  5.1.3 Tense Head
  5.1.4 Complementizer Head
    5.1.4.1 C as a Determinant Head
    5.1.4.2 Grammatical Subject Positions and the Language of C
    5.1.4.3 Additional Evidence for C as the Determinant Head
  5.1.5 Conclusions

5.2 Generalizations Based on the Language of C
  5.2.1 The Language of C
  5.2.2 Code-Switching Between C and T
  5.2.3 Code-Switching Between C and the Wh-Phrase

5.3 Empirical Generalizations

5.4 Implications
  5.4.1 Wh-Questions in Spanish and English
    5.4.1.1 Wh-Questions in English
    5.4.1.2 T-Based Accounts of Spanish Inversion
    5.4.1.3 Other Accounts of Spanish Wh-questions
  5.4.2 Wh-Questions More Generally
  5.4.3 Wh-Questions in Code-Switching
  5.4.4 Code-Switching More Generally
    5.4.4.1 Code-Switching as a Tool of Linguistic Analysis
    5.4.4.2 Theoretical Considerations
    5.4.4.3 Empirical Tests for the Code-switching Framework

5.5 Summary and Conclusions
The goal of this section is to answer the research questions and hypotheses laid out in §3.3 and discuss any implications the results have for wh-questions in Spanish-English code-switching as well as Spanish, English, and the structure of wh-questions in general.

In order to address Research Question 1, repeated below, I begin in §5.1 by reviewing each of the heads or phrases that may most plausibly determine the grammatical subject positions in wh-questions.

RQ1: What factors determine the word order in wh-questions in Spanish-English code-switching?

As discussed in §1.1, these are the following: the subject, the wh-phrase, the tense head, and the complementizer head. The prediction is that if a given head or phrase determines the grammatical subject positions, then the grammatical word orders of the CS sentences should always match those of the monolingual sentences in the language of that head or phrase. We will see that this prediction does not hold for the subject (§5.1.1), the wh-phrase (§5.1.2) or the tense head (§5.1.3).

Since the complementizer is covert in the structures under study here, this prediction cannot be directly tested by inspecting the language of the complementizer and comparing it to the corresponding monolingual sentences. Still, of the four possible heads or phrases that may determine the grammatical word orders for inversion, the complementizer is the only remaining head. Combining this fact with additional evidence in §5.1.4.3 that C determines various other grammatical properties of its complement in code-switching, we conclude it must be the syntactic properties of C that are ultimately responsible for the grammatical subject positions.

I address Research Question 2, repeated below, in §5.2. I argue in §5.1.4.2 that, once we know that C determines the grammatical subject positions, we can ascertain the language of C in a given CS sentence by inspecting the grammatical subject positions for that sentence.

RQ2: Are there restrictions on code-switching between particular constituents in wh-questions in Spanish-English code-switching?
This permits us to test restrictions on code-switching between C and the wh-phrase and between C and the tense head. We will see that, in general, both kinds of code-switches are possible, but there is one important exception. With a simple wh-phrase in an embedded question, C and T cannot be in the same language when T is in Spanish. Since maintaining two heads in the same language should not be ungrammatical, given that it is grammatical in monolingual sentences, there must be another restriction involved. I suggest that the issue, in fact, lies in the relationship between the wh-phrase and C. The normally grammatical configuration in which C and T are in the same language becomes ungrammatical because of a restriction on code-switching between the wh-phrase and C.

In the final section of this chapter, §5.4 I discuss the implications for the results uncovered in the previous sections with respect to properties of wh-questions, wh-questions in code-switching, and code-switching more generally.

5.1 Factors Contributing to Word Order

This section focuses on the first research question and subquestions, repeated below.

RQ1: What factors determine the word order in wh-questions in Spanish-English code-switching?

RQ1-A: Do the syntactic properties of the subject determine which subject positions are grammatical?

RQ1-B: Do the syntactic properties of the wh-phrase determine which subject positions are grammatical?

RQ1-C: Do the syntactic properties of the tense head determine which subject positions are grammatical?

RQ1-D: Do the syntactic properties of the complementizer head determine which subject positions are grammatical?

I consider the role that the subject, the wh-phrase, the tense head, and the complementizer play in word order in Spanish-English code-switching, each in turn, concluding that it is C that ultimately determines the grammatical subject positions in both monolingual and CS wh-questions.
5.1.1 Subject

In order to address Research Question 1 (RQ1), we begin with the first research sub-question (RQ1-A) and the hypothesis for that sub-question (H1-A), repeated below.

**RQ1:** What factors determine the word order in wh-questions in Spanish-English code-switching?

**RQ1-A:** Do the syntactic properties of the subject determine which subject positions are grammatical?

**H1-A:** No, the syntactic properties of the subject will not ultimately determine which subject positions are grammatical.

As discussed in §2.5, I will be using two straightforward premises, repeated below, in order to test each proposed hypothesis, including H1-A, and draw appropriate conclusions from the data.

**Premise 1:** There is a head which has a specific set of syntactic features that ultimately determine the grammatical subject positions in a given wh-question. Call this a determinant head.

**Premise 2:** A head in a given language and in a parallel structure has the same set of syntactic features whether in a monolingual or CS sentence.

RQ1-A is essentially asking whether the subject contains what Premise 1 refers to as a determinant head, and H1-A proposes that it does not. Let us assume that the subject does in fact contain the determinant head, for the sake of argument, see what predictions this makes, and test these predictions with data from our participant groups. If H1-A is correct, these predictions will not hold. Since inversion is a different syntactic process in each language, we will need to test whether the subject is the

---

1While the subjects in this experiment are clearly phrasal, they still will contain a head which is the locus of syntactic features for the phrase, like any other phrase. Whether the potential determinant head is the determiner head or some other head that makes up the subject will not affect the results. Since the subjects for this experiment are always in one language, as laid out in §4.2.2, the language of that head will be the language of the subject.

2Note that these data come from stimuli where the subject and the verb are not in the same language, unlike the rest of the data analysis. This is due to two factors. First, in most cases the grammatical word orders for a given structure match the language of the verb, so these data cannot provide counterevidence against the hypothesis being tested. Second, to simplify exposition, I sought out examples where the largest number of groups possible had the same judgments for the stimuli. This was more feasible with the data where the subject and verb are not in the same language.
determinant head in each language individually. It maybe turn out that each language has a different determinant head.

We will begin by testing a Spanish subject. By Premise 2, it will have the same features in both CS and monolingual sentences. Since the subject’s features determine the grammatical subject positions, as per Premise 1, and those features have to be the same for CS and monolingual, the same subject positions should be grammatical in both monolingual Spanish and CS sentences with a Spanish subject. Now consider the data in (106), where each question has VS word order and the subject of the embedded wh-question appears in bold.

Translation-equivalent sentences are listed in English, Spanish and CS, respectively.

(106) Grammaticality of VS word order with Spanish subjects

a. * No recuerdo cuántas manzanas have eaten tus hijos since Sunday. CS
   not I-remember how many apples your children
   ‘I don’t remember how many apples have eaten your children since Sunday.’

b. No recuerdo cuántas manzanas han comido tus hijos desde el domingo. Sp
   not I-remember how many apples have eaten your children since the Sunday
   ‘I don’t remember how many apples have eaten your children since Sunday.’

c. * I don’t remember how many apples have eaten your children since Sunday. Eng

If the Spanish subject determines the grammatical subject positions in wh-questions, then VS word order should be grammatical whether the sentences are code-switched, as in (106a), or in Spanish, as in (106b). Instead, VS word order is grammatical in Spanish but ungrammatical in CS. Notice that VS word order is also ungrammatical in English, as seen in (106c). This suggests that some English element of the code-switching sentence is actually the determinant head.

Having seen that the subject is not the determinant head in Spanish, let us consider whether English subjects are determinant heads in wh-questions. Again, the prediction is that if English subjects are determinant heads, then the same subject positions should be grammatical, or ungrammatical, whether the subject is in a monolingual or a CS sentence. Take the wh-questions in (107), where each question has VS word order and translation-equivalent sentences are listed in CS, English and Spanish. We focus

---

3 The acceptability judgments for these sentences are the same for all groups except Group 6, which will be discussed below.
here on VS word order for the sake of exposition, but this should be sufficient to test the hypothesis. The determinant head is expected to determine grammatical subject positions across word orders and across all types of wh-questions.\(^4\)

(107) **Grammaticality of VS word order with English subjects**

a. How many poems *han escrito* your *sisters* *este mes*? CS
   ‘How many poems have written your sisters this month?’

b. *How many poems have written* your *sisters* this month? Eng

c. ¿Cuántos *poemas* han *escrito* tus *hermanas* *este mes*? Sp
   ‘How many poems have written your sisters this month?’

Once again, if the subject is in the same language and also is the determinant head, then VS word order should have the same grammatical status whether the sentences are code-switched, as in (107a), or in English, as in (107b). Instead, VS word order is grammatical in CS but ungrammatical in English. As before, the grammatical status of VS word order in the CS example matches the other language, here monolingual Spanish as in (107c), suggesting that some Spanish element of the code-switching sentence is actually the determinant head in the CS sentence.

The examples in (106) and (107) are representative, but there is one group, Group 6, that does not have the same judgments for the sentences in those examples.\(^5\) Still, this group does have parallel data for Spanish subjects, as seen in (108).

(108) **Group 6 Grammaticality of VS word order with Spanish subjects**

a. *Cuántos *poemas* have written tus *hermanas* this month? CS
   ‘How many poems have written your sisters this month?’

b. ¿Cuántos *poemas* han *escrito* tus *hermanas* *este mes*? Sp
   ‘How many poems have written your sisters this month?’

c. *How many poems have written* your *sisters* this month? Eng

\(^4\)Again, keeping assumptions as simple as possible and seeing how far this can take us. See \(^2.5\) for discussion.

\(^5\)In the following examples, unlike the previous examples, the subject and the verb are in the same language.
Like with previous examples, VS word order is predicted to have the same grammatical status in both the CS sentences, as in (108a), and the Spanish sentences, as in (108b), since the subject is in Spanish in both sentences. Instead, VS word order is ungrammatical in CS but grammatical in Spanish, contra the predictions if the subject is the determinant head. On the other hand, VS word order is ungrammatical in the English data, as seen in (108c), suggesting that the determinant head is in English in CS sentences.

In the case of testing an English subject with Group 6, there actually is no complete three-way contrast because the relevant word order has middle ratings in monolingual English, monolingual Spanish, or CS. Still, since the prediction is that an English subject should force English word order, we can look at a structure with a clear contrast between the grammaticality of the CS and monolingual English ratings, as in (109).

(109) Group 6 Grammaticality of VS word order with English subjects

a. No recuerdo cuántas manzanas han comido tus hijos desde el domingo. Sp
   not I-remember how many apples have eaten your children since the Sunday
   ‘I don’t remember how many apples have eaten your children since Sunday.’

b. No recuerdo cuántas manzanas han comido tus hijos desde el domingo. CS
   not I-remember have eaten your children since the Sunday
   ‘I don’t remember how many apples have eaten your children since Sunday.’

c. *I don’t remember how many apples have eaten your children since Sunday. Eng

There is a clear contrast between the grammatical CS sentence in (109b) with an English subject and the ungrammatical monolingual English sentence in (109c). If the English subject were a determinant head, then this would not be the case. While the rating for monolingual Spanish is unclear, as seen in (109a), the prediction for English subjects as determinant heads centers on the whether a given subject position has the same grammatical status in both English and CS, which is clearly not the case here.

Returning to RQ1-A and H1-A, we have found evidence confirming the hypothesis that the syntactic properties of the subject do not ultimately determine which subject positions are grammatical. Rather than the subject determining the grammatical subject positions, we have seen that the grammaticality of a given word order can vary even when the subject is in the same language. Some other head or phrase must determine the grammatical subject positions in wh-questions.
5.1.2 Wh-Phrase

We now turn to the second research sub-question (RQ1-B) of Research Question 1 (RQ1) and the hypothesis for that sub-question (H1-B), repeated below.

**RQ1:** What factors determine the word order in wh-questions in Spanish-English code-switching?

**RQ1-B:** Do the syntactic properties of the wh-phrase determine which subject positions are grammatical?

**H1-B:** No, the syntactic properties of the wh-phrase will not ultimately determine which subject positions are grammatical.

As with the previous section, and as discussed in §2.5, I will be using two straightforward premises, repeated below, in order to test each proposed hypothesis, including H1-B, and draw appropriate conclusions from the data.

**Premise 1:** There is a head which has a specific set of syntactic features that ultimately determine the grammatical subject positions in a given wh-question. Call this a determinant head.

**Premise 2:** A head in a given language and in a parallel structure has the same set of syntactic features whether in a monolingual or CS sentence.

RQ1-B is essentially asking whether the wh-phrase contains what Premise 1 refers to as a determinant head, and H1-B proposes that it does not. Like with the previous section, let us assume that the wh-phrase does in fact contain the determinant head, for the sake of argument. Following Premise 2, a wh-phrase in a given language will have the same features in both CS and monolingual sentences. Since the features of the wh-phrase determine the grammatical subject positions, as per Premise 1, and those features have to be the same for CS and monolingual sentences, this predicts that the same

---

6Depending on the analysis, wh-phrases are not necessarily heads, particularly complex wh-phrases, but they still will contain a wh-head which is the locus of syntactic features for the phrase, like any other phrase. Since the wh-phrases for this experiment are always in one language, as laid out in §2.2.2, the language of the wh-head will be the language of the wh-phrase.
subject positions should be grammatical in both monolingual and CS sentences with a wh-phrase in a
given language. For example, if the wh-phrase in the CS sentences is in Spanish for a given structure,
the grammatical word orders should match those of monolingual Spanish for that structure. If H1-B is
correct, these predictions will not hold. Since inversion is a different syntactic process in each language,
we will need to test whether the subject is the determinant head in each language individually. It may
turn out that each language has a different determinant head.

Consider, for example, matrix questions with a complex wh-phrase, bolded in the examples below,
as in [110], where all participant groups agree on the judgments.

(110) Grammaticality of VS word order with Spanish and English wh-phrases

a. *How many blouses have bought your nieces this fall? Eng

b. ¿Cuántas blusas han comprado tus sobrinas este otoño? Sp
   'How many blouses have bought your nieces this fall?'

c. How many blouses han comprado tus sobrinas this fall? CS
   'How many blouses have bought your nieces this fall?'

d. *Cuántas blusas have bought your nieces este otoño? CS
   'How many blouses have bought your nieces this fall?'

In English, questions like (110a) where the subject appears after the auxiliary and participle (i.e. in VS
word order) are ungrammatical. On the other hand, in Spanish, the same VS word order is perfectly
grammatical, as in (110b). If the language of the wh-phrase predicts the grammatical word orders, as
hypothesized, then VS in a code-switched sentence with an English wh-phrase should be ungrammatical,
since the acceptability should match that of the language of the wh-phrase. However, it turns out that
such sentences are ungrammatical, as seen in (110c). Likewise, in a question with a Spanish wh-phrase,
VS word order is predicted to be grammatical, but that word order is in fact ungrammatical, as seen in
(110d).

Returning to RQ1-B and H1-B, we have found evidence confirming the hypothesis that the syntactic
properties of the wh-phrase do not ultimately determine which subject positions are grammatical. Rather
than the wh-phrase determining the grammatical subject positions, we have seen that the grammaticality
of a given word order can vary even when the language of the wh-phrase is held constant. Some other head or phrase must determine the grammatical subject positions in wh-questions.

5.1.3 Tense Head

We now turn to the third research sub-question (RQ1-C) of Research Question 1 (RQ1) and the hypothesis for that sub-question (H1-C), repeated below.

**RQ1**: What factors determine the word order in wh-questions in Spanish-English code-switching?

**RQ1-C**: Do the syntactic properties of the tense head determine which subject positions are grammatical?

**H1-C**: Yes, the syntactic properties of the tense head will ultimately determine which subject positions are grammatical.

As with the previous section, and as discussed in §2.5, I will be using two straightforward premises, repeated below, in order to test each proposed hypothesis, including H1-C, and draw appropriate conclusions from the data.

**Premise 1**: There is a head which has a specific set of syntactic features that ultimately determine the grammatical subject positions in a given wh-question. Call this a determinant head.

**Premise 2**: A head in a given language and in a parallel structure has the same set of syntactic features whether in a monolingual or CS sentence.

RQ1-C is essentially asking whether the tense head is what Premise 1 refers to as a determinant head, and H1-C proposes that this is in fact the case. Once again, if the syntactic properties of the tense head in a given language determine which subject positions are grammatical, as assumed in Premise 1, then the grammatical subject positions in CS are expected to correspond with the grammatical subject positions of the monolingual sentences in that language, following Premise 2. For example, if the tense
head in the CS sentences is in Spanish for a given structure, the grammatical word orders should match those of monolingual Spanish for that structure.

Because T, unlike the subject and the wh-phrase, is a bound morpheme, it might in principle be in a different language from the rest of the word that it is part of. It is important to be able to establish the language of T in order to interpret its role in the sentence, so I consider several arguments that the language of T corresponds to the language of the verb.

The first argument is based on the different ways in which Spanish and English express aspect morphologically. Because the inflections for tense and aspect of each language mark different aspectual information, such as the preterite/imperfect contrast of Spanish, I take these tense/aspect morphemes (syncretic in both languages) to be a reflection of the language of T. More concretely, the morphology serves as an index to a particular set of features and the language that those features are associated with. Since these morphemes are always in the same language as the verb in my data, I take the language of the verb to reflect the language of T.

The second argument is based on incorporation and additionally suggests that little v is in the same language as T and the verb. As discussed in §2.1.2.2, the verb in Spanish incorporates into little v, and little v in turn incorporates into T. Because these three heads form one morphological word, I think there is good reason to believe that they are in one language, as long as all the individual overt morphemes are realized in one language. In English, T either lowers to little v or is expressed on an auxiliary. In either case, T is incorporated into a verb, so I argue, along similar lines, that T is in the same language as the verb in English.

We are now ready to consider the prediction that the same subject positions will be grammatical in CS and monolingual sentences where T is in the same language. In many cases, this prediction holds true, such as in example (110), repeated here as (111), where the grammaticality of VS word order in CS matches that of the monolingual judgments. When T, bolded in the examples below, is in English in a CS sentence, as in (111a), the grammatical status of the VS word order matches that of monolingual

---

7This may be either the finite main verb, if there is no auxiliary, or the auxiliary verb. In either case, it is the verb that reflects tense, and it is the language of this verb that determines the language of T. Note that in my data the auxiliary and the particle (i.e. the main verb) are always in the same language.

8This is to distinguish from a situation in which the stem of a verb is in one language and the inflection is in another language. For example, as discussed in González-Vilbazo and López (2011), Spanish-German code-switchers can say words such as ‘utilisieren’, meaning ‘use’, composed of the Spanish root ‘utiliz’ and the German infinitive suffix ‘-ieren’.
English, as in (111b). In other words, the word order is ungrammatical. On the other hand, when T is in Spanish in a CS sentence, as in (111c), the grammatical status of the VS word order matches that of monolingual Spanish, as in (111d), and is therefore grammatical.

(111) **Word order can match the language of T (VS word order)**

a. *Cuántas blusas have bought your nieces este otoño?*  
   how many blouses have bought your nieces this fall?  
   ‘How many blouses have bought your nieces this fall?’

b. *How many blouses have bought your nieces this fall?*  
   Eng

c. How many blouses *han comprado tus sobrinas this fall?*  
   have bought your nieces this fall?  
   ‘How many blouses have bought your nieces this fall?’

d. ¿Cuántas blusas *han comprado tus sobrinas este otoño?*  
   how many blouses have bought your nieces this fall  
   ‘How many blouses have bought your nieces this fall?’

However, this is not always the case. There is a general pattern in which the grammatical word orders for one specific structure, embedded questions with simple wh-phrases and a Spanish T, match the grammatical word orders of English. While there is not a parallel pattern with an English T, this still constitutes strong evidence against the hypothesis that syntactic properties of T determine the position of the subject in wh-questions. CS questions with an English T may require an English word order for independent reasons.

Consider the data in (112). Participants find VS to be ungrammatical in English embedded questions with simple wh-phrases, as in (112a), but they find VS to be grammatical for Spanish in the same structure, as in (112b).

(112) **Word order does not match the language of T (comparing VS word order)**

a. *I don’t remember what have bought your colleagues this week.*  
   Eng

b. No recuerdo qué *han comprado tus colegas esta semana.*  
   not I-remember what have bought your colleagues this week  
   ‘I don’t remember what your colleagues have bought this week.’

c. *I don’t remember what han comprado tus colegas this week.*  
   have bought your colleagues  
   ‘I don’t remember what your colleagues have bought this week.’
Likewise, these participants also judge the VS word order to be ungrammatical for the same structure in CS, as in (112c). If the language of T, and the corresponding syntactic features from that language, determined the grammatical word orders as predicted by H1-C, then the CS example with a Spanish T should be grammatical with VS word order, just as it is in Spanish. Instead, the CS sentence is ungrammatical with VS word order, providing counter evidence against the hypothesis in question.\footnote{While we do not have CS evidence that T is not the determinant head in English, there is already good evidence that this is the case in monolingual English. See §3.1.3 for discussion.}

It is important to note that there are two types of apparent exceptions that must be considered, beginning with Group 3. Importantly, this is because this group allows both the English and Spanish word orders, so the English word order is grammatical with a Spanish T for this group just as with the other groups. This can be seen in (113).

(113) **Group 3 allows English and Spanish word order**

a. I don’t remember what your colleagues **have bought** this week. \hspace{1cm} Eng SV  
b. *I don’t remember what **have bought** your colleagues this week. \hspace{1cm} Eng VS  
c. *No recuerdo qué tus colegas **han comprado** esta semana. not I-remember what your colleagues have bought this week  
‘I don’t remember what your colleagues have bought this week.’ \hspace{1cm} Sp SV  
d. No recuerdo qué **han comprado** tus colegas esta semana. not I-remember what have bought your colleagues this week  
‘I don’t remember what your colleagues have bought this week.’ \hspace{1cm} Sp VS  
e. I don’t remember what **tus colegas han comprado** this week. your colleagues have bought \hspace{1cm} CS SV  
‘I don’t remember what your colleagues have bought this week.’  
f. I don’t remember what **han comprado tus colegas** this week. have bought your colleagues \hspace{1cm} CS VS  
‘I don’t remember what your colleagues have bought this week.’

In monolingual English, SV word order is grammatical, but VS word order is ungrammatical, as seen in the contrast between (113a) and (113b). In monolingual Spanish, SV word order is ungrammatical but VS word order is grammatical, as seen in (113c) and (113d). On the other hand, both word orders are grammatical in CS, as seen in (113e) and (113f). In summary, Group 3 accepts both SV, as in English,
and VS, as in Spanish, with a Spanish T in CS questions, exactly as would be predicted if the word orders of both languages were possible.

There is one group that represents a true exception to the generalizations just outlined, Group 6. For this group, the monolingual judgments for each structure follow one of three patterns, two of which prevent directly testing whether the word order matches the language of T. First, for all questions with simple wh-phrases, all word orders for that structure are judged ungrammatical or receive middle ratings in CS, precluding any kind of comparison with monolingual equivalent structures. Second, for the questions with complex wh-phrases, the word orders that are grammatical in Spanish are also grammatical in English. Whenever the grammatical word orders match those of Spanish, it becomes impossible to tell whether the grammatical word orders are based only on Spanish or are also based on English, since they overlap. For example, both SV and VS word orders are grammatical in monolingual Spanish, and SV word order is grammatical in monolingual English. Since both SV and VS word orders are grammatical for the CS sentences, it becomes impossible to tell if T can only be in Spanish or if it can also be in English, precluding testing of H1-C. Finally, there are cases where the word order clearly matches one language but T is also in that language, along the lines of (111) rather than (112). This is the case with complex wh-questions with an English T. Because of this, these data do not provide evidence against the hypothesis that syntactic features of T determine the position of the subject.

5.1.4 Complementizer Head

Before considering the evidence for whether C determines the subject position in wh-questions, I need to address the issue of circularity in the argumentation. A key piece of evidence for evaluating each of the previous research sub-questions and their associated hypotheses has been the language of that head. For example, if T predicted the grammatical subject positions, then when T was in Spanish, we should observe the same grammatical word orders in both monolingual and CS sentences, under the assumption that T has the same features in both contexts and these features are what ultimately determine the grammatical subject positions. Since C is covert in the data analyzed here, and we therefore cannot directly observe the language of C, the evidence for the relationship between C and the subject position must necessarily be more indirect.
In addition, we would like to establish that there is a biconditional, or bidirectional, relationship between C and the grammatical subject positions such that if we inspect the grammatical subject positions, comparing the CS and monolingual data, we can determine the language of the C head. This will allow us to answer RQ2 with respect to the restrictions on code-switching between C and other parts of the wh-question. Without knowing the language of C, we would not be able to determine whether there was a code-switch to begin with, let alone which switches were grammatical.

In essence, then, we are interested in whether C determines the grammatical subject positions, answering RQ1-D, and also whether the grammatical subject positions can tell us the language of C. This sounds suspiciously circular, though notice that in the former we are looking at the syntactic behavior of C and in the latter we are looking at the language of C. To avoid any potential issues of circular reasoning, I would like to carefully layout the arguments for each of the two goals related to C. To do so, I will once again use the same two premises, repeated in the following section, we have been using for each sub-question of RQ1. For each of these goals, we will see that just these two premises, and the empirical evidence reviewed above, are sufficient to draw the indented conclusions while avoiding circular reasoning.

In §5.1.4.1 I walk through the argument that C is a determinant head, and in §5.1.4.2 I argue that, taking C to be the determinant head, we can use the grammatical subject positions to determine the language of C in a given wh-question.

5.1.4.1 C as a Determinant Head

We are now in a position to combine the results from sections 5.1.1, 5.1.2, 5.1.3, and the two premises, repeated below, to argue that it is C that ultimately determines the grammatical subject positions in Spanish and English. Essentially, we are going to apply a process of elimination to a set of possible determinant heads or phrases until we are left with C, avoiding any issues of circularity.

**Premise 1:** There is a head which has a specific set of syntactic features that ultimately determine the grammatical subject positions in a given wh-question. Call this a determinant head.
Premise 2: A head in a given language and in a parallel structure has the same set of syntactic features whether in a monolingual or CS sentence.

From Premise 1 we know that there is some head that determines the grammatical subject positions in each language, and as discussed in §1.1 the heads or phrases that may most plausibly determine the grammatical subject positions are the following: C, T, the wh-phrase, and the subject. Now we need to test each of these heads or phrases in turn.

Let us begin by assuming that T is the determinant head. Since inversion is a different syntactic process in each language, it maybe turn out that each language has a different determinant head. We’ll begin by testing if T is in Spanish. By Premise 2, it will have the same features in both CS and monolingual sentences. Since T’s features determine the grammatical subject positions, as per Premise 1, and those features have to be the same for CS and monolingual sentences, then the same subject positions should be grammatical in both monolingual Spanish and CS sentences with a Spanish T. As we saw in §5.1.3 this is not the case, so T cannot be the determinant head in Spanish. While there is no direct evidence that T is not the determinant head in English, there is independent evidence that C is the determinant head in English based on the role of T-to-C movement in English inversion.

Applying the same reasoning using the data discussed in §5.1.1 on subjects and §5.1.2 on wh-phrases, neither of these phrases contains a determinant head in English or in Spanish. Based on the data, none of the other plausible determinant heads can logically be the determinant head in either language, therefore C must be the determinant head in both Spanish and English. We will discuss some additional evidence for this conclusion in section §5.1.4.3.

5.1.4.2 Grammatical Subject Positions and the Language of C

As mentioned earlier in this section, knowing that C is the determinant head, we would like to be able to argue that if we inspect the grammatical subject positions, comparing the CS and monolingual data, we can determine the language of the C head. This will allow us to address the second research question, as discussed in §5.1.4. The argument proceeds as follows, once again using the following two premises:
Premise 1: There is a head which has a specific set of syntactic features that ultimately determine the grammatical subject positions in a given wh-question. Call this a determinant head.

Premise 2: A head in a given language and in a parallel structure has the same set of syntactic features whether in a monolingual or CS sentence.

We begin with a set of CS sentences with one set of grammatical subject positions (e.g. ✓SV and ✓VS), monolingual sentences in language A with the same grammatical subject positions as the CS sentences, and another set of monolingual sentences in language B with different grammatical subject positions (e.g. *SV and ✓VS). It is important that the monolingual judgements are different, or we cannot distinguish between the two languages. By Premise 1, the determinant head in the CS sentences has a set of features that determine the grammatical subject positions in those sentences. That determinant head can be in either of the two languages, like all words in a CS sentence, and we don’t directly know the language because C is covert.

Let’s test each language, starting with the one where the grammatical subject positions are not the same, language B. If C is in language B, then by Premise 2, it must have the same features as C in the monolingual sentences for language B. Since C is a determinant head, by Premise 1, it must determine the word order in the CS and monolingual sentences for a particular language. Since C has the same features, as argued above, it must determine that the same subject positions are grammatical in both the CS and monolingual sentences, as the same set of features should produce the same outcome, all else being equal. The grammatical subject positions are not the same for the CS and monolingual sentences in language B, so C cannot be in the same language in the CS sentences as it is in the monolingual sentences. If C cannot be in language B, it must be in language A. In addition, following the same logic above with language A does not produce a contradiction. Put concretely, if the grammatical subject positions are different in monolingual Spanish and English and CS grammatical subject positions match one language but not the other language, C must be in the matching language.

We have seen, then, that by inspecting the grammatical subject positions, we can determine the language of the determinant head, C.
5.1.4.3 Additional Evidence for C as the Determinant Head

There is an additional type of evidence supporting the conclusion that C determines the grammatical subject positions in Spanish, English and Spanish-English wh-questions. This comes from a broader hypothesis under which C determines a variety of different grammatical properties that go beyond inversion in wh-questions. If this is correct, then the relationship between the language of C and the properties of inversion for a given wh-question is just a reflection of a more general properties of C where it determines a variety of grammatical properties.

In particular, González-Vilbazo and López (2013) provide evidence that the language of C, and therefore its associated syntactic features, determines the realization of the following properties (i) null subjects, (ii) negation and (iii) whether T is right-headed or left-headed. Before reviewing the data on each property, I need to discuss the data itself. First, all of the data comes from Spanish-German code-switching. Second, González-Vilbazo and López (2013) use evidence from a variety of both declarative and interrogative sentences to argue that, ultimately, we can determine the language of C in sentences like the following by inspecting the language of the simple wh-phrase:

(114) a. No sé a quién Juan amenazó.  
    [neg] know.Is who-ACC Juan threatened  
    ‘I don’t know who Juan threatened.’  
    \[C_{\text{Ger}} \rightarrow \text{SV}\]

b. Ich weiß nicht, a quién drohte Juan.  
    I know not who-DAT threatened Juan-NOM  
    ‘I don’t know who Juan threatened.’  
    \[C_{\text{Spa}} \rightarrow \text{VS}\]

In (114a), since the wh-phrase is in German, C is also in German, and because SV word order is required in German embedded clauses, the sentence is grammatical with SV word order, in spite of the Spanish verb. In (114b), on the other hand, the wh-phrase is in Spanish, so C is also in Spanish, and because VS word order is required in this kind of wh-question, VS word order is grammatical. In this way, the language of the simple wh-phrase indicates the language of C.
We will begin by looking at whether C determines the properties of null subjects. Spanish permits sentences to have unpronounced or null subjects, as seen in (115), while German does not, as seen in the contrast between (116a) and (116b)\(^\text{10}\)

(115) No sé a quién pro amenazó.
[NEG] I know [ACC] who pronominal null subject threatened
‘I don’t know who she/he threatened.’

(116) a. Ich weiß nicht, wem Juan drohte.
I know not who-DAT Juan-NOM threatened.
‘I don’t know who Juan threatened.’

b. *Ich weiß nicht, wem pro drohte.

Since null subjects are traditionally considered to be licensed by the tense head (Barbosa 1995), we will need to consider the language of both C and T in testing the properties of null subjects in CS, giving us four possible combinations:

(117) **Null Subject in Spanish-German Code-switching**

a. Ich weiss nicht a quién hizo drohen pro C\(_{\text{Spa}}\) T\(_{\text{Spa}}\)
I know not who-ACC did threaten null subject
‘I don’t know who she/he threatened.’

b. *No sé, wen/wem pro drohte.
[NEG] know.1s who-ACC/who-DAT null subject threatened
‘I don’t know who she/he threatened.’

c. *No sé wen pro amenazó.
[NEG] know.1s who-ACC null subject threatened
‘I don’t know who she/he threatened.’

d. *No sé a quién pro drohte.
[NEG] know.1s who-ACC null subject threatened
‘I don’t know who she/he threatened.’

It turns out that null subjects are only grammatical when both C and T are in Spanish, as in (117a), suggesting that the requirements for null subjects go beyond properties of T to encompass properties

\(^{10}\text{Note that the Spanish verb for ‘threaten’ assigns accusative case, but the equivalent German verb assigns dative case. For this reason, examples will sometimes have a wh-phrase marked for dative Case, sometimes have a wh-phrase marked for accusative Case, and sometimes have both wh-phrases. The reason for a particular choice is not relevant to the discussion at hand.}\)
of C as well. González-Vilbazo and López (2013), then, have provided novel evidence that C plays a crucial role in null subjects.

Next we turn to the position of T. In German, the tense head is clause-final in embedded clauses, as evidenced by the position of the auxiliary in (118a). In other words, T is right-headed in these clauses and therefore comes after its vP complement, as seen in (118b).

(118)  a. Ich glaube dass Hans dem Professor gedroht hat.
I think that Hans the professor threatened has
‘I think that Hans has threatened the professor.’

b. Ich glaube dass [TP Hans [vP dem Professor gedroht] hat].

Spanish, on the other hand, has a left-headed T in all clauses. What happens if C and T are in different languages? Consider the examples in (119) and (120).

(119)  **Spanish C with Left- and Right-headed T**

a. *Yo no sé a quién hat das Buch verkauft Hans.*
I [NEG] know.1s who-ACC has the book sold Hans
‘I don’t know who Hans sold the book to.’

b. *Yo no sé a quién das Buch verkauft hat Hans.*
I [NEG] know.1s who-ACC the book sold has Hans
‘I don’t know who Hans sold the book to.’

(120)  **German C with Left- and Right-headed T**

a. *No sé wen Juan ha/hat amenazado.*
[NEG] know.1s who-ACC Juan has threatened
‘I don’t know who Juan threatened.’

b. *No sé wen Juan amenazado ha/hat.*
[NEG] know.1s who-ACC Juan has threatened
‘I don’t know who Juan threatened.’

In (119), we can see that only the sentence with left-headed T, example (119a), is grammatical, in spite of the fact that the entire is TP filled with German lexical items. Unfortunately, both (120a) and (120b) are ungrammatical for independent reasons, so we cannot see whether the German C linearizes

---

11 Notice that the subject comes after the verb complex in (119a). This is because we have already determined that embedded wh-questions with a Spanish wh-phrase require VS word order.
the position of T as expected.\textsuperscript{12} Still, the evidence suggests that C determines the position of T, in addition to the availability of null subjects, as we previously saw.

Finally, we turn to the properties of negation. For reasons of space, we will focus on a few illustrative examples of the most important contrasts. Consider the following examples, where \textit{no} is the Spanish negation and \textit{nicht} is the German negation.\textsuperscript{13}

\begin{enumerate}
\item \textbf{Negation with a German C}
\begin{enumerate}
\item *No sé wen Juan \textit{no} amenazó.
\begin{tabular}{l}
\textit{NEG} know.1s \textit{who-ACC} Juan \textit{NEG} threatened. \textit{Spa Neg}
\end{tabular}
\begin{tabular}{l}
'I don’t know who Juan threatened.'
\end{tabular}
\item No sé wen Juan \textit{nicht} amenazó.
\begin{tabular}{l}
\textit{Ger Neg}
\end{tabular}
\end{enumerate}
\end{enumerate}

\begin{enumerate}
\item \textbf{Negation with a Spanish C}
\begin{enumerate}
\item *No sé a quién drohte \textit{nicht} Juan
\begin{tabular}{l}
\textit{NEG} know.1s \textit{who-ACC} threatened \textit{NEG} Juan \textit{Ger Neg}
\end{tabular}
\begin{tabular}{l}
'I don’t know who Juan threatened.'
\end{tabular}
\item No sé a quién drohte \textit{no} Juan
\begin{tabular}{l}
\textit{Spa Neg}
\end{tabular}
\end{enumerate}
\end{enumerate}

We can see in (121) that when C is in German, only German negation is grammatical, as in (121b). On the other hand, when C is in Spanish, as in (122), only Spanish negation is grammatical, as seen in (122b). In both cases, this is in spite of the fact that negation is in the opposite language of T. We can see, then, that the language of C determines at least some of the requirements for negation.

In summary, we have seen that C plays a crucial role in determining a variety of grammatical properties across multiple languages, supporting a parallel claim for another property of C, specifically that C is the determinant head in Spanish, English and Spanish-English wh-questions.

\section{5.1.5 Conclusions}

After carefully reviewing the role of the subject, the wh-phrase, the tense head, and the complementizer head in determining the grammatical subject positions, we have concluded that C is ultimately response

\textsuperscript{12}González-Vilbazo and López (2013) also test a number of other possible CS sentences for the sake of thoroughness, but there are also ungrammatical for the same independent reasons.

\textsuperscript{13}Once again, the position of the subject is based on the language of C.
for the grammatical word orders in a given wh-question. This, in turn, allows us to determine the language of C by inspecting the grammatical subject positions of a given CS sentence, as argued in §5.1.4.2. This is crucial to answering Research Question 2 (RQ2), since C is covert in the structures under study.

Before turning to RQ2, I would like to make two brief comments. First, while we can say that in some sense the language of C determines word order, we are really saying that the language of C indicates the syntactic features of C and the corresponding word order that those features require. Second, notice that the fact that C ultimately determines the position of the subject does not mean that other constituents do not have an effect on the ultimate word order. In particular, since C is covert, the language cannot be directly chosen, unlike all overt elements of the sentence. In this case, particular properties of other constituents, such as the wh-phrase, can affect what language or languages C can be in.

So far we have seen that research on code-switching can provide important insights that go beyond the question of what can be switched where. Still in addressing RQ2 in the next section, we will see that understanding restrictions on specific code-switches is also important and provides significant insights as well.

5.2 Generalizations Based on the Language of C

Knowing that the position of the subject reflects the language of C, we are now in a position to answer RQ2 concerning CS between functional heads and explore what implications these generalizations might have:

**RQ2**: Are there restrictions on code-switching between particular constituents in wh-questions in Spanish-English code-switching?

**RQ2-B**: Are there restrictions on code-switching between the wh-phrase and the complementizer head?

**H2-B**: Yes, there will be restrictions on code-switching between the wh-phrase and the complementizer head.
As we saw in §5.1, the subject, the wh-phrase, and T can each be independently switched and still produce grammatical word orders. For those constituents, then, there does not appear to be any particular restriction on whether any pair of those elements must be in the same language.\footnote{I have not fully explored all possible combinations of pairs of the subject, the wh-phrase and T in different languages to look for restrictions, but there do not appear to be any, given the data reviewed in §5.1.} We will begin with an overview of the overall generalizations concerning the language of C, since this is something that cannot be directly seen in the stimuli, and then we’ll turn to the second research question. This section focuses on generalizations, including those regarding the second research question. Implications of those generalizations will be discussed in §5.4.

5.2.1 The Language of C

There are two key factors which together determine the type of wh-questions examined in this study: (a) the complexity of the wh-phrase (i.e. if it is a simple or complex wh-phrase) and (b) the clause of the wh-question (i.e. if it is in the embedded clause or the matrix clause). Combining these factors produces the following four structures over which we can make generalizations concerning C:

(123) Wh-question Types
    a. matrix questions with a complex wh-phrase
    b. matrix questions with a simple wh-phrase
    c. embedded questions with a complex wh-phrase
    d. embedded questions with a simple wh-phrase

As discussed in §5.1.3, in many cases, the grammatical subject positions in the CS sentences match those of the monolingual sentences based on the language of T. For example, if SV and VS are grammatical in Spanish and T is in Spanish in the CS sentences, then SV and VS are also grammatical in the CS sentences. Since C is not overt in the structures under study here, we will therefore use the language of T as a point of reference. This gives us eight total cases where we can compare the language of C and the language of T, one where T is in Spanish and one where T is in English for each of the four types in (123).
It turns out that, of the eight possible comparisons, C and T are in the same language for all the cases except structure (123d) with a Spanish T. In that specific case, C and T cannot be in the same language, as seen in Table III.

**TABLE III: LANGUAGE OF C: OVERALL GENERALIZATIONS**

<table>
<thead>
<tr>
<th></th>
<th>Spanish T, English wh-phrase</th>
<th>English T, Spanish wh-phrase</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>matrix question</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>complex wh-phrase</td>
<td>C = T</td>
<td>C = T</td>
</tr>
<tr>
<td>simple wh-phrase</td>
<td>C = T</td>
<td>C = T</td>
</tr>
<tr>
<td><strong>embedded question</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>complex wh-phrase</td>
<td>C = T</td>
<td>C = T</td>
</tr>
<tr>
<td>simple wh-phrase</td>
<td>C ≠ T</td>
<td>C = T</td>
</tr>
</tbody>
</table>

Since a detailed review of each of these eight cases for the seven participant groups would be extremely lengthy, I am going to focus on the one case that is exceptional, embedded questions with a simple wh-phrase. Each of the eight cases, including the one to be discussed shortly, follows a similar pattern, where the results are a mix of (a) straightforward results, (b) some less clear but strongly suggestive results, and (c) structures that are fully ungrammatical and therefore not comparable. For detailed tables with the data for evaluating the other cases, see the appendix.

Let’s begin with results of type (a), i.e. the straightforward results. For the structure in question, embedded questions with a simple wh-phrase, Group 4 provides such results, as seen in Table IV. The table is organized as follows. The first column lists each group which has the relevant results and the subject positions tested for the structure in question, here just SV and VS word orders. The remaining columns list the results for each group for CS sentences, monolingual Spanish sentences, and monolingual English sentences. Under each language, there are three columns. First, there is a column for the overall
TABLE IV: C≠T WITH CLEAR RATINGS (EMBEDDED SIMPLE WH-PHRASES)

<table>
<thead>
<tr>
<th>Code-switching</th>
<th>Spanish</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Acceptability</td>
<td>µ</td>
</tr>
<tr>
<td>Group 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>✓</td>
<td>5.25</td>
</tr>
<tr>
<td>VS</td>
<td>*</td>
<td>2.50</td>
</tr>
</tbody>
</table>

acceptability of each word order, where ‘✓’ indicates acceptable, ‘*’ indicates unacceptable, and ‘?’ indicates unclear or middle ratings (those between 3 and 5 on the 7-point Likert scaled used for the present study). Second, there is a column with the mean rating (µ), and finally there is a column with the standard deviation (σ).\(^{15}\)

Notice that only SV word order is acceptable for the CS sentences, just like English but unlike Spanish, where only VS word order is grammatical. As argued in §5.1.4.2 since the grammatical subject positions for the CS sentences match the grammatical subject positions for the monolingual English sentences, we can conclude that C is in English. Since T is in Spanish in these wh-questions, we can further conclude that C and T cannot be in the same language. Note that it is not just the case that C and T do not have to be in the same language. They actually cannot be in the same language for this group. If they could be, then we would expect that, because C would be in Spanish, VS should also be grammatical, as is the case with the monolingual Spanish judgments.

Now let’s turn to the second pattern listed above, (b), where the results are less clear but strongly suggestive. This is the case for Groups 2, 5, and 7, whose results appear below in Table V.

The table is structured in the same way as the previous table. Let’s consider each group in turn. For Group 2, the key difference between monolingual Spanish and monolingual English is that VS is acceptable for Spanish and unacceptable for English, since SV word order is acceptable in both languages. In the case of the CS sentences, while the acceptability of SV word order is unclear, it is clear that VS

\(^{15}\)For some discussion on the standard deviations observed in the results and cases in which they appear to be relatively high, see §4.6.5.
TABLE V: C≠T WITH MIDDLE RATINGS (EMBEDDED SIMPLE WH-PHRASES)

<table>
<thead>
<tr>
<th>Code-switching</th>
<th>Spanish</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Acceptability</td>
<td>μ</td>
</tr>
<tr>
<td>Group 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>?</td>
<td>4.00</td>
</tr>
<tr>
<td>VS</td>
<td>*</td>
<td>2.50</td>
</tr>
<tr>
<td>Group 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>✓</td>
<td>6.75</td>
</tr>
<tr>
<td>VS</td>
<td>?</td>
<td>3.75</td>
</tr>
<tr>
<td>Group 7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>?</td>
<td>4.25</td>
</tr>
<tr>
<td>VS</td>
<td>*</td>
<td>1.88</td>
</tr>
</tbody>
</table>

word order is unacceptable. Importantly, this is also the case for English, suggesting that C is in English in this structure for this group.

Group 5, like the previous group, accepts SV word order in both monolingual Spanish and monolingual English but only accepts VS in monolingual Spanish. For CS, the Group 5 clearly accepts SV word order, but the rating for VS word order is less clear. Still, notice that there is a 3.00 difference between the 6.75 ratings for SV word order and the 3.75 rating for VS word order. I would argue that this contrast is much more like the monolingual English judgments, where SV word order is rated much higher than VS word order, compared to the monolingual Spanish judgments, where SV word order and VS word order have very similar ratings. This also suggests that C is in English in this structure for this group.

Finally, for Group 7, like Group 2, the ratings for SV word order are unclear for CS sentences as well as monolingual Spanish sentences. On the other hand, VS word order is clearly unacceptable for both CS and monolingual English. Once again, the ratings for VS word order pattern with English and not with Spanish, strongly suggesting that C is in English in spite of the fact that T is in Spanish.
The third type of result for the structure we have been focusing on, embedded questions with simple wh-phrases, involves structures that are fully ungrammatical and therefore not comparable. In this case, both SV word order and VS word order are ungrammatical for a number of groups. For example, Group 1 judged both SV word order and VS word order to be grammatical in monolingual Spanish sentences and SV word order to be grammatical in monolingual English sentences. In site of having at least one grammatical subject positions in each language, this group judged both SV word order and VS word order to be ungrammatical for CS sentences. In cases such as this, it is not possible to determine what the language of C is. In all likelihood, both word orders are ungrammatical for independent reasons.

Now we turn to a small number of exceptions to the generalizations outlined in Table III. The first one is an exception to the generalization discussed so far, that C and T cannot be in the same language with simple wh-phrases in an embedded question. More concretely, there is one group, Group 3, where not only can C and T be in different languages, as is the case with all of the other groups, but C and T can also be in the same language for this group, as shown in Table VI.

<table>
<thead>
<tr>
<th></th>
<th>Code-switching</th>
<th>Spanish</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Acceptability</td>
<td>Acceptability</td>
<td>Acceptability</td>
</tr>
<tr>
<td>Group 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>✓</td>
<td>5.50</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>✓</td>
<td>7.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>VS</td>
<td>✓</td>
<td>7.00</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.00</td>
<td>7.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For Group 3 monolingual judgments, SV word order is only grammatical for English, and VS word order is only grammatical for Spanish. On the other hand, both word orders are grammatical for CS sentences. This suggests that C can be in English, making SV word order grammatical, or C can be in Spanish,
making VS word order grammatical. Again, notice that C and T can also be in different languages, as with the remaining groups.

The remaining exceptions involve CS sentences where T is in English, unlike the data covered up to this point in this section, and C could be or is in Spanish. Importantly, the results also suggest that C can be in English in the same structures, fitting in with the overall generalization that wh-questions with an English T allow C to also be in English.

Consider first the acceptability judgments of Group 3 for matrix questions with simple wh-phrases and embedded questions with complex wh-phrases, as seen in Table VII. Once again, these are wh-questions where T is in English and the wh-phrase is in Spanish.

<table>
<thead>
<tr>
<th>Code-switching</th>
<th>Spanish</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceptability</td>
<td>μ</td>
<td>σ</td>
</tr>
<tr>
<td>matrix</td>
<td></td>
<td></td>
</tr>
<tr>
<td>simple</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>*</td>
<td>2.67</td>
</tr>
<tr>
<td>V2</td>
<td>✓</td>
<td>7.00</td>
</tr>
<tr>
<td>VS</td>
<td>?</td>
<td>4.00</td>
</tr>
<tr>
<td>embedded</td>
<td></td>
<td></td>
</tr>
<tr>
<td>complex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>✓</td>
<td>7.00</td>
</tr>
<tr>
<td>VS</td>
<td>✓</td>
<td>5.50</td>
</tr>
</tbody>
</table>

We will begin with the first type of wh-question, matrix questions with simple wh-phrases. Since SV word order is ungrammatical and V2 word order is grammatical whether it is a CS sentence, a
monolingual Spanish sentence, or a monolingual English sentence, these word orders will not help us to determine the language of C. In the case of VS word order, the average acceptability judgment for the CS sentences is exactly 4.00, so it is not clear whether the acceptability rating is more like Spanish, where VS word order is acceptable, or English, where VS word order is unacceptable. Still, the fairly stark contrast between the 7.00 rating for V2 word order and the 4.00 rating for VS word order in CS sentences suggests that V2 is much more acceptable compared to VS, as in English. If that is correct, then C is in English, like with the other groups.

Now we turn to the second exceptional type of wh-question for Group 3 in Table VII: embedded questions with a complex wh-phrase. This group finds both SV word order and VS word order to be acceptable in the CS sentences. This exactly matches the grammatical subject positions of Spanish, where both SV and VS word orders are grammatical, but it is also possible that C is allowed to be in English as well, since both Spanish and English allow SV word order. If that is the case, then once again C can be in English, like with the other groups.

Continuing with possible exceptions to generalizations in Table III, Group 5 has some types of wh-questions where C may be in Spanish with an English T, as seen in Table VIII. Importantly, it is also possible that C is in English for these types of wh-questions, following the general pattern across groups for these structures.

We begin with the first structure listed in the table, matrix questions with complex wh-phrases. In this case, all three word orders are grammatical in the CS sentences. This suggests that C is in Spanish, where all three word orders are also grammatical or at least marginally acceptable, as seen in the judgments for monolingual Spanish. Still, it is also possible that an English C is grammatical in these wh-questions, since the grammatical subject positions in monolingual English are a subset of those of monolingual Spanish.

The second structure in Table VIII: matrix questions with simple wh-phrases, is unclear. Once again, since SV word order is ungrammatical and V2 word order is grammatical whether it is a CS sentence, a monolingual Spanish sentence, or a monolingual English sentence, these word orders will not help us to determine the language of C. Turning to VS word order, the acceptability judgments for the
## TABLE VIII: EXCEPTIONS TO OVERALL GENERALIZATIONS (GROUP 5)

<table>
<thead>
<tr>
<th>Code-switching</th>
<th>Spanish</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceptability</td>
<td>$\mu$</td>
<td>$\sigma$</td>
</tr>
<tr>
<td><strong>matrix</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>complex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>✓</td>
<td>7.00</td>
</tr>
<tr>
<td>V2</td>
<td>✓</td>
<td>6.75</td>
</tr>
<tr>
<td>VS</td>
<td>✓</td>
<td>7.00</td>
</tr>
<tr>
<td><strong>matrix</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>simple</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>*</td>
<td>1.67</td>
</tr>
<tr>
<td>V2</td>
<td>✓</td>
<td>5.50</td>
</tr>
<tr>
<td>VS</td>
<td>?</td>
<td>3.50</td>
</tr>
<tr>
<td><strong>embedded</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>complex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>✓</td>
<td>7.00</td>
</tr>
<tr>
<td>VS</td>
<td>?</td>
<td>4.00</td>
</tr>
</tbody>
</table>

CS sentences are a middle rating, so C may be in English or Spanish. Nevertheless, these results do not contradict the generalization that C and T can be in the same language in this structure.

Finally, the third structure, embedded questions with complex wh-phrases, follows a pattern very similar to the previous structure. One word order, SV word order, is the grammatical for CS, monolingual Spanish and monolingual English sentences, making it impossible to tell what language C is in. As before, VS word order has a middle rating, so C may be in English or in Spanish. However, in contrast to the previous structure, there is a fairly stark difference between the acceptability judgment for SV word order, 7.00, and VS word order, 4.00. This suggests that SV is much more acceptable compared to VS, as is the case in English.
Ultimately, for each exception, it is also possible that C is in English for each structure, matching the overall generalization that C and T are in the same language with an English T, or at least they can be. We have seen, then, that C and T are generally required to be in the same language for most structures but that C and T cannot be in the same language in embedded questions with a simple wh-phrase and a Spanish T.

5.2.2 Code-Switching Between C and T

Now that the generalizations concerning the language of C are in place, we can turn to Research Question 2 (RQ2), starting with the first research sub-question (RQ2-A) and the hypothesis for that sub-question (H2-A), repeated below.

**RQ2:** Are there restrictions on code-switching between particular constituents in wh-questions in Spanish-English code-switching?

**RQ2-A:** Are there restrictions on code-switching between the tense head and the complementizer head?

**H2-A:** Yes, there will be restrictions on code-switching between the tense head and the complementizer head.

In the previous section, we saw that C and T are generally required to be in the same language, with the exception of embedded questions with simple wh-phrases. In and of itself, it is not surprising that it is perfectly grammatical for C and T to be in the same language, as C and T are in the same language in monolingual sentences, of course. More significantly, it is not just that C and T can be in the same language but that, with few exceptions, they must be. This suggests that there are restrictions on code-switching between C and T that require C and T to remain in the same language. This provides support for H2-A.

Interestingly, there is also a structure in which C and T cannot be in the same language, in spite of the fact that this same combination is grammatical in monolingual sentences. Specifically, when the wh-question is in an embedded clause with a simple wh-phrase and a Spanish T, the sentence becomes ungrammatical with a Spanish C. This can be seen in example (124), focusing specifically on
the grammaticality of VS word order. We can see in example (124a) that VS word order is grammatical in monolingual Spanish, while the same word order is ungrammatical in monolingual English, as seen in (124b). Finally, we have the parallel CS structure in (124c). If C can be in Spanish in this structure, then VS word order should be grammatical, as in monolingual Spanish, but, in fact, VS word order is ungrammatical, exactly like with monolingual English.

(124) **Spanish C ungrammatical with an Spanish T (comparing VS word order)**

a. No recuerdo qué han comprado tus colegas esta semana. **Sp**
   ‘I don’t remember what your colleagues have bought this week.’

b. *I don’t remember what have bought your colleagues this week. **Eng**

c. *I don’t remember what han comprado tus colegas this week. **CS**
   ‘I don’t remember what your colleagues have bought this week.’

Since there should not be any syntactic incompatibility between a Spanish C and a Spanish T, as argued above, there must be another code-switch that is actually the source of the ungrammaticality. Assuming, then, that C is still in Spanish, notice that in (124c) the wh-phrase is in English while C, T, and the subject are all Spanish. This suggests that the ungrammaticality instead stems from properties of the English wh-phrase. With that in mind, let us turn to the second research sub-question in the next section.

5.2.3 Code-Switching Between C and the Wh-Phrase

Now we are ready to address the second research sub-question (RQ2-B) of Research Question 2 (RQ2) and the hypothesis for that sub-question (H2-B), repeated below.

RQ2: Are there restrictions on code-switching between particular constituents in wh-questions in Spanish-English code-switching?

RQ2-B: Are there restrictions on code-switching between the wh-phrase and the complementizer head?

H2-B: Yes, there will be restrictions on code-switching between the wh-phrase and the complementizer head.
Again, we begin with the board generalization that C and T are generally required to be in the same language, with the exception of embedded questions with simple wh-phrases. Because the wh-phrase and T are always in opposite languages in the stimuli for the present study and C and T are in the same language, this means that there is a code-switch between C and the wh-phrase in these structures. At first glance, then, it appears that H2-B is not supported, as there do not appear to be restrictions on code-switching between the wh-phrase and C.

However, as discussed in the previous section, C and T are not always in the same language. Specifically, when the wh-phrase is a simple wh-phrase in an embedded question and T is in Spanish, it is grammatical for C to be in English, as seen in example (124) at the end of the previous section. Importantly, it is not just that C can be in English. If C is in Spanish, the sentence is ungrammatical, so C must be in English, the same language as the wh-phrase. Since the wh-phrase is also in English in these structures and C cannot be in Spanish, this means that a code-switch between C and the wh-phrase is not grammatical. This suggests that there is some property or set of properties of English simple wh-phrases in embedded clauses that are not compatible with a Spanish C. Interestingly, the same is not true in matrix questions with a simple wh-phrase nor in embedded questions with complex wh-phrases, as seen in (125) and (126).

(125) Spanish C grammatical in matrix questions with a simple wh-phrase
   a. What han *comprado esos niños* at the mall so far?  
      have bought those children
      ‘What have those children bought at the mall so far?’  
      CS
   b. ¿Qué han comprado esos niños en el centro comercial hasta ahora?  
      what have bought those children in the central commercial until now
      ‘What have those children bought at the mall so far?’  
      Sp
   c. *What have bought those kids at the mall so far?  
      Eng

(126) Spanish C grammatical in embedded questions with a complex wh-phrase
   a. No *recuerdo* how many apples han *comido* your children desde el domingo.  
      not I-remember have eaten since the Sunday
      ‘I don’t remember how many apples your children have eaten since Sunday.’  
      CS
First, consider the sentences in (125), which only differ in structure from those in (124) in that the question is a matrix question. The CS sentence is (125a) is grammatical, exactly like the monolingual Spanish sentence in (125b) and unlike the parallel monolingual English sentence in (125c). Now consider the sentences in (126), which only differ in structure from those in (124) in that the wh-phrase is a complex wh-phrase. The CS sentence in (126a) is grammatical, exactly like the monolingual Spanish sentence in (126b) but unlike the parallel monolingual English sentence in (126c). Again, the contrast between the structures in (124) and these structures indicates that the incompatibility has something to do specifically with the embedded Spanish C and the simple English wh-phrase.

I do not have a definitive analysis, but I would like to sketch out two possibilities, based on two different approaches to the relationship between C and the wh-phrase.

First, consider a proposal in González-Vilbazo (2005) concerning relative pronouns. González-Vilbazo (2005) begins by arguing that relative pronouns have phrasal properties. For example, they stand for verbal arguments that must be a whole phrase and not just a head. On the other hand, relative pronouns trigger subordinate-clause word order in German, where there is no T-to-C movement. It is traditionally assumed that T does not raise to C when C is occupied, suggesting that the relative pronoun is in C. In addition, Spanish-German CS data from González-Vilbazo (2005) also suggest that the relative pronoun is in C. González-Vilbazo (2005) reconciles this contradiction by adopting a proposal in Struckmeier (2005) in which the case, gender, and number properties of the relative pronoun are in C, and the wh or relativizing part is in SpecCP.

Perhaps it is the case that, like relative pronouns in the above proposal, the properties of wh-phrases are sometimes split between SpecCP and C. If this is the case in some Spanish wh-questions, then a code-switch between the simple wh-phrase and the Spanish C may be ungrammatical because of the impossibility of word-internal code-switching, as proposed by Poplack (1980) and MacSwan (1999), among others. While there is some counterevidence suggesting that such code-switches are in fact possible.
González-Vilbazo and López (2011), they appear to be grammatical in only a very circumscribed set of configurations involving a last-resort insertion of a light verb, along the lines of do-support.

A second possibility comes from Van Craenenbroeck (2010)’s proposal discussed in §3.1.4.3 which focuses on a syntactic dependency between the wh-phrase and C, in contrast to the previous morphologically-based account. I will first review the key characteristics of the proposal and what the proposal has to say about the Spanish C and Spanish wh-phrases before turning to the incompatibly between an embedded Spanish C and a simple English wh-phrase.

Van Craenenbroeck (2010) is interested in accounting for a number of syntactic contrasts between a simple wh-phrase and a complex wh-phrase. He proposes that simple wh-phrases have both a Q(uestion)-feature [Q] and an Op(erator)-feature [Op]. The C head is split across two positions, each with one of these two features, so that the simple wh-phrase has to move through (i.e. Agree and Merge with) both the $C_{[Q]}$ and the $C_{[Op]}$. On the other hand, the complex wh-phrase only has a Q-feature and hence moves directly to $C_{[Q]}$.

Interestingly, although Spanish does not have the same syntactic phenomena that Van Craenenbroeck (2010)’s proposal is meant to address, CS data from this dissertation suggest that Spanish wh-phrases have the necessary featural requirements of English C. Consider the following sentences.

(127) Spanish wh-phrases with English C

a. I don’t remember cuántos artículos those researchers have written este semestre.
   ‘I don’t remember how many articles those researchers have written this semester.’

b. No recuerdo qué those bankers have stolen desde la última crisis económica.
   ‘I don’t remember what those bankers have stolen since the last economic crisis.’

In (127a), we can see that a complex Spanish wh-phrase is grammatical with an English C. The judgments for (127b) are less clear because some participant groups find most CS wh-questions with a simple wh-phrase to be ungrammatical, whether the wh-phrase is in Spanish or in English. Setting aside those groups, this structure with a simple Spanish wh-phrase is also grammatical. This suggests that

---

16 The proposal also considers the possibility of the complex wh-phrase directly merging in $C_{[Q]}$.
17 I only include part of the necessary sentences to determine the language of C, for ease of exposition. A complete comparison between the grammatical subject positions of the CS and monolingual sentences makes it clear that C is in English for these structures.
both simple and complex Spanish wh-phrases are compatible with English C. If that is correct, then Spanish wh-phrases must have both the [Q] and [Op] features that English C requires. This in turn suggests that Spanish C has these same features, though perhaps not split across two complementizer heads. If it did not, it would be strange for Spanish wh-phrases to have unnecessary features.

If Spanish C has [Q] and [Op] and so do simple English wh-phrases, then why are these wh-phrases ungrammatical with an embedded Spanish C? Perhaps Spanish C has an additional featural requirement related to the argument status of the wh-phrase. Recall that Spanish wh-questions has an argument-adjunct asymmetry, where questions like (128) require inversion but questions like (129) allow for SV and VS word order.

(128) a. ¿Qué ha comprado Juan?  
what has bought Juan  
‘What has Juan bought?’

b. *¿Qué Juan ha comprado?  
what Juan has bought  
‘What has Juan bought?’

(129) a. ¿Por qué trabaja tanto Juan?  
for what works so much Juan  
‘Why does Juan work so much?’

b. ¿Por qué Juan trabaja tanto?  
for what Juan works so much  
‘Why does Juan work so much?’

If this contrast stems from an additional syntactic dependency that forms between argument wh-phrases and Spanish complementizers, that would explain sentences like (124c) where a simple English wh-phrase and a Spanish C are ungrammatical. Essentially, the English wh-phrase does not have the additional feature required by the Spanish complementizer. While the English wh-phrase is an argument wh-phrase, like its Spanish counterpart, it would not have the additional feature required for the argument-adjunct asymmetry.

Notice, though, that both accounts are only able to partially account for the code-switching restriction we have been discussing. Specifically, these accounts do not address why the CS restriction is specific to embedded wh-questions. Still, it plausible that the difference comes down to having a distinct
complementizer for matrix and embedded clauses. This would not be unprecedented, given a variety of ways in which embedded clauses differ from matrix clauses, including the matrix-embedded asymmetry of English wh-questions and the sentence-final tense head in German embedded clauses.

In summary, we have seen evidence for a restriction on code-switching between a simple English wh-phrase and an embedded Spanish C, and we have discussed a morphological approach and a syntactic approach for accounting for this restriction.

5.3 Empirical Generalizations

Before turning to the implications for the results outlined above, I would like to summarize the empirical generalizations that we have uncovered. The focus here is on the patterns that emerge from the data, rather than on the implications these patterns have for code-switching or language more generally. The implications, both for CS and more generally, will be discussed in the next section.

The first generalization concerns the role of the complementizer head in the phenomenon of inversion. Specifically, we saw evidence that C is ultimately responsible for determining the position of the subject in wh-questions. Future work will have to determine the degree to which this is a broadly cross-linguistic generalization, i.e. whether this goes beyond Spanish and English. Tentative results on inversion in Taiwanese-Spanish (González-Vilbazo et al., 2012) and Japanese-English code-switching (Ebert et al., 2012) suggest that this is the case.

The remaining generalizations revolve around restrictions on code-switching between C and T and between C and the wh-phrase. Let us begin with code-switches between C and T, as outlined in Table IX. In this case, only the language pairings listed are grammatical. That is to say, if it is licit for C and T to be in the same language, as is the case for most structures, then a code-switch between C and T is ungrammatical. Likewise, when it is licit to code-switch between C and T, as is the case for one structure, then it is ungrammatical for C and T to be in the same language. The generalization, then, is that C and T must be in the same language unless the wh-question is an embedded question with a simple English wh-phrase and a Spanish T.

18 Recall that there were some groups for which C could be in both languages with an English T. I will leave for future research the extent to which these groups are truly exceptional.
Now consider Table IX. Unlike with the previous table, wh-questions are generally grammatical whether C and the wh-phrase are in the same language or in the opposite language. Notice, however, one important exception. When wh-question is an embedded question with a simple English wh-phrase and a Spanish T, C and the wh-phrase must be in the same language. In this case, then, the generalization is that code-switches between C and the wh-phrase are perfectly grammatical unless the wh-question is an embedded question with a simple English wh-phrase and a Spanish T.

In summary, we have three empirical generalizations:

1. C determiners the grammatical subject positions in wh-questions.
2. C and T must be in the same language unless the wh-question is an embedded question with a simple English wh-phrase and a Spanish T.
3. Code-switches between C and the wh-phrase are perfectly grammatical unless the wh-question is an embedded question with a simple English wh-phrase and a Spanish T.

\[19\]

While the data for this dissertation cannot directly address this question, data from the pilot experiments suggests that wh-questions in which C and the wh-phrase are in the same language are perfectly grammatical, with the right word order. In addition, such a combination is grammatical in monolingual sentences, of course.
5.4 Implications

So far we have focused on the factors driving inversion in wh-questions and any restrictions on code-switching the complementizer head. We have seen evidence the complementizer head plays a central role in the formation of wh-questions. We have also seen evidence for a restriction on code-switching between a simple English wh-phrase and a Spanish complementizer in embedded questions. In this section, we consider some implications for these results.

Importantly, the results of this study have implications that go beyond the central role of the complementizer head, including various aspects of the morphosyntax of wh-questions and theoretical aspects of code-switching. This is not surprising, given that the properties of wh-questions are the result of the intersection of a variety of different properties, in both monolingual and CS contexts.

5.4.1 Wh-Questions in Spanish and English

5.4.1.1 Wh-Questions in English

We begin with a brief discussion of any possible implications for wh-questions in English. At first glance, it might appear that the central result of this dissertation, that C is a determinant head, does not provide any new insight into the morphosyntax of wh-questions in English. It has already been

<table>
<thead>
<tr>
<th></th>
<th>Spanish T, English wh-phrase</th>
<th>English T, Spanish wh-phrase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complex wh-phrase</td>
<td>$C = \text{Wh}, C \neq \text{Wh}$</td>
<td>$C = \text{Wh}, C \neq \text{Wh}$</td>
</tr>
<tr>
<td>Simple wh-phrase</td>
<td>$C = \text{Wh}, C \neq \text{Wh}$</td>
<td>$C = \text{Wh}, C \neq \text{Wh}$</td>
</tr>
<tr>
<td>Embedded question</td>
<td>$C = \text{Wh}, C \neq \text{Wh}$</td>
<td>$C = \text{Wh}, C \neq \text{Wh}$</td>
</tr>
<tr>
<td>Simple wh-phrase</td>
<td>$C = \text{Wh}$</td>
<td>$C = \text{Wh}, C \neq \text{Wh}$</td>
</tr>
</tbody>
</table>
widely agreed that inversion in English stems from properties of the complementizer, so the suggestion that C determines the grammatical subject positions is not a new finding.

Now consider the following hypothesis (González-Vilbazo and López 2012) in the context of this finding:

**Phase Head Hypothesis (PHH)**

The phase head determines grammatical properties of its complement.

Based on the grammatical property that was the focus of this dissertation, subject inversion, we found support for this hypothesis. This is in addition to evidence in support of the PHH based on other grammatical properties, as discussed in §5.1.4.3. If some version of the PHH is correct, this has implications that go beyond the phenomena discussed up to this point. Take, for example, the phenomenon of superiority effects in multiple wh-questions, illustrated in (130).

(130) a. Who fixed what?
   b. *What did who fix?
   c. Which of these guys fixed which of these cars?
   d. Which of these cars did which of these guys fix?

In English questions with multiple wh-phrases, only one wh-phrase can front. The remaining wh-phrases stay in situ. It turns out that, in most cases, only the higher wh-phrase can front, so a question where the subject wh-phrase has fronted, as in (130a), is perfectly grammatical, while a question where the object wh-phrase has fronted is ungrammatical, as seen in (130b). Notice that all of the wh-phrases involved are simple wh-phrases. If instead both wh-phrases are complex, D-linked wh-phrases, both the subject wh-phrase, as in (130c), and the object wh-phrase, as in (130d), can front. This can be contrasted with Spanish where sentences parallel to (130b) do not show superiority effects:

(131) a. ¿Quién compró qué?
    who bought what
    ‘Who bought what’

   b. ¿Qué compró quién?
    what bought who
    ‘What did who buy?’
This type of multiple wh-question is grammatical whether it is the wh-subject that fronts, as in (131a), or the wh-object, as in (131b) 20 Given the contrast in superiority effects between Spanish and English, CS may be able to shed some light on the source of these differences where the PHH predicts that C should play a central role.

5.4.1.2 T-Based Accounts of Spanish Inversion

The results of this dissertation have significant implications for a class of accounts for inversion in Spanish that I will refer to as T-based accounts. Under these accounts, it is argued that the wh-phrase establishes a dependency with T (Toribio 1993; Barbosa 2001; Goodall 2001; Gutiérrez-Bravo 2005; Zubizarreta 2012, inter alia) 21 The implementation varies, but they have in common the proposal that the wh-phrase moves to, or through, SpecTP (or an equivalent). Since SpecTP has already been filled, the subject remains in situ, after the verb.

At first glance, it might appear that the conclusion that C is the determinant head in Spanish wh-questions, as argued in §5.1 necessarily contradicts any account deriving Spanish inversion from properties of T. This is not quite correct. In fact, it is likely that T does play an important role in inversion in Spanish, even if C is ultimately responsible for determining the grammatical subject positions in Spanish. This is because preverbal subjects in Spanish can occupy SpecTP (cf. §3.1.2.2 for discussion), and any phrase that has moved to the specifier of a head must have some syntactic relationship with that head. For that reason, it is important to consider the T-based approach to inversion in Spanish more closely.

I will do this by reviewing one concrete example of a T-based account, by Goodall (2001), in detail. Though the focus will be on this specific account, the reasoning applies to all of the T-based accounts I have referenced. Goodall (2001) focuses on understanding why Spanish subjects generally

20 Note that this does not mean that Spanish is entirely absent of superiority effects, as observed by López 2012. This can be seen in the contrast between the following two examples:

(i) ¿Quién dijo que compró Juan qué?
   who said that bought Juan what
   ‘Who said that John bought what?’

(ii) *¿Qué dijo quién que compró Juan?

21 Zubizarreta 2012 is not strictly a T-based account but is equivalent with respect to the predictions made based on the language of T.
but, importantly, not always move to a preverbal position (SpecTP in his analysis). He ultimately argues that Spanish, like English and many other languages, has an EPP requirement (i.e. the requirement that a DP or other constituent must move to SpecTP). This predicts that subjects, or other constituents that can satisfy the EPP, must always be preverbal, but subjects can be postverbal in a variety of wh-question structures in Spanish. To account for this and other properties of wh-questions, Goodall (2001) makes use of a Q-feature, essentially a question feature that requires agreement between the head containing the Q-feature and a wh-phrase.

Goodall (2001) accounts for subject inversion by proposing that both C and T have a Q-feature. Unlike his proposed analysis for English, both the Q-feature on C and the Q-feature on T require overt movement. Let’s walk through an example. In (132a), T and its associated Q-feature have merged with vP, and the verb has raised to T. The wh-phrase then raises from its base position to SpecTP in order to check the Q-feature on T, as shown in (132b). Finally, C merges with TP, along with its associated Q-feature, and the wh-phrase raises to SpecCP to check this second Q-feature.

(132)  

a. \[TP \text{ comprar} [u_P \text{ Juan qué}]\] bought Juan what

b. \[TP \text{ qué} \text{ comprar} [u_P \text{ Juan ti}]\] what bought Juan

c. \[CP \text{ qué} \text{ C} [TP \text{ ti comprar} [u_P \text{ Juan ti}]]\] what bought Juan

‘What did John buy?’

To account for the position of the postverbal subject, Goodall (2001) suggests that “[o]nly one feature on a head may have an EPP-feature (force overt movement)” under the assumption that movement to a head involves movement to a specifier of that head and that each head may only have one specifier. Since the Q-feature on T already has the EPP feature, there is no other EPP feature to attract the subject to SpecTP, leaving it in its base-generated position in SpecvP, as seen in the structures in (132).

---

22 The structure and the process of agreement have been simplified in this example and the corresponding example for English in order to keep the focus on the Q-feature and the movement of the wh-phrase.

23 In this account, the EPP is essentially a property of another feature, rather than a feature on its own. When there is no Q-feature, the EPP is generally associated with the phi-features.
Goodall (2001) also accounts for T-to-C movement in English by assuming that both C and T have a Q-feature. English differs from Spanish in that the Q-feature on T does not require overt movement to its specifier. Consider example (133).

(133)  a. \[
\text{TP} \, \text{has}_{\text{uQ, EPP}} \, [\text{vP} \, \text{John bought what today?}]
\]

b. \[
\text{TP} \, \text{John has}_{\text{uQ, EPP}} \, [\text{vP} \, \text{t_i bought what today?}]
\]

c. \[
\text{CP} \, \text{What} \, \text{C has}_{\text{uQ}} \, [\text{TP} \, \text{John has}_{\text{uQ, EPP}} \, [\text{vP} \, \text{t_i bought t_j}]]
\]

d. \[
\text{CP} \, \text{What} \, \text{C has}_{\text{uQ}} \, [\text{has}_{\text{uQ, EPP}} \, [\text{TP} \, \text{John has}_{\text{uQ, EPP}} \, [\text{vP} \, \text{t_i bought t_j}]]]
\]

As with Spanish, T and its associated Q-feature merge with vP, as seen in (133a). In this case there is also a separate EPP feature, since the Q-feature in English does not require the wh-phrase to move to its specifier. Instead, the subject moves to SpecTP, as in (133b). Next, C and its associated Q-feature merge with TP, and the wh-phrase raises to SpecTP to check C’s Q-feature, as seen in example (133c). Finally, T undergoes head movement to C, thus checking its Q-feature under spec-head agreement with the wh-phrase in SpecCP.

Notice that, in this proposal, the properties of T are central to accounting for the differences in the formation of wh-questions in Spanish and English. In particular, T-to-C movement in English occurs in order to check the Q-feature on T. Since the Q-feature does not have an EPP feature, the wh-phrase has no reason to move to SpecTP. Instead, the wh-phrase moves to SpecCP and then T moves to C so that it is in a spec-head relationship with the wh-phrase. In Spanish, on the other hand, the Q-feature on T does have an EPP feature, so the wh-phrase must move to SpecTP before moving to SpecCP, forcing the subject to remain in situ.

Now we are ready to see whether or not the T-based account is supported by the code-switching data from this study. Again, we are interested in code-switched wh-questions where T is in Spanish. It turns out that preverbal subjects are in fact grammatical in some structures. As seen in (134a), a

\[24\] Again, I am leaving out additional operations such as movement to SpecvP as required in some accounts of wh-movement.
preverbal subject is grammatical with a Spanish T (labelled with ‘T_{Sp}’), whereas a postverbal subject, as in \(134b\), is in fact ungrammatical in an otherwise identical structure.

\(134\)  
\[\text{a. } \text{No recuerdo que esos conejos han comido}\]  
\['I don’t remember what those rabbits have eaten since I planted the garden.’\]

\[\text{b. } * \text{No recuerdo que han comido esos conejos}\]  
\['I don’t remember what those rabbits have eaten since I planted the garden.’\]

Parallel data from [Ebert, 2013] on Spanish-German CS follow the same pattern. In \(135a\), a preverbal subject is grammatical, while in \(135b\) a postverbal subject is ungrammatical.

\(135\)  
\[\text{a. } \text{No sé a quién Juan amenazó}\]  
\['I don’t know whom Juan threatened’\]

\[\text{b. } * \text{No sé a quién Juan amenazó}\]  
\['I don’t know whom Juan threatened’\]

The CS data reviewed above show that the subject can in fact raise to SpecTP\(^{26}\) suggesting that there is no Q-feature on Spanish T. Perhaps the [Goodall, 2001] account can be rescued by assuming that there can be more than one EPP feature on a head. In this case, the subject will be able to raise to SpecTP due to the additional EPP feature on T.

While this modification allows the CS data to be accounted for, it ultimately undermines the proposal. With the proposed modification to the theory, the requirement that the wh-phrase move through SpecTP is divorced from whether the subject is required to move or remain in situ, since that simply depends on whether or not there is an extra EPP feature. Wh-questions in which the subject remains in situ no longer have anything to do with the proposed Q-feature on T, removing the need to propose such a feature in the first place. The CS data, then, appear to be counterevidence against [Goodall, 2001]’s T-based proposal to account for inversion in Spanish.

---

\(^{25}\)As discussed in 5.1.4.1 while careful review of the acceptability judgments across groups strongly suggests that C is in English for the wh-questions cited here, embedded questions with simple wh-phrases, the results are not fully uniform. For that reason, the cited acceptability ratings do no directly reflect those of all the participants groups.

\(^{26}\)While there is evidence that preverbal subjects may be in topic position in some wh-questions, rather than SpecTP (for a recent proposal, cf. [Zabizarreta, 2012], this is not the case in the data presented here, as the questions are presented in a neutral context.
Notice also that, if C is in English, as the SV word order suggests, then under Goodall's account, this should not directly change the type of inversion that is grammatical. In Goodall's approach, the difference between Spanish, where VS word order is required, and English, where T-to-C movement is required, comes down to whether or not the Q-feature on T has an EPP feature. This should depend on the language of T, predicting SV word order to be ungrammatical with an Spanish T.

There is another way to potentially rescue the Goodall analysis. Chomsky proposes what is known as feature inheritance, where tense and agreement features of T are inherited from C. Under such an approach, T would inherit any relevant features related to wh-questions, including those involved with inversion. If C is in English, then T should inherit the features of C that allow SV word order to be grammatical, exactly as seen in (134) and (135). Unfortunately, adopting feature inheritance as proposed in Chomsky is problematic for independent reasons. If C is in English, then the Spanish T in these examples would have to inherit English tense and aspect features, but these would not be compatible with a Spanish T because of important differences in the aspectual systems of the two languages.

5.4.1.3 Other Accounts of Spanish Wh-questions

I would like to consider some implications of my results for two additional accounts of wh-questions in Spanish, one based on a phonological requirement for inversion and another one based on projection of the CP layer.

Buesa-García observes that the ungrammaticality of preverbal subjects in many Spanish wh-questions could be looked at as a ban on subjects appearing between the complementizer and the verb. To explain this ban, he proposes that the interrogative Spanish complementizer is a null affix that merges with the verb under PF adjacency, along the lines of Bošković's proposal for the Bulgarian interrogative C. In the context of code-switching, this raises the question of what happens when C and T, essentially the verb, are in different languages, since the process that affixes C to the verb is phonological.

---

27 See §5.1.3 for some additional discussion.
28 Depending on the approach to morphology, this process may also be partially or entirely morphological in nature. In either case, there are implications that go beyond syntax.
Recall that there was some data from Group 3 and Group 5 suggesting that C could be in Spanish with an English T, suggesting that a Spanish C can affix to an English T. If Buesa-García (2008)'s account is going to account for Spanish inversion for these groups, this data raises interesting questions about the process by which C affixes to the verb. For example, since it is a PF process, the process would have to apply cross-linguistically, which has important implications for the architecture of phonology. If instead these data are taken to mean that Buesa-García (2008)'s account does not apply to the variety of Spanish spoken by these groups, then a second account of inversion in Spanish will be required. This is problematic in that inversion for these groups is otherwise similar to the other participant groups and clearly contrasts with inversion in the monolingual English sentences.

The results of this dissertation also have implications for accounts of wh-questions in Spanish in which there is no CP for some wh-questions (Grimshaw, 1997; Zubizarreta, 1998; Gutiérrez-Bravo, 2005). Under these accounts, CP is only projected if there is additional left-peripheral material beyond the wh-phrase, such as a topic. This would predict that CS sentences such as (136a) should be grammatical, since the wh-clause has no left-peripheral material.

(136) **Evidence for a projected CP with Spanish T (comparing VS word order)**

<table>
<thead>
<tr>
<th></th>
<th>CS</th>
<th>Sp</th>
<th>Eng</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>*I don’t remember what han comprado tus colegas this week.</td>
<td>‘I don’t remember what your colleagues have bought this week.’</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>b. No recuerdo qué han comprado tus colegas esta semana.</td>
<td>‘I don’t remember what your colleagues have bought this week.’</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>c. *I don’t remember what have bought your colleagues this week.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

More specifically, if there is no CP and therefore no C, the only functional head left for determining the grammatical subject positions for this sentence is T. We can see in monolingual Spanish sentence in (136b) that VS word order is grammatical with this type of wh-question, so the CS sentence with the Spanish T should also be grammatical. Instead, it is ungrammatical, just as in the monolingual English

---

29 This data is discussed in the last two pages of §5.2.1.

30 Even if embedded sentences are treated differently and instead require a CP, the data from my dissertation suggest that C determines the grammatical subject positions for all wh-questions under study, including matrix wh-questions.
wh-question in (136c). The most straightforward explanation is that the embedded question does have a CP, with an English C, contra the aforementioned accounts.

5.4.2 Wh-Questions More Generally

Though the precise syntactic properties remain to be uncovered, results from the present study provide additional evidence for the distinct syntactic properties of simple wh-phrases and complex wh-phrases. Previous work along these lines includes Reinhart (1986, 1987, 1990), Hornstein and Weinberg (1987), Guéron and May (1987), Aoun et al. (1987), Aoun and Li (2004) and Van Craenenbroeck (2010).

While in most structures a code-switch between C and the wh-phrase is perfectly grammatical, a code-switch between an English simple wh-phrase and a Spanish C is ungrammatical in embedded questions with a Spanish T. Significantly, the contrast between simple and complex wh-phrases only emerges in embedded wh-questions. While I am not able to present an account of this matrix-embedded asymmetry, it is a distinction that future work will need to account for. Note that the contrast between the two types of wh-phrases is likely to go beyond whether or not they are D(iscourse)-linked wh-phrases. First, the stimuli for the present study were specifically designed so that the complex wh-phrase phrases were not D-linked, at least not in the neutral context with which participants are likely to interpret the stimuli. In addition, simple wh-phrases can also be D-linked, as observed in Pesetsky (1987), so the distinction between simple and complex wh-phrases must go beyond issues of discourse.

Finally, there is also evidence of syntactic differences between the two types of wh-phrases from CS in other language pairs, such as Spanish-German code-switching. Consider the following examples from González-Vilbazo and López (2013), where the questions in (137) have simple wh-phrases and the questions in (138) have complex wh-phrases.

(137) a. No sé who Juan amenazó.  
   [neg] know.1s who.acc Juan threatened  
   ‘I don’t know who Juan threatened.’  
   SV

b. *No sé who amenazó Juan.  
   [neg] know.1s who.acc threatened Juan  
   ‘I don’t know who Juan threatened.’  
   VS

31 See §5.2.2 and §5.2.3 for discussion.
When the wh-question contains a simple wh-phrase, as in (137), only the sentence with SV word order is grammatical, while parallel wh-questions in (138) are grammatical whether the subject comes before or after the verb.\footnote{While the matrix clause of the two sets of sentences are in different languages, this does not affect the grammatical subject positions for the structures.} German wh-questions are similar to English wh-questions in that SV word order is generally required in embedded wh-questions, so \( C \) is required to be in German in (137) but may be in German or in Spanish in (138).

The results of this dissertation also have broader implications for the nature of wh-questions. Notice that, with very few exceptions, there are grammatical word orders for every combination of possible code-switches tested. This suggests a great deal of similarity in the underlying syntactic features that drive the formation of wh-questions, at least for Spanish and English. If wh-question formation instead depended on mostly language-specific syntactic features, we would expect that few code-switched wh-questions would be grammatical. Work on wh-questions in Spanish-Taiwanese (González-Vilbazo et al., 2012), Spanish-German (Ebert, 2013) and Japanese-English code-switching (Ebert et al., 2012) suggests that the similarity in the underlying syntactic features is found across a variety of languages, in spite of many contrasts in the properties of wh-questions in each of the individual languages.

### 5.4.3 Wh-Questions in Code-Switching

As we saw in §3.2, little of the research on the formal properties of CS has focused on wh-questions. In addition, theoretical CS research in general has focused on understanding the restrictions on CS, rather than on applying insights from CS to other aspects of linguistic theory. This dissertation contributes to both of those lacunae.

First, the results of this dissertation expand our knowledge of the properties of wh-questions in code-switching, particularly with respect to inversion but also with respect to the relationship between
C and T and C and the wh-phrase. We have seen that proposing that one syntactic head plays a central role in determining the grammatical subject positions has allowed us to have one unified account of inversion across languages, based on the syntactic properties of the complementizer. In addition, we have seen that there is a complex, context-dependent relationship between C, T, and the wh-phrase such that the grammaticality of a given code-switch depends not just on the language of the individual constituents but also on the complexity of the wh-phrase and the clause of the wh-question. This result hints at the complex ways in which the various factors involved in wh-question formation interact.

Second, this dissertation continues ideas proposed in Belazi et al. (1994), Mahootian and Santorini (1996), González-Vilbazo (2005), González-Vilbazo and López (2011) and González-Vilbazo and López (2012) in which code-switching can also provide insight into properties of the computational system. Rather than limiting ourselves to using current theoretical work to inform our understanding of code-switching, we can also use code-switching to inform our understanding of current theoretical work, offering a new and unique source of evidence for linguistic theory. In essence, CS can be a tool of linguistic analysis, rather than a phenomenon to explain. For example, if the complementizer head really is a determinant head, as data from this dissertation suggest, then we have learned something about the nature of wh-questions. This is particularly important for languages like Spanish where there is good reason to believe that the tense head plays a central role in inversion in Spanish. If it is instead C that ultimately determines the presence of inversion, our theoretical understanding of wh-questions in Spanish needs to be reconsidered.

I would like to finish the discussion on wh-questions in code-switching with some suggestions for future work based on wh-phenomena that are particularly amenable to insights from CS as a tool of linguistic analysis. First, as discussed in §5.4.1.1 there are interesting questions concerning cross-linguistic differences for superiority effects that would be worth exploring using CS. What role do different syntactic heads play? Can we tease apart the interactions between the various constituents involved to understand the sources of these cross-linguistic differences?

33While considering the implications of the data from my dissertation or RQ1, we considered whether each head was a determinant head both for Spanish and for English individually, for the sake of completeness. However, in the end, we saw not only that only C could be the determinant head in both languages but that there is additional evidence of the central role that C plays in a variety of other phenomena. Together, this suggests that this result is more general in nature.
Another fruitful line of inquiry concerns the contrast between languages in which the wh-phrase fronts to a left-peripheral position, such as English or Spanish, and languages where the wh-phrase generally remains in situ, such as Japanese, Taiwanese, or Mandarin. Is there a particular head that plays a central role in this contrast? Are there restrictions on the language and type of wh-phrase with respect to this head?

Finally, CS data could help inform our understanding of different types of wh-phrases and both the language-internal and cross-linguistic contrasts in the properties of these wh-phrases. At a language-internal level, there are differences, for example, between simple and complex wh-phrases in English with respect to ellipsis (Van Craenenbroeck 2010). While the results from this dissertation already support the contrast between simple and complex wh-phrases, as discussed in §5.4.2, data from CS could go much further in helping to delimit the kinds of proposals that could account for this contrast. At the cross-linguistic level, the argument status of wh-phrases sometimes affects the process of wh-question formation, as in Spanish inversion, while other times it does not, as in English inversion. There are also effects that are unique to particular wh-phrases, such as ‘why’ (Ko 2005). Using CS data where one language shows such effects and another does not, it would be possible to test the degree to which these effects depend on features of the individual language and features of the specific wh-phrases as well, as how these language-specific features interact with wh-features of another language.

5.4.4 Code-Switching More Generally

The results of this dissertation have a number of implications for code-switching that go beyond wh-questions. I would like to focus on three in particular. First, I will touch on the idea of CS as a tool of linguistic analysis. In addition, the results of this dissertation provide evidence for a broader hypothesis concerning the role of phase heads in CS: the Phase Head Hypothesis (González-Vilbazo and López 2012). Finally, I explore the empirical consequences of data from this dissertation for general assumptions concerning code-switching.
5.4.4.1 Code-Switching as a Tool of Linguistic Analysis

As we saw in §5.4.3, the results of this dissertation continue a line of research in which studying pheno-
momena in CS not only informs our understanding of CS but also individual languages and the human
language faculty more generally (e.g. [González-Vilbazo 2005; González-Vilbazo and López 2011, 2012].
Importantly, this result goes beyond the phenomenon of wh-questions to encompass not only other as-
pects of syntax but also other domains of linguistics, including phonology, morphology and semantics.
Taking advantage of contrasts between languages in the properties of the phenomena of interest, CS
makes it possible to disentangle the role of different components of the linguistic system, whether they
be syllables, morphemes, or words.

5.4.4.2 Theoretical Considerations

With regard to theoretical considerations, I would like to touch on the following hypothesis, proposed
in [González-Vilbazo and López 2012]:

**Phase Head Hypothesis (PHH)**

The phase head determines grammatical properties of its complement.

Let us consider what implications this hypothesis has for inversion in wh-questions in Spanish and
English. We are interested in particular in the complementizer head, as that is the phase head that
this dissertation focused on. In the case of English, the prediction is less interesting because inversion
already clearly involves a relationship between C and a head within C’s complement, T, since inversion
in English involves T-to-C movement.

For Spanish, we first have to consider whether the PHH applies to inversion in Spanish. Recall that
inversion in Spanish wh-questions results from the subject remaining in situ, in which case the subject
is in the specifier of the v. In phase theory, the complement of v will be inaccessible in the next higher
phase, i.e. in the C phase. Notice, though, that the subject is in the specifier and not the complement, so
it continues to be accessible in the C phase. In contrast, when there is no inversion, i.e. when the subject
is preverbal, the subject is in SpecTP. C, then, could influence the presence of inversion in Spanish by
determining whether it is permissible for the subject to raise to SpecTP. The results of this dissertation
suggest that this is exactly what C does. C was found to be the determinant head, so it is C that determines the grammatical subject positions in Spanish wh-questions, supporting the PHH.

5.4.4.3 Empirical Tests for the Code-switching Framework

Finally, the framework for analyzing code-switching that I adopt makes two predictions that can be tested with the experimental data from this dissertation. One of these is based on differences in the grammatical subject positions between groups, and the other is based on whether one structure can have grammatical subject positions that are a blend of the two languages.

More concretely, if CS sentences are built from the syntactic properties available in each individual language using the same sentence building mechanisms as those used in monolingual sentences, the computational system, then this makes two predictions about the patterns of grammatical subject positions that one should see. First, differences in the grammatical subject positions between participant groups for a given monolingual structure should be reflected in the equivalent CS sentences. The contrasting grammatical subject positions reflect differences in the syntactic features of each language. This in turn changes the available syntactic features for building CS sentences and therefore the grammatical subject positions of those sentences. For example, if both SV word order and VS word order are grammatical for one group but only VS word order is grammatical for another group in monolingual Spanish, the same grammatical word orders should occur in CS, unless the sentences are not truly parallel for some independent reason.

Comparing acceptability judgments across groups and across structures, there are in fact three cases where differences in monolingual judgments are reflected in the CS judgments, while there are no cases where such differences are not reflected in the CS judgments. In the remaining structures and for the remaining groups, there are either no differences (i.e. the groups have the same grammatical subject positions) or there are middle ratings that make it unclear whether there are any real differences. Let us take a look at each of these three cases of contrasting grammatical subject positions in turn. Note that each of these contrasts occurs with wh-questions with a Spanish T, likely due to the more uniform judgments of monolingual English.
The first contrast occurs with SV word order in matrix questions with complex wh-phrases. For Group 1, SV word order is grammatical, as seen in (139a), while for Group 2, SV word order is ungrammatical in Spanish in the same structure, as seen in (140a). Likewise, for Group 1, SV word order is grammatical for CS sentences, as in (139b), while SV word order is ungrammatical in CS sentences for Group 2, as in (140b).

(139) **Group 1 SV in matrix questions with complex wh-phrases**

a. ¿Cuántos cables esos empleados han robado en el trabajo?  
   how many cables those employees have stolen in the work  
   ‘How many cables have those employees stolen at work?’

b. How many cables esos empleados han robado at work?  
   those employees have stolen  
   ‘How many cables have those employees stolen at work?’

(140) **Group 2 SV in matrix questions with complex wh-phrases**

a. *¿Cuántos cables esos empleados han robado en el trabajo?  
   how many cables those employees have stolen in the work  
   ‘How many cables have those employees stolen at work?’

b. *How many cables esos empleados han robado at work?  
   those employees have stolen  
   ‘How many cables have those employees stolen at work?’

Along similar lines, there is a contrast in the grammaticality of V2 word order between Group 5, where the word order is grammatical in Spanish as in (141a), and Group 2, where the same word order is ungrammatical in Spanish, as in (142a). Once again, this contrast is reflected in the grammaticality of the CS sentences, where V2 word order is grammatical for Group 5, as seen in (141b), but is ungrammatical for Group 2, as seen in (142b).

(141) **Group 5 V2 in matrix questions with simple wh-phrases**

a. ¿Qué han esos niños comprado en el centro comercial hasta ahora?  
   what have those children bought in the central commercial until now  
   ‘What have those children bought at the mall so far?’

b. What han esos niños comprado at the mall so far?  
   have those children bought  
   ‘What have those children bought at the mall so far?’
(142) **Group 2 V2 in matrix questions with simple wh-phrases**

a. *¿Qué han esos niños comprado en el centro comercial hasta ahora?*  
   **Sp V2**  
   ‘What have those children bought in the central commercial until now?’

b. *What han esos niños comprado at the mall so far?*  
   **CS V2**  
   ‘What have those children bought at the mall so far?’

Finally, there is a suggestive pattern in the contrast between the grammaticality of V2 word order in matrix questions with complex wh-phrases. For Groups 2, 3, 5, and 6, V2 word order is grammatical in both monolingual Spanish and CS sentences, as seen in (143). On the other hand, for Group 1, this same word order only receives middle ratings for both monolingual Spanish and CS sentences, as seen in (144).

(143) **Groups 2, 3, 5 and 6 V2 in matrix questions with complex wh-phrases**

a. *¿Cuántas blusas han tus sobrinas comprado este otoño?*  
   **Sp**  
   ‘How many blouses have your nieces bought this fall?’

b. *How many blouses han tus sobrinas comprado this fall?*  
   **CS**  
   ‘How many blouses have your nieces bought this fall?’

(144) **Group 1 V2 in matrix questions with complex wh-phrases**

a. *¿Cuántas blusas han tus sobrinas comprado este otoño?*  
   **Sp**  
   ‘How many blouses have your nieces bought this fall?’

b. *How many blouses han tus sobrinas comprado this fall?*  
   **CS**  
   ‘How many blouses have your nieces bought this fall?’

While middle ratings are not generally taken as direct evidence, given their often ambiguous nature, the fact that the ratings for all three word orders are exactly parallel to those in monolingual Spanish, as seen in Table XI, suggests that the middle rating could be meaningful.

The second prediction, again based on the CS framework adopted, stems from a few central assumptions already used throughout this dissertation. First, for any structure, the available word orders are
TABLE XI: CODE-SWITCHING RATINGS PARALLEL MONOLINGUAL RATINGS (GROUP 1)

<table>
<thead>
<tr>
<th>Code-switching</th>
<th>Spanish</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceptability</td>
<td>µ</td>
</tr>
<tr>
<td>Complex</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>V2</td>
<td>?</td>
</tr>
<tr>
<td></td>
<td>?</td>
</tr>
<tr>
<td>VS</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>

derived from a combination of the given syntactic features and the ways in which those features are per-
mitted to combine based on the principles of the computational system, whether for CS or monolingual
sentences. Second, there is one element whose syntactic properties ultimately determine the position of
the subject, as previously argued.

This predicts that the grammatical subject positions should be a combination of either the gram-
matical word orders in one language, those of the other language, or both. The determinant head can
only be in one language in a given sentence, and all of syntactic features of that head are present in that
sentence. Thus, all the subject positions produced by those syntactic features should be grammatical.
The actual grammatical word orders should never be a blend (i.e. a subset) of the grammatical word
orders from both languages.

For example, consider what happens if SV word order and VS word order are grammatical in
monolingual Spanish, but only V2 word order is grammatical in monolingual English. In that case, it
should not be possible for only V2 and VS word orders to be grammatical. In order for V2 word order
to be grammatical, it must be possible for C to be in English, since only monolingual English allows V2
word order in this case. On the other hand, VS word order in the CS sentences can only be possible
because C can also be in Spanish, since only monolingual Spanish allows VS word order in this example.
But if C can be in Spanish, SV word order should also be grammatical, which it is not.
To put the results of this prediction in context, let us see how many total structures there are where the prediction might hold. There are 4 different types of wh-questions in the present study, as discussed in §5.2.1, each of which was tested with a Spanish T and an English T, yielding 8 structures. These 8 structures were tested for each of the 7 participant groups, so this means that there are a total of 56 structures where the grammatical subject positions in CS sentences could potentially be a subset of those in monolingual Spanish and monolingual English. Of those 56 structures, two in fact involve grammatical subject positions that are a subset of the available subject positions, in spite of the prediction that subsets should not be possible. I will briefly review each of these structures in turn before discussing the implications.

Both of the exceptions to the prediction occur with a matrix question with a complex wh-phrase and an English T. As seen in Table XII, only V2 word order is grammatical in monolingual English for both groups, while all three word orders are grammatical or marginally grammatical in monolingual Spanish. Group 3 follows the previously described example pattern, where only V2 and VS word orders are grammatical in the CS sentences. The fact that VS word order is grammatical suggests that C is in Spanish, but this would also predict that SV word order would be grammatical. Along similar lines, both SV and V2 word orders are grammatical in the CS sentences for Group 4. Since SV word order is ungrammatical in monolingual English and there is only a 0.37 difference between the ratings for SV word order in CS and Spanish, this suggests that C is in Spanish. Yet if C were in Spanish, this would also predict the VS word order to be grammatical, which it is not.

While these two exceptions to our prediction are potentially problematic, these blends of grammatical subject positions may actually be a reflection of the fact that we just don’t know what participants are testing for a given sentence that they’re judging. They may try English and Spanish C for one structure but only Spanish C for another. In other words, these exceptions are a reflection of performance, rather than competence. The various possible subject positions are presented to the participants individually and far apart from one another, so it is very possible that participants were judging the individual subject positions differently. Notice that this only occurred with two of the 56 possible structures, suggesting that the prediction does hold, in spite of the occasional performance issue. That is,
TABLE XII: SUBSET OF POSSIBLE SUBJECT POSITIONS IS GRAMMATICAL (ENGLISH $T$)

<table>
<thead>
<tr>
<th>Code-switching</th>
<th>Spanish</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Acceptability</td>
<td>$\mu$</td>
</tr>
<tr>
<td>Group 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>*</td>
<td>1.00</td>
</tr>
<tr>
<td>V2</td>
<td>✓</td>
<td>7.00</td>
</tr>
<tr>
<td>VS</td>
<td>✓</td>
<td>7.00</td>
</tr>
<tr>
<td>Group 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>✓</td>
<td>5.00</td>
</tr>
<tr>
<td>V2</td>
<td>✓</td>
<td>6.88</td>
</tr>
<tr>
<td>VS</td>
<td>*</td>
<td>1.75</td>
</tr>
</tbody>
</table>

we find evidence for discrete bundles of syntactic features that ultimately determine the grammatical subject positions for both CS and monolingual sentences.

5.5 Summary and Conclusions

Broadly speaking, this dissertation contributes to our understanding of three aspects of linguistics and linguistic research: (1) the syntactic properties of wh-questions (2) various characteristics of code-switching, both with respect to wh-questions and more generally, and (3) experimental investigation of code-switching.

Beginning with the syntactic properties of wh-questions, we have seen evidence that C is ultimately responsible for determining the position of the subject in wh-questions (answering RQ1, below), but we have also seen that other elements play a role. In particular, there are cases where $T$ must be in the same language as C and cases where the wh-phrase must be in the same language as C in order for the sentence to be grammatical (answering RQ2, below). In other words, while these constituents are not directly responsible for determining the position of the subject, they still play a crucial role in the
formation of wh-questions. Importantly, this is also a complex role, as sometimes T or the wh-phrase must in the same language as C, while other times there is no such restriction.

**RQ1:** What factors determine the word order in wh-questions in Spanish-English code-switching?

**RQ2:** Are there restrictions on code-switching between particular constituents in wh-questions in Spanish-English code-switching?

Looking at wh-questions in Spanish in particular, results from my dissertation provide evidence against the hypothesis that properties of the Spanish T head are fully responsible for the inversion requirements in Spanish (Toribio, 1993; Barbosa, 2001; Goodall, 2001; Gutiérrez-Bravo, 2005; Zubizarreta, 2012, inter alia). These results also have some implications for Buesa-García (2008)’s phonological requirement for inversion and accounts of wh-questions that allow Spanish wh-questions without a CP (e.g. Grimshaw, 1997; Zubizarreta, 1998; Gutiérrez-Bravo, 2005).

More generally, this dissertation provides additional support for the contrast between simple and complex wh-phrases. In particular, it provides evidence that a code-switch between an English simple wh-phrase and a Spanish C is ungrammatical in embedded questions with a Spanish T. This finding will need to be accounted for in any account of wh-questions in Spanish or English.

With regard to wh-questions in code-switching, the results of my dissertation expand our knowledge concerning the role of C in the formation of wh-questions and the relationship between C, T and the wh-phrase. In addition, if the Phase Head Hypothesis (PHH) (González-Vilbazo and López, 2012) is correct, as evidence from this dissertation suggests, this has implications for wh-questions that go beyond the phenomena studied here. For example, the PHH predicts that C, as a phase head, will determine the grammatical properties of its complement. Given the relationship between the wh-phrase and the complementizer, this opens up new avenues of research, including with respect to superiority effects, wh-in situ versus wh-fronting, and different types of wh-phrases.

Turning to code-switching more generally, both the results and the approach to CS from this dissertation extend a line of research (e.g. Belazi et al., 1994; Mahootian and Santorini, 1996; González-Vilbazo, 2005; González-Vilbazo and López, 2011, 2012) in which CS can tease apart the role of individual con-
stituents and their associated properties in a range of syntactic phenomena, thereby providing a unique source of evidence concerning fundamental properties of the human language faculty.

Finally, in terms of methodology and code-switching, this dissertation serves as one example of an experimental approach to investigating code-switching, particularly with respect to syntactic structure. There are a variety of different concerns that need to be addressed in conducting such research, and this dissertation offers one perspective.

In conclusion, the combination of controlled experimental methods and the unique analytical potential of CS to tease apart otherwise opaque syntactic relationships makes for a valuable tool for addressing a wide range of questions in theoretical linguistics.
CITED LITERATURE


166


Montrul, Silvina. 2004. Subject and object expression in Spanish heritage speakers: A case of mor-


tro Internacional de Lingüística del Noroeste de la Universidad de Sonora*, 131–150. Universidad de Sonora, México.


Prince, Alan, and Paul Smolensky. 1993. Optimality Theory: Constraint Interaction in Generative Gram-

Reinhart, Tanya. 1986. Wh-in-situ: WHO vs. WHICH N. Ms. Tel Aviv University.

Reinhart, Tanya. 1987. Wh-in-situ: who vs. which N. Ms. Tel Aviv University.


A APPENDIX

A.1 Dissertation Materials

A.1.1 English Cloze Test

**Cloze Test - English**

_for each blank in the following passage, please select one of the three options provided. Please, choose the option appropriate for the context._

Joe came home from work on Friday. It was payday, but he wasn't ______ excited about it. He knew that ______ he sat down and paid his ______ and set aside money for groceries, ______ for the car, and a small ______ in his savings account, there wouldn't be ______ much left over for a good ______.

He thought about going out for ______ at his favorite restaurant, but he ______ wasn't in the mood. He wandered ______ his apartment and ate a sandwich. ______ a while, he couldn't stop ______ worrying about the money situation. Finally, ______ got into his car and started ______.

He didn't have a destination in ______, but he knew that he wanted ______ far away from the city ______ he lived. He turned onto a quiet country ______. The country sights made him feel ______. His mind wandered as he drove ______ small farms and he began to ______ living on his own piece of ______ and becoming self-sufficient. It had always ______ a dream of his, but he ______ never done anything to make it ______ reality. Even as he was thinking, ______ logical side was scoffing at his ______ imaginings. He debated the advantages and ______ of living in the country and ______ his own food. He imagined his ______ equipped with a solar energy panel ______ the roof to heat the house ______ winter and power a water heater. ______ envisioned fields of vegetables for canning ______ preserving to last through the winter. ______ the crops had a good yield, ______ he could sell the surplus and ______ some farming equipment with the extra ______.

Suddenly, Joe stopped thinking and laughed ______ loud, "I'm really going to go ______ with this?"

_Adapted from American Kernal Lessons: Advanced Students' Book. O'Neill, Cornelius and Washburn (1981)"
Cuestionario Lingüístico – Parte 1: Selección Múltiple

En cada una de las siguientes oraciones hay una palabra o una frase que ha sido omitida. Seleccione la opción que mejor complete la oración.

1. Al oír del accidente de su buen amigo, Paco se puso ____________________.
2. No puedo comprarlo porque me ____________________ dinero.
3. Tuvo que guardar cama por estar ____________________.
4. Aquí está tu café, Juanito. No te quemes, que está muy ____________________.
5. Al romper los anteojos, Juan se asustó porque no podía ____________________ sin ellos.
6. ¡Pobrecita! Está resfriada y no puede ____________________.
7. Era una noche oscura sin ____________________.
8. Cuando don Carlos salió de su casa, saludó a un amigo suyo, «Buenos días, ____________________»
9. ¡Qué ruido había con los gritos de los niños y el ____________________ de los perros!
10. Para saber la hora, don Juan miró el ____________________.
11. Yo, que comprendo poco de mecánica, sé que el auto no puede funcionar sin ____________________.
12. Nos dijo mamá que era hora de comer y por eso ____________________.
13. ¡Cuidado con ese cuchillo o vas a ____________________ el dedo!
14. Tuvo tanto miedo de caerse que se negó a ____________________ con nosotros.
15. Abrió la ventana y miró: en efecto, grandes lenguas de ____________________ salían llameando de las casas.
16. Compró ejemplares de todos los diarios pero en vano. No halló ____________________
17. Por varias semanas acudieron colegas del difunto profesor a el dolor de la viuda.

18. Sus amigos pudieron haberlo salvado pero lo dejaron .

19. Al salir de la misa me sentía tan caritativo que no pude menos que a un pobre mendigo que había allí sentado.

20. Al lado de la Plaza de Armas había dos limosneros pidiendo .

21. Siempre maltratado por los niños, el perro no podía acostumbrarse a de sus nuevos amos.

22. ¿Dónde estará mi cartera? La dejé aquí mismo hace poco y parece que el necio de mi hermano ha vuelto a .

23. Permaneció un gran rato abstraído, los ojos clavados en el fogón y el pensamiento .

24. En vez de dirigir el tráfico estabas charlando, así que tú mismo del choque.

25. Posee esta tierra un clima tan propio para la agricultura como para .

26. Aficionado leal de obras teatrales, Juan se entristeció al saber del gran actor.

27. Se reunieron a menudo para efectuar un tratado pero no pudieron .

28. Se negaron a embarcarse porque tenían miedo de .

29. La mujer no aprobó el cambio de domicilio pues no le gustaba .

30. Era el único que tenía algo que comer pero se negó a .
El sueño de Joan Miró

Hoy se inaugura en Palma de Mallorca la Fundación de Joan Miró, en el mismo lugar en donde el artista vivió sus últimos treinta y cinco años. El sueño de Joan Miró se ha [__________]. Los fondos donados a la ciudad por el pintor y su esposa en 1981 permitieron que el sueño se [__________]; más tarde, en 1986 el Ayuntamiento de Palma de Mallorca decidió [__________] al arquitecto Rafael Moneo un edificio que [__________] a la vez como sede de la entidad y como museo moderno. El proyecto ha tenido que [__________] múltiples obstáculos de carácter administrativo. Miró, coincidiendo [__________] los deseos de toda su familia, quiso que su obra no quedara expuesta en ampulosos panteones de arte o en [__________] de coleccionistas acaudalados; por ello, en 1981, creó la fundación mallorquina. Y cuando estaba [__________] punto de morir, donó terrenos y edificios, así como las obras de arte que en ellos [__________].

El edificio que ha construido Rafael Moneo se enmarca en [__________] se denomina «Territorio Miró», espacio en el que se han [__________] de situar los distintos edificios que constituyen la herencia del pintor.

El acceso a los mismos quedará [__________] para evitar el deterioro de las obras. Por otra parte, se [__________], en los talleres de grabado y litografía, cursos [__________] las distintas técnicas de estampación. Estos talleres también se cederán periódicamente a distintos artistas contemporáneos, [__________] se busca que el «Territorio Miró» [__________] un centro vivo de creación y difusión del arte a todos los [__________].

La entrada costará 500 pesetas y las previsiones dadas a conocer ayer aspiran [__________] que el centro acoja a unos 150.000 visitantes al año. Los responsables esperan que la institución funcione a [__________] rendimiento a principios de la [__________] semana, si bien el catálogo completo de las obras de la Fundación Pilar y Joan Miró no estará listo hasta dentro de dos años.
A.1.3 Training

Training: Información General

In this study we are interested in finding out cómo funciona el code-switching de inglés a español y vice-versa. Partimos de la idea que code-switching is a form of linguistic expression like any other and therefore it is subject to rules and restrictions like any other. Las reglas y restricciones que nos interesan no tienen nada que ver con lo que hayas aprendido en la escuela, but with the linguistic structures that you have in your mind as a native speaker. Let us give you an example in English:

(1) There is likely that John likes Mary.

You'll notice that the sentence above is accompanied by a scale. We will use this scale to rate sentences from 1 to 7, where 1 means completely impossible and 7 means completely possible. Here we selected 1 (in yellow) because native speakers of English find this sentence very strange. It is perfectly understandable but there is something about its structure that sounds un-English.

En cambio, the following sentence sounds perfect, hence selecting 7:

(2) There is someone in the garden.

Veamos un ejemplo en español:

(3) Esa película le ha gustado a nadie.

De nuevo, esta oración se lleva un 1. Aunque la entendemos perfectamente, hay algo que suena raro en esta oración, no es una oración en español. Para comparar, considere la siguiente oración:
(4) Juan compró el periódico.

Esta oración es normalísima. Por lo tanto, le damos un 7.

Y ahora comparemos two examples in code-switching:

(5) John le bought una casa.

(6) Mis primos nadaron in the pool for three hours.

Muchos hablantes bilingües de español/inglés who practice code-switching regularly or occasionally agree that sentence (5) sounds strange, just like (1) and (3), while (6) sounds completely fine.

In the survey, you might find that some sentences are neither perfect nor totally awful. Therefore, you can evaluate the sentences you will read as either 1 or 7, pero también te damos varias opciones por si hay dudas.

**Training: More Details**

Lee la siguiente oración and then give it a rating from 1 to 7.

(7) I know my neighbor chewed gum yesterday.

This sentence may sound odd because it doesn't seem like something anyone would ever say or care about. The question is whether this is una oración posible en inglés. Si es el caso, recibe un 7. Now, take a moment to rate the following sentence:
(8) He surprise that no one called yesterday.

On the other hand, this sentence is comprehensible and the situation seems plausible (una posible situación: ayer él esperaba una llamada, pero al final no lo llamaron), but the sentence is not possible to say in English. It's just not English, for whatever reason, and therefore it receives a rating of 1.

Para juzgar cada oración, the question, then, is whether the sentence is possible in English, Spanish or code-switching, even if you don’t know why anyone would actually say the sentence.

And let’s take a look at one final pair of sentences:

(9) Compró Juan un nuevo libro for the party this week?

(10) I want ir al mercado.

En cuanto al significado, the first one is a little odd. You don’t usually buy a book for a party. Pero como no hay nada raro en la forma de la oración, it would be rated a 7. La otra oración, on the other hand, se puede entender pero no se puede decir. Because of that, it would be rated 1.

Training: Práctica

Let’s have you practice with a few more sentences. Por favor, rate the following sentences on a scale from 1 to 7.

(11) Dijo que, los niños, los iban a traer mañana.
(12) The child seems sleeping.

(13) What do you think that John forgot at the office?

(14) She gave me el libro de ejercicios.

(15) They used to serve bebidas alcohólicas en ese restaurante.

(16) Eres una mujer proud.

(17) It is believed to be obvious by everybody that Fred is crazy.

(18) What do you worry if John forgets at the office?
(19) Nunca nosotros hemos visto esa película.

(20) Tom owns a Mercedes y Pedro también.

(21) John has been durmiendo demasiado hoy.
A.1.4 Background Questionnaire

Demographic Information

1. First name:
2. Apellido (Last name):
3. Año de nacimiento (Year of birth):
4. Highest level of education completed (Some high school, High school, Some college, College graduate)
5. Place of birth (city, state, country):
6. Lugar de residencia actual (ciudad, estado, país) (Place of current residence (city, state, country)):
7. If you were not born in the US
   a) What ages did you live in your birth country?:
   b) How long have you lived in the US?:

Familia (Family)

1. Where are your parents (or caregivers) from? (city, state, country)
   a) Parent 1:
   b) Parent 2:
2. Qué lengua(s) hablan sus padres (o las personas encargadas del cuidado)? (marque todas las que apliquen) (What languages do your parents speak (or the people in charge of your care)? (check all that apply))
   a) Persona 1 (Español, Inglés, Mixto, Otra(s): ______) (Person 1 (Spanish, English, Mixed, Other(s)))
   b) Persona 2 (Español, Inglés, Mixto, Otra(s): ______) (Person 1 (Spanish, English, Mixed, Other(s)))
3. What language(s) do your parents (or caregivers) use when speaking to you? (check all that apply)
   a) Parent 1 (Spanish, English, Mixed, Other(s): ______)
   b) Parent 2 (Spanish, English, Mixed, Other(s): ______)
4. Qué lengua(s) usa usted cuando habla con sus padres (o las personas encargadas del cuidado)? (marque todas las que apliquen) (What language(s) do you use when you speak with your parents (or the people in charge of your care) (check all that apply))
   a) Persona 1 (Español, Inglés, Mixto, Otra(s): ______) (Person 1 (Spanish, English, Mixed, Other(s)))
b) Persona 2 (Español, Inglés, Mixto, Otra(s): _____) (Person 1 (Spanish, English, Mixed, Other(s)))

5. If you have siblings, what language(s) do you use when speaking to each other? (check all that apply)
   a) Spanish, English, Mixed, Other(s): _____

Linguistic History

1. A qué edad empezó usted a aprender las siguientes lenguas? (At what age did you start learning the following languages?)
   a) Inglés (Desde nacer, Desde tener 1 año, Desde tener 2 años, Desde tener 3 años, Desde tener 4 años, Desde tener 5 años, Desde tener 6 años, Después de tener 7 años) (English (Since birth, Since 1 year old, Since 2 year old, Since 3 year old, Since 4 year old, Since 5 year old, Since 6 year old, After 7 years old))
   b) Español (Desde nacer, Desde tener 1 año, Desde tener 2 años, Desde tener 3 años, Desde tener 4 años, Desde tener 5 años, Desde tener 6 años, Después de tener 7 años) (Spanish (Since birth, Since 1 year old, Since 2 year old, Since 3 year old, Since 4 year old, Since 5 year old, Since 6 year old, After 7 years old))

2. Did you begin to speak both English and Spanish by age 6? (Yes, No)

3. Qué lengua(s) oía usted en casa desde que nació hasta los 5 años? (marque todas las que apliquen) (What language(s) did you hear at home from birth until 5 years old? (check all that apply))
   a) Español, Inglés, Mixto, Otra(s): _____ (Spanish, English, Mixed, Other(s))

4. What language(s) did your parents use mostly when speaking to you while growing up? (check all that apply)
   a) Spanish, English, Mixed, Other(s): _____ (Spanish, English, Mixed, Other(s))

Dominio Lingüístico (Linguistic Competence)

1. Rate your current overall language ability in English (5 - Understand and speak fluently like a native speaker, 4 - Understand and speak comfortably, with little difficulty, 3 - Understand and speak, but with some difficulty, 2 - Understand and speak, but with great difficulty, 1 - Understand but cannot speak)

2. Califique su dominio general del español (5 - Comprender y hablar con fluidez como un nativo hablante, 4 - Comprender y hablar con facilidad, sin muchas dificultades, 3 - Comprender y hablar, pero con algunas dificultades, 2 - Comprender y hablar, pero con muchas dificultades, 1 - Comprender, pero no hablar)  

3. Rate your English proficiency in the following areas

---

1The translation is the same as above but for Spanish.
a) Reading (5 - Native speaker command, 4 - Very good, 3 - Good, 2 - Needs work, 1 - Poor)
b) Writing (5 - Native speaker command, 4 - Very good, 3 - Good, 2 - Needs work, 1 - Poor)
c) Speaking (5 - Native speaker command, 4 - Very good, 3 - Good, 2 - Needs work, 1 - Poor)
d) Understanding (5 - Native speaker command, 4 - Very good, 3 - Good, 2 - Needs work, 1 - Poor)

4. Califique su dominio del español en las siguientes áreas:

a) Lectura (5 - Native speaker command, 4 - Very good, 3 - Good, 2 - Needs work, 1 - Poor)
b) Escritura (5 - Native speaker command, 4 - Very good, 3 - Good, 2 - Needs work, 1 - Poor)
c) Habla (5 - Native speaker command, 4 - Very good, 3 - Good, 2 - Needs work, 1 - Poor)
d) Comprensión (5 - Native speaker command, 4 - Very good, 3 - Good, 2 - Needs work, 1 - Poor)

5. Indicate the percentage of time you regularly use English and/or español en las siguientes lugares/situaciones (...Spanish in the following places/situations)

a) En casa (At home) (100% English - 0% Spanish, 90% English - 10% Spanish, 80% English - 20% Spanish, 70% English - 30% Spanish, 60% English - 40% Spanish, 50% English - 50% Spanish, 40% English - 60% Spanish, 30% English - 70% Spanish, 20% English - 80% Spanish, 10% English - 90% Spanish, 0% English - 100% Spanish)
b) Visiting family members (if not at home)
c) En la escuela o el trabajo (At school or at work)
d) Con amigos (With friends)
e) On the telephone
f) De vacaciones (On vacation)
g) While shopping
h) En fiestas y eventos sociales (At parties and social events)

Bilingüismo (Bilingualism)

1. Is there someone you speak both English and Spanish with? (Yes, No)
   a) If yes, please, indicate who:

2. Cuando hablas con esta(s) persona(s) usan ustedes las dos lenguas en la misma conversación? (Sí, No) (When you speak with that person or those people, do you use two languages in the same conversation? (Yes, No))

3. When speaking with this person (or people), do you ever use both languages in the same sentence? (Yes, No)

4. If yes, please, think of an example sentence that you could possibly use con esta(s) persona(s): (... with that person or with those people)
A.1.5 Stimuli

This section presents the stimuli used in the dissertation experiment. Because the final list of CS stimuli and the majority of the monolingual stimuli are generated from 32 monolingual base sentences, the entire set of stimuli is not listed. Instead, I present the monolingual base sentences, a sample CS sentences generated from a translation-equivalent pair of sentences, and a sample of monolingual sentences based off of the same pair of base sentences. Following that is a table of some additional monolingual stimuli that were included with the experiment.

Table [XIII] is organized as follows. The first column lists the monolingual sentences, with the English version in one row and then the translation equivalent in Spanish in the next row. Note that the subject position of each sentence is based on the canonical subject position for that sentence for that language. The actual stimuli manipulate the final subject position based on the subject positions included for a given type of wh-question. The ‘Lexical Unit’ column indicates the pairs of translation-equivalent sentences, and the ‘Language’ column indicates the language of the given monolingual base sentence. The ‘Complexity’ column indicates whether the wh-phrase of the wh-question is a simple or a complex wh-phrase, and the final column indicates whether the wh-question is a matrix or embedded question.

Table [XIV] lists all of the CS sentences generated from the Lexical Unit 2 monolingual base sentences. Because the lexical unit is an embedded question with a simple wh-phrase, all of the generated sentences also have these characteristics. The subject and the adverb are also in the opposite language, as are the wh-phrase and the verb. In addition, the same sentences appear in SV and VS word order.

Table [XV] lists all of the monolingual sentences generated from the Lexical Unit 2 monolingual base sentences. Once again, all of the sentences are embedded questions with a simple wh-phrase. Being monolingual sentences, only the word order is varied.

Finally, Tables [XVI] and [XVII] list all of the monolingual stimuli that were not generated from the monolingual base sentences in Table [XIII]. The T-to-C movement tests were for testing whether a given participants allowed T-to-C movement in Spanish, and the natural examples consisted of sentences from McCloskey (2006) for testing whether participants judged embedded wh-questions with inversion as acceptable. Finally, a number of stimuli were included to test superiority effects.
<table>
<thead>
<tr>
<th>SentenceLexicalUnit</th>
<th>Language</th>
<th>Complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>I don't remember how many beers those students have bought in the last hour.</td>
<td>English</td>
<td>complex embedded</td>
</tr>
<tr>
<td>No recuerdo cuántas cervezas han comprado esos estudiantes en la última hora.</td>
<td>Spanish</td>
<td>complex embedded</td>
</tr>
<tr>
<td>I don't remember what your colleagues have bought this week.</td>
<td>English</td>
<td>simple embedded</td>
</tr>
<tr>
<td>No recuerdo qué han comprado tus colegas esta semana.</td>
<td>Spanish</td>
<td>simple embedded</td>
</tr>
<tr>
<td>How many blouses have your nieces bought this fall?</td>
<td>English</td>
<td>complex matrix</td>
</tr>
<tr>
<td>Cuántas blusas han comprado tus sobrinas este otoño?</td>
<td>Spanish</td>
<td>complex matrix</td>
</tr>
<tr>
<td>What have those kids bought at the mall so far?</td>
<td>English</td>
<td>simple matrix</td>
</tr>
<tr>
<td>Qué han comprado esos niños en el centro comercial hasta ahora?</td>
<td>Spanish</td>
<td>simple matrix</td>
</tr>
<tr>
<td>I don't remember how many apples your children have eaten since Sunday.</td>
<td>English</td>
<td>complex embedded</td>
</tr>
<tr>
<td>No recuerdo cuántas manzanas han comido tus hijos desde el domingo.</td>
<td>Spanish</td>
<td>complex embedded</td>
</tr>
<tr>
<td>I don't remember what those rabbits have eaten since I planted the garden.</td>
<td>English</td>
<td>simple embedded</td>
</tr>
<tr>
<td>No recuerdo qué han comido esos conejos desde que planté el jardín.</td>
<td>Spanish</td>
<td>simple embedded</td>
</tr>
<tr>
<td>How many french fries have those guys eaten since they arrived?</td>
<td>English</td>
<td>complex matrix</td>
</tr>
<tr>
<td>Cuántas papas fritas han comido esos muchachos desde que llegaron?</td>
<td>Spanish</td>
<td>complex matrix</td>
</tr>
<tr>
<td>What have your classmates written since the homework was assigned?</td>
<td>English</td>
<td>simple matrix</td>
</tr>
<tr>
<td>Qué han escrito tus compañeros desde que se asignó la tarea?</td>
<td>Spanish</td>
<td>simple matrix</td>
</tr>
<tr>
<td>I don't remember how much money those criminals have stolen in the last decade.</td>
<td>English</td>
<td>complex embedded</td>
</tr>
<tr>
<td>No recuerdo cuánto dinero han robado esos delincuentes en la última década.</td>
<td>Spanish</td>
<td>complex embedded</td>
</tr>
<tr>
<td>I don't remember what those bankers have stolen since the last economic crisis.</td>
<td>English</td>
<td>simple embedded</td>
</tr>
<tr>
<td>No recuerdo qué han robado esos banqueros desde la última crisis económica.</td>
<td>Spanish</td>
<td>simple embedded</td>
</tr>
<tr>
<td>How many cables have those employees stolen at work?</td>
<td>English</td>
<td>complex matrix</td>
</tr>
<tr>
<td>Cuántos cables han robado esos empleados en el trabajo?</td>
<td>Spanish</td>
<td>complex matrix</td>
</tr>
<tr>
<td>What have your neighbors stolen while you weren't there?</td>
<td>English</td>
<td>simple matrix</td>
</tr>
<tr>
<td>Qué han robado tus vecinos mientras no estabas?</td>
<td>Spanish</td>
<td>simple matrix</td>
</tr>
<tr>
<td>Sentence</td>
<td>Subject</td>
<td>Wh-phrase</td>
</tr>
<tr>
<td>----------</td>
<td>---------</td>
<td>-----------</td>
</tr>
<tr>
<td>I don't remember what your colleagues have bought this week.</td>
<td>English</td>
<td>Spanish</td>
</tr>
<tr>
<td>I don't remember what tus colegas have bought this week.</td>
<td>Spanish</td>
<td>English</td>
</tr>
<tr>
<td>I don't remember qué your colleagues have bought esta semana.</td>
<td>English</td>
<td>Spanish</td>
</tr>
<tr>
<td>I don't remember qué tus colegas have bought this week.</td>
<td>Spanish</td>
<td>English</td>
</tr>
<tr>
<td>Sentence</td>
<td>Language Complexity</td>
<td>Word Order</td>
</tr>
<tr>
<td>----------</td>
<td>---------------------</td>
<td>------------</td>
</tr>
<tr>
<td>I don't remember what your colleagues have bought this week.</td>
<td>English simple</td>
<td>SV embedded</td>
</tr>
<tr>
<td>No recuerdo qué tus colegas han comprado esta semana.</td>
<td>Spanish simple</td>
<td>VS embedded</td>
</tr>
<tr>
<td>I don't remember what have bought your colleagues this week.</td>
<td>English simple</td>
<td>VS embedded</td>
</tr>
<tr>
<td>No recuerdo que han comprado tus colegas esta semana.</td>
<td>Spanish simple</td>
<td>V2 embedded</td>
</tr>
</tbody>
</table>

TABLE XV: MONOLINGUAL LEXICAL UNIT 2 SENTENCES
<table>
<thead>
<tr>
<th>Sentence</th>
<th>Language</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Esos estudiantes jamás han terminado el examen.</td>
<td>Spanish</td>
<td>T-to-C movement tests</td>
</tr>
<tr>
<td>2. Susana todavía estudia historia del arte.</td>
<td>Spanish</td>
<td>T-to-C movement tests</td>
</tr>
<tr>
<td>3. Mi vecino nunca ha corrido por las tardes.</td>
<td>Spanish</td>
<td>T-to-C movement tests</td>
</tr>
<tr>
<td>4. Yo sencillamente no aguanto a ese mocoso.</td>
<td>Spanish</td>
<td>T-to-C movement tests</td>
</tr>
<tr>
<td>5. ¿A quién jamás has ofendido tú con tus acciones?</td>
<td>Spanish</td>
<td>T-to-C movement tests</td>
</tr>
<tr>
<td>6. ¿Qué idioma todavía estudia Pepita en su tiempo libre?</td>
<td>Spanish</td>
<td>T-to-C movement tests</td>
</tr>
<tr>
<td>7. ¿Dónde nunca ha corrido tu hermana?</td>
<td>Spanish</td>
<td>T-to-C movement tests</td>
</tr>
<tr>
<td>8. ¿A cuál de esos mocosos sencillamente no aguantas?</td>
<td>Spanish</td>
<td>T-to-C movement tests</td>
</tr>
<tr>
<td>9. ¿Quién compró qué?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>10. ¿Quién escribió qué?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>11. ¿Quién robó qué?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>12. ¿Quién comió qué?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>13. ¿Qué compró quién?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>14. ¿Qué escribió quién?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>15. ¿Qué robó quién?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>16. ¿Qué comió quién?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>17. ¿Quién crees que compró qué?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>18. ¿Quién crees que escribió qué?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>19. ¿Quién crees que robó qué?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>20. ¿Quién crees que comió qué?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>21. ¿Qué crees que compró quién?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>22. ¿Qué crees que escribió quién?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>23. ¿Qué crees que robó quién?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>24. ¿Qué crees que comió quién?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>25. ¿Quién crees que compró cuáles libros?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>26. ¿Quién crees que escribió cuáles libros?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>27. ¿Quién crees que robó cuáles libros?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>28. ¿Quién crees que comió cuáles dulces?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>29. ¿Cuáles libros crees que compró quién?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>30. ¿Cuáles libros crees que escribió quién?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>31. ¿Cuáles libros crees que robó quién?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>32. ¿Cuáles dulces crees que comió quién?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>33. ¿Qué niño compró cuál bicicleta?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>34. ¿Qué niño escribió cuál poema?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>35. ¿Qué niño robó cuál bicicleta?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>36. ¿Qué niño comió cuales dulces?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>37. ¿Cuál bicicleta compró qué niño?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>38. ¿Cuál poema escribió qué niño?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>39. ¿Cuál bicicleta robó qué niño?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>40. ¿Cuáles dulces comió qué niño?</td>
<td>Spanish</td>
<td>Superiority tests</td>
</tr>
</tbody>
</table>
**TABLE XVII: ADDITIONAL MONOLINGUAL ENGLISH STIMULI**

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Language</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Who bought what?</td>
<td>English</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>2. Who wrote what?</td>
<td>English</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>3. Who stole what?</td>
<td>English</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>4. Who ate what?</td>
<td>English</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>5. What did who buy?</td>
<td>English</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>6. What did who write?</td>
<td>English</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>7. What did who steal?</td>
<td>English</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>8. What did who eat?</td>
<td>English</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>13. What did Susan fix how?</td>
<td>English</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>14. What did Sandy buy how?</td>
<td>English</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>15. What did Steve sell how?</td>
<td>English</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>16. What did Sam make how?</td>
<td>English</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>17. Which of these cars did which of these guys fix?</td>
<td>English</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>18. Which of these boats did which of these guys sell?</td>
<td>English</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>19. Which of these cases did which of these lawyers win?</td>
<td>English</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>20. Which of these books did which of these lawyers buy?</td>
<td>English</td>
<td>Superiority tests</td>
</tr>
<tr>
<td>21. I found out how did they get into the building.</td>
<td>English</td>
<td>Natural examples</td>
</tr>
<tr>
<td>22. I usually know who might they hire.</td>
<td>English</td>
<td>Natural examples</td>
</tr>
<tr>
<td>23. I remember clearly how many people did they arrest.</td>
<td>English</td>
<td>Natural examples</td>
</tr>
<tr>
<td>24. The police discovered who had they beaten up.</td>
<td>English</td>
<td>Natural examples</td>
</tr>
<tr>
<td>25. You’d be better off asking why did he marry me.</td>
<td>English</td>
<td>Natural examples</td>
</tr>
<tr>
<td>26. I asked him from what source could the reprisals come.</td>
<td>English</td>
<td>Natural examples</td>
</tr>
<tr>
<td>27. I wonder what is he like.</td>
<td>English</td>
<td>Natural examples</td>
</tr>
<tr>
<td>28. The baritone was asked what did he think of Mrs Kearney’s conduct.</td>
<td>English</td>
<td>Natural examples</td>
</tr>
<tr>
<td>29. You’d be better off asking why he married me.</td>
<td>English</td>
<td>Natural examples</td>
</tr>
<tr>
<td>30. I asked him from what source the reprisals could come.</td>
<td>English</td>
<td>Natural examples</td>
</tr>
<tr>
<td>31. I wonder what he is like.</td>
<td>English</td>
<td>Natural examples</td>
</tr>
<tr>
<td>32. The baritone was asked what he thought of Mrs Kearney’s conduct.</td>
<td>English</td>
<td>Natural examples</td>
</tr>
</tbody>
</table>
A.2 Dissertation Data Summary

This section contains the acceptability judgments for each of the participants groups. To facilitate comparing the CS and monolingual acceptability judgments, the monolingual Spanish and monolingual English ratings are listed next to the ratings for each CS structure. The CS ratings are listed for those structures where T is in Spanish (and therefore the wh-phrase is in English) and then where T is in English (and therefore the wh-phrase is in Spanish). The subject is in the same language as T for all structures. Because the same structures are listed when T is in Spanish and when T is in English in the CS sentences, the monolingual acceptability judgments are listed twice.

The results tables are organized as follows. The first column is organized first by the language of the tense head (Spanish T then English T) and then by the factors determining the structure of each type of wh-question (clause, complexity and subject position). The remaining columns list the results for each word order for CS sentences, monolingual Spanish sentences, and monolingual English sentences. Under each language, there are three columns. First, there is a column for the overall acceptability of each word order, where ‘✓’ indicates acceptable, ‘*’ indicates unacceptable, and ‘?’ indicates unclear or middle ratings (those between 3 and 5 on the 7-point Likert scaled used for the present study). Second, there is a column with the mean rating (\( \mu \)), and finally there is a column with the standard deviation (\( \sigma \)).
<table>
<thead>
<tr>
<th></th>
<th>Spanish</th>
<th></th>
<th>English</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Acceptability</td>
<td>µ</td>
<td>σ</td>
<td>Acceptability</td>
<td>µ</td>
</tr>
<tr>
<td><strong>Code-switching</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Spanish T</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>matrix</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>complex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>✓</td>
<td>6.13  0.83</td>
<td>✓</td>
<td>5.38  2.07</td>
<td>*</td>
</tr>
<tr>
<td>V2</td>
<td>?</td>
<td>3.13  2.42</td>
<td>?</td>
<td>3.88  2.10</td>
<td>✓</td>
</tr>
<tr>
<td>VS</td>
<td>✓</td>
<td>5.63  2.00</td>
<td>✓</td>
<td>6.75  0.46</td>
<td>*</td>
</tr>
<tr>
<td><strong>simple</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>*</td>
<td>1.00  0.00</td>
<td>*</td>
<td>2.38  2.00</td>
<td>*</td>
</tr>
<tr>
<td>V2</td>
<td>*</td>
<td>1.88  1.36</td>
<td>*</td>
<td>2.50  1.85</td>
<td>✓</td>
</tr>
<tr>
<td>VS</td>
<td>?</td>
<td>3.75  2.43</td>
<td>✓</td>
<td>6.13  0.99</td>
<td>*</td>
</tr>
<tr>
<td><strong>embedded</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>complex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>✓</td>
<td>5.38  2.72</td>
<td>✓</td>
<td>6.88  0.35</td>
<td>✓</td>
</tr>
<tr>
<td>V2</td>
<td>?</td>
<td>4.88  2.75</td>
<td>✓</td>
<td>6.63  0.74</td>
<td>*</td>
</tr>
<tr>
<td>VS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>simple</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>*</td>
<td>3.00  1.69</td>
<td>✓</td>
<td>5.00  2.14</td>
<td>✓</td>
</tr>
<tr>
<td>V2</td>
<td>*</td>
<td>1.88  1.64</td>
<td>✓</td>
<td>6.63  0.74</td>
<td>*</td>
</tr>
<tr>
<td>VS</td>
<td>*</td>
<td>2.38  2.33</td>
<td>✓</td>
<td>6.13  0.99</td>
<td>*</td>
</tr>
<tr>
<td><strong>English T</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>matrix</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>complex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>*</td>
<td>2.75  2.31</td>
<td>✓</td>
<td>5.38  2.07</td>
<td>*</td>
</tr>
<tr>
<td>V2</td>
<td>✓</td>
<td>6.25  1.16</td>
<td>?</td>
<td>3.88  2.10</td>
<td>✓</td>
</tr>
<tr>
<td>VS</td>
<td>*</td>
<td>1.50  1.41</td>
<td>✓</td>
<td>6.75  0.46</td>
<td>*</td>
</tr>
<tr>
<td><strong>simple</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>*</td>
<td>1.00  0.00</td>
<td>*</td>
<td>2.38  2.00</td>
<td>*</td>
</tr>
<tr>
<td>V2</td>
<td>*</td>
<td>2.75  1.91</td>
<td>*</td>
<td>2.50  1.85</td>
<td>✓</td>
</tr>
<tr>
<td>VS</td>
<td>*</td>
<td>1.75  2.12</td>
<td>✓</td>
<td>6.13  0.99</td>
<td>*</td>
</tr>
<tr>
<td><strong>embedded</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>complex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>✓</td>
<td>6.63  1.06</td>
<td>✓</td>
<td>6.88  0.35</td>
<td>✓</td>
</tr>
<tr>
<td>V2</td>
<td>*</td>
<td>1.75  2.12</td>
<td>✓</td>
<td>6.63  0.74</td>
<td>*</td>
</tr>
<tr>
<td>VS</td>
<td>*</td>
<td>2.25  2.12</td>
<td>✓</td>
<td>5.00  2.14</td>
<td>✓</td>
</tr>
<tr>
<td><strong>simple</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>*</td>
<td>1.00  0.00</td>
<td>✓</td>
<td>6.13  0.99</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Code-switching</td>
<td></td>
<td>Spanish</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>---------------</td>
<td>------------</td>
<td>---------</td>
<td>------------</td>
<td>-------------</td>
</tr>
<tr>
<td></td>
<td>Acceptability</td>
<td>µ</td>
<td>σ</td>
<td>Acceptability</td>
<td>µ</td>
</tr>
<tr>
<td><strong>Spanish T</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>matrix</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>complex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>*</td>
<td>1.00 0.00</td>
<td>*</td>
<td>2.25 2.50</td>
<td>*</td>
</tr>
<tr>
<td>V2</td>
<td>✓</td>
<td>5.50 3.00</td>
<td>✓</td>
<td>5.50 3.00</td>
<td>✓</td>
</tr>
<tr>
<td>VS</td>
<td>✓</td>
<td>5.50 3.00</td>
<td>✓</td>
<td>7.00 0.00</td>
<td>*</td>
</tr>
<tr>
<td><strong>simple</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>*</td>
<td>3.00 3.46</td>
<td>?</td>
<td>4.00 3.46</td>
<td>*</td>
</tr>
<tr>
<td>V2</td>
<td>*</td>
<td>2.50 3.00</td>
<td>*</td>
<td>2.50 3.00</td>
<td>✓</td>
</tr>
<tr>
<td>VS</td>
<td>?</td>
<td>4.00 3.46</td>
<td>✓</td>
<td>7.00 0.00</td>
<td>*</td>
</tr>
<tr>
<td><strong>embedded</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>complex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>✓</td>
<td>7.00 0.00</td>
<td>✓</td>
<td>7.00 0.00</td>
<td>✓</td>
</tr>
<tr>
<td>V2</td>
<td>?</td>
<td>4.00 3.46</td>
<td>✓</td>
<td>7.00 0.00</td>
<td>*</td>
</tr>
<tr>
<td>VS</td>
<td>?</td>
<td>4.00 3.46</td>
<td>✓</td>
<td>7.00 0.00</td>
<td>✓</td>
</tr>
<tr>
<td><strong>simple</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>*</td>
<td>2.50 3.00</td>
<td>✓</td>
<td>7.00 0.00</td>
<td>*</td>
</tr>
<tr>
<td>V2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>English T</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>matrix</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>complex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>*</td>
<td>2.50 3.00</td>
<td>*</td>
<td>2.25 2.50</td>
<td>*</td>
</tr>
<tr>
<td>V2</td>
<td>✓</td>
<td>7.00 0.00</td>
<td>✓</td>
<td>5.50 3.00</td>
<td>✓</td>
</tr>
<tr>
<td>VS</td>
<td>✓</td>
<td>1.00 0.00</td>
<td>✓</td>
<td>7.00 0.00</td>
<td>*</td>
</tr>
<tr>
<td><strong>simple</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>*</td>
<td>1.00 0.00</td>
<td>?</td>
<td>4.00 3.46</td>
<td>*</td>
</tr>
<tr>
<td>V2</td>
<td>✓</td>
<td>5.50 3.00</td>
<td>*</td>
<td>2.50 3.00</td>
<td>✓</td>
</tr>
<tr>
<td>VS</td>
<td>*</td>
<td>1.00 0.00</td>
<td>✓</td>
<td>7.00 0.00</td>
<td>*</td>
</tr>
<tr>
<td><strong>embedded</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>complex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>✓</td>
<td>7.00 0.00</td>
<td>✓</td>
<td>7.00 0.00</td>
<td>✓</td>
</tr>
<tr>
<td>V2</td>
<td>*</td>
<td>1.25 0.50</td>
<td>✓</td>
<td>7.00 0.00</td>
<td>*</td>
</tr>
<tr>
<td>VS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>simple</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>✓</td>
<td>5.25 2.87</td>
<td>✓</td>
<td>7.00 0.00</td>
<td>✓</td>
</tr>
<tr>
<td>V2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**TABLE XX: GROUP 3 ACCEPTABILITY JUDGMENTS**

<table>
<thead>
<tr>
<th>Code-switching</th>
<th>Spanish</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Acceptability</td>
<td>$\mu$</td>
</tr>
<tr>
<td><strong>Spanish T</strong></td>
<td>matrix</td>
<td></td>
</tr>
<tr>
<td>complex</td>
<td>SV</td>
<td>?</td>
</tr>
<tr>
<td></td>
<td>V2</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>VS</td>
<td>✓</td>
</tr>
<tr>
<td>simple</td>
<td>SV</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>V2</td>
<td>?</td>
</tr>
<tr>
<td></td>
<td>VS</td>
<td>✓</td>
</tr>
<tr>
<td>embedded</td>
<td>complex</td>
<td>SV</td>
</tr>
<tr>
<td></td>
<td>V2</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>VS</td>
<td>✓</td>
</tr>
<tr>
<td>simple</td>
<td>SV</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>V2</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>VS</td>
<td>✓</td>
</tr>
<tr>
<td><strong>English T</strong></td>
<td>matrix</td>
<td></td>
</tr>
<tr>
<td>complex</td>
<td>SV</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>V2</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>VS</td>
<td>✓</td>
</tr>
<tr>
<td>simple</td>
<td>SV</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>V2</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>VS</td>
<td>?</td>
</tr>
<tr>
<td>embedded</td>
<td>complex</td>
<td>SV</td>
</tr>
<tr>
<td></td>
<td>V2</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>VS</td>
<td>✓</td>
</tr>
<tr>
<td>simple</td>
<td>SV</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>V2</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>VS</td>
<td>*</td>
</tr>
<tr>
<td>Code-switching</td>
<td>Spanish</td>
<td>English</td>
</tr>
<tr>
<td>------------------------</td>
<td>--------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td></td>
<td>Acceptability</td>
<td>μ</td>
</tr>
<tr>
<td>Spanish T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>matrix</td>
<td></td>
<td></td>
</tr>
<tr>
<td>complex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>✓</td>
<td>5.38</td>
</tr>
<tr>
<td>V2</td>
<td>?</td>
<td>4.75</td>
</tr>
<tr>
<td>VS</td>
<td>✓</td>
<td>7.00</td>
</tr>
<tr>
<td>simple</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>*</td>
<td>1.00</td>
</tr>
<tr>
<td>V2</td>
<td>*</td>
<td>1.75</td>
</tr>
<tr>
<td>VS</td>
<td>*</td>
<td>2.50</td>
</tr>
<tr>
<td>embedded</td>
<td></td>
<td></td>
</tr>
<tr>
<td>complex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>✓</td>
<td>5.50</td>
</tr>
<tr>
<td>V2</td>
<td>✓</td>
<td>7.00</td>
</tr>
<tr>
<td>VS</td>
<td>✓</td>
<td>5.33</td>
</tr>
<tr>
<td>simple</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>✓</td>
<td>5.25</td>
</tr>
<tr>
<td>V2</td>
<td>*</td>
<td>2.00</td>
</tr>
<tr>
<td>VS</td>
<td>*</td>
<td>2.50</td>
</tr>
<tr>
<td>English T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>matrix</td>
<td></td>
<td></td>
</tr>
<tr>
<td>complex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>✓</td>
<td>5.00</td>
</tr>
<tr>
<td>V2</td>
<td>✓</td>
<td>6.88</td>
</tr>
<tr>
<td>VS</td>
<td>*</td>
<td>1.75</td>
</tr>
<tr>
<td>simple</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>*</td>
<td>1.00</td>
</tr>
<tr>
<td>V2</td>
<td>?</td>
<td>3.13</td>
</tr>
<tr>
<td>VS</td>
<td>?</td>
<td>3.25</td>
</tr>
<tr>
<td>embedded</td>
<td></td>
<td></td>
</tr>
<tr>
<td>complex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>✓</td>
<td>7.00</td>
</tr>
<tr>
<td>V2</td>
<td>✓</td>
<td>7.00</td>
</tr>
<tr>
<td>VS</td>
<td>*</td>
<td>2.50</td>
</tr>
<tr>
<td>simple</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>*</td>
<td>2.75</td>
</tr>
<tr>
<td>V2</td>
<td>*</td>
<td>2.00</td>
</tr>
<tr>
<td>VS</td>
<td>*</td>
<td>1.00</td>
</tr>
<tr>
<td>Code-switching</td>
<td>Spanish</td>
<td>English</td>
</tr>
<tr>
<td>----------------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td></td>
<td>Acceptability</td>
<td>$\mu$</td>
</tr>
<tr>
<td>Spanish T matrix</td>
<td></td>
<td></td>
</tr>
<tr>
<td>complex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>✓</td>
<td>7.00</td>
</tr>
<tr>
<td>V2</td>
<td>✓</td>
<td>7.00</td>
</tr>
<tr>
<td>VS</td>
<td>✓</td>
<td>7.00</td>
</tr>
<tr>
<td>simple</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>*</td>
<td>3.00</td>
</tr>
<tr>
<td>V2</td>
<td>✓</td>
<td>6.50</td>
</tr>
<tr>
<td>VS</td>
<td>✓</td>
<td>7.00</td>
</tr>
<tr>
<td>embedded</td>
<td></td>
<td></td>
</tr>
<tr>
<td>complex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>✓</td>
<td>7.00</td>
</tr>
<tr>
<td>V2</td>
<td>✓</td>
<td>7.00</td>
</tr>
<tr>
<td>VS</td>
<td>✓</td>
<td>7.00</td>
</tr>
<tr>
<td>simple</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>✓</td>
<td>6.75</td>
</tr>
<tr>
<td>V2</td>
<td>✓</td>
<td>6.75</td>
</tr>
<tr>
<td>VS</td>
<td>✓</td>
<td>7.00</td>
</tr>
<tr>
<td>English T matrix</td>
<td></td>
<td></td>
</tr>
<tr>
<td>complex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>✓</td>
<td>7.00</td>
</tr>
<tr>
<td>V2</td>
<td>✓</td>
<td>6.75</td>
</tr>
<tr>
<td>VS</td>
<td>✓</td>
<td>7.00</td>
</tr>
<tr>
<td>simple</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>*</td>
<td>1.67</td>
</tr>
<tr>
<td>V2</td>
<td>✓</td>
<td>5.50</td>
</tr>
<tr>
<td>VS</td>
<td>?</td>
<td>3.50</td>
</tr>
<tr>
<td>embedded</td>
<td></td>
<td></td>
</tr>
<tr>
<td>complex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>✓</td>
<td>7.00</td>
</tr>
<tr>
<td>V2</td>
<td>✓</td>
<td>4.00</td>
</tr>
<tr>
<td>VS</td>
<td>✓</td>
<td>4.00</td>
</tr>
<tr>
<td>simple</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>?</td>
<td>4.00</td>
</tr>
<tr>
<td>V2</td>
<td>✓</td>
<td>4.00</td>
</tr>
<tr>
<td>VS</td>
<td>*</td>
<td>3.00</td>
</tr>
<tr>
<td>Code-switching</td>
<td>Spanish</td>
<td>English</td>
</tr>
<tr>
<td>----------------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td></td>
<td>Acceptability</td>
<td>μ</td>
</tr>
<tr>
<td>Spanish T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>matrix</td>
<td></td>
<td></td>
</tr>
<tr>
<td>complex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>?</td>
<td>4.00</td>
</tr>
<tr>
<td>V2</td>
<td>✓</td>
<td>7.00</td>
</tr>
<tr>
<td>VS</td>
<td>✓</td>
<td>7.00</td>
</tr>
<tr>
<td>simple</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>*</td>
<td>1.00</td>
</tr>
<tr>
<td>V2</td>
<td>*</td>
<td>1.00</td>
</tr>
<tr>
<td>VS</td>
<td>?</td>
<td>4.00</td>
</tr>
<tr>
<td>embedded</td>
<td></td>
<td></td>
</tr>
<tr>
<td>complex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>✓</td>
<td>7.00</td>
</tr>
<tr>
<td>V2</td>
<td>*</td>
<td>1.00</td>
</tr>
<tr>
<td>VS</td>
<td>✓</td>
<td>7.00</td>
</tr>
<tr>
<td>simple</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>*</td>
<td>1.00</td>
</tr>
<tr>
<td>V2</td>
<td>*</td>
<td>1.00</td>
</tr>
<tr>
<td>VS</td>
<td>*</td>
<td>1.00</td>
</tr>
<tr>
<td>English T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>matrix</td>
<td></td>
<td></td>
</tr>
<tr>
<td>complex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>*</td>
<td>2.50</td>
</tr>
<tr>
<td>V2</td>
<td>✓</td>
<td>5.50</td>
</tr>
<tr>
<td>VS</td>
<td>*</td>
<td>1.00</td>
</tr>
<tr>
<td>simple</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>*</td>
<td>1.00</td>
</tr>
<tr>
<td>V2</td>
<td>*</td>
<td>1.00</td>
</tr>
<tr>
<td>VS</td>
<td>*</td>
<td>1.00</td>
</tr>
<tr>
<td>embedded</td>
<td></td>
<td></td>
</tr>
<tr>
<td>complex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>✓</td>
<td>7.00</td>
</tr>
<tr>
<td>V2</td>
<td>*</td>
<td>1.00</td>
</tr>
<tr>
<td>VS</td>
<td>*</td>
<td>1.00</td>
</tr>
<tr>
<td>simple</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>*</td>
<td>1.00</td>
</tr>
<tr>
<td>V2</td>
<td>*</td>
<td>1.00</td>
</tr>
<tr>
<td>VS</td>
<td>*</td>
<td>1.00</td>
</tr>
<tr>
<td>Code-switching</td>
<td>Spanish</td>
<td>English</td>
</tr>
<tr>
<td>----------------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td></td>
<td>Acceptability</td>
<td>μ</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spanish T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>matrix</td>
<td></td>
<td></td>
</tr>
<tr>
<td>complex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>?</td>
<td>4.88</td>
</tr>
<tr>
<td>V2</td>
<td>*</td>
<td>2.50</td>
</tr>
<tr>
<td>VS</td>
<td>√</td>
<td>6.75</td>
</tr>
<tr>
<td>simple</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>*</td>
<td>1.00</td>
</tr>
<tr>
<td>V2</td>
<td>*</td>
<td>2.25</td>
</tr>
<tr>
<td>VS</td>
<td>?</td>
<td>4.50</td>
</tr>
<tr>
<td>embedded</td>
<td></td>
<td></td>
</tr>
<tr>
<td>complex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>√</td>
<td>5.38</td>
</tr>
<tr>
<td>V2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VS</td>
<td>√</td>
<td>6.88</td>
</tr>
<tr>
<td>simple</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>?</td>
<td>4.25</td>
</tr>
<tr>
<td>V2</td>
<td>*</td>
<td>1.88</td>
</tr>
<tr>
<td>VS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>English T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>matrix</td>
<td></td>
<td></td>
</tr>
<tr>
<td>complex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>?</td>
<td>3.38</td>
</tr>
<tr>
<td>V2</td>
<td>√</td>
<td>6.38</td>
</tr>
<tr>
<td>VS</td>
<td>*</td>
<td>1.75</td>
</tr>
<tr>
<td>simple</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>*</td>
<td>1.00</td>
</tr>
<tr>
<td>V2</td>
<td>?</td>
<td>3.75</td>
</tr>
<tr>
<td>VS</td>
<td>?</td>
<td>3.25</td>
</tr>
<tr>
<td>embedded</td>
<td></td>
<td></td>
</tr>
<tr>
<td>complex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>√</td>
<td>7.00</td>
</tr>
<tr>
<td>V2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VS</td>
<td>*</td>
<td>1.75</td>
</tr>
<tr>
<td>simple</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>*</td>
<td>2.50</td>
</tr>
<tr>
<td>V2</td>
<td>*</td>
<td>2.33</td>
</tr>
<tr>
<td>VS</td>
<td>*</td>
<td>1.00</td>
</tr>
</tbody>
</table>
VITA

EDUCATION

Aug 2014  University of Illinois at Chicago
Ph.D., Hispanic Studies – Descriptive and Theoretical Linguistics

May 2008  University of Illinois at Chicago
M.A., Hispanic Studies – Descriptive Linguistics

May 2002  Beloit College
B.A., Majors in Spanish and Computer Science
Study abroad: Quito, Ecuador (Fall 2000 – Spring 2001)

PUBLICATIONS


PRESENTATIONS


LABORATORY EXPERIENCE

June 2012 – June 2013
Bilingualism Research Laboratory, Univ. of Illinois at Chicago
Graduate Student Director

June 2008 – May 2012
Bilingualism Research Laboratory, Univ. of Illinois at Chicago
Director of Information Technology

TEACHING - UNIVERSITY OF ILLINOIS AT CHICAGO

Aug 2006 – Present
Teaching Assistant
Instructor of record, fully responsible for teaching:
Linguistics 150 – Introduction to the Study of Language Fall 2010, Fall 2011
Spanish 101 – Elementary Spanish I Su 2010
Spanish 102 – Elementary Spanish II Spr 2011
Spanish 103 – Elementary Spanish III Su 2007
Spanish 104 – Topics in Spanish Language and Culture Su ’08, Su ’09, Su ’12, Su ’13
Spanish 110 – Elementary Spanish Review Fall 2006, Spr 2007 (2), Fall 2007
Spanish 202 – Spanish Grammar in Practice Fall 2009 (2), Spr 2013, Spr 2014

Traditional Teaching Assistant:
Spanish 206 – Introduction to Hispanic Linguistics Spr 2010
Spanish 502 – Theoretical & Research Foundations of Communicative Language Teaching Fall 2008

MENTORING OF UNDERGRADUATE RESEARCH ASSISTANTS

Fall 2012 – Spring 2013
Mirissa Kiko Burten

Fall 2012
Connie Luo

RESEARCH TOOL EXPERIENCE

Amazon’s Mechanical Turk (Online recruitment and data collection tool)
Coded and ran of a variety of experiments; integrated with Ibex

Ibex (Web-based tool for self-paced reading and acceptability judgment tasks)
Ran both SPR and AJT experiments; added custom code where necessary; integrated with Mechanical Turk
Qualtrics (Online survey tool)
Set up survey for sub-experiments over multiple days; integrated with Ibex.

Standard research tools:
SPSS, Microsoft Excel

**Administrative Experience**

<table>
<thead>
<tr>
<th>Time</th>
<th>Role</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall 2010 –</td>
<td>Course Coordinator for the Spanish Basic Language Program</td>
<td>University of Illinois at Chicago</td>
</tr>
<tr>
<td>Spring 2012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fall 2011</td>
<td>Interim Assistant Director of the Spanish Basic Language Program</td>
<td>University of Illinois at Chicago</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Summer 2010</td>
<td>Interim Director of the Spanish Basic Language Program</td>
<td>University of Illinois at Chicago</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Summer 2009</td>
<td>Interim Director of the Spanish Basic Language Program</td>
<td>University of Illinois at Chicago</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fall 2007 –</td>
<td>Course Coordinator for the Spanish Basic Language Program</td>
<td>University of Illinois at Chicago</td>
</tr>
<tr>
<td>Spring 2009</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**University Service**

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aug 2007 –</td>
<td>UIC Talks in Linguistics</td>
<td>Organizing Committee</td>
</tr>
<tr>
<td>May 2011</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aug 2011 –</td>
<td>2012 UIC Bilingualism Forum</td>
<td>Head Organizer</td>
</tr>
<tr>
<td>Dec 2012</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Awards and Honors**

<table>
<thead>
<tr>
<th>Time</th>
<th>Award</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall 2012 –</td>
<td>Audrey Lumsden-Kouvel Fellowship</td>
<td></td>
</tr>
<tr>
<td>Spring 2013</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Travel Awards**

- Graduate College Presenter’s Award          Fall 2013
- LAS Ph.D. Student Travel Award              Fall 2013
- LCSL Support for Graduate Students          Spring 2012
- Graduate College Presenter’s Award          Spring 2012
- LAS Ph.D. Student Travel Award              Fall 2011
- Graduate Student Council Travel Award       Fall 2011
- LCSL Support for Graduate Students          Fall 2010
LANGUAGES

English – Native Speaker
Spanish – Near-native proficiency
French – Intermediate high proficiency
American Sign Language – Intermediate high proficiency
German – Basic Reading Proficiency