Derivational Complexity Effects in Bilingual Adults: Instances of Interrogative Inversion in Spanish

Joshua Frank

1. Introduction

The influence of one language on another is frequently observed in a bilingual's grammar. Many researchers in the field of language acquisition have developed theories that predict which linguistic domains are more susceptible to cross-linguistic influence (CLI) than others. Some scholars argue that structural complexity is a critical factor in the conditioning of CLI (e.g., Cuza, 2012; Cuza & Frank, 2011, 2010; Hulk & Müller, 2000; Yip & Matthews, 2009). Derivational Complexity Hypothesis (DCH) predicts that structures from one language with fewer internal merge operations taken before spell-out can transfer to the other more complex language (e.g., Jakubowicz, 2011; Jakubowicz & Strik, 2008; Slavkov, 2011; Strik & Pérez-Leroux, 2011; Strik, 2012). The following study investigates this issue further. Specifically, I examine and compare the extent to which advanced heritage language (HL) and second language (SL) learners of Spanish produce subject-verb inversion in matrix and embedded *wh*-questions.

2. Derivational Complexity Hypothesis

Jakubowicz (2011) defines derivational complexity by a metric, which states that fewer merges constitute a less complex derivation. The assumption is that learners are in some way aware of the number of times a copy must be merged and sensitive to the number of constituents that undergo internal merge. The exact metric as stated in Jakubowicz (2005) is represented in (1) below:

(1) Derivational Complexity Metric
   a. Merging α_i n times gives rise to a less complex derivation than merging α_i (n+1) times.
   b. Internal Merge of α gives rise to a less complex derivation than Internal Merge of $\alpha + \beta$.

The author goes on to explain that DCH applies to several areas of language acquisition, which includes but is not limited to L1A, L2A, and SLI, as well as adult language processing in general. Still, few studies have considered derivational complexity effects in the adult bilingual population.

One such study is Slavkov (2011). He investigated the acquisition of long distance (LD) *wh*-questions in native Bulgarian and French speakers learning English as a second language. Slavkov argues that complex *wh*-questions in English, due to the associated high degree of complexity (successive cyclical application of *wh*-movement), should be difficult to acquire by the aforementioned populations. French allows various complex *wh*-question constructions: *wh*-movement, *wh* in situ,

* Joshua Frank, The University of Texas at Austin, jfrank18@utexas.edu. Thank you to Alejandro Cuza and Chiyo Nishida for their helpful feedback on earlier drafts of this work. I am also grateful to the audience of the 2012 Hispanic Linguistics Symposium at the University of Florida, particularly the helpful suggestions of Jason Rothman and Joyce Bruhn de Garavito. Furthermore, the comments of two anonymous reviewers proved extremely helpful in the development of this paper. All the remaining errors are my own.

 Selected Proceedings of the 16th Hispanic Linguistics Symposium, ed. Jennifer Cabrelli Amaro et al., 143-155. Somerville, MA: Cascadilla Proceedings Project.

wh +esk, and cleft structures. Bulgarian, like English, relies exclusively on long distance *wh*-movement. However, unlike English, it licenses multiple *wh*-fronting. Slavkov specifically asks if adult language learners resort to alternative, potentially less complex constructions, before they have fully acquired long distance *wh*-movement. He offers examples (2) and (3) below:

(2) What$_i$ do you think t_i Mary is eating t_i (Long Distance *wh*-movement)
(3) Do you know what$_i$ Mary is eating t_i (Biclausal with short *wh*-movement)

Example (2) demonstrates *wh*-movement outside the clause, while (3) provides an alternative and less complex construction which permits a shorter movement. Other potential question formation strategies include monoclausal substitutes and medial-copying. It has been argued that spell-out of an overt intermediate copy (a 'bridge') makes the derivation simpler, where example (2) might be read as *What*$_i$ *do you think what*$_i$ *Mary is eating* t_i.

Slavkov finds that 66 second language (SL) learners of English resort to alternative constructions with a lower degree of complexity. Furthermore, while both groups' avoidance strategies are similar, the rates at which they turn to each strategy differ. The Bulgarians produce more monoclausal and medial constructions, while the French produce more LD and biclausal constructions. Example responses are represented in (4-7) respectively:

(4) Where is he? (Monoclausal)
(5) What do you think what he's doing now? (Medial-copying)
(6) What do you think she's reading? (Long distance *wh*-movement)
(7) Do you know who buy eggs? (Biclausal with short or no *wh*-movement)

Slavkov concludes that persistent difficulties are found in the production of complex *wh*-questions in L2 English. Furthermore, a high production of LD *wh*-movement by both populations may be the result of positive transfer. Both French and Bulgarian permit such constructions. To further support this analysis, he explains that constructions not licensed in the L2 are infrequently produced. Finally, some SL learners do resort to constructions not supported by either the L1 or the L2 grammar. He argues that such constructions may alleviate working memory load, providing an overt intermediate 'bridge' (i.e., medial copying).

3. Interrogative Inversion in *Wh*-questions

As Cuza (2012) explains, Spanish and English diverge in the way they represent *wh*-questions. In Spanish, there is subject-verb inversion (VS) in both matrix and embedded *wh*-questions (with the exception of the *por qué* "why" *wh*-extraction site). On the other hand, in English, there is only subject-verb inversion in matrix *wh*-questions and never with the lexical verb (e.g., Goodall, 2011; Radford, 2004). Furthermore, English matrix *wh*-questions require insertion of the finite auxiliary *do*.

(8) a. ¿Qué quería el vendedor? (SPAN Matrix)
wh-phrase + lexical verb + subject

b. What did the salesman want? (ENG Matrix)
wh-phrase + auxiliary *do* + subject + lexical verb

(9) a. María le preguntó a su padre [qué quería el vendedor]. (SPAN Embedded)
wh-phrase + lexical verb + subject

b. Mary asked her dad [what the salesman wanted]. (ENG Embedded)
wh-phrase + subject + lexical verb

Examples (8) and (9) summarize the typological differences across these two languages. The remaining paragraphs in this section motivate and discuss the various merge operations within each construction.[1]

3.1. The Syntax of Spanish and English

In order to understand the differences in Spanish and English question formation it is first important to consider the strength metaphor (e.g., Pesetsky & Torrego, 2001; Rizzi, 1996). For example, in Spanish, there is V-to-T movement because Spanish is a strong finite T language. This specifically means that the Tense (Tns) affix has a strong V-feature. This feature requires the Tns affix to have a verb attached to it. In the absence of an auxiliary merger, the main verb must be attracted to the T position, as represented by (10) below.

On the other hand, a weak finite T language, such as English, contains a Tns affix with a weak V-feature (e.g., Radford, 2004). By virtue of its weakness, this feature cannot attract the main verb. As a result, in the absence of auxiliary merger, the Tns affix is attached to the verb by being lowered to it during a process called affix hopping. That is to say, the verb remains in situ, as demonstrated by example (11). The important contrast between strong and weak finite T languages is how the Tns-feature attaches to the main verb (assuming the absence of an auxiliary).

(10) Tú tirar+Tns ~~tirar~~ la pelota. (Spanish: Strong finite T)

(11) You ~~Tns~~ throw+Tns the ball. (English: Weak finite T)

Assuming V-to-T movement in Spanish, as in (10), it is easy to visualize the subject in the [Spec, TP] position (i.e., Rizzi, 1996). This is the assumption that I am making for the purpose of this study[2].

In order to convert a Spanish declarative clause into an interrogative clause one inverts the verb+Tns element with the subject (e.g., Rizzi, 1996). This inversion can be analyzed as T-to-C head movement which is triggered by a strong C[3]. That is to say, C contains a question particle Q with a strong T-feature. Once arriving in C, the verb+Tns element attaches to Q and becomes verb+Tns+Q, as in example (12). As demonstrated by auxiliary inversion in questions, English is also a strong C language. However, auxiliary verb insertion is required in order to demonstrate the interrogative inversion. (Think about the weak finite T and how T will otherwise be empty because of affix hopping.) See example (13) below:

(12) Tirar+Tns+Q tú ~~tirar+Tns~~ ~~tirar~~ la pelota (Spanish: Strong finite T and Strong C)

(13) Do+Tns+Q you ~~do+Tns~~ throw the ball (English: Weak finite T and Strong C)

Thus far, all the examples have considered main clauses. Interestingly, embedded clauses behave differently. In English, interrogative inversion only occurs in main clauses (Radford, 2004). That is to say, in embedded clauses, C is without the strong T feature that attracts the verb+Tns element. This explains why the embedded *wh*-question in example (14) is grammatical. If one converts this sentence

---

[1] For an extra-syntactic analysis on subject-verb inversion see Goodall (2004). He argues that *wh*-questions strain working memory capacity because the *wh*-phrase must be held in working memory until it can be assigned to a gap. This occurs when the head (that subcategorizes for the gap) is processed. Thus, there is a greater strain on working memory when the subject is pre (as opposed to post) verbal (i.e., the head is processed later).

[2] For an alternative view see Radford (2004) and Zagona (2002). Radford argues that in English, the subject originates in [Spec, VP] and is then attracted to the [Spec, TP] because of an [EPP] feature of T. This feature requires the [Spec, TP] to project a subject with person and number properties. Zagona (2002) on the other hand argues that subject movement in Spanish is optional depending on the functional features of the particular sentence.

[3] See Goodall (2001) or Barbosa (2001) for an argument against T-to-C movement in Spanish.

into a direct question, interrogative inversion is required. Hence, there is 'dummy do' or auxiliary merger in T, followed by T-to-C movement, as demonstrated by (15).

(14) He asks when you throw the ball.

(15) a. He asks: "When do you ~~do~~ throw the ball?"
b. *He asks: "When you throw the ball?"

On the other hand, Spanish does display subject-verb inversion in embedded *wh*-questions. Without this inversion the sentence sounds odd, as in example (16). This inversion is triggered by the same strong T feature of interrogative C displayed in Spanish and English matrix questions.

(16) a. María me preguntó [cuándo **regresa Juan**]
b. ?María me preguntó [cuándo **Juan regresa**]
"Mary asks me when John returns"

A final movement one must consider in *wh*-questions is *wh*-movement. Unlike the V-to-T and T-to-C movements, *wh*-movement is a maximal projection because it projects into the specifier as opposed head of the phrase, specifically the [Spec, CP] (e.g., Rizzi, 1996). This movement is motivated by both the extended projection principle [EPP], as well as the [wh] feature within C. The [EPP] feature "…requires C to project, as its specifier, an expression which matches some feature of C" (Radford 2004: 198). It follows that the *wh*-expression, which matches the [wh] feature of C, will be projected to [Spec, CP].

In summary, I assume that subjects in English and Spanish are in [Spec, TP]. Furthermore, inversion in Spanish embedded clauses is motivated by a Strong C (i.e., a strong T feature), which is noticeably absent in English. This yields Wh-V-Subj and Wh-Subj-V word orders in Spanish and English respectively. The analysis is further supported by the fact that there is no auxiliary movement in English embedded clauses but mandatory movement in matrix clauses. Thus, Spanish and English matrix *wh*-questions are both [+inversion] due to a strong C, while embedded questions are [+inversion] in Spanish (excluding all exceptional dialects) and [-inversion] in English due to varying strengths of the T feature in C.

3.2. The Derivational Complexity of Wh-questions

In this study, which analyzes the acquisition of interrogative inversion by advanced English SL and HL learners of Spanish, I apply a strict interpretation of the Derivational Complexity Metric (DCM). Specifically, the following two conditions must be met in order for derivational complexity to condition cross-linguistic influence from language A to language B:

(1) The construction in language A must be less complex (i.e., fewer external and internal merge operations) than the equivalent construction in language B.

(2) A final step (spell-out) in the derivational analogue of language A must be an initial step in the sequence of language B.

If both of these conditions are met, then I predict the construction from language A will transfer into language B. Specifically, the final step (i.e., the spell out) in language A will serve as an avoidance strategy (alternative construction) in language B. As discussed in the previous section, I am assuming that in both Spanish and English *wh*-movement follows the same derivational analogue and the subject merges in [Spec, TP]. Thus, the complexity rating system in Table (1), where zero is the least complex, will not take these merges into account.

Table 1. Spanish and English *Wh*-question Complexity Scale

Verb movement	Complexity	Sentence type	Language
1. In-situ	0	Embedded	ENG
2. V-to-T	1	NA	NA
3. V-to-T-to-C	2	Matrix and Embedded	SPAN
3. Auxiliary merge and T-to-C	2	Matrix	ENG

The strict two condition interpretation of DCH, which I am adopting for this study, does not predict transfer of the English matrix constructions into Spanish. Recall that the difference between the two languages is that Spanish displays V-to-T-to-C movement while English displays auxiliary merge and then T-to-C movement. Thus, the sum of the external and internal merges for both languages is two. Therefore, one language's construction cannot be said to be more complex than the other. Hence, condition (1) is not met.

The metric is quite different for embedded constructions. While the English verb remains in situ (within the verb phrase), with no other merges, the Spanish verb moves from V-to-T-to-C. Thus the sum of external and internal merges in English is zero as compared to two for Spanish. Therefore, Spanish is the more complex construction, satisfying condition (1). Condition (2) is also met, where the verb merging in the [head, VP] is the final step in English and the initial step in the Spanish V-to-T-to-C movements. Thus, DCH predicts transfer of the English construction into Spanish, which includes a failure to adjust (or reset) both the inversion parameter and the feature strength properties.

3.3. Evidence from Adult Heritage Language Learners

As discussed in Footnote (1), Goodall (2004) offers a non-syntactic perspective on what motivates interrogative inversion in Spanish. He specifically argues that inversion minimizes working memory load. One prediction that falls from his hypothesis is that the complexity associated with embedding is not an important factor in determining the acceptability of the verb in the pre and post subject position. As previously discussed, Derivational Complexity Hypothesis predicts just the opposite. Recent studies on second language learners and heritage language learners, adults and children alike have provided support for DCH (i.e., Cuza, 2012; Cuza and Strik, 2012; Frank, 2013).

Cuza (2012) investigated the acquisition of interrogative (subject-verb) inversion in both matrix and embedded contexts by adult Spanish heritage language (HL) learners. Results from an acceptability judgment task (AJT) suggest that HL learners only reject ungrammatical embedded questions (questions without subject-verb inversion) 12% of the time. On the other hand, they have much less trouble with matrix questions (70% rejection). See examples (17) and (18) respectively:

(17) *No sé a quién Rosa le entregó el violín. (Embedded "ungrammatical" item)
"I don't know to whom Rosa gave the violin"

(18) *¿A quién Luis conoció en Paris? (Matrix "ungrammatical" item)
"Who did Luis meet in Paris?"

Results from a written production task follow the same trend. These results support his hypothesis that embedded constructions will be more difficult to acquire because this is where "…English and Spanish entail a different grammatical mechanism: obligatory inversion in Spanish (T-to-C) and no movement in English" (27).

In summary, the present study is motivated by the thin body of work that has investigated derivational complexity effects in adult bilingual populations (e.g., Slavkov, 2011). Furthermore, a strict interpretation of DCH demonstrates that matrix and embedded *wh*-questions in Spanish and English are an interesting linguistic construction to analyze. Finally, recent literature that has investigated interrogative inversion in Spanish and English bilinguals has shown persistent difficulties in producing subject-verb inversion in embedded contexts. The aforementioned literature has motivated the following research questions, which will guide the remainder of the study.

(1) Do advanced adult HL and SL learners (i.e., independent variable *population*) demonstrate similar subject-verb inversion production patterns in matrix and embedded *wh*-questions?
(2) Do the two populations produce more instances of inversion in the matrix or the embedded contexts (i.e., independent variable *sentence type*)?
(3) Do the results provide evidence for or against a strict Derivational Complexity Hypothesis?

I predict that the HL learners will be less affected by cross-linguistic influence than the SL learners. HL learners have acquired Spanish to some degree as an L1 (e.g., Montrul, 2008). I speculate that the HL learner advantage stems from their earlier age of exposure and greater use of Spanish outside of the classroom, in a more natural setting (e.g., Cuza & Frank, 2010).

Furthermore, in accordance with Cuza (2012), I predict that the sentence type will be a significant factor. Specifically, both groups will produce fewer instances of inversion in the embedded than the matrix *wh*-question tokens. Importantly, it is in the former construction that the conditions of the DCH are met. That is to say, Spanish is the more complex language and the external merge of the lexical verb in the English embedded constructions is an initial step in the syntactic derivation of the Spanish ones. This is not the case for matrix constructions. Thus, I predict that the English features and parameters will transfer into Spanish embedded but not matrix contexts, providing support for DCH.

4. The Study

4.1. Participants

A total of 27 (n=27) bilingual adults participated in this study, eleven of them HL learners and sixteen SL learners. Participants were first asked to fill out a language history questionnaire previously used by Cuza & Frank (2011), which contained questions on language use, exposure, and education.

The questionnaire was followed by a written proficiency test. This test consisted of a cloze passage (with three multiple-choice response options for each blank) from a version of the *Diploma de Español como Lengua Extranjera* (DELE) and a multiple choice vocabulary section from an MLA placement test. The maximum possible score on the proficiency test was 50. Following previous research, scores between 40 to 50 points were considered an advanced proficiency level, while scores between 30 to 39 and 0 to 29 were considered intermediate and low respectively (e.g., Cuza & Frank, 2011; Montrul & Bowles, 2010; Montrul & Slabakova, 2003). Table (2) below summarizes the participants' profile compiled on the basis of these two sets of questionnaires.

Table 2. Participant Profile

Elicited Information	HLL (n=11)	SLL (n=16)
1. Place of birth	10 USA, 1 Mexico	16 USA
2. Level of education	Undergraduate Student	Graduate Student
3. Language spoken as a child	SPAN (8/11), SPAN and ENG (3/11)	ENG (16/16)
4. Mean Age	20	27
5. Mean Linguistic Proficiency	41/50 (8 advanced, 3 high intermediate)	44/50 (15 advanced, 1 high intermediate)

These two bilingual populations were specifically selected in order to compare how their different language histories and contrasting ages of initial exposure would affect the results. Proficiency (high intermediate to advanced), age (>18), and level of education (University) were controlled for so that these factors could be held as constants throughout the analysis. The difference in the number of participants in each group (eleven versus sixteen) is simply a product of recruitment efforts.

4.2. Methodology

Data elicitation consisted of a written production task. Specifically, participants were presented with several sentences, which included matrix and embedded *wh*-questions with linguistic fragments in a random order. They were then asked to unscramble the fragments and to rewrite them in the order they preferred (Abing, 2013). A total of 12 test tokens (6 matrix and 6 embedded) along with 10 fillers were utilized in this study. Two versions of the test with the fragments scrambled in different orders were created so as to avoid priming effects. *Wh*-extraction sites included the locative adjunct *dónde* "when", the temporal adjunct *cuándo* "where", and the prepositional phrase *con quién* "with whom". See examples (19) and (20) below:

19. ¿/con quién / Julia y Marco/ a Texas/ volvieron/?
Mi oración: ¿Con quién volvieron Julia y Marco a Texas (Expected Response)

20. /la esposa/le preguntó a José/ su hermano/con quién/ bailó/.
Mi oración: La esposa le preguntó a José con quién bailó su hermano. (Expected Response)

Excluding exceptional dialects, Goodall (2004) argues that non subject-verb inversion in adjunct (with the exception of *por qué* "why") and argument *wh*-extraction sites is not accepted by native Spanish speakers. Furthermore, he claims that arguments are even less acceptable than adjuncts. *Wh*-type was not considered as an independent variable for the purposes of the present study. The objective was to compare the instances of inversion across the population (SL and HL learner) and sentence type (matrix and embedded) independent variables. However, a future study might consider whether the *wh*-extraction site is a significant factor in the production of interrogative inversion.

5. Results

5.1. Group Analysis

As previously mentioned, the two independent variables of interest in this study are *sentence type* (two levels: matrix and embedded) and *population* (two levels: HL and SL learner). The dependent variable is subject-verb inversion realized. The measurement of this variable is a zero for no inversion and a one for inversion. I have borrowed Cuza's (2012) monolingual control group results from a similar dehydrated sentence task in order to assist with the following analysis. A quick glance at the mean subject-verb inversion realized by group indicates that the control group produced subject-verb inversion in both sentence types over 95% of the time. This is ceiling performance (Table 3).

The HL learner group produced inversion 71% of the time in embedded contexts and 86% of the time in matrix ones. These results support Cuza's (2012) findings that HL learners produce non-inversion constructions, specifically in embedded contexts, at a greater rate than the monolingual norm. The SL learners followed a similar trend, 60% and 87% respectively.

Table 3. Mean Subject-verb Inversion Realized by Group

Population	Embedded	Matrix
1. *Control (n=10) (Cuza, 2012)*	*(>.95)*	*(>.95)*
2. HL learners (n=11)	0.71	0.86
3. SL learners (n=16)	0.60	0.87

Because all variables are categorical, and specifically the dependent variable (DV) is a binary outcome, a binary logistic regression was run. Results from this statistical analysis show a main effect for *sentence type,* but not for *population*. Furthermore, there is a significant interaction between these two variables, as summarized in Table (4) below:

Table 4: Binary Logistic Regression Analysis for Subject-Verb Inversion

Independent Variable	*z*-value	DoF	Significance
1. Sentence Type (MAT, EMB)	-0.997	1	p=0.002
2. Population (HLL, SLL)	-1.202	1	p=0.461
3. Sentence Type * Population	0.564	1	p=0.003

In Figure (1) below, the independent variable (IV) *population* has been collapsed. This demonstrates that the two groups (as one unit) produced more instances of inversion in matrix than in embedded contexts. Specifically, the results from this collapsed IV resemble the control group results much more in the matrix (0.87 versus 0.95) than in the embedded constructions (0.66 versus 0.95). Moreover, the embedded results are right around chance performance.

The logistic regression was then respecified for the embedded sentence type. This returned no significant difference between learner groups ($z= -1.01$, $p>0.05$). Not surprisingly, specifying for the matrix sentence type also did not return a significant difference between groups ($z= 0.06$, $p>0.05$). In sum, neither of the adult bilingual populations is producing significantly more instances of inversion than the other in either sentence type.

In Figure (2), the independent variable *sentence type* (two levels: matrix and embedded) has been collapsed. This graph indicates that there is very little overall performance difference (instances of inversion) between the HL and SL learner test groups, 0.79 and 0.74 respectively.

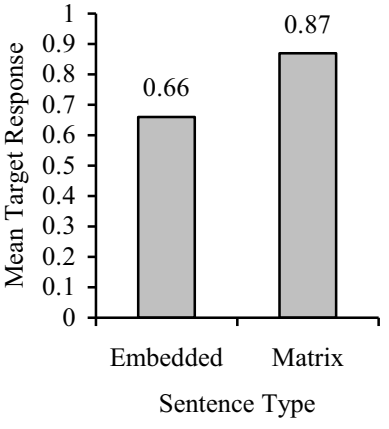

Figure 1. Collapsed Population

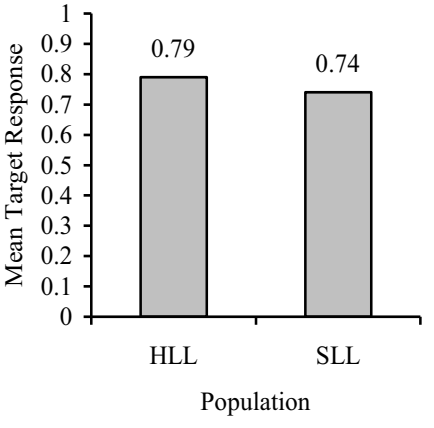

Figure 2. Collapsed Sentence Type

This motivated a second respecification of the logistic regression, specifically for the two levels of the independent variable *population* (SL and HL learner). Results indicate that *sentence type* is significant for SL learners ($z= -2.75$, $p<0.05$). Interestingly, it misses significance for HL learners ($z= -1.34$, $p>0.05$).

In summary, collapsing the independent variable *population* demonstrates that the bilinguals as one unit produce significantly more instances of inversion in matrix than in embedded contexts. Furthermore, the two bilingual populations performed similarly to each other in each of the sentence types. That is to say, statistical significance was not reached between the two populations in either the matrix or embedded contexts. However, preliminary data do suggest that HL learners may be more sensitive to inversion in embedded contexts than SL learners (see Individual Analysis section).

Surprisingly, only the SL learner population actually produced a significantly different number of inversions across the independent variable *sentence type*. This indicates that the SL learners are carrying more of the statistical weight in the main effect found for sentence type. Though the HL learner data missed significance, they do follow a trend that mirrors the SL learner data. Specifically, both populations produce more instances of inversion in matrix contexts.

5.2. Individual Analysis

The following Tables (5) and (6) were created in order to analyze the individual performance of the participants. They isolate the instances of inversion within each group in the matrix and embedded constructions respectively. In both cases the participants were separated into four categories, which were dependent upon how many examples of inversion they produced. Category one is labeled *high* inversion. *High* means the participant produced 100% of the tokens with inversion. Category two is *medium* inversion. *Medium* means the participant produced at least 50% of the tokens (but not all of them) with inversion. Category three is labeled *low* inversion. *Low* means the participant produced less than 50% of the tokens (but at least one) with inversion. Finally, category four is *null* inversion. *Null* is defined as no instances of inversion produced.

Table 5. Matrix Subject-verb Inversion Realized, Individual Analysis

	High	Medium	Low	Null
Control (Cuza, 2012)[4]	*10 of 10*	*0 of 10*	*0 of 10*	*0 of 10*
HLL	7 of 9	2 of 9	0 of 9	0 of 9
SLL	13 of 16	0 of 16	3 of 16	0 of 16

Table 6. Embedded Subject-verb Inversion Realized, Individual Analysis

	High	Medium	Low	Null
Control (Cuza, 2012)[5]	*9 of 10*	*1 of 10*	*0 of 10*	*0 of 10*
HLL	8 of 11	0 of 11	1 of 11	2 of 11
SLL	5 of 14	2 of 14	3 of 14	4 of 14

As demonstrated in Table (5), in terms of matrix tokens, each of Cuza's (2012) controls falls into category one, high inversion. The majority of the bilingual participants perform similarly well. Specifically, 7/9 heritage language learners and 13/16 second language learners display high inversion. The remainder of the HL learners (2/9) fall into the medium category, while the remainder of the SL learners (3/16) fall into the low category. Without this individual analysis the distinction between instances of inversion in the bilingual groups in these matrix contexts is lost (mean=0.86 and 0.87 respectively).

One reviewer wondered why the learners categorized as "low" with respect to subject-verb inversion in matrix clauses were included in the analysis. S/he argues that knowledge of inversion in matrix clauses should be a prerequisite to the study of inversion in lower clauses. I argue that all of the learners did demonstrate knowledge of inversion in matrix clauses. Specifically, they all produced at least one instance of inversion. This is in stark contrast with the six learners who fall into the "null" category for inversion in embedded clauses. Furthermore, all participants are advanced learners of Spanish, as demonstrated by the DELE proficiency exam. This was the determining factor for whether the participant was included in the study or not.

In Table (6), which represents the embedded tokens, 9/10 of Cuza's (2012) controls fall into the high category, while the last one falls into the medium category. In terms of the bilingual populations,

[4] Cuza (2012) used a slightly different cutoff for category one. Specifically, high inversion included five or six instances of inversion out of a total of six tokens.

[5] See Footnote (4) above.

8/11 of the HL learners produce high instances of inversion, while 1/11 produces low inversion and 2/11 produce no inversion. The SL learners display the most variation in these embedded contexts. They are almost equally distributed across all four categories. 5/14 have high production, 5/14 fall into a combined medium and low category, and 4/14 display no instances of inversion.

This individual analysis demonstrates that the majority of the HL learners display a native-like inversion production in embedded contexts. However, three of them do not. 9/14 SL learners do not fall into the high inversion category. This demonstrates a clear contrast between the two bilingual populations, despite the fact that statistical significance was not found in the group analysis ($z= -1.01$, $p>0.05$, and mean= 0.71 and 0.60 respectively).

An explanation for the individual variation within each group has proven elusive. There was no evidence of a direct association between instances of inversion in *wh*-questions and proficiency level or time spent abroad. Furthermore, those in the low categories in the matrix contexts were not the same as those who were in the low or null categories in the embedded contexts.

The reader may have noticed from Tables (5) and (6) that two HL learners have been excluded from the matrix token analysis and two SL learners from the embedded token analysis. Importantly, two HL learners placed the subject in the sentence initial position in the matrix clauses simply because it was the only word that was capitalized (i.e., proper name). Because both participants explained that they misunderstood the exercise, they were not included in the matrix analysis.

Furthermore, two SL learners produced their embedded tokens as direct questions. As one reviewer explains, direct question formation can certainly be seen as an alternative strategy to embedded questions. However, the present study strictly analyzes the typological differences associated with equivalent Spanish and English matrix and embedded clauses. Producing an embedded token as a direct question entails converting the sentence to a matrix construction. Thus, the embedded tokens from these two participants would have to be analyzed as matrix tokens, while being treated as embedded ones for every other participant. This inconsistency is why the direct question responses were not factored in to the overall analysis.

In summary, while the group analysis did not suggest that HL learners have an overall more complete representation of subject-verb inversion in *wh*-questions, the individual analysis tells a slightly different story. Specifically, HL learners produce more instances of inversion in embedded contexts than SL learners. This is not the case for matrix contexts, where both groups display many instances of inversion, though less than the monolingual norm. These results are supported by the fact that the SL learners produced significantly more instances of inversion in matrix than in embedded contexts, while the HL learners missed significance in the *sentence type* independent variable.

6. Discussion

6.1. Research Questions Revisited

In this study I asked whether derivational complexity as a condition for CLI provides a possible explanation for the results. Both of the conditions which I have used to guide the strict interpretation of the DCH are met in embedded contexts. Specifically, (1) Spanish is the more complex language and (2) the external merge of the lexical verb ([head, VP]) in the English constructions is an initial step in the syntactic derivation of the Spanish ones (V-to-T-to-C). This is not true for matrix contexts. DCH therefore predicts that the English embedded question word order (SV) will transfer into the Spanish embedded construction but not into the matrix construction.

Given that English is the *other* language of both test groups, I hypothesized that if derivational complexity conditions cross-linguistic influence, then both bilingual groups would display fewer instances of inversion in embedded contexts. This prediction was supported by the results. Both bilingual populations produce more instances of inversion in matrix than in embedded contexts, though only SL learners reached statistical significance. Furthermore, preliminary results do suggest that HL learners produce more instances of inversion in embedded contexts than do SL learners, despite similar proficiencies, ages, and levels of education.

Radford (2004) defines a feature as a language specific device used to describe a particular grammatical property. He goes on to explain that a parameter is a dimension of grammatical variation

within or across languages. Importantly, transfer from English into Spanish in embedded contexts entails that the [+inversion] parameter is not in place in bilingual Spanish. This is because the English weak feature of T (in C) is transferred into Spanish. Specifically, bilinguals have trouble adjusting the features, or resetting the parameter. I argue that this is triggered by derivational complexity, as strictly defined.

Furthermore, this persistent difficulty is partially conditioned by the input, where by virtue of being bilingual both the SL and HL learners receive less input in each language than a monolingual would. Specifically, the use of, and exposure to, embedded *wh*-questions in a minority language or classroom setting may be insufficient to offset transfer effects. Thus, the conditions for native-like attainment are not ideal, where English would be infrequently used and dominance would have shifted to Spanish (e.g., Birdsong, 2009).

What complicates the learner's task still more is that subject-verb inversion is not categorical. For one, Spanish is a pro (drop) language, where an over subject is not always obligatory (e.g., Zagona, 2002). Furthermore, in terms of the [verb movement] parameter, Spanish is a mixed language (e.g., Ayoun, 2005). Specifically, Ayoun argues that Spanish also instantiates constructions that are both plus and minus verb movement with respect to adverb placement and floating quantifiers. Finally, the *wh*-extraction site *por qué* "why" (adjunct) adds more variability and "noise" to the input. In this context, subject-verb inversion is argued to be optional (e.g., Goodall, 2004).

In terms of processing constraints, the role of working memory load in determining subject and verb word orders is either minimal or trumped by derivational complexity effects. Goodall (2004) argues that Spanish speakers produce interrogative inversion because it lowers the processing cost. Specifically, there is a greater strain on working memory when the subject is pre (as opposed to post) verbal. That is to say, the head is processed later. Nevertheless, bilinguals frequently produce subject-verb word orders in embedded contexts. What's more, Goodall (2004) claims that embeddedness is not an important factor in determining the acceptability of subject and verb word orders. The results of this study suggest that the complexity associated with embedded constructions is at the very least correlated with a decrease in interrogative inversion.

The results do support my expectation that the heritage language learners would behave more like the control group than would the second language learners. The age of initial exposure seems to be a relevant factor in language performance. Montrul (2008) suggests, in what she titles the Weaker Language L1 Hypothesis, that simultaneous bilinguals acquire both languages with the same cognitive and linguistic means regardless of dominance. Taking this hypothesis into account, it is safe to claim that many heritage speakers have acquired Spanish to some degree as an L1. Thus, they have access to the implicit learning mechanism, specifically Spanish features and parameters (e.g., Montrul, 2008). The implicit learning mechanism is an innate language faculty that a child is born with. This access allows children to select for the particular features and parameters of his/her particular language, amongst all of the options in human language.

On the other hand, the SL learner group, though more advanced (Proficiency rating of 44 versus 41), only began speaking Spanish after puberty. The results from this study support the maturational effects theory, where adults learning a second language do not have access to the implicit learning mechanism (i.e., the features and parameters) of the L2. A native speaker of English learning Spanish will use English knowledge as the basis for acquiring Spanish. This will be to their detriment when the features and parameters diverge across the two languages. Specifically, they must rely on their explicit knowledge and must practice a lot in order to overcome cross-linguistic influence. This theory can account for why the two bilingual populations did not perform identically.

6.2. Conclusion

In conclusion, cross-linguistic influence conditioned by derivational complexity, input conditions, the complexity associated with an embedded construction, and maturational effects together easily provide an explanation for the results of this study. These four factors need to be further investigated in bilingual populations in order to determine to what extent they are directly associated (as opposed to correlated) with bilingual language acquisition. For example, an anonymous reviewer explains that if

derivational complexity triggers incomplete representation in bilingual Spanish (i.e., at the level of features and parameters), then one should expect to see derivational complexity effects in other Spanish properties, and in other language pairs in general.

An anonymous reviewer also explains that complementing a production task with a representation task would strengthen the study. H/she argues that one is left wondering whether there is a competence/performance divergence, and whether the production results can simply be explained by processing constraints (i.e., bilinguals have the proper representations without performance data to match). Cuza (2012) did incorporate both types of tasks (production and representation) and found persistent difficulties in embedded contexts in the HL learner population. Thus it does not appear that the difficulties can be reduced to limitations on working memory.

However, I do agree that not including a task that elicits mental representations is a limitation in my study. If it can be determined the extent to which the linguistic representations and the processing capabilities in Spanish differ between the HL and SL learner groups, this would inform us as to the nature of problems related to derivational complexity.

Other limitations in the present study include the number of participants (n=27), as well as, the absence of my own control group. While this study primarily aims to compare the two bilingual populations (SL and HL learners), I do at times turn to a monolingual control group to assist in the discussion (see Cuza, 2012). His task design and tokens are similar to mine but not entirely the same.

A future study should investigate the reverse language combination. For example, one might ask if SL learners of English (L1 Spanish) have a complete representation of English matrix and embedded *wh*-questions (Frank, 2013). The conditions for Derivational Complexity Hypothesis are not met in this direction. I would hypothesize that *sentence type* will not be a significant factor in English constructions. Specifically, the participants will not produce more instances of inversion in matrix than in embedded contexts. If they do, and behave much like the participants in this study, then this will be evidence against DCH. A similar study can be designed around a different language combination altogether, such as Catalán or Romanian with Spanish. Because these languages have analogous *wh*-question constructions, DCH would similarly not apply.

Appendix

(List of Tokens: 1-6 embedded, 7-12 matrix)

1. Jorge le preguntó a Celina cuándo lava Marta los platos.
2. La esposa le preguntó a José con quién baila su hermano.
3. María le preguntó a su padre dónde aprendió Susana la gramática.
4. La esposa le preguntó a José cuándo encontró las llaves Sammy.
5. Susana le preguntó a Ana dónde ve Eric la televisión.
6. Felipe le preguntó a José con quién ganó el hijo la competencia.

7. ¿Dónde dio Cristo el regalo a su madre?
8. ¿Cuándo termina Pilar con la clase?
9. ¿Con quién volvieron Julia y Marco a Texas?
10. ¿Con quién cantó la canción María?
11. ¿Cuándo le echó Dani azúcar al café?
12. ¿Dónde escuchó Jorge la música?

References

Abing, Jesse. (2013). Resolving Spanish word-order issues: An analysis of learner-learner interactions between heritage and second language learners. Presented at the Dialogue in Multilingual, Multimodal, Multicompetent Communities of Practice Workshop, International Association of Dialogue Analysis. The University of Texas. Austin, TX.

Barbosa, Pilar. (2001). On inversion in wh-questions in Romance. *Subject inversion in Romance and the theory of universal grammar*, 20-59.

Birdsong, David. (2009). Age and the end state of second language acquisition. In W.C. Ritchie &T.K. Bhatia (Eds.), *The New Handbook of Second Language Acquisition* (pp. 401-424). Bingley, UK: Emerald Group.

Cuza, Alejandro. (2013). Crosslinguistic influence at the syntax proper: interrogative subject-verb inversion in heritage Spanish. *The International Journal of Bilingualism*, *17*(1), 71-96.

Cuza, Alejandro, & Strik, Nelleke. (2012). Patterns of morphosyntactic convergence and child L1 attrition: Evidence from subject-verb inversion in Spanish-English bilingual children. Presented at the 42nd Linguistics Symposium on Romance Languages, April 19-22. University of Southern Utah.

Cuza, Alejandro, & Frank, Joshua. (2011). Transfer effects at the syntax-semantics interface: The case of double-*que* questions in heritage Spanish. *The Heritage Language Journal*, *8*(2), 66-89.

Cuza, Alejandro, & Frank, Joshua. (2010). The acquisition of double *QUE* questions in Heritage and L2 Spanish. In *Proceedings of 2010 Canadian Linguistic Association Conference.*

Frank, Joshua. (2013). "Good enough" representation in L2 Spanish and English wh-questions. Presented at the 43rd Linguistics Symposium on Romance Languages (LSRL). City University of New York (CUNY). New York, NY.

Goodall, Grant. (2011). Syntactic Satiation and the Inversion Effect in English and Spanish Wh-Questions. *Syntax*, *14*(1), 29-47.

Goodall, Grant. (2004). On the syntax and processing of Spanish wh-questions. In B. Schmeiser, V. Chand, A. Kelleher, & A. Rodriguez (Eds.), Proceedings of WCCFL (Vol. 23, pp. 101–114). Somerville, MA: Cascadilla Press.

Goodall, Grant. (2001). The EPP in Spanish. In *Objects and other subjects* (pp. 193-223). Springer Netherlands.

Hulk, Aafke. & Müller, Natascha. (2000). Bilingual first language acquisition at the interface between syntax and pragmatics. *Bilingualism: Language and Cognition,* 3(3), 227-244.

Jakubowicz, Celia. (2011). Measuring derivational complexity: New evidence from typically developing and SLI learners of L1 French. *Lingua, 121,* 339-351.

Jakubowicz, Celia. (2005). The Language Faculty: (Ab)normal development and interface constraints. Presentation, GALA 2005, Siena.

Jakubowicz, Celia, & Strik, Nelleke. (2008). Scope-marking strategies in the acquisition of long distance wh-questions in French and Dutch. *Language and Speech*, 51, 101–132.

Montrul, Silvina, & Bowles, Melissa. (2010). Is grammar instruction beneficial for heritage language learners? Dative case marking in Spanish. *The Heritage Language Journal, 7*(1), 47–63.

Montrul, Silvina. (2008). Incomplete acquisition in bilingualism. Re-examining the age factor. Amsterdam: John Benjamins.

Montrul, Silvina, & Slabakova, Roumyana. (2003). Competence similarities between native and near-native speakers: An investigation of the Preterite/Imperfect contrast in Spanish. *Studies in Second Language Acquisition, 25*(3), 351–398.

Pesetsky, David, & Torrego, Esther. (2001). T-to-C movement: Causes and consequences. In M. Kenstowicz (Ed.), *Ken Hale: A life in language* (pp. 355–418). Cambridge, MA: MIT Press.

Radford, Andrew. (2004). *Minimalist Syntax: Exploring the Structure of English*. Cambridge: Cambridge University Press.

Rizzi, Luigi. (1996). Residual verb second and the wh-criterion. In A. Belletti & L. Rizzi (Eds.), *Parameters and functional heads* (pp. 63–90). Oxford, UK: Oxford University Press.

Slavkov. Nikolay. (2011). Derivational Complexity Effects in L2 Acquisition. In M. Pirvulescu et al (Eds.), *Proceedings of* GALANA 2010, 227-240. Somerville, MA: Cascadilla Proceedings Project.

Strik, Nelleke. (2012). Wh-questions in child bilingual acquisition of French: Derivational Complexity and cross-linguistic influence. *Canadian Journal of Linguistics, 57(1),* 133-151.

Strik, Nelleke. & Pérez-Leroux, Ana Teresa. (2011). Jij doe wat girafe? *Wh*-movement and inversion in Dutch-French bilingual inversion. *Linguistic Approaches to Bilingualism, 1(2),* 175-205.

Yip, Virginia. & Matthews, S. (2009). Cross-linguistic influence in bilingual and multilingual contexts. Invited paper at the *2009 International Symposium on Bilingualism* (ISB7). University of Utrecht, The Netherlands.

Zagona, Karen. (2002). *Sintaxis Generativa del Español*. Madrid: Visor Libros.

Selected Proceedings of the 16th Hispanic Linguistics Symposium

edited by Jennifer Cabrelli Amaro, Gillian Lord, Ana de Prada Pérez, and Jessi Elana Aaron

Cascadilla Proceedings Project Somerville, MA 2013

Copyright information

ISBN 978-1-57473-459-1 library binding

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Frank, Joshua. 2013. Derivational Complexity Effects in Bilingual Adults: Instances of Interrogative Inversion in Spanish. In *Selected Proceedings of the 16th Hispanic Linguistics Symposium*, ed. Jennifer Cabrelli Amaro et al., 143-155. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #2931.