THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF SPANISH, ITALIAN, AND PORTUGUESE

PROCESSING COSTS WHILE BILINGUALS READ SPANISH-ENGLISH CODE SWITCHES: AN EYE TRACKING STUDY

MARIE-CHRISTINE THEBERGE
Spring 2012

A thesis
submitted in partial fulfillment
of the requirements
for baccalaureate degrees in
Spanish and Communication Sciences and Disorders
with honors in Spanish

Reviewed and approved* by the following:

Dr. Paola Dussias
Associate Professor of Spanish, Linguistics and Psychology
Thesis Supervisor

Dr. John Lipski
Edwin Erle Sparks Professor of Spanish & Linguistics
Honors Adviser

* Signatures are on file in the Schreyer Honors College

ABSTRACT

The ability of a speaker to switch from one language to another, often several times in a single utterance, is referred to as code-switching. Code-switches can occur inter-sententially, between sentences (e.g., My sister arrived. **Ya podemos empezar a comer**/My sister arrived. We can start eating now), or intra-sententially, within one single sentence (e.g., Mi amiga dijo que **her sister** va a tomar **several science courses** este semestre/My friend said that her sister will take several science courses this semester). One might imagine that switching back and forth between two languages, especially intra-sententially, where we can see an interaction of the two grammatical systems, would be very difficult; yet we see bilingual speakers code-switch seamlessly, without pauses or hesitation, suggesting that the use of two languages is a natural process.

Although a very common occurrence among bilinguals, code-switching has been a largely unexplored area of study from a psycholinguistic perspective, and particularly from the perspective of the language comprehender. To fill this gap in the literature, this study examines whether there are processing costs incurred while Spanish-English bilinguals read code-switched sentences. Code-switching is thought to occur mostly in speech, but it has also permeated written texts, especially more informal correspondence, such as email and text messages. During production (i.e., while speaking or writing), bilingual speakers are in control of where a code-switch will occur. However, for the bilingual listener or reader, code-switches might be unexpected and thus potentially difficult to process.

Using eye-tracking technology, bilinguals' eye moments were while they read different types of intra-sentential code-switches, and analyses consisted of comparing the time taken to read different conditions. Longer reading times are normally associated with some type of processing cost. Results show that switches that are produced more frequently in bilingual speech are also easier for readers to comprehend. This finding is the first type of evidence in the code-switching literature showing a link between production and comprehension. It is significant because it is in line with current theories of monolingual sentence processing that emphasize the close relationship between comprehension and production (Gennari & MacDonald, 2009). As such, these findings add crucial data to our understanding of the architecture underlying human language processing.

TABLE OF CONTENTS

ACKNOLEDGEMENTS

I would first and foremost like to thank Dr. Paola Dussias, not only helping me complete this thesis, but also for allowing me the privilege of working in her laboratory for the last three years. She is one of the main reasons I completed my undergraduate studies at Penn State, and will forever be in her debt. I would also like to thank Rosa Guzzardo for all of her help and guidance over the past three years, and for always being a patient and knowledgeable mentor; Jorge Valdes Kroft for sparking my interest in linguistics and introducing me to research through the Center for Language Science at Penn State; and all of the participants that took the time to take part in this study because it would not have been possible without you. Lastly, I would like to thank my parents for helping me through it all.

This research was made possible by the generous funding of an Undergraduate Research Summer Discovery Grant; a Schreyer Honors College Scholarship; and a National Science Foundation grant BCS-0821924 to Paola E. Dussias and Chip Gerfen, NSF Grants BCS-0955090 and OISE-0968369 to Judith F. Kroll and Paola E. Dussias. All errors are my own.

CHAPTER 1: Introduction

There are many bilingual communities in which speakers regularly switch from one language to another, often several times in a single utterance; this phenomenon is referred to as *code-switching*. Although many second language learners often switch between languages because they have yet to gain full mastery of the second language, code-switching in balanced bilinguals does not occur due to lexical gaps, but rather because switches better express what speakers want to convey in one language or the another. In fact, code-switching often happens seamlessly, spoken without hesitation, pauses, or corrections, suggesting that it is not random interference of one language with the other, but rather a natural process that reflects a systematically controlled integration of two linguistic systems (Dussias, 2003). Code-switching, rather than representing a debasement of linguistic skill, is actually a sensitive indicator of bilingual ability (Poplack, 1980).

It is important to note, however, that code-switching does not occur with all balanced bilinguals, but rather in specific bilingual communities. Dutch-English bilinguals, for example, are very balanced but have not been found to code-switch. Code-switching with Spanish-English bilinguals, on the other hand, is common, but this does not mean that all individuals that speak both English and Spanish will code-switch. In Spain, for example, code-switching is not a common linguistic practice, whereas in Puerto Rico, it is commonplace to hear Spanish and English used interchangeably, especially among the upper class (Pérez Casas, 2008). This variance is partly due to the fact that language is as much a means of communication as it is a product of social and cultural norms, and it is therefore not surprising that there exist different attitudes towards code-switching. For some, the intermixing of two languages is informal, while for others it is a strong marker of cultural identity. This discrepancy may be due to the fact that code-switching has been stigmatized in

certain educational settings, but fully embraced in other bilingual communities (Bullock, & Toribio 2009).

One important characteristic to note with regard to code-switching is that there are two types: inter-sentential and intra-sentential switches. Inter-sentential code-switches occur at sentence boundaries, as shown in the following example:

(1) Acaba de llegar un paquete. *Let's open it.*

'A package has just arrived. Let's open it.

Conversely, intra-sentential code-switches occur at clausal boundaries within a sentence:

(2) The teacher said *que los niños están* playing outside.

'The teacher said that the kids are playing outside.'

The distinction between the two is important because although inter-sentential switches still require the ability to shift quickly between the two languages, the two systems remain separate; it is only at the intra-sentential level that the interaction between two grammatical systems can be observed (Dussias, 2003). This interaction requires greater simultaneous control of both languages, and thus provides a window into the processing of code-switches.

Code-switching is often thought of as occurring in speech, and while its ubiquitous use in bilingual speech has been extensively studied, few studies have looked at code-switching from the perspective of the bilingual reader or listener. Given the increasing presence of code-switching in printed text (particularly email and social media environments), it is important that researchers also turn their attention to how code-switches are processed by the comprehener. Although looking at auditory comprehension would be more difficult because of lack of adequate methodologies, it is still possible to examine comprehension via reading, because code-switching has also been found to permeate written

language, especially informal modes such as email. Below is an example of a code-switch used during an email correspondence, taken from Montes-Alcalá, 2005:

(3) *Después me quedé pensando por lo que dijiste que* that would make the other guy mad.

'Later I was left wondering what you had said that would make the other guy mad.'

There is limited comprehension data, and even less looking at reading comprehension, but since there is broad consensus that reading activates the system employed in auditory language processing (e.g., Perfetti, 1999), there are also advantages to looking at reading comprehension, rather than speech. When a listener hears speech, they hear a constant stream of sounds that they have to interpret and retain long enough to hear the end of the utterance. With written text, the reader can go back and reread the words, because they have the orthographic representations in from of them. Using an eye tracker allows us to trace these eye-movement regressions, giving us a visual window into the processing costs of switches. Using a comprehension questions after each sentence also allows us to elicit more natural reading comprehension.

Why do we study code-switching?

We study code-switching because it gives us a window into the interaction of the two languages. Code-switching is different than looking at both languages of a bilingual separately, because while there are set grammatical rules for both languages, there are no determined rules for how those languages interact. Even so, when code-switching data is analyzed, certain switches occur much more frequently than others, which suggests that bilinguals unknowingly code-switch under certain linguistic constraints (Poplack, 1980).

There is abundant data in the literature on production, yet very little on the comprehension of code-switching, but looking at the comprehension of code-switching is

important for several reasons. For one, in production, code-switching is often motivated by a word or phrase in the other language that better conveys pragmatic intentions (Myers-Scotton, 2005). For bilingual listeners or readers, however, switches can be unexpected and potentially more difficult to process than within-language sentences. For any communication to take place, there must be a sender and a receiver of the message, and although the sender can plan what they want to say, the listener has to be ready to interpret the message, which could potentially be more challenging if the message is expressed using more than one language.

Relating the comprehension of code-switches to production: The Production-Distribution-Comprehension Model

Some current psycholinguistic theories of language comprehension place an emphasis on language exposure and the experience that the speaker has with language. One model that argues for the existence of a close correspondence between comprehension and production patterns is the Production-Distribution-Comprehension (PDC) framework (MacDonald & Thornton, 2009). The premise of the model rests on the idea that sentence complexity effects observed during reading comprehension derive from particular distributional patterns in production, which in turn create distributional regularities that shape comprehender's interpretations. In other words, certain constructions are easier to produce, which makes them used with more frequency, and therefore, are easier to comprehend because they are more expected.

The role of frequency is greatly emphasized in the PDC model: frequent constructions are more easily activated, and therefore easier for the receiver of the message to understand than less frequent constructions. Studies have shown, for example, that comprehension difficulty is influenced by the match between syntactic structure and the frequency with

which verbs appear in that structure (i.e., verb bias). Thus, transitive verbs (e.g., believe) most often used by speakers with sentential complements cause less comprehension difficulty when followed by a sentential complement than by a noun phrase (NP) complement (Dussias, 2003; Garnsey, Pearlmutter, Myers, & Lotocky, 1997; Spivey-Knowlton and Sedivy,1995; Tanenhaus & Trueswell, 1995; Trueswell & Kim, 1998). Returning to code-switching, when we examine the verb+ the complemetizer as a syntactic site for code-switching, we see that switches in which the complementizer appears in Spanish are very frequent but switches in which the complementizer appears in English are rare.

The variable that I examined in the present study is *verb bias*. Verb bias refers to the tendency of a verb to be used by speakers with a certain type of complement. In the present study, two types of verb were selected. *Sentence-complement* biased verbs and *Equi-biased* verbs. The verb 'believe' is considered to be a sentence-complement biased verb. This is so because although 'believe' can be used with a noun phrase complement (e.g., I believe *the boy*), when speakers use the verb 'believe' in conversations with other interlocutors, they most often use it with a sentence complement (e.g., I believe the boy was telling the truth). Equi-biased verbs are verbs that are equally likely to be followed by one complement or another. For example, when the verb 'confirmed' is used by English speakers, it is equally likely to be followed by a noun phrase complement (e.g., The man confirmed the appointment) as by a sentence complement (e.g., The man confirmed the appointment was tomorrow).

In English, when verbs that require a sentence complements are used, it is common to see a complementizer (a conjunction which marks a complement clause) heading the sentence complement clause. An example of a complementizer in English is 'that', as illustrated in the example below:

The man hoped *that* the plane would arrive on time.

In English, the presence of the complementizer following a sentence complement bias verb is optional. Returning to the example above, the sentence "the man hoped the plane would arrive on time" would be just as grammatically correct as without 'that'. In Spanish, however, the presence of a complementizer is obligatory. The following is an example of a Spanish sentence with an SC verb (taken from the Mark Davies Corpus de Español, http://www.corpusdelespanol.org/x.asp):

> Me *confesó que* él creía que estaba viendo un fantasma
> 'He confessed to me that he thought he was seeing a ghost.'

Here, *confesó* has to be followed by the complementizer 'que' for the sentence to be grammatically acceptable. Given the obligatory presence of the complementizer in Spanish, one can assume that SC biased verbs in Spanish and their complementizer should act as a unit. With respect to code-switching, the prediction would be that less switching should be permitted between an SC biased verb and its accompanying complementizer. With EQ biased verbs, the complementizer is not as strongly linked to the verb—because the verb can appear with other types of complements, such as direct objects, prepositional phrases, etc. Because the verb and the complement are breakable, we should expect code-switching to be better here.

One way we find patterns in production is to analyze different corpora. For the present study, I analyzed the code-switches from a column called *La Calentita,* which appears in the newspaper *Gibraltar's National Dish. La Calentita* is an editorial column included in the online version of the Gibraltar newspaper; *Panorama* (http://www.panorama.gi/). The column revolves around an informal discussion of politics and social events between two fictional characters, Cynthia and Cloti, and features many instances of Spanish-English code-switching. Since the written corpus is generated by one columnist, it may seem a stretch to use these data to generalize to the code-switching production of other bilingual Gibraltarians, but the code-switching patterns found in the

Calentita column have, for the most part, proven to coincide with those of fluent bilingual Gibraltarians (Moyer, 1995). Consequently, the examples of code-switching patterns presented in the column can be considered to be consistent with those of the code-switching community in Gibraltar. The following are a few examples from the corpus that exemplify Spanish-English switches involving the Spanish complementizer. The first two examples ((a) and (b)) show a code-switch from Spanish to English after the complementizer. Examples ((c) through (k)) show an English verb with the Spanish complementizer. Finally, in example (l) we see the verbs 'dice' and 'say' both in Spanish and English, but even when the verb is in English, it appears with the Spanish 'que'.

(a) y por otro lado Natwest *dice que* they are going to expand all that is going to happen
(and on the other hand, Natwest says that they are going to expand all that is going to happen)

(b) *dice que* we are treading on dangerous ground to mix up with the other side…
(he says we are treading on dangerous ground to mix up with the other side…)

(c) ….they have **decided que** los shops pueden abrir una hora mas los Thursdays...
(they have decided that the shops can be opened on hour longer on Thursdays…)

(d) Anyway, let's **hope que** no nos afecte el crunch...
(Anyway, let's hope that the crunch doesn't affect us…)

(e) I say, haven't you **heard que** el Chief Minestra has gone to Madrid?
(I say, haven't you heard that the Chief Minister has gone to Madrid?)

(f) Now that he has gone to Madrid he will have **realised que** he could not fly on Iberia
(Now that he has gone to Madrid he will have realized that he could not fly on Iberia)

(g) pero El Picaro has **said que** if that applies to the Hopposition it must also apply to el Governation.
(But the Picaro has said that if that applies to the Hopposition it must also apply to the Governation)

(h) I **thought que** he had joined our campaign to claim the Campo de Gibraltar
(I thought that he had joined our campaign to claim the Camp of Gibraltar)

(i) …and we **agreed que** el Governation had agreed to raise the age to 18.
(…and we agreed that the Governation had agreed to raise the age to 18.)

(j) Lo Spanish have **decided que** they are taking their claim to our waters…
(The Spanish have decided that they are taking their claim to our waters…)

(k) Did you hear el Chief Minestra **say que** se espera un financial time bomb? *Dice que* está ticking.
(Did you hear the Chief Minister say that a financial time bomb is expected? He says it's ticking)

(l) Buddy el Speaker *dice que* he will be rationing los comments del Hopposition en Question Time, pero El Picaro *has* **said que** if that applies to the Hopposition…
(Buddy the Speaker says that he will be rationing the comments of the Hopposition in Question Time, but the Picaro has said that if that applies to the Hopposition…

What this examples show is that in code-switching, Spanish complementizers are very frequent but English complementizers are not. Although there were many examples of English verbs taking 'que' there were no instances where a Spanish verb was followed by 'that'.

Given the corpus data, the assumption is that faster switches that occur more frequently in production would be read faster to process by our participants. Based on the PDC model and what is found in production, as exemplified in the Gibraltar Corpus, we would expect that SC-bias verbs would be faster to read with the Spanish complementizer than with the English 'that', both because they act as a syntactic unit but also because these is the predominant pattern in production.

CHAPTER 2: Method

Participants

Seventeen participants that took part in this experiment, all between the ages of 19 and 23, with the average being 20 years old. Each participant was given financial compensation for participating in this study. To determine the language proficiency of our participants, we used a series of behavioral tasks, including a language history questionnaire, a Spanish grammar test, English grammar test and the Boston Naming Test. Each task is explained below.

1. *Language history questionnaire.* With the language history questionnaire, participants self-reported information relating to their age of acquisition of both languages, which language they feel they understand, speak, write, and read better, as well as their usage and exposure to written and spoken code-switches. The majority of participants reported that they understand and speak both languages about the same, but their preference for using either depends on who their talking or writing to. All participants acquired Spanish before the age of 5, but there was some variance as to when they acquired English, with the average age being 6 ½ years of age. The majority reported that they read in English more than 3 hours a day, reading newspapers, magazines, textbooks, and literary works. Reading in Spanish on the other hand was only about 30 minutes for most participants, mostly reading newspapers and magazines. All participants were exposed more to English on a daily basis, but speak both English and Spanish regularly, and felt they understood both at a very proficient level. When asked how often they code-switch, most participants said sometimes or most of the time, and usually code-switch with friends or family members, and about half believed code-switching was an important part of their identity. The full language history questionnaire is provided in Appendix A.

2. *Boston Naming Test (BNT).* In the BNT, participants were shown 30 pictures in Spanish and 30 in English and asked to verbally identify each picture. The pictures were presented in order from least difficult to most difficult. Examples in English include tree and house (least difficult) to trellis and protractor (most difficult). Spanish examples of least difficult words include cama (bed) and lápiz (pencil) to most difficult \, paleta (paint palette) and ábaco (abacus). Participants received 1 point for each expected response, and 0 for an incorrect response or lack of response. For the entire list of picture names, see Appendix B. Two scores are reported, one as more "conservative" and the other in parentheses more "liberal" to take into account picture ambiguity for a few of the stimuli. The mean for the Spanish set was 17 (19), and fo the English set was 19 (20).

3. *Diploma de Español como Lengua Extranjera (Diploma of Spanish as a Foreign Language, DELE)*

Spanish knowledge was also assessed by administering a section of the DELE. The DELE is a standardized test of Spanish issued by the Ministry of Education, Culture, and Sport of Spain, which tests proficiency in Spanish at seven levels (http://diplomas.cervantes.es/en). The test had a maximum score of 50 points and was comprised of three sections: a cloze test, vocabulary section, and a multiple-choice grammar test. Participants received 1 point for each correct answer and 0 points for incorrect answers. The average score for this assessment measure was 34.5 out of 50.

4. *The Michigan English Language Institute College Entrance Test (MELICET).* Similarly to the Spanish, English knowledge was assessed by administering the MELICET, an advanced level English language test. The part of the test that we used looked language comprehension. This test also had a maximum sore of 50 points, and consisted of conversational grammar, cloze, vocabulary and reading comprehension.

The average score for this grammar test was 40.78. In all, the assessment measures indicate that these participants were proficient speakers of English and Spanish.

Materials

Using a verb-norming study from Dussias, Martful, Gerfen, Bajo Molina (2010), I divided verbs according to their subcategorization bias into SC-bias verbs and EQ-bias verbs. Twelve verbs of each type were used, and for every block, there was an SC-bias verb with the Spanish and English complementizer, and an EQ-bias verb with the same complementizers, for a total of four conditions. With crossing of two factors, verb bias and language of complementizer, four conditions emerge. A sample item set is provided below:

Stimulus examples:

1. La embajadora *dijo que the* delegates would be discussing global warming.
   (condition 1: SC+que)

2. La embajadora *dijo that the* delegates would be discussing global warming.
   (condition 2: SC+that)

3. La embajadora *confirmó que the* delegates would be discussing global warming.
   (condition 3: EQ+que)

4. La embajadora *confirmó that the* delegates would be discussing global warming.
   (condition 4: EQ+that)

Participants were shown one of four files, containing a total of 152 stimulus items, consisting of 6 practice, 48 experimental sentences, and 96 fillers. For a complete list of experimental stimuli, see Appendix C. Fillers included sentences entire in either Spanish or English, as well as sentences that started in Spanish and ended in English, and vise-versa. All stimulus items were counterbalanced across files to ensure that all participants saw items in each condition but no participant saw the same sentences representing each condition.

Procedure

Stimuli were presented on a color monitor using an EyeLink 1000 desktop-mounted eye-tracker, interfaced with an IBM-compatible PC. A chinrest was used to prevent head movement. Participants were instructed to read the sentences silently at their own pace. After reading each sentence, participants were asked to complete a comprehension question task. The tasks was added to ensure that participants remained attentive while reading the sentences. In the comprehension task, participants were asked to answer a 'yes/no' question related to the content of the sentence they had just read. In order to avoid participant fatigue, the experimenter would periodically offer the participant a break, as to give their eyes a chance to rest. The participant was recalibrated every time they took a break, or anytime their eye moments strayed too far from the sentence region.

CHAPTER 3: Analysis

We could not solely compare the fixation durations on the language of the complemetizer (known as the critical region), because in Conditions 2 and 4, the location of the complementizer, 'that', coincided with the switch site. In order to ensure that times were not affected by a switch cost, we took the average means of words 4 and 5, with word 4 was always the complementizer and word 5 always being a determiner, so that any difference between conditions could be explained by differences due to verb bias, and not from having to switch from one language to another.

Four different measures were analyzed; gaze duration, right-bound duration, regression path duration, and total time. *Gaze duration* is considered to be the total duration of all fixations in a target region until the eyes fixated a region of text that was either progressive or regressive to the target region, provided that the first fixation on the target region does not occur after any fixations on words further along in the text. The second measure is *right-bounded duration*, which is the total duration of all fixations in a target region until the eyes fixated a region of text that was progressive to the target region, provided that the first fixation on the target region did not occur after any fixations on words further along in the text. Although seemingly very similar, the difference between right-bound duration and gaze duration is that right-bound duration does not take into account any movement backwards. The third measure is *regression-path duration*, which is the total duration of all fixations that occurred from the first fixation on a target region until the target region was exited in a progressive manner (including fixations on the target region and fixations on words regressive to the target region), provided that the first fixation on the target region did not occur after any fixations on words further along in the text. Lastly, total duration is the total of all fixations in a target region.

A 2x 2 analysis of the variance (ANOVA) on *gaze duration* with verb bias as one factor (SC bias vs. EQ bias) and language of the complementizer (que vs. that) as the second factor reveals no main effect on verb bias, no interaction between verb bias and language of the complementizer (all $p > 1$), but a main effect of language of the complementizer ($F(1,16) = 16.21$, $p = .001$), indicating that reading the complementizer 'que' plus a determiner took less time (M=353.81 , SD=26.5) than reading the complementizer 'that' with the same determinter (M= 411.01, SD= 35.3). Table 1 below shows the average gaze duration time for each condition. If we compare reading times for Conditions 1(SC+que+DET.) and 3 (EQ+que+DET.) with Conditions 2 (SC+that+DET.) and 4 (EQ+that+DET.), we see that reading a sentence with 'que' was faster than with 'that'.

Table 1

Condition	Mean and Standard Deviation
Average of Gaze-Word 4 QUE + Word 5 THE- SC verb	M =350.83 (SD= 120.85)
Average of Gaze-Word 4 THAT + Word 5 THE- SC verb	M =403.17 (SD= 164.97)
Average of Gaze-Word 4 QUE + Word 5 THE- EQ verb	M =356.68 (SD= 104.43)
Average of Gaze-Word 4 THAT + Word 5 THE- EQ verb	M =418.83 (SD= 136.99)

A 2x 2 analysis of the variance (ANOVA) on *right-bounded duration* with verb bias as one factor (SC bias vs. EQ bias) and language of the complementizer (que vs. that) as the second factor revealed no main effect on verb bias, no interaction between verb bias and language of the complementizer (all $p > 1$), but a main effect of language of the complementizer ($F(1,16) = 17.98$, $p = .001$) such that reading the complementizer 'que' plus the determiner "the" took less time (M=397.37 , SD= 34.8) than reading the complementizer 'that' (M=473.32 , SD=45.3). Table 2 below shows the average right-bound duration for each condition. If we compare reading times for Conditions 1(SC+que+DET.) and 3

(EQ+que+DET.) with Conditions 2 (SC+that+DET.) and 4 (EQ+that+DET.), we again see that reading a sentence with ‘que’ was faster than with ‘that’.

Table 2

Condition	Mean and Standard Deviation
Average of RB Duration-Word 4 QUE + Word 5 THE- SC verb	M =397.09 (SD= 158.88)
Average of RB Duration-Word 4 THAT + Word 5 THE- SC verb	M =456.59 (SD= 199.09)
Average of RB Duration-Word 4 QUE + Word 5 THE- EQ verb	M =397.63 (SD= 128.47)
Average of RB Duration-Word 4 THAT + Word 5 THE- EQ verb	M =490.04 (SD= 185.78)

An additional 2x 2 analysis of the variance (ANOVA) on *regression path duration* with verb bias as one factor (SC bias vs. EQ bias) and language of the complementizer (que vs. that) as the second factor revealed no main effect on verb bias, no interaction between verb bias and language of the complementizer (all $p > 1$), but a main effect of language of the complementizer ($F(1,16) = 19.81$, $p = .001$), indicating that such that reading the complementizer ‘que’ plus the determiner “the” also took less time (M=503.67 , SD=49.1) than reading the complementizer ‘that’ (M=611.97 , SD= 65.81). Table 3 below shows the average regression path duration for each condition. If we compare reading times for Conditions 1(SC+que+DET.) and 3 (EQ+que+DET.) with Conditions 2 (SC+that+DET.) and 4 (EQ+that+DET.), we see that reading a sentence with ‘que’ was faster than with ‘that’.

Table 3

Condition	Mean and Standard Deviation
Average of RP Duration-Word 4 QUE + Word 5 THE- SC verb	M =503.55 (SD= 229.40)
Average of RP Duration-Word 4 THAT + Word 5 THE- SC verb	M =592.44 (SD= 311.15)
Average of RP Duration-Word 4 QUE + Word 5 THE- EQ verb	M =503.79 (SD= 182.48)
Average of RP Duration-Word 4 THAT + Word 5 THE- EQ verb	M =631.49 (SD= 252.19)

The final analysis involved a 2x 2 analysis of the variance (ANOVA) for *total duration* with verb bias as one factor (SC bias vs. EQ bias) and language of the complementizer (que vs. that) as the second factor. Results revealed no interaction between verb bias and language of the complementizer (all $p > 1$), but a main effect of language of the complementizer ($F(1,16) = 21.70$, $p = .001$), indicating that such that reading the complementizer 'que' plus the determiner "the" took less time (M=881.97 , SD=76.3) than reading the complementizer 'that' (M= 975.41, SD=85.1). There is also a main effect of verb bias in total duration, indicating that SC-bias verbs were read faster (M= 902.49, SD= 77.9) than EQ-bias verbs (M=954.88, SD= 83.5). Table 4 below shows that the average times for Conditions 1 (SC+que+DET.) and 2 (SC+that+DET.) was faster than Conditions 3 (EQ+que+DET.) and 4 (EQ+that+DET.), meaning that SC-bias verbs were read faster than EQ-bias verbs, but also that Conditions 1 and 3, with 'que', were read faster than Conditions 2 and 4 with 'that'.

Table 4

Condition	Mean and Standard Deviation
Average of Total Duration-Word 4 QUE + Word 5 THE- SC verb	M =863.17 (SD= 322.60)
Average of Total Duration-Word 4 THAT + Word 5 THE- SC verb	M =941.81 (SD= 333.06)
Average of Total Duration-Word 4 QUE + Word 5 THE- EQ verb	M =900.75 (SD= 320.49)
Average of Total Duration-Word 4 THAT + Word 5 THE- EQ verb	M =1008.99 (SD= 375.02)

Interestingly, unlike the other three measures, total duration showed that it was easier to read a complementizer that is preceded by an SC-bias verb than a complementizer that is preceded by an EQ-bias verb. One possible reason for this could be that since an SC-bias verb normally appears followed by a complementizer to be part of its syntactic unit, and therefore is processed faster than a complementizer following an EQ-bias verb.

Overall, results show that there is little difference between sentential complement verbs and equi-bias verbs, but the language of the complementizer does have an effect. Participants consistently read sentences with ‘que’ much faster than they did when the complementizer ‘that’ appeared, which is consistent with what we see in production.

CHAPTER 4: Discussion

The purpose of this study was to look at the reading comprehension of Spanish-English bilinguals when reading code-switches. Manipulating verb bias (SC-bias or EQ-bias) and the language of the complementizer following the verb (que or that), we expected that verbs preceding the Spanish complementizer 'que' would be read faster than verbs preceding the English complementizer 'that'. We also expected switches to be easier with EQ-bias verbs, because they don't act as a syntactic unit and therefore, should be more able to take the English complementizer than SC-bias verbs.

Participants were all native Spanish speakers that have lived and studied in the United States, therefore having had exposure to both languages, and when recruiting participants, there was an emphasis on finding people that regularly code-switch. Language proficiency was determined using several behavioral tasks, including a self-reported language history questionnaire. Using an eye-tracker that recorded their eye moment, each participant was shown a series of sentences with Spanish-English code-switches, which was then followed by a comprehension question, as to elicit more natural reading.

Through several measures, results showed that regardless of the verb-bias, verbs followed by the Spanish 'que' as opposed to the English 'that', were consistently read faster, indicating that verbs+que are easier to comprehend for Spanish-English bilinguals. Only when looking at total duration did we see an effect in verb-bias, where SC-bias verbs, regardless of the language of the complementizer, were read faster than EQ-bias verbs. This study supports the PDC model, in that what is frequent in production is more easily comprehended than less frequent constructions. Future research should look at sentences starting in English and switching into Spanish, to compare if there is a difference for comprehension for starting in one language or another.

REFERENCES

Bullock, B. E., & Toribio, A. J. (2010). *Cambridge handbook of linguistic code-switching*. Cambridge Univ Pr.

Christiansen, M. H., & MacDonald, M. C. (2009). A usage-based approach to recusion in sentence processing. *Language Learning*, *59*, 126-161.

Dussias, Paola E. (2003): "On the relationship between comprehension and production data in codeswitching", in:Wiltshire, Caroline & Camps, Joaquim (eds.): *Romance Phonology and Variation*. Amsterdam: John Benjamins, 27-38.

Dussias, Paola E., Martful, Sandra, Gerfen, Chip, Bajo Molina, Teresa (2010). Usage frequencies of complement-taking verbs in Spanish and English: Data from Spanish monolinguals and Spanish-English bilinguals. *Behavioral Research Methods*.

Gennari, S. P., & MacDonald, M. C. (2008). Linking production and comprehension processes: The case of relative clauses. *Cognition*, *111*, 1-23.

MacDonald, M. C., & Thornton, R. (2009). When lagnauge comprehension reflects production constraints: Resolving ambiguities with the help of past experience. *Memory and Cognition*, *37*(8), 1177-1186.

Montes-Alcala, C. (2005). Dear amigo: Exploring code-switching in personal letters. In L. A. Ortiz López & M. Lacorte (Eds.), Contactos y contextos lingüísticos: El español en los Estados Unidos y en contacto con otras lenguas (pp. 173-185). Madrid: Lingüística Iberoamericana.

Myers-Scotton, C. (2005). Supporting a differential access hypothesis: Codeswitching and other contact data. In J. F. Kroll and A. M. B. de Groot (Eds.), Handbook of Bilingualism: Psycholinguistic Approaches, pp. 326-48. New York: Oxford University Press.

Perfetti, C. A. (1999). Comprehending written language: A blueprint of the reader. In P. Hagoort & C. Brown (Eds.), Neurocognition of language processing (pp. 167-208).

Oxford University Press.

Pfaff, C. W. (1979). “Constraints on language mixing: Intrasentential code-switching and borrowing in Spanish/English. *Language*, *55*(2), 291-318.

Perez Casas, Marisol. (2008). Codeswitching and identity among island Puerto Rican bilinguals

Poplack, S. (1980). “Sometimes I’ll start a sentence in Spanish y termino en Espanol: toward a typology of code-switching.” *Linguistics* 18, 581-618.

Appendix A: Language History Questionnaire

1. First Name
2. Last Name
3. Sex
4. Age
5. Handedness
6. Email address
7. Telephone number
8. Are you currently enrolled at Penn State University?
9. At PSU, you are currently a...
    a. Freshman
    b. Sophomore
    c. Junior
    d. Senior
    e. Graduate Student
10. Were you born in the continental U.S.?
11. If you were not born in the continental U.S., where were you born?
12. If you were not born in the continental U.S., during what ages did you live in your country?
13. If you were not born in the continental U.S., how long (in years and months) have you lived in the U.S.?
14. List the cultures with which you identify. For each culture listed, rate the extent to which you identify with it, using a scale from 1 (very low identification) to 10 (complete identification).
15. In general, which language do you prefer to use?
16. In general, which language do you feel that you SPEAK better?
17. In general, which language do you feel that you UNDERSTAND better?
18. In general, which language do you feel that you READ better?
19. In general, which language do you feel that you WRITE better?
20. Did you being to speak both English and Spanish before age 5?
21. At what age did you start ACQUIRING English?
22. At what age did you start LISTENING to English?
23. At what age did you start SPEAKING English?
24. At what age did you start READING English?
25. At what age did you become FLUENT in English?
26. At what age did you start ACQUIRING Spanish?
27. At what age did you start LISTENING to Spanish?
28. At what age did you start SPEAKING Spanish?
29. At what age did you start READING Spanish?
30. At what age did you become FLUENT in Spanish?
31. List the number of years and months you have spent in a country where English is spoken.
32. List the number of years and months you have spent with a family that speaks English.
33. List the number of years and months you have spent in a school and/or work environment where English is spoken.
34. List the number of years and months you have spent in a country where Spanish is spoken.
35. List the number of years and months you have spent with a family that speaks Spanish.
36. List the number of years and months you have spent in a school and/or work environment where Spanish is spoken.
37. How many hours per day do you read English?
38. On a scale from 1 (almost never) to 10 (always), rate to what extent you are currently exposed to English while reading.
39. What type of materials do you read in English? (choose all that apply)
    - Newspapers
    - Magazines
    - Textbooks
    - Literary works

- Other _____

40. How many hours per day do you read in Spanish?
41. On a scale from 1 (almost never) to 10 (always), rate to what extent you are currently exposed to Spanish while reading.
42. What type of materials do you read in Spanish? (choose all that apply)
    - Newspapers
    - Magazines
    - Textbooks
    - Literary works
    - Other _____
43. How many hours per day do you watch TV/movies in English?
44. On a scale from 1 (never) to 10 (always), rate to what extent you are currently exposed to English while watching TV/movies.
45. On a scale from 1 (never) to 10 (always), rate to what extent you are currently exposed to English while listening to radio/music.
46. On a scale from 1 (never) to 10 (always), rate to what extent you are currently exposed to English with family.
47. On a scale from 1 (never) to 10 (always), rate to what extent you are currently exposed to English with friends.
48. How many hours per day do you watch TV/movies in Spanish?
49. On a scale from 1 (never) to 10 (always), rate to what extent you are currently exposed to Spanish while watching TV/movies.
50. On a scale from 1 (never) to 10 (always), rate to what extent you are currently exposed to Spanish while listening to radio/music.
51. On a scale from 1 (never) to 10 (always), rate to what extent you are currently exposed to Spanish with family.
52. On a scale from 1 (never) to 10 (always), rate to what extent you are currently exposed to Spanish with friends.
53. List what percentage of time you are, on average, exposed to English and Spanish (must total 100%).
54. When choosing to read a text available in all your languages, in what percentage of cases would you choose to read it in each of your languages? Assume that the original version of the text was written in another language that is unknown to you (must total 100%).
55. Do you speak both English and Spanish on a regular basis?
56. What are your parents' (or caretakers') and siblings' (if you have any) native language?
57. What language or languages do you speak at home?
58. Specify with whom you speak each of the languages selected in question #57.
59. When choosing a language to speak with a person who is equally fluent in all your languages, what percentage of time would you choose to speak each of your languages? (must total 100%)
60. On a scale from 1 (never) to 10 (always), rate how frequently others identify you as a non-native speaker based on your accent in English.
61. In your opinion, on a scale from 1 (none) to 10 (pervasive), how much of a foreign accent do you have in English?
62. On a scale from 1 (never) to 10 (always), rate how frequently others identify you as a non-native speaker based on your accent in Spanish.
63. In your opinion, on a scale from 1 (none) to 10 (pervasive), how much of a foreign accent do you have in Spanish?
64. On a scale from 1 (very low) to 10 (perfect), rate your level of proficiency SPEAKING English.
65. On a scale from 1 (very low) to 10 (perfect), rate your level of proficiency UNDERSTANDING English.
66. On a scale from 1 (very low) to 10 (perfect), rate your level of proficiency READING English.
67. On a scale from 1 (very low) to 10 (perfect), rate your level of proficiency WRITING English.
68. On a scale from 1 (very low) to 10 (perfect), rate your level of proficiency SPEAKING Spanish.
69. On a scale from 1 (very low) to 10 (perfect), rate your level of proficiency UNDERSTANDING Spanish.
70. On a scale from 1 (very low) to 10 (perfect), rate your level of proficiency READING Spanish.
71. On a scale from 1 (very low) to 10 (perfect), rate your level of proficiency WRITING Spanish.

72. Code-switching means using both Spanish and English in the same sentence when you are talking to someone else. Do you ever code-switch?
73. How often do you code-switch?
74. When you code-switch, who are you usually talking to?
75. Think about some of the instances in which you have found yourself code-switching. When you code-switch, how often do you start the sentence in Spanish and finish in English?
76. When you code-switch, how often do you start the sentence in English and finish in Spanish?
77. Besides code-switching in speech, how often do you code-switch in writing?
78. In which type of written materials do you code-switch? (choose all that apply)
    - Email
    - Online chatrooms
    - Online blogs
    - Online IM
    - Letters on paper
    - Notes on paper
    - Journal diary
    - other
79. Why do you think you code-switch (in speech and/or written language)?
80. Do you think that code-switching is an important part of your identity?
81. Besides code-switching yourself, how often are you exposed to oral code-switching from other people?
82. Besides code-switching yourself, how often are you exposed to written code-switching?
83. In what type of materials are you exposed to written code-switching? (choose all that apply)
    - Email
    - Online chatrooms
    - Online blogs
    - Online IM
    - Letters on paper
    - Notes on paper
    - Journal diary
    - other

Appendix B: Boston Naming Task

ENGLISH

tree	racket	pelican
house	seahorse	pyramid
whistle	wreath	funnel
flower	globe	accordion
saw	harmonica	compass
broom	acorn	latch
mushroom	igloo	tongs
pretzel	cactus	sphinx
camel	harp	trellis
bench	knocker	protractor

SPANISH

cama	caracol	estetoscopio
lápiz	volcán	bozal
peine/peinilla	dardo	unicornio
tijeras	canoa	soga/dogal/horca
cepillo de dientes	castor	espárrago
pulpo	rinoceronte	trípode
helicóptero	zancos	pergamino/rollo
perchero/percha/gancho	dominós	yugo/yunta
silla de ruedas	escalera eléctrica	paleta
máscara	hamaca	ábaco

Appendix C: Experimental Items

El entrenador dijo que the players needed to start training earlier.

La actriz afirmó que her role had been difficult to perform.

El juez sugirió que the court should take a short recess.

La cantante creía que the concert would be her last performance.

El estudiante admitió que his effort on the project was poor.

La estudiante pensó que the meeting had been rescheduled for Wednesday.

El artista aseguró que the painting would be completed on time.

La madre sospechó que her son was lying about his grades.

El periodista confesó que his sources had been completely made up.

El arquitecto advirtió que the blueprints wouldn't be finished on time.

El florista declaró que the bouquet needed more bright pink lillies.

El científico prometió que his colleague would be given proper recognition.

El abuelo sugirió that the children should go outside to play.

El turista sospechó that his wallet had been stolen that night.

La profesora advirtió that the students should review all their notes.

El político prometió that his taxcut would provide relief for all.

El nadador dijo that this race had been the most challenging.

El padre admitió that his son had been too severely punished.

El escritor confesó that the protagonist in the novel was imaginary.

El farmaceuta declaró that the medicine was causing harmful side effects.

El estudiante creía that the professor would postpone the last project.

La enfermera pensó that the patient needed more time to recover.

El cocinero afirmó that his recipe was an old family secret.

El director aseguró that the musical would be a Broadway hit.

El fotógrafo negó que the photo had been taken at night.

El jefe anunció que the workers needed more time for vacation.

La embajadora garantizó que the delegates would be discussing global warming.

El camarero comentó que the wait to be seated was long.

El conductor respondió que the train would be arriving on time.

La vecina recordó que her garden had been destroyed by weeds.

El ingeniero comprendió que the building should be built facing southwest.

El sastre garantizó que the dress would be made of silk.

El peluquero reconoció que the customer was unsatisfied with her haircut.

El camarero negó que the coffee he was using had expired.
El novio sostuvo que the date would be an unforgetable experience.
La masajista confirmó que her client would need extensive massage therapy.
El científico esperó that his results were going to change history.
El guía aconsejó that the group should take everything with them.
La niña comentó that her teacher had confiscated her favorite toy.
El cura garantizó that the service on Sunday would be short.
La policía anunció that the gang had committed the terrible crime.
El capitán reconoció that the team should practice before the game.
El detective confirmó that the house had been robbed by professionals.
El piloto sostuvo that the flight would take longer than expected.
El jardinero comprendió that the flowers would start blooming in April.
El jugador negó that his teammate had purposely missed the shot.
El cazador reconoció that his shot had injured an innocent bystander.
El fotógrafo comentó that the pictures had all been digitally retouched.

ACADEMIC VITA

Marie-Christine Theberge

Education

The Pennsylvania State University, University Park, PA

The Liberal Arts College
The Schreyer Honors College

- Bachelor's of Arts in Spanish, Expected May 2012

The College of Health and Human Development
The Schreyer Honors College

- Bachelor's of Sciences in Communication Sciences and Disorders, Expected May 2012

Publications & Presentations

Rosa Elisa Guzzardo Tamargo, Rosa E., Paola E. Dussias, Chip Gerfen, Christine Theberge, Jorge Valdés Kroff, Jason Gullife: *Linking comprehension costs to production patterns during the processing of mixed language*. Co-presenter at the International Symposium on Bilingualism in Oslo, Norway: Summer 2011.

Research Experience

Undergraduate Research Assistant for the Center for Language Studies, August 2009-May 2011
Honors Thesis, August 2011-present

Award and Honors

Summer Discovery Grant. Summer 2011
Schreyers Honors Scholarship. Fall 2010
Liberal Arts Enrichment Award. Summer 2010
Volunteer of the Semester, Fall 2010

Other Activities

Student Red Cross Club
CPR Instructor (August 2010 – May 2012)
Health and Safety Officer (August 2010 – May 2012)
On-Site Coordinator Trainer (August 2010 – May 2012)
President (August 2011 – May 2012)

Council of Lionhearts
Board Member (August 2010 – May 2012)

Paterno Fellow Program
PFP Advisory Board (August 2009 – May 2012)