That-trace effects in Spanish-English Code-Switching*

Shane Ebert & Bradley Hoot

Abstract

Many languages, including English, exhibit a restriction on subject extraction over complementizers called the *that*-trace effect. Although extensively studied, this phenomenon remains a puzzle. Not all languages exhibit the effect; Spanish does not. Spanish also allows postverbal subjects, while English does not, which has been linked to the *that*-trace effect. Because English and Spanish differ in these properties, combining lexical items from both languages in a single derivation, as in code-switching, offers additional insight into the nature of the restriction. Two acceptability judgment tasks of Spanish/English code-switching reveal that a single Spanish functional head is insufficient to license either post-verbal subjects or subject extraction. Instead, we argue, the *that*-trace effect and related properties arise from the interaction of two heads.

Keywords: *that*-trace, complementizers, subjects, code-switching

*Acknowledgments: We would like to thank the two anonymous reviewers and the volume's editor, Luis López, for their valuable comments, which undoubtedly improved this chapter. Similarly, we thank the members of the UIC Bilingualism Research Laboratory for helpful feedback on previous versions of this work and Jennifer Cabrelli Amaro for insightful discussions of data analysis. We are also grateful to the Spanish faculty in DePaul University's Department of Modern Language and UIC's Department of Hispanic and Italian Studies for assistance with data collection, and to all our participants. Finally, we owe a debt of gratitude to the honoree of this volume, Kay González-Vilbazo, for his all his encouragement and support over the years.

1. The problem

A variety of languages exhibit a contrast between subject and object extraction over certain complementizers, as seen in (1), a restriction known as the *that*-trace effect.

(1) ***that*-trace effect**

a. ✓ Who do you think that Sue met __?

b. ✓ Who do you think Sue met __?

c. * Who do you think that __ met Sue?

d. ✓ Who do you think __ met Sue?

Concretely, when a *wh*-object is extracted from an embedded clause, the sentence is acceptable whether the complementizer is overtly expressed, as in (1a), or null, as in (1b). In contrast, extraction of a *wh*-subject from an embedded clause is unacceptable when the complementizer is overt, as in (1c), but acceptable when the complementizer is null, as in (1d). In other words, there is a restriction on the extraction of a subject *wh*-phrase from an embedded clause when the complementizer of that embedded clause is overtly expressed.[1] The *that*-trace effect is part of a broader set of phenomena known as complementizer-trace effects, which include not just

[1] For the sake of exposition, we will continue referring to the embedded complementizer as either overt or null, but depending on the analysis, sentences without a phonologically overt complementizer may in fact have no complementizer at all.

wh-movement but Ā-extractions more generally (Bresnan, 1977, p. 178-182), and which have been the subject of considerable interest in linguistic theory.

While research on the *that*-trace effect stretches back to the 1970s, including seminal work by Perlmutter (1968, 1971), Bresnan (1972, 1977) and Chomsky and Lasnik (1977), the phenomenon remains a puzzle in many regards (see Pesetsky, 2017, for a review). Syntactic accounts of the effect assume that it can be traced to fundamental principles of the computational system. This is in part due to evidence that the effect is cross-linguistic, including evidence from French (Perlmutter, 1971), Russian (Pesetsky, 1982), Wolof (Martinović, 2017), and Nupe (Kandybowicz, 2006, p. 220-221), among others. Still, the specific mechanism continues to be debated. Some accounts appeal to linear order, proposing that extraction from a position to the right of the complementizer is banned (e.g., de Chene, 2000; Kandybowicz, 2006, 2009). Other accounts focus on hierarchical structure, instead proposing that extraction from a position immediately dominated by an overt complementizer is banned. To the degree that these latter accounts are correct, understanding comp-trace effects has important implications for key properties of syntax, such as restrictions on the locality of syntactic movement.

Crucially, while there is evidence for the *that*-trace effect in a variety of languages, some languages do not exhibit it. For example, Spanish, like a few other Romance languages, allows extraction of a *wh*-phrase over an overt complementizer whether the *wh*-phrase is an object, as in (2a) or a subject, as in (2b).

(2) **Subject extraction in Spanish**

a. ¿Qué crees **que** compró Susana?

'What do you think that Susana bought?'

b. ¿Quién crees **que** compró el libro?

'Who do you think bought the book?'

We propose to take advantage of this contrast between Spanish and English and combine it with the unique properties of code-switching (CS) to provide novel evidence regarding the *that*-trace effect. Specifically, we use CS to combine different sets of lexical items from both languages and observe the effect that each combination of syntactic properties has on the acceptability of the resulting sentence. In this way, CS data can provide insight that is not available when considering only monolingual data.

This paper is organized as follows. Section 2 provides the relevant background, including our baseline assumptions, the relevant facts about Spanish and English complementizers, and the theoretical motivation of

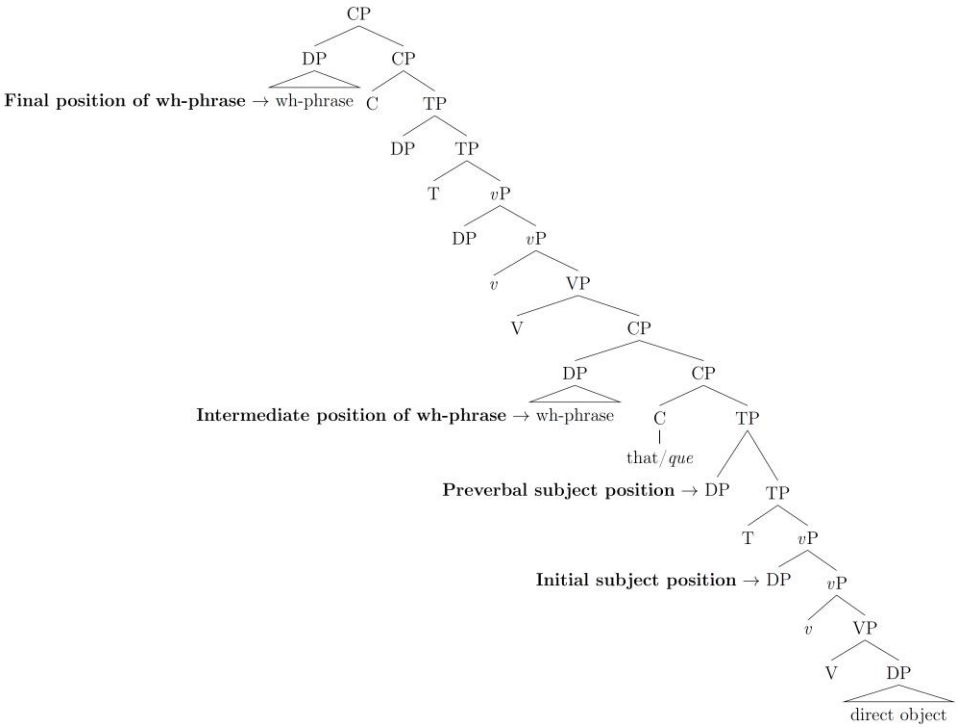
the project. Section 3 makes explicit our research questions and makes predictions based on the theoretical approaches discussed. Section 4 outlines the general methods used to test the predictions, and sections 5 through 8 present the design and results of each experiment in turn. Section 9 discusses the implications of our results.

2. Background and motivation

2.1. Assumptions

We assume a view of syntax compatible with the Minimalist Program (Chomsky, 1995 et seq.). We further assume derivation by phases (Chomsky, 2001), where the phase heads include at least the complementizer (C) and little *v*. We make no specific assumptions about the fine structure of the C-domain (Fin, Force, etc.), instead simplifying by referring to any such head as C. We assume the English word *that* and the Spanish word *que* are realizations of a C head, and thus head a phase. The complement of C is a Tense Phrase (TP), headed by T. Preverbal subjects are in Spec,TP, and English always requires preverbal subjects due to the EPP. Spanish allows postverbal subjects; when Spanish subjects are postverbal, they are in situ in Spec,*v*P, below the verbal complex in T. Example (3) illustrates our structural assumptions, including the key positions of subjects and the landing sites of moved *wh*-phrases.

(3) **Structural assumptions**



With regard to CS, we take a "no-third-grammar" approach, under which nothing governs the possible combinations of the two languages beyond the grammars of the two contributing languages. This perspective, first laid out in Woolford (1983) and updated to a Minimalist framework in MacSwan (1999), contrasts with approaches such as those of Myers-Scotton (1993) and Jake et. al. (2002) which propose CS-specific principles for governing CS, sometimes known as a third-grammar approach. By adopting a no-third-grammar approach, patterns of acceptability can be taken to reflect the interaction of language-specific properties within one underlying computational system. Under this view,

the acceptability of a given sentence depends on the features of the lexical items in the derivation, regardless of the language from which they are taken. It is because of this that combining two languages which differ on the relevant syntactic property can provide new insights. In the present case, contrasting the acceptability of sentences with lexical items from Spanish and English, which differ with regard to the *that*-trace effect, can inform our understanding of this phenomenon (for further discussion of the utility of CS for linguistic theory, see González-Vilbazo et al., 2013; González-Vilbazo & López, 2011, 2012).

2.2. Contrasts between Spanish and English

Complementizers in English and Spanish contrast in two relevant ways. First, as previously mentioned, only Spanish allows extraction of a subject *wh*-phrase over an overt complementizer (4a)[2], while English has the *that*-trace effect: subjects cannot be extracted over *that* (5c) but can be extracted when the complementizer is null (5d). Second, Spanish always requires an overt complementizer—compare (4a) to (4b) and (4c) to (4d)—while only English allows null complementizers, as seen in (5b) and (5d), repeated from example (1) for the sake of exposition.

(4) **Spanish: Obligatory complementizers**

[2] Note that the subject in (4a) is postverbal, unlike in English. See footnote 3 for some discussion.

a. ✓ ¿A quién crees **que** conoció Susana __?

‘Who do you think that Susana met?’

b. * ¿A quien crees **Ø** conoció Susana __?

c. ✓ ¿Quién crees **que** conoció __ a Susana?

‘Who do you think met Susana?’

d. * ¿Quién crees **Ø** conoció __ a Susana?

(5) **English: *that*-trace effect**

a. ✓ Who do you think **that** Sue met __?

b. ✓ Who do you think **Ø** Sue met __?

c. * Who do you think **that** __ met Sue?

d. ✓ Who do you think **Ø** __ met Sue?

In addition to differing with regard to subject extraction, subjects in Spanish and English differ in another way: in general, only Spanish subjects can be postverbal. In fact, there appears to be a correlation between the availability of postverbal subjects and the availability of subject extraction. As originally noted by Rizzi (1982), languages such as Spanish and Italian, which allow subjects to remain postverbal even in declarative sentences, also do not demonstrate the *that*-trace effect.

Rizzi (1982) provides direct evidence that in one of these Romance languages, Italian, extraction of a subject *wh*-phrase from an embedded clause does in fact take place from a postverbal position. In addition,

Menuzzi (2000) provides evidence that *wh*-subject extraction in another Romance language without the *that*-trace effect, Brazilian Portuguese, also takes place from a postverbal position. Importantly, this is in spite of the fact that postverbal subjects in Brazilian Portuguese are otherwise less free than in Spanish or Italian (Chao, 1981). It is not the general availability of inversion but the specific availability of postverbal subjects with *wh*-extraction which obviates the *that*-trace effect. Given the evidence from other Romance languages that successful subject extraction over a complementizer takes place from a postverbal position, it is plausible to assume that the same will hold true for Spanish.[3]

While it is relatively uncontroversial that English subjects are in Spec,TP in the relevant structures, we are not aware of any direct evidence regarding the position from which *wh*-subjects are extracted in Spanish. However, with regard to matrix questions, evidence based on adverb position (Suñer, 1994), sentence negation (Suñer, 1994), sluicing (Suñer, 1994), the interpretation of adverbs (Buesa-García, 2008) and the

[3] An anonymous reviewer points out that it is somewhat curious that the literature has focused to a significant degree on subject position when another obvious distinction between languages like English and languages like Spanish or Italian is that complementizers are optional in the former but obligatory in the latter, as we noted. This correlation is not absolute, though: French has obligatory complementizers but shows the *that*-trace effect. However, French also allows subject extraction over a special form of the complementizer (*qui* instead of *que*), which further points to the complementizer, not the subject, as the locus of the relevant restriction. Because the previous literature has focused on subject position, we follow that tack as well; however, we recognize that this question is not settled and that alternatives should be considered. If it is the case that the *that*-trace effect ultimately comes down to a property of C, we contend that research like the present study is that much more useful, because it allows us to isolate the properties of C, although a complete investigation of this issue is beyond the scope of this chapter.

historical origins of V2 phenomenon (Goodall, 1993) suggests that postverbal subjects remain in situ, in Spec,*v*P. Given that fact, along with the wide range of contexts in which postverbal subjects are available in Spanish, it seems likely that they would also be available in *wh*-subject extraction, and Suñer (1994) claims that VS order is optionally available in embedded questions.[4]

Given this link, if the results of this study find that participants accept postverbal subjects in *wh*-questions in Spanish while also having structures in which there is no *that*-trace effect, this would suggest that this assumption is correct. Thus, determining the available subject positions in each language also plays an important role in understanding the degree to which each language evidences a *that*-trace effect.

[4] There is some disagreement in the literature regarding when verb-subject (VS) word order (i.e., postverbal subjects) is optional or required, though that may be due to dialect differences. In long-distance extraction through multiple clauses, Torrego (1984) claims that VS word order is required, so (a), with VS word order in each clause, is acceptable but (b) with VS word order only in the matrix clause is unacceptable:

Multiclausal long-distance extraction in Spanish

a. Un viaje a las Canarias dice *Juan* que quería *la gente* que hiciera *Antonio* este verano.
b. *Un viaje a las Canarias dice *Juan* que *la gente* quería que *Antonio* hiciera este verano.
c. Un viaje a las Canarias dice *Juan* que quería *la gente* que *Antonio* hiciera este verano.

Importantly, though, the sentence becomes acceptable if all higher clauses have postverbal subjects, even if the most embedded one does not, as seen in (c). Suñer (1994) makes similar claims regarding the optionality of VS word order in the most embedded clause. The present paper focuses on extraction from one clause, so the most embedded clause will also be the only embedded clause, predicting that both SV and VS word order should be acceptable.

With these basic contrasts in mind, we turn now to the theoretical accounts of the *that*-trace effect and of issues related to C and T which inform the design of the present study.

2.3. Theoretical background

There are two principal approaches to accounting for the *that*-trace effect (see Pesetsky, 2017, for an overview). According to one class of approaches, the restriction on subject extraction over an overt complementizer is based on linear word order. Early accounts of this type include Bresnan's (1972) Fixed Subject Condition and Chomsky and Lasnik's (1977) That-trace filter, though both accounts were already largely abandoned by the 1980's (Pesetsky, 2017). A more recent linear approach, argued for in particular by de Chene (2000) and Kandybowicz (2006), is based on the idea that the complementizer-trace effects, including the *that*-trace effect, are due to a restriction not only on word order but also the prosodic boundaries that arise from those word orders. These analyses are promising, successfully accounting for some *that*-trace data[5] that other approaches cannot, but they remain somewhat speculative given that a clear consensus on the structure of the prosody/syntax

[5] In particular, these approaches address the observation by Bresnan (1977, 194 footnote 6) that placing an adverbial expression between the complementizer and the verb noticeably ameliorates the *that*-trace effect. If the *that*-trace effect is a purely syntactic phenomenon, it is not immediately clear why an intervening adverbial expression would affect the underlying syntactic restriction.

interface and the nature of its rules has not yet emerged, especially with regard to code-switching.

The second class of approaches, which represent the majority of the work on the *that*-trace effect, take a structure-based approach, where the restriction on extraction of the subject over an overt complementizer stems from syntactic principles. This is the family of accounts that is most relevant to the present paper, as the CS evidence and the theoretical framework we use to interpret that evidence are particularly well-suited for syntactic analysis. We focus here on two types of accounts: those which emphasize the role of C and those which stress the role of T.

Within the first type, Pesetsky and Torrego (2001) combine a feature-based theory of movement and economy conditions on movement to account for complementizer-trace effects, including the *that*-trace effect. Their proposal stems from an observation by Koopman (1983) that English T-to-C movement is similar to complementizer-trace effects, essentially constituting a *do*-trace effect:

(6) ***do-trace* effect in matrix questions**

a. ✓ What did Mary buy __?

b. * What Mary bought __?

c. * Who did __ buy the book? [*unless *did* is emphatic]

d. ✓ Who __ bought the book?

Pesetsky and Torrego suggest that the same underlying principle might connect the *do*-trace effect and the *that*-trace effect, explaining the parallels in the restrictions on T-to-C movement and subject extraction. They propose that the putative complementizer *that* is really an invariant allomorph of a tense or aspect element raised to C, similar to *do* in interrogative clauses. English declarative C, then, is always a null morpheme. Elements that were previously thought to instantiate C, including *that* and *for*, have actually moved from T to C. This movement is triggered by two probes on C: one that triggers Ā-movement (e.g., a *wh*-probe) and one that triggers T-to-C movement (a tense probe). Crucially, nominals, including subject *wh*-phrases, also bear tense features that can satisfy a tense probe. Since a preverbal subject with the relevant Ā-features, such as a *wh*-phrase, is capable of satisfying the needs of both probes at once, it is required to do so for economy reasons, preventing T-to-C movement and therefore the overt *that*. In this way, economy requirements and the properties of nominals create the *that*-trace effect. This same economy condition prevents *do*-support with subject *wh*-phrases, explaining the *do*-trace effect in matrix *wh*-questions.

For Pesetsky and Torrego, then, the *that*-trace effect is the result of unvalued features on C and how they are satisfied by features on the external argument or on T. The difference between Spanish and English

ultimately stems from C itself: English C is a null morpheme that is phonologically realized as *that* when T is raised to it but not otherwise, while Spanish C is a true complementizer, always realized as *que*.[6]

The idea that the difference between Spanish and English could ultimately come down to a difference on C is promising because C is a phase head, and phase heads have proven to be important for a number of syntactic processes. However, C's role as a phase head is rarely a central part of proposals aiming at accounting for the *that*-trace effect. There are proposals in other arenas, though, which argue that phase heads should be understood to play a central role in determining the grammatical properties of their complements, so perhaps by appealing to such a proposal the *that*-trace effect could be unified with other phenomena.

In an analysis based mostly on evidence from Spanish-German CS, González-Vilbazo and López (2012) focus on the little *v* phase head, which, they argue, determines several properties of the selected VP, including word order. For example, since Spanish requires verb-object (VO) word order in most contexts while German requires object-verb (OV) word order, when little *v* is in Spanish, VO word order is required, and when little *v* is in German, OV word order is required, regardless of

[6] Spanish does not receive a full treatment by Pesetsky and Torrego, although they do suggest a possible analysis. However, some details of their analysis rest on assumptions regarding Spanish that differ from ours, so for reasons of space we go into less detail about their proposal for Spanish. While it is plausible to adapt their analysis to our assumptions, the important point is that the difference between the two languages ultimately comes down to the features of C.

the language of the verb or the object. These facts lead González-Vilbazo and López to propose the following hypothesis:

(7) **The phase head hypothesis (PHH)**

The phase head determines grammatical properties of its complement.

This proposed relationship between the phase head and its complement also extends to other phase heads, including C. In a subsequent pilot study, González-Vilbazo and López (2013) found that C does appear to determine properties of TP in Spanish-German CS, including some aspects of word order. It is important to note that, although this hypothesis was proposed largely based on CS data, it is not a 'third-grammar' constraint on code-switching, but rather a proposal about the role of phase heads in the syntax more broadly, one which just happens to be easiest to observe when combining lexical items from languages with different properties.

Importantly for our purposes, this hypothesis makes predictions regarding the *that*-trace effect. If the phase head C determines properties of its complement, including properties related to word order, this fact may play a role in determining whether subject extraction is possible or not. Although the PHH does not address the question of why it should be

the case that extraction from subject position should be barred, it does coincide with Pesetsky and Torrego's approach in that it locates cross-linguistic variation on C: the features of C will determine whether postverbal subjects are possible or not, and therefore will determine whether subjects can be extracted or not.

Instead of putting the emphasis on C, another plausible tack to take would be to analyze the difference between Spanish and English as ultimately stemming from T. If the *that*-trace effect is a ban on extracting from subject position—i.e., from Spec,TP—then it is logical to consider that maybe it is the head that requires subjects to raise in the first place that is the issue. Rizzi and Shlonsky (2007) take such an approach, accounting for the *that*-trace effect by adopting two ideas from Rizzi (2006):

(8) **Rizzi (2006)**

a. An element moved to a position dedicated to some scope-discourse interpretive property, a criterial position, is frozen in place (Criterial Freezing).

b. Classical EPP, the requirement that clauses have subjects, can be restated as a criterial requirement, the Subject Criterion.

Central to these ideas is the concept of a *criterial position*, a landing site for syntactic movement which indicates a particular scope-discourse interpretive property for the moved phrase. For example, the scope of a question is indicated by moving a wh-phrase to Spec,CP of the relevant clause. That is, the scope of a question must be marked by moving a wh-phrase to a criterial position, which Rizzi refers to as the Q-Criterion. Along similar lines, Rizzi proposes that there is a Subject Criterion which requires a DP to occupy Spec,TP. The DP occupying this criterial position is interpreted as the subject. In other words, there is a particular position which, when occupied, indicates a particular interpretation, in this case the subject. As (8b) suggests, this is a restatement of the classical EPP. Criterial positions have the additional property that the phrase which occupies it cannot move further, which is known as Criterial Freezing. For example, while focus movement can move an embedded phrase to the main clause, as in (9) , a *wh*-phrase satisfying the Q-Criterion cannot undergo focus movement, as seen in (10) (Rizzi & Shlonsky, 2007).

(9) a. Pensavo che avessero scelto la RAGAZZA, non il ragazzo.
'I thought they had chosen the GIRL, not the boy.'

b. La RAGAZZA pensavo che avessero scelto, non il ragazzo.
'The GIRL I thought they had chosen ___, not the boy.'

(10) a. Mi domandavo quale RAGAZZA avessero scelto, non quale ragazzo.
'I wondered which GIRL they had chosen, not which boy.'

b. * Quale RAGAZZA mi domandavo avessero scelto, non quale ragazzo.
'Which GIRL I wondered they had chosen, not which boy.'

Rizzi and Shlonsky (2007) suggest that subject extraction is disallowed due to this process of Criterial Freezing. More concretely, subject extraction is disallowed because subjects, including *wh*-subjects, are required to raise to Spec,TP due to the Subject Criterion (8b) but they can raise no further because Spec,TP is a criterial position and subjects are frozen in place due to Criterial Freezing (8a).[7] Object extraction, however, is unaffected by Criterial Freezing because objects do not raise from a criterial position. Criterial Freezing applies whether or not the complementizer is null, so Rizzi and Shlonsky must make some additional assumptions regarding complementizers to account for the *that*-trace effect, appealing to a split CP and some independent facts about complementizers in order to explain why extraction from subject position is possible when C is null. With regard to the present paper, though, the

[7] Following Rizzi (2006), Rizzi and Shlonsky (2007) introduce an additional functional head specifically for subjects, but for the sake of simplicity, we will use T and raise subjects to Spec,TP. This does not change the aspects of the account relevant to the current paper.

relevant aspect of their proposal is that the difference between subjects and objects with regard to extraction is that subject position is a criterial position, and so subjects are frozen in Spec,TP. Instead of being triggered by features on C, for Rizzi and Shlonsky the *that*-trace effect is ultimately the result of the Subject Criterion, which is a restatement of the classical EPP, that is, a feature of T.

3. Research Questions and Predictions

To summarize so far, Spanish and English differ in two ways: Spanish requires overt C while English has the *that*-trace effect, and Spanish allows postverbal subjects, which are correlated with allowing subject extraction. In order to account for these facts, we examined proposals that emphasized either the role of C or the role of T. Pesetsky and Torrego coincide with González-Vilbazo and López in proposing that the relevant differences between Spanish and English are due to the features of C. They differ with regard to the role of subject position. For Pesetsky and Torrego, subject position does not play a crucial role in whether a subject can be extracted; only the question of whether the unvalued features on C are valued is important. For González-Vilbazo and López, on the other hand, C determines word order in its complement, which we have interpreted to mean that Spanish C will license postverbal subjects and will therefore license subject extraction. Finally, for Rizzi and Shlonsky, the most important factor is the Subject Criterion, which freezes

subjects in Spec,TP: English T requires subjects to raise, where they freeze, while Spanish T allows subjects to stay in situ in Spec,*v*, which is not a criterial position, and thus they can be extracted.

In light of this discussion, the goal of this project is to discover which head determines whether subject extraction over a complementizer is possible, but because this property appear to be closely tied to subject position, we must first address another, preliminary question:

(11) **Research Question 1**

What determines subject position in Spanish-English CS?

We consider two possible hypotheses. González-Vilbazo and López's (2012, 2013) Phase Head Hypothesis predicts that the phase head, in this case C, determines the word order in its complement. They would thus predict postverbal subjects will be available when C is Spanish but not when C is English. This view is also compatible with Pesetsky and Torrego's (2001) view that features of C manipulate the position of T and the external argument (although it is not part of their proposal). On the other hand, it is commonly assumed that the position of the subject is determined by properties of T, not only in *wh*-questions (e.g., Barbosa, 2001; Goodall, 2001; Gutiérrez-Bravo, 2002, 2008) but also more generally (specifically, the EPP, as in, e.g., Chomsky, 1981). This predicts

that postverbal subjects will be available when T is Spanish but not when T is English. This view is also compatible with Rizzi and Shlonsky's (2007) analysis of the *that*-trace effect due to Criterial Freezing, in which the Subject Criterion is a restatement of the EPP. In order to test these predictions, we test the extraction of objects from embedded questions in CS sentences in which the language of C and of T is varied. We test each combination of C and T with both preverbal and postverbal subjects, as in (12)-(15), in order to determine which functional head's presence correlates with acceptability of postverbal subjects.

(12) **English C, Spanish T, Preverbal Subject**

What did the teachers assume that *el niño había leído antes del examen*?

(13) **English C, Spanish T, Postverbal Subject**

What did the teachers assume that *había leído el niño antes del examen*?

(14) **Spanish C, English T, Preverbal Subject**

Qué asumieron los maestros que the child had read before the test?

(15) **Spanish C, English T, Postverbal Subject**

Qué asumieron los maestros que had read the child before the test?

Once we establish the subject position facts, we then return to the primary goal of the project with our second research question:

(16) **Research Question 2**

What determines whether subject extraction over a complementizer is possible?

Given the previous evidence for a correlation between subject extraction and postverbal subjects, we predict that the functional head that determines subject position will also determine whether or not subject extraction is possible. That is, if C determines subject position, then the cases in which C allows postverbal subjects will also be the cases in which subject extraction is possible over C; in other words, subject extraction will be possible over Spanish C but not English C. If it is T that determines subject position, then those cases in which T allows postverbal subjects will also allow subject extraction; in other words, subject extraction will be possible even over English C in the presence of Spanish T, but it will not be possible with English T. In order to test these predictions, we test the extraction of subjects from embedded questions over English C and Spanish C with both English and Spanish T in the embedded clause, as in (17)-(20).

(17) **English C, Spanish T**

Who did the teachers assume that *había leído el texto antes del examen*?

(18) **English C, English T**

Quién asumieron los maestros that had read the text before the test?

(19) **Spanish C, Spanish T**

Who did the teachers assume *que había leído el texto antes del examen*?

(20) **Spanish C, English T**

Quién asumieron los maestros que had read the text before the test?

For each of these research questions, we also first test participants' monolingual Spanish and English, in order to be sure that these participants' grammars include functional heads with the expected properties, as recommended by González-Vilbazo et al. (2013).

4. Methods

There were four experiments. There are two main research questions, and for each of these there were two experiments—one with monolingual stimuli and one with CS stimuli. Because our predictions for subject extraction depend on the availability of postverbal subjects, we consider first the question of subject position and then the question of

subject extraction over a complementizer. Thus the first experiment tests subject position in embedded questions with object extraction in monolingual sentences, the second tests the same phenomenon in CS sentences, the third tests subject extraction from embedded questions in monolingual sentences, and, finally, the fourth tests subject extraction in CS.

All the experiments were acceptability judgment tasks (AJTs). AJTs are a valuable tool for investigating code-switching because they allow researchers to isolate particular syntactic features of a sentence by controlling the language (and thus the features) of each lexical item (see Schütze & Sprouse, 2013 for a discussion of the value of AJTs for linguistics; see González-Vilbazo et al., 2013 for a discussion of AJTs in CS research). Because all the experiments had the same format and participants, this section describes the general design and procedure, while the following sections (5 through 8) present the details of each experiment in turn, including its design, results, and implications.

4.1. Participants

The participants were simultaneous or early sequential Spanish-English bilinguals, who acquired both English and Spanish before the age of 7. Participants were either born and raised in the United States or arrived before age 7, and they were educated in the United States. All reported at least one parent or caregiver who spoke Spanish. Participants

who had significant exposure to a language other than Spanish or English were excluded, as were participants who reported no Spanish-speaking family members or whose scores on the proficiency test indicated very low Spanish proficiency.

The proficiency test was an adaptation of part of the *Diploma del español como lengua extranjera* (DELE), which is commonly used in bilingualism research of all types (e.g., Montrul, 2002). It consists of a 30-question multiple choice test and a 20-question cloze test. The maximum score is 50, and participants who scored below 24, which indicates "low" proficiency, were excluded.

Because all participants were raised and educated in the U.S., it was expected that English would be the dominant language for these speakers, as is generally the case for heritage speakers of Spanish in the U.S. And indeed, their self-reports bear this out. As such, and for the sake of time, no direct English proficiency measure was included, as it was expected that most would score at or near ceiling.

Of the 27 participants who met the criteria for inclusion, 11 were determined to be 'non-cooperative' (Juzek & Häussler, 2015); that is, it was clear they did not complete the task appropriately. Following Juzek (2016), who based his calculations on Bader and Häussler's (2010) estimate of 225ms per word plus 25ms per character, we estimated that the expected reading time for our shortest sentence was 1600ms. Additionally,

we assumed, based on Bader and Häussler's (2010) results that a minimum of 500ms was needed to make a judgment, giving an expected minimum time per trial of 2100ms. Following Juzek (2016), we took half this time as an absolute minimum cutoff point for responses. Any participant who judged more than 20% of their sentences faster than 1000ms was excluded, on the assumption that these participants were merely 'clicking through' the stimuli without registering real judgments. Indeed, it was clear that most participants spent significantly more time, as the median reaction time per item of the remaining participants was around 7 seconds. Excluding the non-cooperative participants left a final sample of 16 participants. Relevant characteristics of these participants are presented in Table 1.

Table 1. Participant Characteristics

Age range	18-40
Mean age	22.4
Mean age of acquisition of Spanish	Birth
Mean age of acquisition of English	4.25
Overall self-reported Spanish proficiency (1-5, 1=low, 5=high)	4.5
Overall self-reported English proficiency (1-5, 1=low, 5=high)	4.8
Mean Spanish proficiency test score (max. 50)	40.9

4.2. Procedure

The experiment was conducted via the Internet using Ibex 0.3.8 (Drummond, 2007) and hosted at http://spellout.net/ibexfarm/. (On the validity of internet-based research, see Sprouse (2011).) The central task of the experiment was an AJT with a seven-point Likert scale, with the endpoints of the scale marked "good" and "bad". There were four main sections of the experiment, and participants had the opportunity to take a break between each one. The whole experiment lasted slightly more than an hour on average.

First, participants read instructions and practiced the task. They read an explanation of the task written in Spanish-English CS, following the advice of González-Vilbazo et al. (2013), who argue that this helps put participants in bilingual mode (Grosjean, 1985 et seq.) and to establish the acceptability of CS in general. They then completed a training session that included anchoring items in monolingual English, monolingual Spanish, and CS to establish the range of the scale (Schütze & Sprouse, 2013), and finally they completed 10 announced practice items in CS to get used to using the scale to record judgments.

Second, they completed a single block of 96 CS sentences (critical stimuli for the two CS experiments, plus fillers). Sentences were pseudo-randomized for each trial, such that target stimuli were always separated by two fillers, but the order in which stimuli were presented varied for each participant. Tokens were presented one by one: after participants

recorded their judgment by clicking on a number in the scale or by pressing the number on their keyboard, the next sentence appeared.

Third, participants completed the monolingual judgments in two blocks: one of 44 English sentences and one of 40 Spanish sentences. As with the CS block, all sentences were pseudo-randomized so that target stimuli were always separated by fillers but the order was different for each participant, and tokens were presented one by one. For each of the eight lists, the order was rotated, so half the participants saw the English block first and half saw the Spanish block first. Regardless of which order a participant did them in, the Spanish block was always preceded by the Spanish proficiency test.

Finally, participants completed the background questionnaire.

4.3. Design and materials

All the tasks shared a common format and were controlled in the same manner for several potential confounds. All the target tokens consisted of questions in which the *wh*-word was extracted from within an embedded clause. The matrix clause was headed by a verb that takes a clausal complement and were unlikely to take a human direct object (following Ritchart, Goodall, & Garellek, 2016), such as *think, assume, insist, admit, explain, claim,* etc., and its subject was always definite, human, and plural. The embedded clause consisted of a transitive verb, and either the subject or object was extracted. To control for frequency

effects, the verb and all arguments were taken from the 5,000 most common Spanish words (Davies, 2006) or the 5,000 most common English words (Davies, 2008). In the embedded clause, subjects were always animate, human, and singular; objects were always inanimate and singular. Both were always definite, and the verbs were chosen such that only the intended interpretation (human subject, inanimate object) was plausible. Because we assume that an overt tensed auxiliary verb is a realization of T and because establishing the language of T unambiguously was important to the experiment design, the embedded verbs were always presented in the pluperfect, consisting of a past tense auxiliary plus a past participle (e.g., *había ganado* / *had won*). The pluperfect was chosen over the present perfect because it sounds more natural with past tense main clause verbs, which in turn sound more natural in information-seeking questions than present tense verbs. Because in Spanish both the verbs and the *wh*-word *quién/quiénes* 'who' are marked for number, the fact that the matrix subject was plural and the embedded subject was singular avoided problems of ambiguous reference. Every sentence ended with an adverbial or PP adjunct that modified the TP/VP (rather than an argument).

Thirty-two lexicalizations were created for the CS experiments, distributed across eight lists. Each list contained four tokens in each cell of each experiment—a total of 32 tokens—distributed with a Latin square so that no lexicalization repeated in any list. Participants saw each of the

eight structure types four times but never saw more than one sentence from a given lexicalization (following Cowart, 1997).

In addition to the target stimuli, 64 fillers were included in each list (in total, participants judged 96 CS sentences). Fillers had different lexicalizations than the target stimuli, but all were CS. Some fillers were adjunct questions with *how* or *why* (e.g., *Why did the critic conclude that the artist didn't draw well?*) and some were embedded questions (e.g., *Due to their symptoms, the doctor wonders what the patients have eaten.*), so that all the stimuli were somewhat similar, to obscure the purpose of the experiment. As recommended by Cowart (1997), fillers ran the full range of acceptability, including CS stimuli that are known to be clearly unacceptable (e.g., switches between negation and the verb) and those known to be clearly acceptable (e.g., switches between an adjective and noun).

As in the CS experiments, monolingual tokens were distributed across eight lists. For the monolingual experiments, though, the English stimuli and Spanish stimuli were presented separately, although they shared the same 36 lexicalizations, with English and Spanish translation-equivalent versions of each. Each list contained four tokens in each cell of each experiment, distributed with a pseudo-Latin square so that no lexicalization repeated in any list within a given language (because of the uneven number of conditions, it was not possible to do a true Latin

square). Participants saw each of the five English structure types four times (total 20 sentences) and each of the four Spanish structure types four times (total 16 sentences) but within each language never saw more than one sentence from a given lexicalization. Each monolingual list also contained 24 fillers, created in the same way as for the CS experiment. In total, participants judged 44 sentences in English and 40 sentences in Spanish.

In addition to the main experimental tasks, participants completed a training before beginning the experiment (see section 4.2), as well as a Spanish proficiency test (described in section 4.1) and brief background questionnaire modified from one used in previous studies (Ebert, 2014; Koronkiewicz, 2014) that included information about language history, use, and attitudes.

4.4. Data pre-processing and statistical analysis

Prior to analysis, all responses were examined visually to check for outliers (defined as judgments more than 2 standard deviations below or above the mean judgment for that sentence type) and other potential problems (such as inversion of the scale, missing responses, or participants who rated every stimulus the same); none were found. Additionally, reaction times for all judgments were examined, and the same cutoff of 1000ms that was used to exclude non-cooperative participants was used to exclude judgments recorded in error (such as accidental button presses or

mistaken double-clicking). Of the 1088 judgments, 7 individual judgments were excluded as errors using this criterion, leaving a total of 1081 judgments for analysis.

Each of these judgments was then z-score transformed, as suggested by Schütze and Sprouse (2013), which allows each score to be expressed on a standardized scale. (For the sake of completeness, the raw ratings and z-scores for the full data set are given in Appendix A, but we will otherwise report only the z-scores.) A given participant's overall mean rating is calculated, and then each judgment is transformed so that it is expressed as a number of standard deviations from that mean. This transformation reduces scale bias and can help alleviate scale compression (an issue especially relevant for CS research, where overall ratings can be compressed at the low end of the scale, particularly by those who have negative attitudes toward CS; see Badiola, Delgado, Sande & Stefanich, 2016), while maintaining the relationships within the data. In the present experiment, because we expected that a given participant's mean rating may differ among CS, monolingual English, and monolingual Spanish stimuli, each participant's z-score transformation was separately calculated for each of the three language types.

After pre-processing, a linear mixed model was fit to the data for each of the experiments. Each experiment received a separate analysis due to the fact that each concerned different independent variables. The

specifics of each model are described below, but in each case, the fixed effects in the model were decided based on the experiment design, while the random effects specification and their covariance structure were chosen by comparing measures of goodness of fit, as advised by Eddington (2015). Whenever possible, repeated measures were modeled by including a random intercept and slope by participant and variation among items was modeled by including a random intercept by lexicalization.[8] The random effects structures we report in the following sections are those that provided the best fit for each model.

5. Experiment 1: Monolingual Object Extraction

5.1. Experiment 1 Design

The monolingual object extraction stimuli tested the availability of postverbal subjects in participants' English and Spanish, because postverbal subjects have been found to be related to the *that*-trace effect. It used a 2 x 2 design. The first factor was Language (Spanish or English)

[8] Given that the purpose of including by-item random effects in the model is to abstract away from the particular words used to instantiate the syntactic structures of interest, we chose to model this variation by lexicalization (sets of words, used across conditions) rather than at the level of individual items (a particular combination of words and experimental condition) because it reflects the fact that the words used in a given sentence were not independent of the words used for the same lexicalization in a different condition, and also because in principle it would allow us to model, via the random slope, the possibility that lexicalizations may also differ in how they vary across conditions, due to the fact that a number of lexical and other factors are involved in the acceptability of overt complementizers in English. However, in most cases, including a random effect by lexicalization did not converge, and in the two cases in which it did converge, only a random intercept was included in the final model because including a random slope decreased the model's fit.

and the second was the Subject Position (preverbal or postverbal). The English and Spanish sentences were translation equivalents. The four types and example sentences are presented in Table 2.

Table 2. Monolingual Object Extraction

Language	Subject	Example
Eng	Pre-V	What did your parents think that your sister had broken in the living room?
Eng	Post-V	What did your parents think that had broken your sister in the living room?
Spa	Pre-V	*¿Qué pensaron tus padres que tu hermana había roto en la sala de estar?*
Spa	Post-V	*¿Qué pensaron tus padres que había roto tu hermana en la sala de estar?*

5.2. Experiment 1 Results

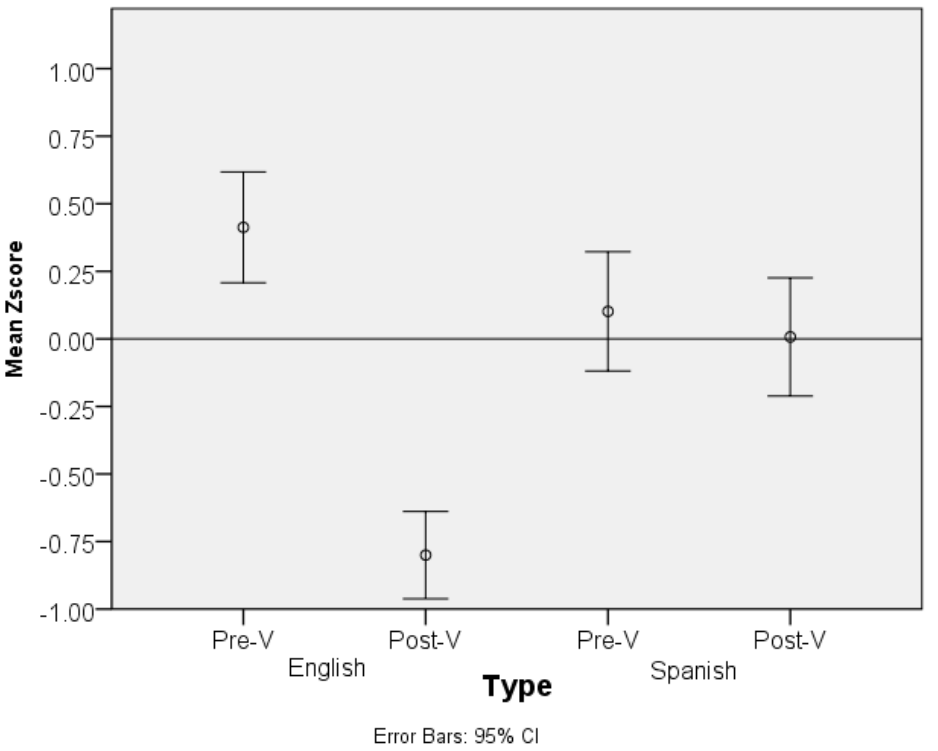
For this experiment, a linear mixed model was fit with two fixed factors—Language and Subject Position—along with a random intercept for Participant and for Lexicalization (including random slopes resulted in models with worse fit). A Type III test of fixed effects revealed a significant effect for Language ($F_{1,225.4} = 6.12$, $p = .014$), a significant effect for Subject Position ($F_{1,225.9} = 42.38$, $p < .001$), and a significant interaction ($F_{1,221.0} = 35.30$, $p < .001$). Estimated marginal mean scores for each condition are presented in Table 3 and Figure 1.

Table 3. Monolingual Object Extraction Results

Lang	Subject	Example	Mean Z-Score	SE
Eng	Pre-V	What did your parents think that your sister had broken in the living room?	.426	.110

Eng	Post-V	What did your parents think that had broken your sister in the living room?	-.792	.109
Spa	Pre-V	*¿Qué pensaron tus padres que tu hermana había roto en la sala de estar?*	.091	.109
Spa	Post-V	*¿Qué pensaron tus padres que había roto tu hermana en la sala de estar?*	.031	.110

Figure 1. Monolingual Object Extraction Results



The significant main effect of Language indicates that, overall, participants rated Spanish sentences higher than English sentences, and the significant main effect of Subject Position indicates that, overall, preverbal subjects are rated higher than postverbal subjects. However, the result of interest is the interaction between the two, which indicates that the acceptability of the two subject positions is different in each language. To

concretely examine this difference, pairwise comparisons were conducted, with the Bonferroni correction for multiple comparisons, which revealed that in English, preverbal subjects are significantly better than postverbal subjects ($p < .001$), while in Spanish, no difference was found between the two subject positions ($p = .665$).

5.3. Experiment 1 Discussion

Before being able to address what determines subject position in CS, it was necessary to verify that subjects behave as expected in the monolingual grammars of our bilingual participants, especially in light of the fact that cross-linguistic influence can affect many aspects of a bilingual's grammar, including subject position (Cuza, 2012). Our data indicate that these participants do indeed possess grammars that have the same features expected for monolinguals. Concretely, their English allows only preverbal subjects, while their Spanish allows both preverbal and postverbal subjects.

Finding that these participants accept both preverbal and postverbal subjects in monolingual Spanish sentences means that they accept preverbal subjects in embedded clauses in both English and Spanish. Importantly, there is nonetheless a difference in the feature specification of their Spanish and English lexical items. Crucially only Spanish licenses postverbal subjects, so it must be some Spanish lexical item or items that license that subject position, as expected.

6. Experiment 2: CS Object Extraction

6.1. Experiment 2 Design

After establishing that subjects do indeed behave as expected in participants' Spanish and English, the CS object extraction experiment attempted to isolate which functional head determines the availability of postverbal subjects. These sentences always had an overt C and a switch site between C and T, and this experiment also had a 2 x 2 design. The first factor was the Language of C/T (English C with Spanish T vs. Spanish C with English T) and the second factor was Subject Position (preverbal or postverbal). The CS experiment did not use the same lexicalizations as the monolingual experiment. The four types and example sentences are presented in Table 4.

Table 4. CS Object Extraction

C	T	Subject	Example
Eng	Spa	Pre-V	What did the teachers assume that *el niño había leído antes del examen*?
Eng	Spa	Post-V	What did the teachers assume that *había leído el niño antes del examen*?
Spa	Eng	Pre-V	*Qué asumieron los maestros que* the child had read before the test?
Spa	Eng	Post-V	*Qué asumieron los maestros que* had read the child before the test?

6.2. Experiment 2 Results

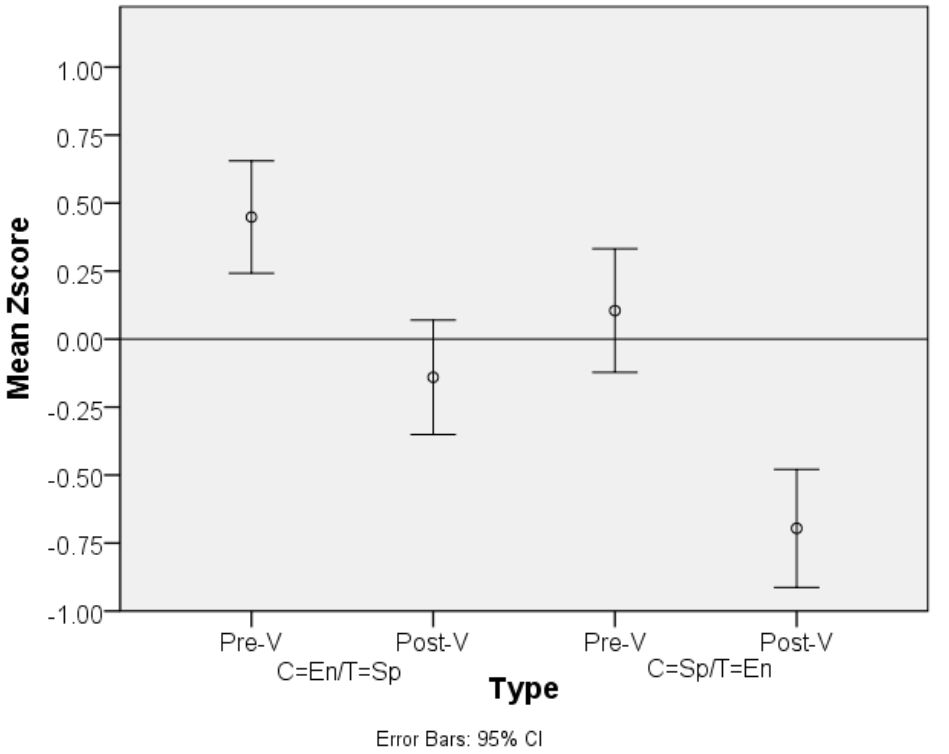
For this experiment, a linear mixed model was fit with two fixed factors—Language of C/T and Subject Position—along with a random intercept and slope for Participant over the interaction of Language by

Subject Position. In this case, including a random effect by lexicalization caused the model not to converge, so only random effects by participant were included. A Type III test of fixed effects revealed a significant effect for Language ($F_{1,45.4}$ = 10.20, p = .003), a significant effect for Subject Position ($F_{1,45.5}$ = 24.35, p < .001), but no significant interaction ($F_{1,45.4}$ = 0.57, p = .453). Estimated marginal mean scores for each condition are presented in Table 5 and Figure 2.

Table 5. CS Object Extraction Results

C	T	Subject	Example	Mean Z-Score	SE
Eng	Spa	Pre-V	What did the teachers assume that *el niño había leído antes del examen*?	.450	.144
Eng	Spa	Post-V	What did the teachers assume that *había leído el niño antes del examen*?	-.140	.143
Spa	Eng	Pre-V	*Qué asumieron los maestros que* the child had read before the test?	.105	.144
Spa	Eng	Post-V	*Qué asumieron los maestros que* had read the child before the test?	-.698	.144

Figure 2. CS Object Extraction Results



The significant main effect of Language indicates that, overall, participants rated sentences with English C and Spanish T higher than sentences with Spanish C and English T. The significant main effect of Subject Position indicates that, overall, preverbal subjects are rated higher than postverbal subjects. However, the result of interest is the lack of interaction between the two, which indicates that the acceptability of the two subject positions was not found to vary depending on the language of C or T. Instead, regardless of which functional head is present, there is a consistent effect of reduced acceptability when the sentence has a postverbal subject. Similarly, for each subject position, there is a

consistent effect of reduced acceptability when switching between Spanish C and English T.

6.3. Experiment 2 Discussion

We originally considered two possible predictions, based on previous theoretical accounts, for which functional head determines whether subjects can be postverbal: the licensor could be either C or T. However, our CS results suggest that neither functional head can license postverbal subjects alone; instead, it may be that only when both C and T are Spanish that postverbal subjects are possible. The evidence supporting this claim is that, although we observe the effects of subject position and the language of the relevant functional heads on acceptability, we find no interaction between them. Preverbal subjects are more acceptable overall, and sentences with English C and Spanish T are more acceptable overall, but neither Spanish functional head is connected with increased relative acceptability of postverbal subjects.

When interpreting the results of a judgment experiment, it is necessary to keep in mind that acceptability is a complex psychological percept, and judgments may reflect several factors beyond the grammaticality of a sentence. We have taken a number of steps to deal with this fact (multiple lexicalizations, factorial designs, z-score transformations), but it is worth considering it again when interpreting the results. Sometimes, when there are only two options which differ

minimally and whose difference in acceptability is predicted to be relatively stark, it is enough to establish that the two receive different judgments in the appropriate direction. That is the case, for instance, with our monolingual Spanish judgments to establish that *que* is indeed obligatory for these speakers in experiment 3. Most of the time, though, just knowing that one sentence type is more acceptable than another may not tell you what you need to know about the syntactic feature under investigation, which is why we employed a factorial design, as recommended by Schütze and Sprouse (2013). In these designs, the relevant question is not whether a given sentence is rated higher than another, but rather whether the *difference* in acceptability between two sentences has the same magnitude and direction as the difference in acceptability between two otherwise similar sentences without the relevant feature. In other words, we need to look at the interaction between the two factors in the design.

In the present case, for example, if we compared only (21a) and (21b), we might come to the conclusion that C determines subject position, because C is English, which should disallow postverbal subjects, and (21b) is rated lower than (21a). It thus may appear that postverbal subjects are rated lower in the presence of English C.

(21) **CS subject position**

a. What did the teachers assume that *el niño había leído antes del examen*? (SV)

b. What did the teachers assume that *había leído el niño antes del examen*? (VS)

c. *Qué asumieron los maestros que* the child had read before the test? (SV)

d. *Qué asumieron los maestros que* had read the child before the test? (VS)

However, it may be the case that postverbal subjects provoke reduced acceptability across the board, such that any sentence with a postverbal subject will be lower in acceptability than an equivalent sentence with a preverbal subject, in the absence of an intervening factor. And, in fact, if we compare (21c) and (21d), where C is in Spanish, we find the same pattern: the postverbal subject is rated lower there too.

Regardless of the language of C or T, then, postverbal subjects produce the same drop in acceptability. If one of the two Spanish heads were able to license postverbal subjects, we would expect the difference between postverbal and preverbal subjects to be reduced when that head is present, much like what was found for the monolingual Spanish results, in which there was no significant difference between postverbal and

preverbal subjects. Since we know these speakers' Spanish allows postverbal subjects, if the licensing of postverbal subjects depended on either C or T alone, that head's presence should ameliorate the decrease in acceptability. That is, we should see an interaction between the two factors, with postverbal order producing a less severe drop in acceptability in the presence of the relevant Spanish functional head. That this interaction is missing suggests that the presence of either Spanish C or Spanish T alone is not enough to license postverbal subjects. In contrast, the monolingual data, with C and T both in Spanish, do show the relevant interaction, which we take to suggest that postverbal subjects are only available when both C and T are in Spanish.

Of course, we are aware that absence of evidence is not evidence of absence; it may be that there is an interaction and we have simply failed to detect it. We are not aware of any widely accepted method for determining the power of a linear mixed model post hoc, but these results are based on a respectable sample size of 64 judgments per sentence type, so we believe the experiment has reasonable power. Additionally, we judge the pattern to be relatively clear: each factor—the language of C/T and the subject position—individually affects acceptability, but the effect of subject position doesn't vary according to the language of C/T. For this reason we are arguing that this result shows that neither head is enough to

improve the acceptability of postverbal subjects, and thus that postverbal subjects are licensed only by some combination of C and T.

One possible objection to our conclusion that it is C and T together that determine subject position might be that these speakers could simply be demonstrating transfer from English in their CS judgments, which is why the pattern of judgments looks English-like regardless of the language of C/T. They are English-dominant, after all, and we know that there is cross-linguistic influence in bilinguals on many features, including subject position (Cuza, 2012). We believe this is unlikely in the present case. First, we know these speakers have not transferred English features to their Spanish generally, as evidenced from the monolingual Spanish results, which were as expected for monolinguals. Nonetheless, it could be that CS sentences allow for greater influence from English, since they trigger bilingual mode, with greater activation of English. However, if that were the case, we would expect the CS results for subject extraction to uniformly parallel those of monolingual English. Instead, as we will see in section 8, this is not what is observed. We thus conclude that these results suggest that both C and T are involved in licensing postverbal heads.

7. Experiment 3: Monolingual Subject Extraction

7.1. Experiment 3 Design

The purpose of the monolingual subject extraction experiment was to verify that these participants' Spanish and English had the expected

features, namely the presence of the *that*-trace effect in English and obligatory overt C in Spanish. In order to confirm the *that*-trace effect, it was necessary to compare subject and object extraction across overt and null C. The English stimuli thus had a 2 x 2 design: Wh-Type (subject or object) and Realization of C (overt *that* or null). The four sentence types and example sentences are presented in Table 6.

Table 6. Monolingual English *that*-trace

Wh	C	Example
Object	That	What did your parents think that your sister had broken in the living room?
Object	Null	What did your parents think your sister had broken in the living room?
Subject	That	Who did your parents think that had broken the statue in the living room?
Subject	Null	Who did your parents think had broken the statue in the living room?

To verify that C is obligatory in participants' Spanish, we also tested Spanish sentences with subject extraction and with overt C *que* or a null C, as shown in Table 7. The Spanish and English sentences were translation equivalents and used the same lexicalizations as the other monolingual experiment.

Table 7. Monolingual Spanish Subject Extraction

Lang	C	Example
Spa	Que	*¿Quién pensaron tus padres que había roto la estatua en la sala de estar?*
Spa	Null	*¿Quién pensaron tus padres había roto la estatua en la sala de estar?*

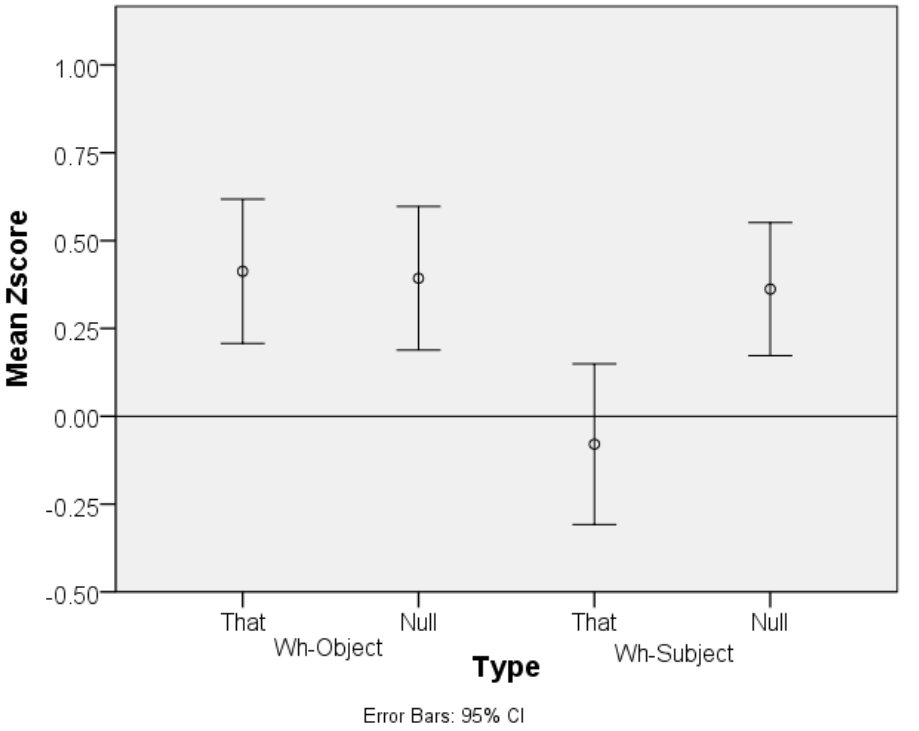
7.2. Experiment 3 Results

For this experiment, the two languages were analyzed separately. For the English sentences, which tested whether these participants have the *that*-trace effect in their English, a linear mixed model was fit with two fixed factors—Wh-Type and Realization of C—along with a random intercept for Participant. Including a random slope for participant or any random effect for lexicalization either did not improve the model fit or caused it not to converge. A Type III test of fixed effects revealed a significant effect for Realization of C ($F_{1,237.0}$ = 4.41, p = .037), a significant effect for Wh-Type ($F_{1,237.0}$ = 6.79, p = .010) and a significant interaction ($F_{1,237.0}$ = 5.29, p = .022). Estimated marginal mean scores for each condition are presented in Table 8 and Figure 3.

Table 8. Monolingual English *that*-trace Results

Wh	C	Example	Mean Z-Score	SE
O	That	What did your parents think that your sister had broken in the living room?	.413	.114
O	Null	What did your parents think your sister had broken in the living room?	.393	.114
S	That	Who did your parents think that had broken the statue in the living room?	-.079	.114
S	Null	Who did your parents think had broken the statue in the living room?	.362	.114

Figure 3. Monolingual English *that*-trace Results



The significant main effect of Wh-Type indicates that, overall, participants rated sentences with object extraction better than subject extraction, and the significant main effect of Realization of C indicates that participants rated null C higher than *that*. However, examination of Figure 3 makes clear that both of these results are merely artifacts of the result of interest, which is the interaction between the two. The interaction indicates that the acceptability of the two possible realizations of C is different depending on the type of *wh*-word extracted, with extraction of a *wh*-subject over *that* resulting in significantly worse ratings than the other three structures. To confirm this difference, pairwise comparisons were

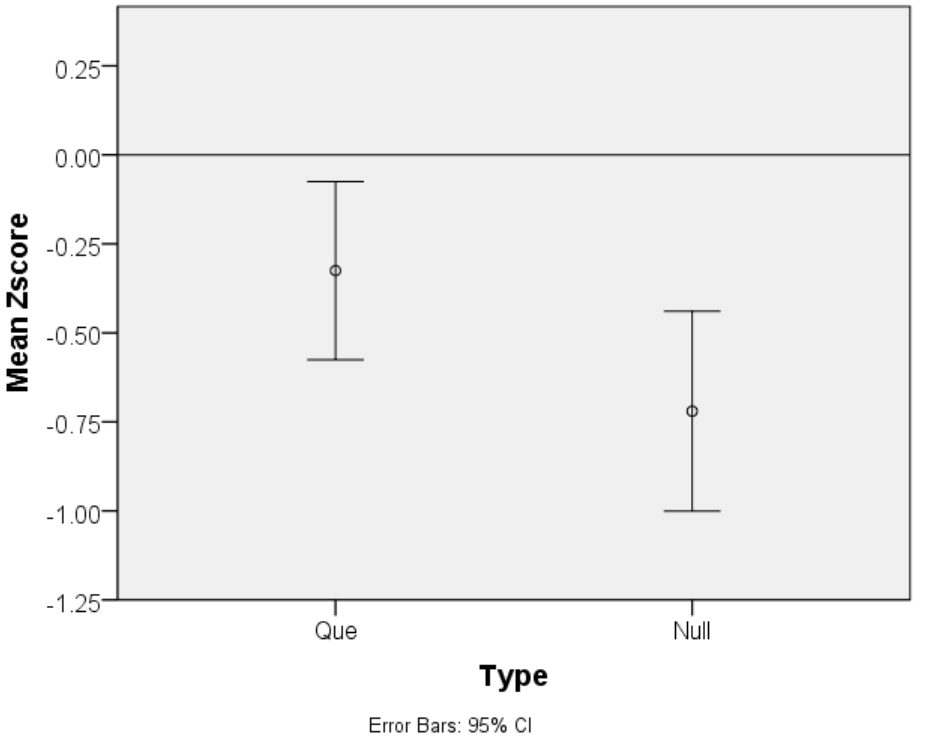
conducted, with the Bonferroni correction for multiple comparisons. These revealed that when the subject is extracted, null C is significantly better than overt *that* (p = .002), while when the object is extracted, no difference was found between the two realizations of C (p = .889). Additionally, when there is an overt *that*, subject extraction is significantly worse than object extraction (p = .001), while with null C, the two extraction types were not found to differ (p = .829).

For the Spanish monolingual sentences, which tested whether *que* was obligatory in their Spanish, a linear mixed model was fit with a single fixed factor—Realization of C—along with a random intercept for Participant. Including a random slope for Participant or any random effect by lexicalization either caused the model not to converge or resulted in a model with worse fit, so only the random intercept for Participant was included. A Type III test of fixed effects revealed a significant effect for Realization of C ($F_{1,111}$ = 6.31, p = .013). Estimated marginal mean scores for each condition are presented in Table 9 and Figure 4.

Table 9. Monolingual Spanish Subject Extraction

Lang	C	Example	Mean Z-Score	SE
Spa	Que	*¿Quién pensaron tus padres que había roto la estatua en la sala de estar?*	-.325	.186
Spa	Null	*¿Quién pensaron tus padres había roto la estatua en la sala de estar?*	-.720	.186

Figure 4. Monolingual Spanish Subject Extraction



The significant main effect of the realization of C indicates that participants rate sentences with overt *que* higher than those with a null complementizer in Spanish (p = .013).

7.3. Experiment 3 Discussion

As with the previous monolingual experiment, Experiment 3 serves as a first step to establish that the properties of C and T in these speakers' grammars are as expected in order to be able to then consider these heads in CS. For experiment 3, finding C and T with the expected features means finding that the *that*-trace effect is attested in their

monolingual English judgments and that *que* is obligatory in their monolingual Spanish. This is indeed what we find. Starting with English, the pattern is exactly as expected: subject extraction over an overt C is worse than all other options. Given the factorial design of the stimuli, we can conclude that it is neither the presence of *that* nor the extraction of subjects alone that results in the low ratings of subject extraction over an overt complementizer but rather the interaction of those two factors, which is also borne out by our pairwise comparisons.

This clear evidence of the *that*-trace effect in English is worth highlighting. Some previous studies have failed to find evidence of the *that*-trace effect with monolingual English speakers (Sobin, 1987) or have found variation in the judgments speakers give (Cowart, 1997, 2003). The present study constitutes robust evidence for the presence of the *that*-trace effect in the English of these speakers.

For Spanish, the preference for *que* over a null C is equally clear, conforming to expectations. We can thus conclude that the relevant features of these speakers' functional heads are the same as those expected for monolingual speakers.

8. Experiment 4: CS Subject Extraction

8.1. Experiment 4 Design

After establishing the properties of participants' Spanish and English separately, the CS subject extraction CS experiment attempted to

isolate which functional head determines whether subject extraction over an overt C is possible. These tokens always had an overt C.

The first step was to be sure that the results found in Experiment 3, where C and T were in the same language because the sentences were in one language, hold as well for CS when the two functional heads are in the same language. That is, it was necessary to test CS sentences with no switch between C and T to determine that any observed effects on the patterns of acceptability were not simply due to CS. So we first tested the sentences in Table 10.

Table 10. CS Subject Extraction, No Switch

Wh	C	T	Example
Subj	Eng	Eng	*Quién asumieron los maestros* that had read the text before the test?
Subj	Spa	Spa	Who did the teachers assume *que había leído el texto antes del examen*?

After that preliminary step, the main stimuli for this experiment were designed to isolate the effect of a specific functional head on subject extraction, so it was necessary to compare subject and object extraction across both Spanish and English C. These stimuli thus had a 2 x 2 factorial design: Wh-Type (subject or object) and Language of C/T (English C with Spanish T vs. Spanish C with English T). The lexicalizations used in this CS experiment were the same as the other CS experiment but different

from the monolingual experiments. The four types and example sentences are presented in Table 11.

Table 11. CS Subject and Object Extraction

Wh	C	T	Example
Obj	Eng	Spa	What did the teachers assume that *el niño había leído antes del examen*?
Obj	Spa	Eng	*Qué asumieron los maestros que* the child had read before the test?
Subj	Eng	Spa	Who did the teachers assume that *había leído el texto antes del examen*?
Subj	Spa	Eng	*Quién asumieron los maestros que* had read the text before the test?

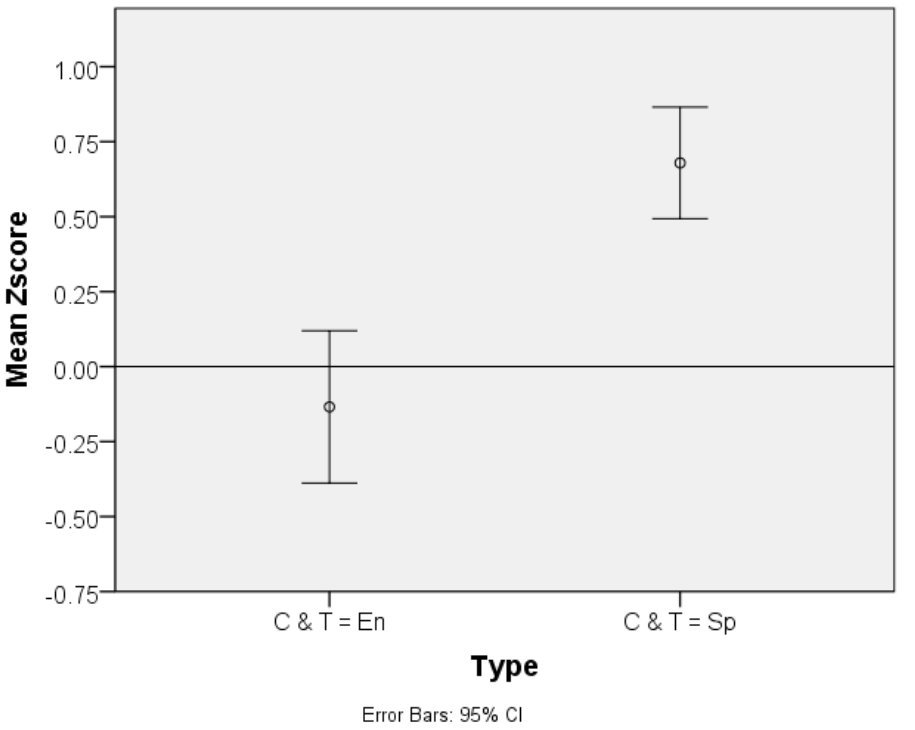
8.2. Experiment 4 Results

First, in order to exclude the possibility that the effects observed in this experiment could be due to CS itself, we also considered CS sentences in which C and T were in the same language, to verify that these behaved the same way as the monolingual sentences. A linear mixed model was fit with a single fixed factor—Language of C/T—along with a random intercept for Participant and for Lexicalization (including random slopes did not improve the model fit). A Type III test of fixed effects revealed a significant effect for Language of C/T ($F_{1,102.6} = 33.19$, $p < .001$). Estimated marginal mean scores for each condition are presented in Table 12 and Figure 5.

Table 12. CS Subject Extraction with Congruent C & T

Wh	C	T	Example	Mean Z-Score	SE
Subj	Eng	Eng	*Quién asumieron los maestros* that had read the text before the test?	-.158	.137
Subj	Spa	Spa	Who did the teachers assume *que había leído el texto antes del examen*?	.681	.137

Figure 5. CS Subject Extraction with Congruent C and T



These CS stimuli behave as expected, with extraction from an embedded clause with both C and T in English rated low and extraction from an embedded clause with both C and T in Spanish rated high.

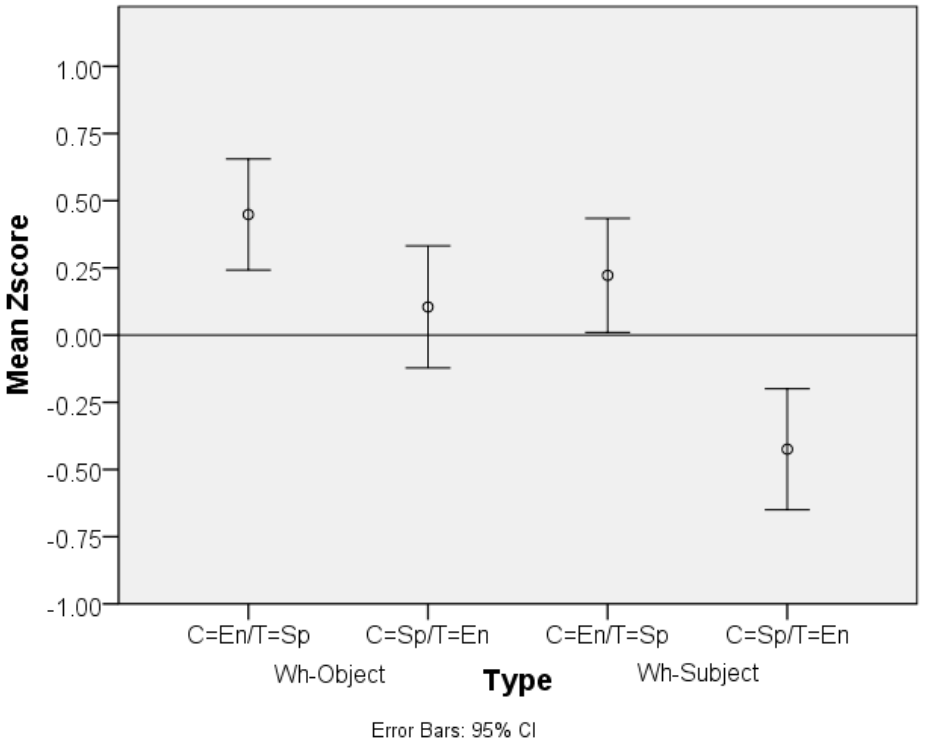
Turning to switches between C and T, a linear mixed model was fit with two fixed factors—Wh-Type and Language of C/T—along with a random intercept and slope by Participant over the interaction of Wh-Type

by Language. Including a random effect by lexicalization caused the model not to converge, so only the random effects for participant were included. A Type III test of fixed effects revealed a significant effect for Wh-Type ($F_{1,45.5}$ = 8.02, p = .007), as well as a significant effect for Language of C/T ($F_{1,45.4}$ = 13.69, p = .001), but no significant interaction between the two conditions ($F_{1,45.4}$ = 1.25, p = .270). Estimated marginal mean scores for each condition are presented in Table 13 and Figure 6.

Table 13. CS Subject and Object Extraction Results

Wh	C	T	Example	Mean Z-Score	SE
Obj	Eng	Spa	What did the teachers assume that *el niño había leído antes del examen*?	.451	.146
Obj	Spa	Eng	*Qué asumieron los maestros que* the child had read before the test?	.107	.146
Subj	Eng	Spa	Who did the teachers assume that *había leído el texto antes del examen*?	.222	.146
Subj	Spa	Eng	*Quién asumieron los maestros que* had read the text before the test?	-.420	.147

Figure 6. CS Subject and Object Extraction Results



The significant main effect of Wh-Type indicates that, overall, participants rated sentences with object extraction better than subject extraction, and the significant main effect of Language indicates that, overall, participants rated sentences with English C and Spanish T higher than sentences with Spanish C and English T (as in the previous CS experiment). However, the result of interest is the lack of interaction between the two: we did not find evidence that the acceptability of the two types of extraction varies depending on the language of C or T. Instead, regardless of which argument is extracted, there is a consistent effect of

reduced acceptability when the switch is from Spanish C to English T when compared to switches from English C to Spanish T.

8.3. Experiment 4 Discussion

Experiment 4 addresses the question of which head determines whether subject extraction is available over overt C. The prediction we made, based on previous work that has shown a relationship between the availability of postverbal subjects and the availability of subject extraction over C (Rizzi, 1982; Menuzzi, 2000), was that the same head(s) found to be responsible for subject position would be responsible for allowing or prohibiting subject extraction. We concluded in section 6.3 that subject position was determined by C and T together rather than one of them alone, which would lead us to predict that subject extraction would only be possible when both C and T were in Spanish, because this was the only case in which postverbal subjects appear to be possible. The subject extraction results align with this prediction: extraction over an overt complementizer appears to be licensed only when both C and T are in Spanish, whereas the presence of only one Spanish functional head is insufficient.

As with the previous research question, the main evidence supporting this conclusion is the lack of interaction between the extracted argument and the language of C/T. As before, if we look only at (22a) and (22b), we might conclude that it is T that makes subject extraction is

possible because (22a), with Spanish T, is rated higher than (22b), with Spanish C. Because Spanish allows subject extraction, it stands to reason that the Spanish head of the more acceptable sentence must be the one making the extraction of the subject possible. Since the sentence with Spanish T is rated higher than the sentence with Spanish C, one might conclude from these data alone that T is what allows subject extraction.

(22) **CS subject extraction**

a. Who did the teachers assume that *había leído el texto antes del examen*?

b. *Quién asumieron los maestros que* had read the text before the test?

However, we previously observed that for object extraction there is an asymmetry in these CS results, wherein sentences with English C and Spanish T were overall rated more highly than those with Spanish C and English T. Because of this, the fact that we found (22a) was rated more highly than (22b) may be due simply to this asymmetry and not to one sentence allowing subject extraction while the other does not.

The only way to rule out the effect of the switch between C and T is to use the factorial design to compare the subject extraction sentences above to parallel object extraction sentences.

(23) **CS object extraction**

a. What did the teachers assume that *el niño había leído antes del examen*?

b. *Qué asumieron los maestros que* the child had read before the test?

These sentences have the same switch sites as (22) but extract objects instead of subjects. (We chose to compare object extraction sentences with preverbal subjects because those were found to be more acceptable in the previous experiment.) The relevant comparison is whether the difference between subject extraction sentences with a given head is smaller than, larger than, or the same as the object extraction sentence with the same head. So we compare (22a) and (23a) and then compare (22b) and (23b), and the result of interest is whether one difference is greater than or less than the other. We found that subject extraction was in general worse than object extraction, so we expect a decrease in acceptability from (23) to (22), but if the decrease is less for one particular configuration of C/T, then we have reason to believe that one of those heads (by assumption, whichever one is in Spanish) facilitates subject extraction. In other words, if we had found that the difference between (22a) and (23a) showed a large decrease in acceptability for

(22a), when the subject was extracted over *that*, but the difference between (22b) and (23b) was small or nonexistent, we would be led to conclude that Spanish C, the Spanish functional head present in the (b) examples, reduced the decrease in acceptability provoked by extracting a subject over an overt C, thus indicating that it was this functional head that permitted subject extraction.

However, that's not what we found; instead, we found only consistent effects for language of C/T and for extraction type, but no interaction. We observe a decrease in acceptability whenever there is Spanish C, and we observe a decrease in acceptability whenever we extract a subject, but neither functional head makes extracting a subject any better. This indicates that the presence of C or T from one language or the other does not affect acceptability differently for subject or objects. Subject extraction over an overt C isn't made better by the presence of either the Spanish C or Spanish T, but switches between English C and Spanish T are always better than the reverse.

As before, of course we are aware that merely not finding an interaction does not mean that there is no effect, and this may not allow us to conclude that both functional heads are needed. It may be that we have simply failed to detect the relevant interaction. As before, though, we believe the experiment has reasonable power, and we nonetheless found no evidence to support the idea that either functional head alone is able to

produce an increase in acceptability for subject extraction relative to parallel object extraction cases. What's more, in this case we do have some direct evidence that subject extraction is possible when both heads are in Spanish.

In order to control for the possibility that we would observe decreased acceptability for subject extraction merely due to CS itself, we also tested cases like (24), where C and T were in the same language.

(24) **CS subject extraction with congruent C and T**

a. *Quién asumieron los maestros* that had read the text before the test?

b. Who did the teachers assume *que había leído el texto antes del examen*?

Here we find that (24b) is rated substantially higher than (24a). While we do not have the full factorial design with these sentences and thus our analysis is limited, it seems fair to say that (24a) is behaving like our monolingual English examples (disallowing subject extraction) while (24b) is behaving like monolingual Spanish (allowing it). So it appears that neither functional head alone is enough to allow subject extraction, but both together are.

Why didn't we compare (22) directly to (24)? Because the switch site for (22) was different than the one in (24), with only the former having a switch between C and T, which, as we have seen, has a definite effect on acceptability.[9] Comparing (22) and (24) would have been misleading, as what might appear to be the effect of one functional head on subject extraction was actually only the effect of where the code-switch was. Instead, the comparison we made—between (22) and (23)—is the one that allows us to conclude that neither functional head alone ameliorates the *that*-trace effect. That data plus our results from (24) lead us to conclude that both C and T together play a role in permitting subject extraction.

9. General Discussion

This section discusses the results of the experiments in relation to our research questions and explores their implications.

9.1. What determines subject position in Spanish/English CS?

If subject position must be accounted for based on an interaction or combination of C and T, as our results suggest, this has a number of

[9] In fact, some previous work on CS between C and T has suggested that such switches are not ever possible (Belazi, Rubin and Toribio, 1994; González-Vilbazo, 2005). However, some previous work with AJTs (González-Vilbazo & López, 2013; Hoot, 2011) has found they are acceptable in at least some circumstances, and they are attested in corpora of naturalistic speech (Callahan, 2004). In our results, roughly half of the CS ratings with switches between C and T were at or above the midpoint of the scale in the raw ratings (see Appendix A), which, considering that CS ratings tend to be depressed overall due to negative attitudes toward or unfamiliarity with CS, we take to indicate that switching between C and T is not ruled out *a priori*.

implications. First, many accounts of subject position in Spanish *wh*-questions assume that subject position is exclusively a matter of properties of the tense head (e.g., Barbosa, 2001; Goodall, 2001; Gutiérrez-Bravo, 2002, 2008; Zubizarreta, 2012), but our results suggest that C also plays an important role.[10] Likewise, Rizzi (2006) proposes that subject position is determined by the functional head dominated by C, but according to their account, C is not involved.[11]

These results are also not compatible with accounts that give C a central role in determining the position of the subject. For example, González-Vilbazo and López's (2012) phase head hypothesis (PHH) predicts that C, as a phase head, should determine the grammatical properties of the complement, including subject position. While González-Vilbazo and López (2013) did find evidence that C alone determined subject position in Spanish-German CS, our results instead suggest that both C and T are necessary, at least in Spanish-English CS.

[10] Strictly speaking, Zubizarreta's (2012) account involves both T and a new functional head, *phi*, that immediately dominates T, but this account would make the same unsupported predictions as accounts based on T alone. The tense head eventually incorporates into this *phi* head, along with the accompanying little *v* and V heads that previously incorporated into T, so the language of the verb is also the language for both T and *phi*.

[11] In Rizzi's model, this head is a dedicated projection for the subject, rather than T, but the important point is that C does not play a role.

9.2. What determines whether subject extraction over a complementizer is possible?

Our results support an account of the *that*-trace effect that depends on the interaction or combination of C and T, rather than either one alone. This is problematic, for example, for Rizzi and Shlonsky's (2007) account of the *that*-trace effect, which is based on a combination of a general syntactic principle, Criterial Freezing, and a particular property of T, specifically the EPP. The complementizer does not play a role in this account, which is incompatible with our results.

Similarly, we did not find clear support for González-Vilbazo and López's PHH (2013) which posits that essential features of TP, including subject position, depend on features of C. This would predict that properties of C should be sufficient to determine when there is a *that*-trace effect, but this is not what we found.

At first glance, Pesetsky and Torrego's (2001) account appears consonant with our results in that both C and T play a role in the account. C provides the syntactic triggers, with the *wh*-probe and the tense probe, and T can satisfy the tense probe in certain structures. Crucially, however, T plays no role in subject extraction under their account. They assume that the subject raises to Spec,TP in both languages, so the *wh*-subject is able to satisfy both of C's probes and is required to do so for economy reasons. The difference between Spanish and English ultimately stems from C itself: English C is a null morpheme that is phonologically realized as *that*

when T is raised to it but not otherwise, while Spanish C is a true complementizer, always realized as *que*. Spanish does not manifest the *that*-trace effect because it has a true complementizer that is always overtly realized. This predicts that English C and Spanish C should have differential effects on subject extraction with an overt complementizer compared to object extraction with an overt complementizer, regardless of the language of T, but this is not what we found.

However, a few small changes in their assumptions bring the account much closer to aligning with our data. Pesetsky and Torrego assume that postverbal subjects in Spanish appear postverbally for reasons similar to English, where the subject raises to Spec,TP and T later raises above the subject to C. As discussed in section 2.2, there is strong evidence that postverbal subjects in Spanish are not in Spec,TP but rather in a lower position. If we adopt this assumption along with Pesetsky and Torrego's proposal of the feature specification of C, their account can be modified to align more closely with our data and to straightforwardly account for the correlation between subject position and subject extraction, as well as the differences between Spanish and English.

Recall that Pesetsky and Torrego claim that C probes for two unvalued features: a *wh*-feature and a tense feature. In English, the *wh*-subject is able to satisfy both the *wh*-probe and the tense probe because the *wh*-subject is raised from Spec,TP. Since the tense probe on C agrees with

the closest head with the appropriate features, the economy principle Pesetsky and Torrego adopt forces the *wh*-subject to satisfy the tense probe. This makes it impossible for *that* to surface when subjects are extracted, since *that* is simply the phonological realization of T raised to C. In this way they explain the asymmetry between subject and object extraction: the former is one movement, the latter is two. Spanish C has the same unvalued features, but the difference is that *que* is a true complementizer. Again, subject extraction is one movement, and object extraction is two: one to raise the *wh*-object and another in which T raises and right-adjoins to C. This latter move is needed to produce post-verbal subjects, because the subject is otherwise in Spec,TP.

However, assuming that Spanish post-verbal subjects are in situ rather than raising them to Spec,TP simplifies things. If the *wh*-subject is below T, then T will be the closest head with a T feature, and both types of extraction will require two movements: one that raises T to C to value the tense feature and one that raises the *wh*-word to value the *wh*-feature. Thus there would be no difference in Spanish between subject and object extraction, which is what we observe, and this fact would be tied to subject position in Spanish, which is also what we observe. Furthermore, this would even allow us to take their account one step farther: it could be the case that, just like in English, the putative Spanish complementizer *que* is really the phonological realization of a feature of T raised to C. Such a

possibility would unify the behavior of complementizers across the two languages and also removes from their proposal the other assumption that we view as problematic—obligatory T-to-C movement in Spanish—because the verbal complex would no longer need to adjoin to (and thus be phonologically realized with) *que*, while at the same time straightforwardly tying the difference between English and Spanish C realization to the other independently necessary difference in subject position. Of course, there are likely additional wrinkles that would need to be worked out, but this suggestion may be a way to integrate our findings with one of the most prominent accounts of the *that*-trace effect, and it seems like the most promising theoretical account to match our results.

9.3. Subject Position and That-trace

Our results provide further support for the notion that the availability of subject extraction is linked directly to the availability of postverbal subjects. As previously noted, Rizzi (1982) and Menuzzi (2000) provide evidence that *wh*-subjects are extracted from a postverbal position in Italian and Portuguese, respectively, but we are not aware of any direct evidence that Spanish *wh*-subjects are extracted from a postverbal position. Our CS results, however, do provide some indirect evidence that this extends to Spanish. Specifically, the results for object extraction suggest that both C and T are necessary for determining the position of the subject, and the results for subject extraction suggest that

both heads are also involved in determining when there is a *that*-trace effect. The fact that both appear to involve C and T, particularly when many accounts of each phenomenon would predict either C or T alone would be sufficient, strengthens the empirical support for linking these two properties.

9.4. The EPP

Our results also have implications for the EPP, in the sense of a requirement on T that subjects raise to Spec,TP.[12] Traditionally, the EPP has been conceptualized as a feature on T that requires preverbal subjects in languages like English, while languages like Spanish which allow post-verbal subjects have either no EPP or an optional EPP. However, if our results are on the right track, it appears that the EPP cannot be a feature of T alone. Instead, the presence of either English functional head is enough to force a preverbal subject, since post-verbal subjects are only licensed when both heads are in Spanish, which implies an unexpected role for C in determining subject position. For example, even sentences with English C and Spanish T behave like English, allowing only preverbal subjects, a surprising result if the EPP is a feature of T. This may indicate that the classical EPP needs to be revised.

[12] The EPP is sometimes used as a more general requirement that a phrase agreeing with a given head must raise to the specifier of that head (e.g., Chomsky, 2000, 2001), but here we focus specifically on the position of the subject.

Additional support for this suggestion comes from previous studies looking at null subjects in CS. Cross-linguistically, the availability of post-verbal subjects is correlated with the existence of null subjects; for instance, Italian and Spanish allow both null subjects and postverbal subjects, while English and German allow neither. Previous work on Spanish-German CS (González-Vilbazo & López, 2013) and Spanish-English CS (Sande, 2017) have found that in order to license null subjects in CS, both C and T must be in the null subject language. If null subjects and postverbal subjects are taken to be the result of the same property, our findings follow these previous studies in suggesting a role for C in determining properties of subjects that it is not usually assumed to have.

9.5. CS and Experimental Syntax

Finally, our study has implications for the experimental study of syntax and of CS. We have shown evidence that using experimental methods can be a valuable source of data to test and verify theoretical claims, including finding clear evidence supporting the general accounts of how complementizers behave in Spanish and English, and showing a clear example of the *that*-trace effect. We also show how investigating CS under a no-third grammar approach can bear fruit by revealing insights about each of the languages involved that may not be available with monolingual stimuli.

Empirically, it is worth pointing out that our results echo previous work on CS is in finding that there is a "directional" preference in these CS sentences: sentences with English C and Spanish T were always better than the reverse, regardless of subject position. We do not have sufficient evidence to offer a theoretical explanation of why this may be; we limit ourselves to pointing out that within other CS pairs it has been attested that certain switch sites accept only one of the two logically possible switch "directions." For example, in Spanish-German switches between little *v* and VP, only Spanish little *v* + German VP is possible while German little *v* + Spanish VP is unattested (González-Vilbazo & López, 2011).[13]

10. Conclusions and future directions

This project set out to use CS to shed new light on theoretical accounts of the *that*-trace effect. The principal findings were: (a) both C and T appear to be involved in determining subject position in Spanish-English CS; (b) both C and T appear to be involved in determining whether subject extraction from an embedded clause is possible in Spanish-English CS. These findings support theoretical approaches that

[13] For example, (a), with Spanish little v + German VP, is acceptable but (b), with German little v + Spanish VP, is not:

Asymmetry in little *v* + VP code-switching (González-Vilbazo & López, 2011)

a. *Juan hace* nähen das Hemd ('Juan sews the shirt.')
b. *Hans tut *coser la camisa*. ('Hans sews the shirt.)

unify subject position and the *that*-trace effect and suggests that it is necessary to take both C and T into account in order to determine the availability of those properties. Ample avenues for future research remain, however, including a more thorough investigation of subject position in Spanish-English CS with and without switches between C and T, an investigation of C and T switches in non-interrogative contexts, tests of other features related to the C domain in CS, and many others. Future work on CS has the potential to make many more contributions to linguistic theory.

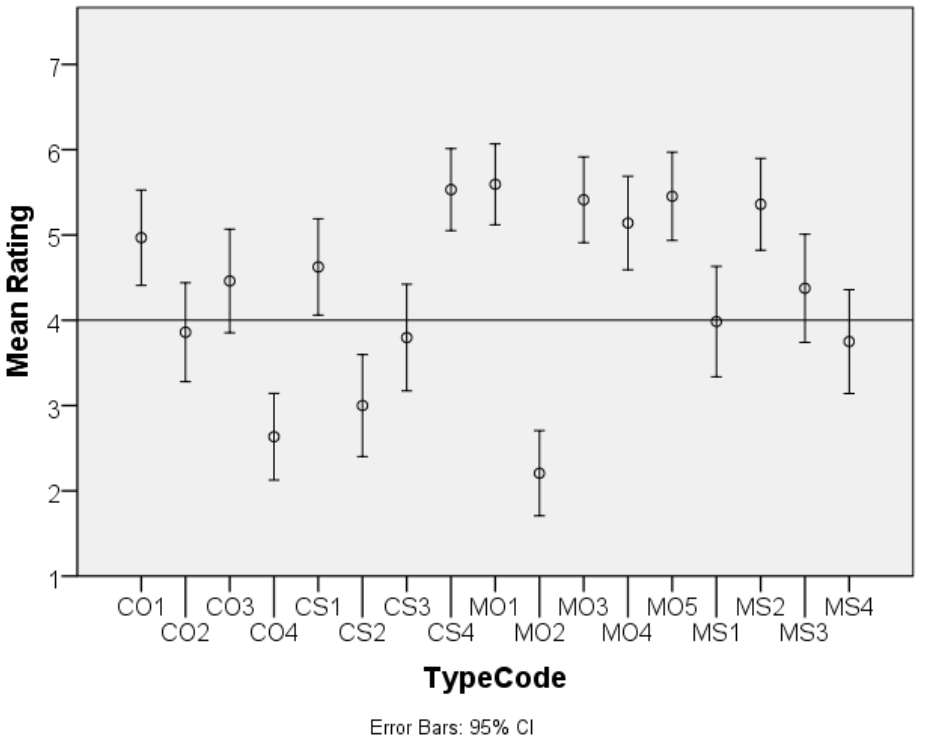
Appendix A: Raw descriptive results

The raw mean ratings and SDs are listed in Table 14 and presented in Figure 7. These ratings show that participants assigned a range of acceptability to the sentences in the experiment. Furthermore, although the CS responses were slightly lower than monolingual responses overall, some CS responses—including those with switches between C and T (like CO1, CO3, and CS1)—are at or above the midpoint of the scale. This indicates that these switches do not necessarily produce low ratings in and of themselves but rather that participants are responding to the experimental manipulation.

Table 14. Raw mean ratings

Type	Wh	C	T	Subject	Example	Mean Rating	SD
CO1	Obj	Eng *that*	Spa	Pre-V	What did the teachers assume that *el niño había leído antes del examen*?	4.97	2.2
CO2	Obj	Eng *that*	Spa	Post-V	What did the teachers assume that *había leído el niño antes del examen*?	3.86	2.3
CO3	Obj	Spa	Eng	Pre-V	*Qué asumieron los maestros que* the child had read before the test?	4.46	2.4
CO4	Obj	Spa	Eng	Post-V	*Qué asumieron los maestros que* had read the child before the test?	2.63	2.0
CS1	Subj	Eng *that*	Spa	Pre-V	Who did the teachers assume that *había leído el texto antes del examen*?	4.63	2.3
CS2	Subj	Spa	Eng	Pre-V	*Quién asumieron los maestros que* had read the text before the test?	3.00	2.4
CS3	Subj	Eng *that*	Eng	Pre-V	*Quién asumieron los maestros* that had read the text before the test?	3.80	2.5
CS4	Subj	Spa	Spa	Pre-V	Who did the teachers assume *que había leído el texto antes del examen*?	5.53	1.9
MO1	Obj	Eng *that*	Eng	Pre-V	What did your parents think that your sister had broken in the living room?	5.59	1.9
MO2	Obj	Eng *that*	Eng	Post-V	What did your parents think that had broken your sister in the living room?	2.21	2.0
MO3	Obj	Spa	Spa	Pre-V	¿Qué pensaron tus padres que tu hermana había roto en la sala de estar?	5.42	2.0
MO4	Obj	Spa	Spa	Post-V	¿Qué pensaron tus padres que había roto tu hermana en la sala de estar?	5.14	2.2
MO5	Obj	Eng Null	Eng	Pre-V	What did your parents think your sister had broken in the living room?	5.45	2.1
MS1	Subj	Eng *that*	Eng	Pre-V	Who did your parents think that had broken the statue in the living room?	3.98	2.6
MS2	Subj	Eng Null	Eng	Pre-V	Who did your parents think had broken the statue in the living room?	5.36	2.2
MS3	Subj	Spa	Spa	Pre-V	Quién pensaron tus padres que había roto la estatua en la sala de estar?	4.38	2.5
MS4	Subj	Spa	Spa	Pre-V	Quién pensaron tus padres había roto la estatua en la sala de estar?	3.75	2.4

Figure 7. Raw mean ratings



We present the raw ratings here for the sake of completeness, to demonstrate that the scale was used properly by participants, and to demonstrate that some CS ratings with switches between C and T receive ratings at or above the midpoint of the scale, but these results also show some of the expected effects of the fact that acceptability judgments are necessarily complex psychological percepts affected by many factors, including lower ratings for CS sentences overall, lower ratings for grammatical Spanish sentences compared to similar English sentences, and ratings which are lowered by different degrees depending on the type

of violation (compare ungrammatical MO2 with equally ungrammatical MS1, for instance). For this reason, as mentioned in the main text, we carry out our main analysis on the z-score transformations of the raw ratings (Schütze & Sprouse, 2013), which ameliorates these problems to some extent.

References

Bader, M., & Häussler, J. (2010). Toward a model of grammaticality judgments. *Journal of Linguistics, 46*(2), 273–330. https://doi.org/10.1017/S0022226709990260

Badiola, L., Delgado, R., Sande, A., & Stefanich, S. (2016). *Towards methodological standards in code-switching research*. Paper presented at the Bilingualism Forum, Chicago, IL.

Barbosa, P. (2001). On inversion in wh-questions in Romance. In A. Hulk & J.-Y. Pollock (Eds.), *Subject Inversion in Romance and the Theory of Universal Grammar* (pp. 20–59). Oxford, New York: Oxford University Press.

Belazi, H. M., Rubin, E. J., & Toribio, A. J. (1994). Code Switching and X-Bar Theory: The Functional Head Constraint. *Linguistic Inquiry, 25*(2), 221–237.

Bresnan, J. (1972). *Theory of complementation in English syntax* (Doctoral dissertation). Massachusetts Institute of Technology,

Cambridge, MA. Retrieved from http://www.ai.mit.edu/projects/dm/theses/bresnan72.pdf

Bresnan, J. (1977). Variables in the theory of transformations. In P. W. Culicover, T. Wasow, & A. Akmajian (Eds.), *Formal syntax* (pp. 157–196). New York: Academic Press.

Buesa García, C. (2008). *The subject-gap restriction in Spanish wh-questions: Some empirical and theoretical consequences*. Second General Examination, University of Connecticut, Storrs, CT.

Callahan, L. (2004). *Spanish/English Codeswitching in a Written Corpus*. Amsterdam: John Benjamins Publishing.

Chao, W. (1981). PRO-drop languages and nonobligatory control. In W. Chao & D. Wheeler (Eds.), *University of Massachusetts Occasional Papers in Linguistics 7* (pp. 46–74). Amherst, MA: GLSA.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris Publications.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. (2000). Minimalist inquiries: The framework. In R. Martin, D. Michaels, & J. Uriagereka (Eds.), *Step by step: essays on Minimalist syntax in honor of Howard Lasnik* (pp. 89–156). Cambridge, MA: MIT Press.

Chomsky, N. (2001). Derivation by phase. In M. J. Kenstowicz (Ed.), *Ken Hale: a life in language* (pp. 1–52). Cambridge, MA: MIT Press.

Chomsky, N., & Lasnik, H. (1977). Filters and Control. *Linguistic Inquiry*, *8*(3), 425–504. https://doi.org/10.2307/4177996

Cowart, W. (1997). *Experimental syntax: Applying objective methods to sentence judgements*. Thousand Oaks, CA: Sage Publications.

Cowart, W. (2003). *Detecting syntactic dialects: The that-trace phenomenon*. Paper presented at the the 39th Annual Meeting of Chicago Linguistic Society, Chicago, IL.

Cuza, A. (2012). Crosslinguistic influence at the syntax proper: Interrogative subject–verb inversion in heritage Spanish. *International Journal of Bilingualism*, *17*(1), 71–96. https://doi.org/10.1177/1367006911432619

Davies, M. (2006). *A Frequency dictionary of Spanish: core vocabulary for learners*. New York/London: Routledge.

Davies, M. (2008). The Corpus of Contemporary American English (COCA): 520 million words, 1990-present. Retrieved from http://corpus.byu.edu/coca/

de Chene, B. (2000). *Prosody and subject traces*. Unpublished manuscript, Waseda University. Tokyo, Japan. Retrieved from http://www.f.waseda.jp/dechene/OP-PST.pdf

Ebert, S. (2014). *The morphosyntax of wh-questions: Evidence from Spanish-English code-switching* (Doctoral dissertation). University of Illinois at Chicago, Chicago, IL.

Eddington, D. (2015). *Statistics for linguists: A step-by-step guide for novices*. Newcastle upon Tyne: Cambridge Scholars Publishing.

González-Vilbazo, K. (2005). *Die Syntax des Code-Switching. Esplugisch: Sprachwechsel an der Deutschen Schule Barcelona.* (Doctoral dissertation). Universität zu Köln, Cologne, Germany.

González-Vilbazo, K., Bartlett, L., Downey, S., Ebert, S., Heil, J., Hoot, B., … Ramos, S. (2013). Methodological Considerations in Code-Switching Research. *Studies in Hispanic & Lusophone Linguistics*, *6*(1), 119–138.

González-Vilbazo, K., & López, L. (2011). Some properties of light verbs in code-switching. *Lingua*, *121*(5), 832–850. https://doi.org/10.1016/j.lingua.2010.11.011

González-Vilbazo, K., & López, L. (2012). Little v and parametric variation. *Natural Language & Linguistic Theory*, *30*(1), 33–77. https://doi.org/10.1007/s11049-011-9141-5

González-Vilbazo, K., & López, L. (2013). *Phase switching*. Presented at Code-switching in the bilingual child: Within and across the clause, Center for General Linguistics (ZAS), Berlin.

Goodall, G. (1993). SPEC of IP and SPEC of CP in Spanish wh-questions. In W. J. Ashby, M. Mithun, & G. Perissinotto (Eds.), *Linguistic Perspectives on Romance Languages: Selected Papers from the XXI Linguistic Symposium on Romance Languages, Santa Barbara, February 21–24, 1991* (pp. 199–209). Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/cilt.103

Goodall, G. (2001). The EPP in Spanish. In W. Davies & S. Dubinsky (Eds.), *Objects and Other Subjects*. Dordrecht: Kluwer Academic Publishers.

Grosjean, F. (1985). The bilingual as a competent but specific speaker-hearer. *Journal of Multilingual and Multicultural Development*, *6*(6), 467–477. https://doi.org/10.1080/01434632.1985.9994221

Gutiérrez-Bravo, R. (2002). *Structural markedness and syntactic structure: A study of word order and the left periphery in Mexican Spanish* (Doctoral dissertation). University of California - Santa Cruz, Santa Cruz, CA.

Gutiérrez-Bravo, R. (2008). Topicalization and preverbal subjects in Spanish wh-interrogatives. In J. Bruhn de Garavito & E. Valenzuela (Eds.), *Selected proceedings of the 10th Hispanic Linguistics Symposium* (pp. 225–236). Somerville, MA: Cascadilla Proceedings Project.

Hoot, B. (2011). *Complementizer asymmetry in Spanish-English code-switching*. Paper presented at the 8th International Symposium on Bilingualism, Oslo, Norway.

Jake, J. L., Myers-Scotton, C., & Gross, S. (2002). Making a minimalist approach to codeswitching work: Adding the Matrix Language. *Bilingualism: Language and Cognition*, *5*(1), 69–91. https://doi.org/10.1017/S1366728902000147

Juzek, T. (2016). *Acceptability Judgement Tasks and Grammatical Theory* (Doctoral dissertation). University of Oxford, Oxford, UK.

Juzek, T., & Häussler, J. (2015). *Non-cooperative behaviour & acceptability judgement tasks*. Paper presented at the Methods and Linguistic Theories Symposium, Bamberg, Germany.

Kandybowicz, J. (2006). Comp-trace effects explained away. In D. Baumer, D. Montero, & M. Scanlon (Eds.), *Proceedings of the 25th West Coast Conference on Formal Linguistics* (pp. 220–228). Somerville, MA: Cascadilla Proceedings Project.

Kandybowicz, J. (2009). Embracing edges: Syntactic and phono-syntactic edge sensitivity in Nupe. *Natural Language & Linguistic Theory*, *27*(2), 305–344. https://doi.org/10.1007/s11049-009-9064-6

Koopman, H. (1983). ECP Effects in Main Clauses. *Linguistic Inquiry*, *14*(2), 346–350. https://doi.org/10.2307/4178332

Koronkiewicz, B. (2014). *Pronoun categorization: Evidence from Spanish/English code-switching* (Doctoral dissertation). University of Illinois at Chicago, Chicago, IL.

MacSwan, J. (1999). *A Minimalist Approach to Intra-sentential Code-switching*. New York: Garland.

Martinović, M. (2017). Wolof wh-movement at the syntax-morphology interface. *Natural Language & Linguistic Theory*, *35*(1), 205–256. https://doi.org/10.1007/s11049-016-9335-y

Menuzzi, S. (2000). That-trace effects in Portuguese. *Fórum Lingüístico*, *2*(1), 13–39.

Montrul, S. (2002). Incomplete acquisition and attrition of Spanish tense/aspect distinctions in adult bilinguals. *Bilingualism: Language and Cognition*, *5*(01). https://doi.org/10.1017/S1366728902000135

Myers-Scotton, C. (1993). *Duelling languages: grammatical structures in codeswitching*. Oxford: Clarendon.

Perlmutter, D. (1968). *Deep and surface structure constraints in syntax* (Doctoral dissertation). Massachusetts Institute of Technology, Cambridge, MA.

Perlmutter, D. (1971). *Deep and surface structure constraints in syntax*. New York: Holt, Reinhart, and Wilson.

Pesetsky, D. (1982). Complementizer-trace phenomena and the nominative island condition. *The Linguistic Review*, *1*(3), 297–343. https://doi.org/10.1515/tlir.1982.1.3.297

Pesetsky, D. (2017). Complementizer-trace effects. In M. Everaert & H. van Riemsdijk (Eds.), *The Wiley Blackwell Companion to Syntax* (2nd ed.). Wiley Blackwell.

Pesetsky, D., & Torrego, E. (2001). T-to-C movement: Causes and consequences. In M. Kenstowicz (Ed.), *Ken Hale: A life in language* (pp. 355–426). Cambridge, MA: MIT Press.

Ritchart, A., Goodall, G., & Garellek, M. (2016). Prosody and the That-Trace Effect: An Experimental Study. In Kyeong-min Kim et al. (Ed.), *Proceedings of the 33rd West Coast Conference on Formal Linguistics* (pp. 320–328). Somerville, MA: Cascadilla Proceedings Project. Retrieved from http://www.lingref.com/cpp/wccfl/33/paper3251.pdf

Rizzi, L. (1982). *Issues in Italian syntax*. Dordrecht: Foris Publications.

Rizzi, L. (2006). On the form of chains: criterial positions and ECP effects. In L. L. S. Cheng & N. Corver (Eds.), *Wh-movement: moving on* (pp. 97–134). Cambridge, MA: MIT Press.

Rizzi, L., & Shlonsky, U. (2007). Strategies of subject extraction. In U. Sauerland & H.-M. Gärtner (Eds.), *Interfaces + recursion =*

language? Chomsky's minimalism and the view from syntax-semantics (pp. 115–160). Berlin: Mouton de Gruyter.

Sande, A. (2017). *C plus T as a necessary condition for pro-drop: evidence from code-switching* (Doctoral dissertation). University of Illinois at Chicago, Chicago, IL.

Schütze, C. T., & Sprouse, J. (2013). Judgment data. In R. J. Podesva & D. Sharma (Eds.), *Research methods in linguistics* (pp. 27–50). Cambridge: Cambridge University Press.

Sobin, N. (1987). The Variable Status of Comp-Trace Phenomena. *Natural Language & Linguistic Theory*, *5*(1), 33–60. https://doi.org/10.2307/4047574

Sprouse, J. (2011). A validation of Amazon Mechanical Turk for the collection of acceptability judgments in linguistic theory. *Behavior Research Methods*, *43*(1), 155–167. https://doi.org/10.3758/s13428-010-0039-7

Suñer, M. (1994). V-Movement and the Licensing of Argumental Wh-Phrases in Spanish. *Natural Language & Linguistic Theory*, *12*(2), 335–372. https://doi.org/10.2307/4047796

Torrego, E. (1984). On Inversion in Spanish and Some of Its Effects. *Linguistic Inquiry*, *15*(1), 103–129. https://doi.org/10.2307/4178369

Woolford, E. (1983). Bilingual Code-Switching and Syntactic Theory. *Linguistic Inquiry*, *14*(3), 520–536. https://doi.org/10.2307/4178342

Zubizarreta, M. L. (2012). A note on the Spanish left periphery. In L. Brugé, A. Cardinaletti, G. Giusti, N. Munaro, & C. Poletto (Eds.), *Functional heads: The cartography of syntactic structures, Volume 7* (pp. 112–128). Oxford, New York: Oxford University Press.