THE INTEGRATION OF LONE ENGLISH NOUNS INTO BILINGUAL SONORAN SPANISH

by

Ryan M. Bessett

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF SPANISH AND PORTUGUESE

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY
WITH A MAJOR IN SPANISH

In the Graduate College

THE UNIVERSITY OF ARIZONA

2017

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Ryan M. Bessett, titled The Integration of Lone English Nouns into Bilingual Sonoran Spanish that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

_____ Date: (12/7/16)
Dr. Ana M. Carvalho

_____ Date: (12/7/16)
Dr. Sonia Colina

_____ Date: (12/7/16)
Dr. Rena Torres Cacoullos

_____ Date: (12/7/16)
Dr. Lillian Gorman

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

_____ Date: (12/7/16)
Dissertation Director: Dr. Ana M. Carvalho

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of the requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that an accurate acknowledgement of the source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the copyright holder.

SIGNED: Ryan M. Bessett

ACKNOWLEDGEMENTS

First and foremost, I would like to thank Professor Ana Carvalho without whom this dissertation would not be possible. Thank you for all the lessons, guidance, mentorship and for always pushing me to my full potential. I would also like to thank Professor Sonia Colina for her support and insights that led to making this dissertation what it is. Thank you to Professors Rena Torres Cacoullos and Lillian Gorman for being part of my committee and for providing constructive feedback. To my colleagues who shared with me this long journey that is graduate school, thank you for sharing in the many highs and lows and for being there when I needed you most. I would like to thank my participants for spending the time to sit with me and provide the corpora that made this dissertation possible. Lastly, thank you to my family for all your love and support.

DEDICATION

To my parents for their constant love and support.

TABLE OF CONTENTS

ABSTRACT

Using data from Arizona, United States, the present study seeks to further our understanding of lone other language items (LOLIs) in bilingual discourse and their status as either borrowings or codeswitches by measuring the degree of incorporation that can indicate a LOLI's status as a borrowing or codeswitching. To accomplish this aim, nouns from 40 sociolinguistic interviews from 8 Spanish monolingual speakers from Sonora, Mexico, 8 English monolingual speakers from Arizona, and 24 Spanish-English bilinguals from Arizona (from Sonoran families) are compared.

Codeswitching can be defined as the "juxtaposition of sentences or sentence fragments, each of which is internally consistent with the morphological and syntactic (and optionally, phonological) rules of the language of its provenance" (Poplack, 1993, p. 255). Borrowing involves the incorporation of LOLIs from a donor language incorporated into a recipient language and need to be morphologically and syntactically adapted into the recipient language (Poplack, Sankoff, and Miller, 1988; Sankoff, Poplack, and Vanniarajan, 1990). Accordingly, the key difference between codeswitching and borrowing is that borrowings are morphosyntactically incorporated into the recipient language while codeswitches are not incorporated. It is important to note that in terms of LOLIs' status, phonological integration has been discarded for being too variable and therefore not a reliable factor in discerning one-item codeswitches from borrowings (Poplack and Sankoff, 1984; Poplack, Sankoff, and Miller, 1988).

In order to measure the degree of incorporation that can in turn indicate the LOLI's status as a borrowing or a codeswitch, the present study applies a sociolinguistic comparative method to loanwords, following Poplack and Meechan (1995, 1998) by comparing nouns from Spanish (recipient language), nouns from English (donor language), and LOLIs from English in Spanish discourse. Since phonology has not been applied to the method of analysis, this study also seeks to explore if phonological integration is correlated to morphosyntactic integration of determiner realization of LOLIs.

The results show, in accordance to previous studies, that the LOLIs overall act morphosyntactically like patrimonial Spanish words in terms of the variables that condition determiner usage. In terms of how phonological integration interacts with morphosyntactic integration, it does seem that the two correlate. LOLIs with Spanish morphology are more morphosyntactically similar to Spanish patrimonial nouns and LOLIs with English phonology are more morphosyntactically similar to English patrimonial nouns in both overall frequencies and the factors that condition determiner usage, leading to the hypothesis that LOLIs that are integrated phonologically are established borrowings and LOLIs that are not integrated phonologically are either codeswitches or nonce borrowings. We provide further evidence for this hypothesis by examining the pauses and false starts that are present before LOLIs with Spanish versus English phonology. The results indicate that LOLIs with English phonology are more often preceded with pauses and false starts than LOLIs with Spanish phonology.

The findings of this study suggest that phonological integration is a factor that should be brought back to the discussion on discerning LOLIs' status as a borrowing or a codeswitch.

Chapter 1:

Introduction

1. Introduction

"I spoke to him in English and Spanish because it is a habit" said a participant in Bessett, Casillas and Ramírez Martínez (2017). This statement made by a bilingual from the border community of Nogales, Arizona demonstrates a unique linguistic practice that bilingual speakers possess, the ability to switch between the two languages they speak in the same discourse. This particular participant is also expressing the naturalness of this process, a widely used linguistic resource in the community. But what happens when bilingual speakers introduce one-word other language items into the discourse of another language? This dissertation seeks to better our understanding of such one-word other language items by assessing the way in which they are integrated, or unintegrated, into the Spanish discourse of Spanish/English bilinguals in Southern Arizona. Before detailing the specific aims of the dissertation, this chapter will explore contact in the Southwest United States and global issues relating to language contact in general. First, in section 2, we describe this population in macro-sociolinguistic terms. Secondly, we provide a discussion of contact induced changes that are frequent and a discussion of whether any type of change is possible, in section 3. Next, we will examine more closely the process of codeswitching (section 4) and lexical borrowing (section 5). We will narrow our focus in section 6 by discussing how variationist studies have attempted to discern one-item codeswitches from

lexical borrowings. Finally, in section 7 we will summarize the material explored in the previous sections, outline the main goals of the dissertation, and specify how the chapters are divided.

2. Macro-sociolinguistic description of bilinguals in the United States

First, it is important to contextualize the bilingual population that will be examined in this dissertation. The specific region of these bilingual speakers is the Southwest United States, and the languages they speak are Spanish and English. While the percentage of Hispanics in the United States continually increases, the number of Spanish speakers in the US does not increase at nearly the same rate (Hidalgo, 2001; Potowski and Carreira, 2010; Silva-Corvalán, 2005). This is due to gradual loss of Spanish via a shift to English (Silva-Corvalán, 2005). Nevertheless, Spanish is the second most spoken language in the US (Silva-Corvalán, 2005). The number of Hispanics and Spanish Speakers in the Southwest went through two major historical losses in terms of emigration from the area; the first took place directly after the Mexico-US War and the second during the Depression Era, when massive deportation was undertaken (Hidalgo, 2001). However, recent census data (from 1980, 1990, 2000) has shown an increase in immigration patterns and it is expected that these numbers will continue to increase exponentially in the future (Silva-Corvalán, 2005). In the US overall, but most especially in the Southwest, Hispanics of Mexican origin far outnumber any other country of origin (Potowski and Carreira, 2010; Silva-Corvalán, 2005).

One common consideration for Spanish in the US is how the language is maintained or shifted to English. It is generally consented that the path to language shift among Hispanics in the United States takes about three generations, in which the third generation of speakers are very dominant in English if not monolingual (Bills, 2010; Potowski and Carreira, 2010; Silva-Corvalán, 2005). While Bills (2010) suggests that assessing language maintenance and loss is difficult and that those who wish to make a claim for maintenance will find an argument in favor, and vice versa, it still is a well-studied and important question when describing a bilingual community. A major factor in Spanish language maintenance in the United States is the location of where people settle, the closer to the border correlating with higher rates of language maintenance (Bills, Hernández Chávez, and Hudson, 1995). Continuous influxes of Spanish speakers also maintain the language in the community (Bills, Hernández Chávez and Hudson, 1995; Potowski and Carreira, 2010). Furthermore, attitudes toward Spanish are good predictors of maintenance or loss, where positive attitudes correlate to higher probability of maintenance and are more likely when there is a larger percent of the population in the community that is Spanish speaking (Bills, 2010; Bills, Hernández Chávez, and Hudson, 1995; Rivera-Mills, 2000).

It is also important to consider the linguistic behavior of the bilinguals being discussed in this section. In order to better understand the linguistic aspects of Hispanics in the US, it can be advantageous to separate distinct generations of speakers, as their linguistic behavior tends to differentiate based on generation. Silva-Corvalán (2005) provides a clear separation of generations: Group 1 (foreign born, arriving in the US before age 12), Group 2 (US born or arrived before age 12), and Group 3 (have at least one parent in Group 2) (p. 215). Among the linguistic behaviors of these speakers, there is a general tendency for loss of morphosyntactic distinctions, for example Groups 2 and 3 begin to converge uses of the preterite and imperfect

tenses (Silva-Corvalán, 2005). In addition to loss of structure, it is also postulated that Group 2 and 3 bilinguals accelerate language internal change, again, converging independent morphosyntactic distinctions, for instance in the case of the extension of *estar* (Gutiérrez, 2003; Silva-Corvalán, 1994, 2005, 2008), a notion Bessett (2015) argues against. Semantic and pragmatic changes also take place when Group 2 and 3 speakers extend meanings of Spanish words to incorporate English uses (i.e. *descansar* 'to rest' can take on the meaning of *'lay off'*, or *librería 'bookstore'* is calqued to 'library') (Potowski and Carreira, 2010; Silva-Corvalán, 2005). A dissenting view on the characterization of linguistic outcomes based on generations comes from Valdés (2000) who states that "so varied indeed are the different types of English-Spanish bilinguals found in Mexican American communities that it is impossible to conjecture about language strengths and weaknesses based on generation, age, schooling, period of residence in the United States, or any other such criteria" (p. 102).

After discussing the broad sociological and linguistic make up of bilinguals in the US, particularly in the Southwest region, it is important to now consider what contact induced language change is and how it is incorporated in bilingual communities such as the one described in Section 2. We will turn now to language contact theory in general in Section 3.

3. Contact theory and methods of studying contact induced language change

It is commonly believed that when languages are in contact linguistic change is possible at all levels (lexical, phonological, morphological, syntactic, semantic/pragmatic) as long as the

contact is more than superficial (Meyerhoff, 2009; Myers-Scotton, 2006; Thomason, 2001; Thomason and Kaufman, 1988; Weinreich, 1953). However, Sankoff (2002) states that it is still questionable whether or not syntax can be borrowed, in agreement with King (2000), Lefebvre (1985), Prince (1988), among others.

Many argue that there is an order for what types of change will occur, suggesting that lexical change is most easily borrowed, followed by phonological, morphological, and lastly syntactic change being least probable (Muysken, 1985; Sankoff, 2002; Thomason, 2001[1]; Thomason and Kaufman, 1988; Van Coestem, 1988). Illustrating the importance of this order, Sankoff (2002) points out that lexical and phonological change are the "major 'gateways' to all other aspects of contact-influenced change" (p. 642). Furthermore, Thomason (2001) argues in favor of a borrowing scale that takes into account the type of contact and predicts linguistic change accordingly. In "casual contact" situations only lexical borrowing of content words is expected; subsequently, under "slightly more intense contact" lexical borrowings of function words begin to appear as well as phonological borrowings; later with "more intense contact" morphological borrowings begin to appear including inflectional morphemes, as well as syntactic changes including word order from (i.e. SOV -> SVO); and finally, under "intense contact" anything goes (Thomason, 2001, p. 70-71).

Among the possible types of contact induced change, Thomason and Kaufman (1988) specify that changes take place in two fundamentally distinct ways: interference and borrowing. Interference refers to changes that take place in what can be described as a speaker's L2, and is often the result of incomplete acquisition. In this type of change, structural elements are most

---

[1] Thomason (2001) places the caveat here that the opposite is true if those bringing about the innovation are L2 speakers of the language, in which case it is more likely that they undergo structure interference (p. 64).

often incorporated (phonology, morphology, or syntax). On the other hand, borrowings are incorporated to a speaker's L1 and most often consist of lexical items (but structural borrowings can occur based on the type of contact, see Thomason's (2001) borrowing scale above). In the context of bilinguals in the United States, we could consider members of Silva-Corvalán's Group 1 to be most likely to incorporate interference into English, Group 3 into Spanish, and borrowings to occur in Spanish by members of Groups 1 and 2.

We will make a special reference to lexical borrowings here, since they are of particular importance for this dissertation. When languages are in contact, the most common type of contact induced change, by far, are lexical borrowings. Borrowings are most often taken from the majority language and borrowed into the minority language. However, there are also cases of substratum[2] influence where words from the minority language are incorporated into the majority language (Sankoff, 2002, p. 648). Within lexical loans, the most easily or often borrowed are content words such as nouns, verbs, and adjectives (Muysken, 2000; Poplack and Meechan, 1998; Thomason, 2001; Sankoff, 2002).

3.1 Variationist approaches to language contact

When discussing the possible effects of language contact, it would be remiss not to include the many contributions that variationist sociolinguistics has brought to the discussion. First and foremost, Poplack and Levy (2010) provide a warning that it is the tendency of many linguists who see variation in a contact situation to assume that the variation is the result of

---

[2] Substratum interference, or imposition, occurs when foreign language speakers impose their own first language phonological habits on their own use of the second language (Thomason and Kaufman, 1988).

contact (p. 391). While Poplack and Levy do not mean to negate the existence of contact induced change, they merely call for a scientific examination of the variation in question to ascertain whether or not the particular variable could in fact be the result of internal change.

A further consideration in contact theory is made as to what type of structural changes can take place. Contact theorists, Thomason and Kaufman, propose that "anything goes" when it comes to possible structural contact induced change (p. 14). Thomason (2001) reaffirms this idea and provides counter examples for the types of change that are generally reported as being impossible (p. 63-66). On the other hand, Silva-Corvalán (2008) maintains that, in order for a structural change to take place, the new structure must be possible in the receptive language. To strengthen this argument, Silva-Corvalán (2008) provides quantitative evidence that the structures that suffer change in California Spanish, for example subject pronoun expression, elision of the complementizer *que 'that'*, and lexical-syntactic calquing, are all structures that are grammatically allowed in Spanish. In the author's opinion, these changes represent accelerations of changes that are already internally in progress in the language, more than directly the result of a new structure in the language due to contact.

Along similar lines to the intensity scale of contact suggested by Thomason (2001), Meyerhoff (2009) outlines a series of stages of contact induced change based on how similar the variable grammars of the two languages become in terms of the linguistic factors that condition a variable. Meyerhoff (2009), through the comparative sociolinguistic method, discusses three types of similarities that can be found between communities, weak transfer/replication, strong transfer, and calquing. Weak transfer/replication is described as having the same factor groups in both communities. Strong transfer not only has the same factor groups in both communities

but also shares the same order of factor groups. Lastly, calquing is defined as having the same factor groups in the same order, as well as the same order of factors within a factor group.

Another contribution that variationists have made to the study of language contact comes from Poplack and Levey urging linguists to compare the contact variety with "an appropriate reference dialect" (2010, p. 395). The idea is to compare a bilingual community with the closet related monolingual (non-contact) variety. Additionally, Tagliamonte (2003) outlines the comparative sociolinguistic method in which two communities are compared. This comparison is not made based solely on the overall usage rates, which the author points out can be an unreliable account of similarities[3], but also the factor weights and within factor weights of the linguistic factors included in a regression analysis. If the linguistic variables that condition the bilingual community are similar to the monolingual community, then it can be concluded that the two are related, and therefore an internal change is a more accurate explanation for the variable. This method of comparison, including the appropriate reference dialect is illustrated in work on subject pronoun expression by Hurtado (2001) who compares recent Colombian Miami arrivals to bilinguals of Colombian parents/grandparents born in Miami, and the extension of *estar* by Bessett (2015) who compares speakers from Sonora, Mexico to bilingual speakers born in Arizona and who have Sonoran parents/grandparents.

This dissertation combines the variationist perspective with contact language phenomena in order to better understand the differences between two particularly interesting contact processes, namely codeswitching and borrowing. To do this, we follow the approach of Poplack and Meechan (1995, 1998), which will be thoroughly explained in Section 6, after we outline the

---

[3] Poplack, Zentz and Dion (2012) further the notion that overall frequencies can be "spurious and misleading" when looking at similarities between two communities (p. 251).

basic understanding of codeswitching as it is represented in the literature (Section 4) and discuss lexical borrowing (Section 5).

4. Codeswitching: An overview

"Es un lenguaje mezclado, pues es un espanglish. No sabes si te están hablando en español o inglés o en qué te están hablando…" 'It's a mixed language, well it's Spanglish. You don't know if they are talking to you in Spanish or in English or in what they are talking to you' -(T12). This comment from a participant in Southern Arizona in the Bessett (2012) corpus, demonstrates how bilinguals are conscious of the community practice of codeswitching, a process that is very common among bilingual speakers.

Varying definitions of codeswitching place focus on different aspects of the phenomena. For example, Poplack (1993) defines the term as the "juxtaposition of sentences or sentence fragments, each of which is internally consistent with the morphological and syntactic (and optionally, phonological) rules of the language of its provenance" (p. 255), placing an emphasis of the grammaticality of the combination of the two languages. Toribio (2011) explains that codeswitching is "the linguistic phenomena whereby a speaker uses two linguistic codes within a single speech event, alternating between phrases, clauses, and utterances, often with no change in addressee or subject" (p. 530), referencing that the two languages are spoken by the same interlocutor in the same utterance. And Myers-Scotton (2009) asserts that codeswitching is "the use of two language varieties in the same conversation, not counting established borrowed words

or phrases from one variety into the other" (p. 473), the importance of separating loans from codeswitching in this definition will be referenced in more detail in Section 6 of this chapter.

The degree to which languages are combined to form codeswitching is often divided into three main types, depending on where, syntactically, the switch takes place: inter-sentential switching, intra-sentential switching, and intra-clause switching (Myers-Scotton, 2006, 2009; Poplack, 1993, 2001; Toribio, 2011; Weinreich, 1953). Inter-sentential switches involve a change at the sentence level, one sentence is in language A and the subsequent in language B, as can be seen in example (1) below from the *Corpus del español en el sur de Arizona (CESA)*:

> (1) ¿Quién sabe la verdad? Who knows but… (CESA013)
>   'Who knows really?'

Intra-sentential codeswitching is a switch that happens within a sentence but possibly between two clauses, as shown in example (2):

> (2) He wasn't as involved in it as before pero fíjate que se comportó bien cuando- cuando nos casamos. (CESA013)
>   'but notice that he behaved when we got married.'

The third time, and most integrated, is intra-clause switching and involves a switch from language A to language B within a single clause, as shown in example (3):

> (3) Pero sí, dice que fue una fue un culture shock in the beginning. (CESA013)

A more in-depth analysis of the types of switches that occur comes from Muysken (2000, 2005) who suggests a structural hierarchy for the types of alternations bilinguals include in codeswitching, see (4) bellow:

> (4) separate sentences < coordinate clauses < adverbial clauses < adverbial phrases < dislocated arguments < … (Muysken, 2005, p. 152)

When within the bilingual community *X* type of switching is observed, all types to the left on the hierarchy should also be present in the community (p. 152). In addition to where switches take

place, Muysken (2000, 2005) suggests hierarchies for the grammatical categories of elements

that can be inserted as part of the switch, as shown in (5):

> (5) a. **Category:** nouns < adjectives < adverbs < verbs < adpositions < conjunctions < …
> b. **Complexity:** stems < compounds < fixed < modifier + head combinations <
> discontinuous idioms < …
> c. **Morphology:** nominal plural < participle ending < derivational morphology < …
> (Muysken 2005, p. 153)

While Auer and Muysken also consider the structural complexity of the site of

codeswitching when separating bilingual speech by type, the authors place the emphasis for

classification on different criteria. Auer (1999) provides three distinct classifications that fall

along a continuum: codeswitching (clear to discern insertion from alternation; used intentionally

by the speaker to convey a specific semantic function), language mixing (difficult to tell

insertion from alternation), and fused lects (the two languages converge grammatically, losing

the option to speak with only LA or LB, i.e. mixed languages). Muysken (2005) separates

codemixing into three types: alternation (small constituents switched), insertion (diverse

categories, including function words inserted), and congruent lexicalization (irregular patterns

involving unrelated languages) (p. 153).

Similarly to research dedicated to contact induced language change, much debate in the

literature has been dedicated to possible constraints on codeswitching. It is widely believed that

some constraints do exist that limit the possibilities of where switches can occur (Gingras, 1974;

Lipski, 1978; MacSwan, 2013; MacSwan and Colina, 2014; Pfaff, 1979; Poplack, 1980, 1993,

2001, 2004; Muysken, 2000[4]; Myers-Scotton, 1993, 2006; 2009; Sankoff, 1998; Toribio, 2011;

among others). Specifically, the "equivalence constraint" states that switches occur when the

switch will not violate the grammar of either language (Lipski, 1978; Pfaff, 1979; Poplack, 1980,

---

[4] Muysken (2000) deviates from the other studies in this list in that while he maintains that
constraints do exist, he affirms that in congruent lexicalization anything goes.

1993, 2001, 2004; Muysken, 2000; Sankoff, 1998). This constraint seems to hold true since less than 1%[5] of 1,835 tokens of codeswitching between Spanish-English bilinguals were grammatical in both languages in a study on Puerto Ricans in New York (Poplack, 1980, p. 600). An additional proposed constraint is the "free morpheme constraint" states that switching can occur after any element as long as the element is not a bound morpheme (Poplack, 1980, p. 585).

Dissenting views from the idea that codeswitching is universally constrained are also present within the literature (Berk-Seligson, 1986 and Bokamba, 1989). Bokamba (1989) presents counter examples for seven syntactic constraints (size-of-constituent, conjunction, complementizer, adjective [noun], clitic pronoun, free morpheme, and equivalence constraints), demonstrating that bilinguals can, and do, codeswitch even in these positions. Berk-Seligson (1986) specifically demonstrates that Spanish-Hebrew bilinguals quite often violate the equivalence constraint. Instead of linguistic constraints on codeswitching, Bokamba (1989) argues that there are socio-psycholinguistic factors that are better predictors of the type of codeswitching that will occur, among them degree of bilingualism, and social norms of *who* codeswitches and in *which* situations.

In addition to the linguistic outlook of codeswitching, it is also important to consider extralinguistic concerns, namely *who* codeswitches and *why/when*. In terms of who codeswitches, Toribio (2011) provides a detailed chart of types of codeswitching and the speakers who engage in them, see (6) below:

---

[5] Of the 11 cases that violated the equivalence constraint, the majority (7) dealt with noun-adjective placement, since Spanish prefers a N-Adj order and English Adj-N. These codeswitches are one-word switches and thus presents the possibility that they are in fact borrowings, and as such, the equivalence constraint is not being violated.

(6)   Monolingual   Bilingual   Monolingual
         Spanish (S)   (E) English
              S….S_e….Se….S^e….*SE*….**SE/ES**….*ES*….E^s….Es….E_s….E
         Borrowing   Code-switching   Borrowing
         Tag-switching   inter- intra-sentential inter-   Tag-switching
         Emblematic switching   Emblematic switching
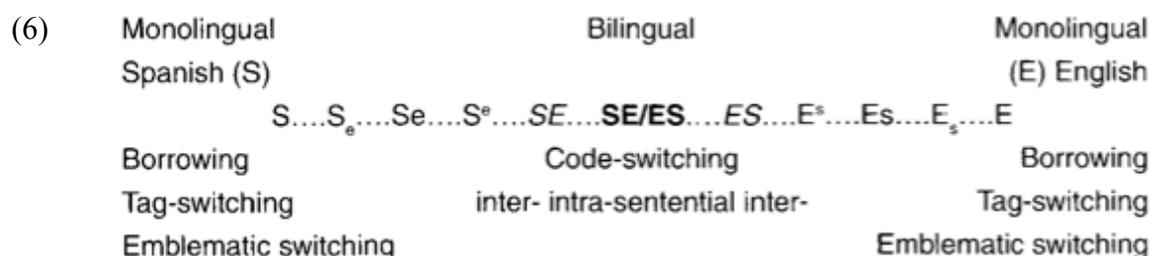
<div align="right">(Toribio, 2011, p. 536)</div>

The chart in (6) explains that the closer a speaker approaches being bilingual, the more complex his/her codeswitching can become, where true (balanced) bilinguals are the most creative in being able to construct intra-sentential codeswitching. Muysken (2000) agrees with this position, stating complex mixing involves high bilingual competence (p. 154).

Codeswitching is generally considered an in-group activity, bilinguals codeswitch with those they perceive as being members of the codeswitching community (Poplack, 1993, 1980; Meyers-Scotton, 2006, 2009; Toribio, 2011). Toribio (2011) outlines some of the motivations behind using codeswitching: quotation, addressee specification, interjections, reiteration, message, topic shift, persuasion, emphasis, performance of identity, in-group membership, or as a conventionalized norm of social behavior. In terms of style, Myers-Scotton (2009) discusses the Markedness Model, which postulates that speakers in a bilingual community place specific types of codeswitching into marked (unexpected) and unmarked (expected) uses (p. 477). Poplack (1980) also notices speaker variability in the style of codeswitching practiced.

As seen in this section, codeswitching is a process that involves structurally incorporating a (group of) word(s) from one language into the discourse of another; however, another common effect of contact mentioned in Section 3 is the incorporation of loanwords. This phenomenon will now be explored more thoroughly in Section 5.

5. Lexical borrowing

One-word insertions of a language A into a language B are by far the most frequent elements bilinguals use to mix the two languages in their speech; specifically, in Wolof-French and Fongbe-French bilingual speech, approximately 90% of French insertions into Wolof/French were single words (Poplack and Meechan, 1995). Loanwords can be single words or compound and need to be morphologically and syntactically adapted into the recipient language (Poplack, Sankoff, and Miller, 1988; Sankoff, Poplack, and Vanniarajan, 1990). It is generally accepted that the longer a borrowing has been part of the community grammar, the more stable its phonological integration into the recipient language will be (Haugen, 1950; Ortigosa and Otheguy, 2007; Poplack, 2012; Poplack, Sankoff, and Miller, 1988). In reference to what types of loanwords are borrowed, English words in Canadian French discourse were predominantly nouns (64%), followed by verbs (14%), interjections and frozen expressions (12%), adjectives (8%), and conjunctions (1%) (Poplack, Sankoff, and Miller, 1988, p. 63). This finding is also upheld in subsequent quantitative research (Poplack, 2012; Poplack and Dion, 2012; among others). This order confirms the assertions made by contact theory explored in Section 3 that lexical elements are the most common to be borrowed (Muysken, 2000; Sankoff, 2002; Thomason, 2001).

Loanwords can be classified into three categories: loans (where the meaning and the form come from the donor language), calques (where the form comes from the recipient language and the meaning from the donor language), and phrasal calques (compound calques) (Smead, 2000).

Calques can be further subcategorized as follows: similar-sense (meaning is similar to an already existing one) versus different-sense (meanings is substantially different from existing ones); merged-form (when the borrowed word is phonologically similar in the two languages) versus independent-form (non-similar phonologically); duplicating (a similar patrimonial word exists) versus innovating (no patrimonial word exists) (Otheguy, García and Fernández, 1989).

In terms of why bilinguals incorporate loanwords, Otheguy and García (1993) suggest that the majority of loanwords are incorporated out of a need to reference objects, concepts, or ideas that do not have a specific name in the recipient language or the term used in the recipient language does not have the same semantic representation. Furthermore, they discover that more loanwords are used when participants are describing life and culture of the donor language country (86%) versus the life and culture in the country of the recipient language (22%). In the same study, it was found that gender, age, nationality, and education-level are not good predictors of loanword incorporation. Lapidus Shin (2010) finds that bilinguals strive for efficiency, often borrowing words that have fewer and shorter syllables than particular patrimonial words.

6. Differentiating one-item codeswitches from borrowings

Lastly, there has been much attention placed on the separation between one-word codeswitches and lexical borrowings (Aaron, 2014; Blas Arroyo and Tricker, 2000; Ghafar and Meechan, 1998; Myers-Scotton, 1993, 2006, 2009; Poplack, 2012; Poplack and Meechan, 1995;

Poplack, Sankoff, and Miller, 1988; Poplack and Vanniarijan, 1990; Torres Cacoullos and

Aaron, 2003a, 2003b; Torres Cacoullos and Vigil, 2002, 2003; among countless others[6]). In

order to accomplish the goal of discerning these two distinct language mixing mechanisms of

bilinguals, studies have compared the frequency rates of linguistic factors (phonological,

morphological, and syntactic adaptation) between nonce borrowings, established borrowings,

patrimonial words of the matrix language, and words in the donor language in the midst of

unambiguous codeswitches. Only three studies follow the variationist comparative method

outlined in Section 3, using comparisons of group factor and within group factor rankings of the

linguistic variables (Torres Cacoullos and Vigil, 2003; Torres Cacoullos and Aaron, 2003a and

2003b). As discussed earlier, a multivariate analysis that explores the factors that condition a

variable is better equipped as a tool to make comparisons between groups, as opposed to pure

frequencies because frequencies can be misleading. The use of overall percentages can be biased

if participants of one group have more tokens in contexts that favor one variant and thus show a

higher overall rate even though they have the same constraint hierarchy (see work on subject

pronoun expression, for example, Torres Cacoullos and Travis, 2010; Travis, 2007; among

others).

      In an effort to use phonological adaptation as a predictor of codeswitching versus

borrowing, Poplack, Sankoff, and Miller (1988) compare words borrowed by more than 10

speakers, words used by fewer than five 199 (99 used by only 1), patrimonial words, and

codeswitches and find that codeswitches are much less frequently adapted phonologically than

nonce or established borrowings. Poplack (2012) furthers the analysis and notices that

---

[6] Reported here are only variationist studies that provide methods to explore this difference;
however, virtually every study on codeswitching makes a direct reference that there is a need to
separate the two due to theoretical implications (e.g. formulating the type of constraints that
govern codeswitching).

codeswitching results in phonological adaptation 27% of the time, while unattested (47%) and

attested (68%) loanwords were phonologically integrated more often (p. 647). Yet, due to the

variability, the authors conclude that phonological adaptation is not the best factor to discern

borrowings from codeswitches. Many other studies have excluded the phonological factor

altogether (Poplack and Dion, 2012; Poplack, Sankoff, and Vanniarijan, 1990; Torres Cacoullos

and Aaron, 2003a, 2003b; among many others). The reasons for excluding phonological

integration from studies on lone other language items is given in response to the work of Poplack

and associates, for example Torres Cacoullos and Aaron (2003a) point out that "there is

substantial consensus that phonology is not a reliable gauge for distinguishing between

borrowing and code-switching" (p. 290), while Poplack and Dion (2012) state that "recipient-

language phonology often colors donor-language items even in the absence of mixing ("foreign

accent"), so phonological criteria are not reliable indicators of loanword integration" (p. 284).

Considered to be far better predictors are the morphosyntactic features that have also

been implemented. A list of all features considered in the literature includes:

- Integration of case: (Ghafar and Meechan, 1998; Sankoff, Poplack and Vanniarijan, 1990)

- Word order: (Ghafar and Meechan, 1998; Poplack and Dion, 2012; Sankoff, Poplack and Vanniarijan, 1990,1990)

- Plural marking: (Poplack and Dion, 2012; Sankoff, Poplack and Vanniarijan, 1990; Torres Cacoullos and Aaron, 2003a)

- Verb inflection: (Poplack and Dion, 2012)

- Realization of determiners: (Aaron 2014; Blas Arroyo and Tricker, 2000; Ghafar and Meechan, 1998; Poplack and Dion, 2012; Poplack and Meechan, 1995; Torres Cacoullos and Aaron, 2003b; Torres Cacoullos and Virgil, 2002, 2003)

- Gender assignment/agreement: (Aaron 2014; Blas Arroyo and Tricker, 2000; Poplack and Dion, 2012)

- Formation of contractions: (Blas Arroyo and Tricker, 2000)

- Complements: when complement formation is a different process (possibly order) in the two languages (Blas Arroyo and Tricker, 2000; Sankoff, Poplack and Vanniarijan, 1990)

- Specificity: specific, nonspecific, generic (Torres Cacoullos and Aaron, 2003a, 2003b; Torres Cacoullos and Virgil, 2003)

- Syntactic position: (Torres Cacoullos and Aaron, 2003a, 2003b)

- Semantic class: (Aaron 2014; Torres Cacoullos and Aaron, 2003a, 2003b; Torres Cacoullos and Virgil, 2003)

- Use of a modifier: (Torres Cacoullos and Aaron, 2003a)

- String position: noun position within a series of nouns (Torres Cacoullos and Aaron, 2003a)

- Polarity: negated vs non negated structure (Torres Cacoullos and Aaron, 2003a)

    Among studies that have as their central research question whether or not one-item switches should be considered borrowings or codeswitches, all conclude that based on the comparisons of the linguistic factors outlined above, these one-items switches are indeed borrowings (Aaron 2014[7]; Blas Arroyo and Tricker, 2000; Ghafar and Meechan, 1998; Poplack

---

[7] One distinction in Aaron (2014) is that while nouns used by three or more speakers follow the pattern of Spanish nouns, borrowings only used once pattern like English nouns in English discourse.

and Dion, 2012; Poplack and Meechan, 1995; Sankoff, Poplack and Vannianajan, 1990; Torres

Cacoullos and Aaron, 2003a, 2003b; Torres Cacoullos and Virgil, 2002, 2003). While these

results are extremely conclusive, having left behind the linguistic factor of phonological

integration in their analyses raises the question as to how that factor would contribute to the

discussion.

7. Summary of previous research and aims of the dissertation

In this chapter we have explored patterns of bilingualism. Specifically, we have

discussed the macro-sociolinguistic profile of bilinguals (Spanish-English) living in the United

States, with an emphasis on the Southwest region. It was found that there is a tendency for

language shift from Spanish to English to occur by the third generation of immigrant families.

Additionally, due to large waves of new immigrants and via trans-frontier interactions in border

communities, there is evidence of language maintenance. It was also suggested that, in terms of

generation, Groups 2 and 3 tend to show signs of linguistic convergence, specifically with

morphosyntactic variables (i.e. subject pronoun expression and imperfect/preterite distinctions).

Later, it was discussed that contact induced change can occur at all levels (with the possible

exception of syntax) and that a hierarchy of change exists where lexical borrowing is most likely,

followed by borrowings at the level of phonology, then morphology, and finally syntax. When

describing codeswitching, it was suggested that there are differences in the linguistic forms of

the switches depending on type of bilingualism, community norms, and individual styles. Also,

in general, there was consensus that there are some constraints on codeswitching, usually involving the grammar structures of the two languages involved in the switching; however, some maintained that no such limits exist based on counter examples that violated the constraints. Furthermore, it was discussed that loanwords are the most commonly borrowed material in contact situations and that nouns are the prominent grammatical class that is borrowed. Bilinguals incorporate loanwords for various reasons including to address new concepts/ideas that did not previously exist in the recipient language and for efficiency. Lastly, we explored variationist research that attempted to discern one-item codeswitches from lexical borrowings. Through comparative analysis of linguistic factors that condition word usage in the matrix language, the donor language, unambiguous codeswitching, and one-item switches, these variationist studies discovered that one-item switches behave like patrimonial words/established loans, and therefore should be considered loans and not codeswitches.

Having summarized the findings of the previous research, we are now able to propose future avenues of study. One such avenue comes from the fact that, with the exception of Poplack and Sankoff (1984) and Poplack, Sankoff, and Miller (1988), a more in-depth analysis as to the types of phonological adaptation found in codeswitching is absent from the literature. Such an endeavor has been explored outside the field of variationist sociolinguistics in phonetic/phonological research (based on laboratory speech) and has yielded mixed results. Bullock (2009) reports that there is a phonetic overlap at the onset of codeswitches for both voice onset time and the English dark-l ([ɫ]). On the other hand, MacSwan and Colina (2014) find that there is an abrupt and absolute switch in the phonological grammar at the onset of codeswitching. These mixed results may be due to the type of bilinguals included in the analysis. Whereas MacSwan and Colina (2104) only consider balanced bilinguals who grew up

speaking both languages from childhood, Bullock (2009) includes late learner bilinguals who are dominant in their L1, and the effect of VOT and dark-l in codeswitching may be a question of interference from their L1 to their L2. Additionally, Poplack (2012) finds that proficiency is an important factor in loanword phonological adaptation, and includes in her analysis bilinguals that are dominant in French. It stands to reason, then, that phonology as an accurate predictor of codeswitching/borrowing in balanced bilinguals (learning both languages from childhood) has not been sufficiently explored for variationist research to have arrived at the conclusion that phonology should be disregarded in comparative studies.

While the unanimous findings that one-item switches are indeed borrowings is extremely robust (as shown above), it is important not to forget the concerns about omitting phonological adaptation from the analysis. It is important to separate one-item switches that are phonologically integrated from those that are not, and compare them to the various other groups. If phonologically integrated items behave more like established loans of the matrix language, while unadapted items behave more like the donor language, a strong argument could be made that some one-item switches are borrowings, as suggested in the literature, while others are indeed codeswitches. This is precisely the endeavor we will undertake in this dissertation.

In order to further explore lone other language items in bilingual discourse, this dissertation consists of three studies based on the spontaneous Spanish speech of bilingual Spanish/English speakers in Southern Arizona (Tucson and Nogales), as well as monolingual Spanish speech in Hermosillo, Sonora, Mexico and monolingual English speech in Tucson, Arizona. The first study, presented in Chapter 2, explores the morphosyntactic integration of one-word English items in Spanish discourse. Following the methodology of Torres Cacoullos and Aaron (2003a, 2003b), we test the linguistic factors that condition the probability that a noun

phrase will be produced with a determiner or with a bare noun. We provide comparisons, not only with the English and Spanish of the bilingual speakers, but also with monolingual comparisons in both English and Spanish as controls. We find that determiner realization is conditioned very similarly between monolingual and bilinguals, that Spanish and English have three key conflict sites in terms of how the factors that condition determiner realization, and that lone other language nouns as a whole are conditioned more similarly to Spanish nouns than English nouns. The second study, Chapter 3, brings phonology back to the analysis of lone other language items. In this chapter we provide an in-depth breakdown of how lone English nouns are phonologically adapted, or not, into Spanish phonology. The results indicate that English nouns are either adapted into Spanish phonology or maintain English phonology; in the vast majority of the cases phonological adaptation occurs at the word level, meaning there are very few cases of words having mixed phonology. We then compare the morphosyntactic integration of lone English nouns that are phonologically adapted into Spanish to those that are not phonologically adapted into Spanish and find that there is a correlation between phonological and morphosyntactic adaptation. Lone English nouns are that are phonologically integrated into Spanish are also morphologically integrated into Spanish. On the other hand, lone English nouns that are not phonologically integrated into Spanish are morphologically more English like in terms of the factors that condition determiner realization. In the last study, presented in Chapter 4, we explore the connection between pausing and false starts, which may be signs of a higher cost in cognitive processing and as such may be signs of codeswitching, in order to further understand the importance of phonological adaptation of lone other language nouns in bilingual discourse. The results show that lone English items that are phonologically adapted into Spanish have fewer pauses and shorter pauses and fewer false starts than lone English nouns that are not

phonologically adapted into Spanish. If pauses and false starts are taken to indicate a higher

processing cost in codeswitching, these results may provide more evidence that lone English

items that are not phonologically adapted into Spanish are codeswitches and not borrowings.

Finally, in Chapter 5, we provide conclusions, limitations, and suggestions for future research.

Chapter 2:

Morphosyntactic integration of LOLI nouns

1. Introduction

Within the literature on language contact, it is widely attested that the lexicon is the part of grammar most permeable to the effects of contact (Muysken, 2000, 2005; Sankoff, 2002; Thomason, 2001; Thomason and Kaufman, 1988; among many others). Language contact usually results in the use of one-item lexical elements from the donor language in the recipient language. These one-item lexical elements have been referred to in the literature as 'lone other language items' (Poplack and Meechan, 1995; Poplack, Sankoff, and Miller, 1988) or LOLIs (Poplack, 2012). In terms of lexical class, contact theory maintains that nouns are most abundant items to be borrowed (Muysken, 2000; Sankoff, 2002; Thomason, 2001), an assertion that was quantified among English words borrowed into Canadian French discourse which were predominantly nouns (64%), followed by verbs (14%), interjections and frozen expressions (12%), adjectives (8%), and conjunctions (1%) (Poplack, Sankoff, and Miller, 1988, p. 63). Because of their abundance in the data, most studies limit the analysis of LOLIs to nouns. Following suit, the present investigation will analyze the classification of LOLI nouns in bilingual discourse.

Virtually all studies that deal with the effects of language contact on the lexicon make a point that it is necessary to distinguish LOLIs as either codeswitches or borrowings (Aaron,

2014; Blas Arroyo and Tricker, 2000; Ghafar and Meechan, 1998; Myers-Scotton, 1993, 2006,

2009; Poplack, 2012; Poplack and Meechan, 1995; Poplack, Sankoff, and Miller, 1988; Sankoff,

Poplack and Vanniarijan, 1990; Torres Cacoullos and Aaron, 2003a, 2003b; Torres Cacoullos

and Vigil, 2002, 2003; among countless others[8]). Codeswitching can be defined as the

"juxtaposition of sentences or sentence fragments, each of which is internally consistent with the

morphological and syntactic (and optionally, phonological) rules of the language of its

provenance" (Poplack, 1993, p. 255). Borrowing involves the incorporation of LOLIs from a

donor language incorporated into a recipient language. These loanwords can be single words or

compounds and need to be morphologically and syntactically adapted into the recipient language

(Poplack, Sankoff, and Miller, 1988; Sankoff, Poplack, and Vanniarajan, 1990). Accordingly, the

key difference between codeswitching and borrowing is that borrowings are incorporated into the

recipient language while codeswitches are not incorporated. Using data from the Southwest

United States, the present study seeks to further our understanding of LOLIs and their status as

either borrowings or codeswitches by measuring the degree of incorporation that can indicate a

LOLI's status as a borrowing or codeswitch. Before delving into the specific research questions,

however, it is crucial that we review the literature that pertains to the study of LOLIs as

codeswitches or borrowings.

---

[8] Reported here are only variationist studies that provide methods to explore this difference; however, the great majority of studies on codeswitching make a direct reference to the need to separate codeswitches and borrowings due to theoretical implications (e.g. formulating the type of constraints that govern codeswitching).

1.1 Theoretical background and previous research

In order to measure the degree of incorporation that can in turn indicate the LOLI's status as a borrowing or a codeswitch, Poplack and associates have proposed applying a sociolinguistic comparative method to loanwords (Poplack and Meechan, 1995, 1998). This method involves making comparisons between various groups of nouns. First, a control group of patrimonial nouns from the recipient language is identified (nouns that are natively from the recipient language). Secondly, a group of established borrowings, or lexical borrowings that have been in the community for a period of time, is assembled. Third, a group of nouns from the donor language is gathered from multi-word codeswitches in bilingual discourse. It is important to establish the morphosyntactic patterns of these groups in order to ensure that there are clear separations from the patterns of the recipient and the donor language. It is expected that the established borrowings should pattern after the nouns from the recipient language, since they are expected to be morphosyntactically integrated into that language. Lastly, a final group of ambiguous LOLIs, (in the sense that it is unclear if they are unincorporated codeswitches or incorporated borrowings), is compared to the other groups. Previously, studies have only compared morphosyntactic patterns of nouns in bilingual speech, no comparisons have been made to how monolinguals pattern using the same morphosyntactic features. Specifically, Aaron (2014) argues we should only consider bilingual speech in order to avoid the "comparative fallacy" brought forth in Bley-Vroman (1983). By using a comparative sociolinguistic method, the investigators hypothesize that if ambiguous LOLIs pattern like nouns from recipient language, then LOLIs should be considered borrowings, since this patterning would be evidence of incorporation. Conversely, if LOLIs pattern after nouns from codeswitches in the donor

language, then LOLIs should be considered codeswitching, given that this lack of incorporation indicates a total switch to the other language, albeit a short one.

Various morphosyntactic features have been used to establish the patterns of the recipient and donor languages so that the comparisons described above can be carried out to test whether ambiguous LOLIs are integrated borrowings or non-integrated codeswitches. One of the most frequently included factors is the use of determiners (Aaron, 2014; Blas Arroyo and Tricker, 2000; Ghafar and Meechan, 1998; Poplack and Dion, 2012; Poplack and Meechan, 1995). By measuring the percentage of usage of definite articles, indefinite articles, demonstratives, possessives, and zero determiners among the various groups, these studies establish that the determiner usage of LOLIs patterns after the recipient language and therefore LOLIs are most likely borrowings because they are morphosyntactically integrated. Other factors considered include word order (Ghafar and Meechan, 1998; Poplack and Dion, 2012; Sankoff, Poplack and Vanniarijan, 1990,1990), formation of contractions (Blas Arroyo and Tricker, 2000), complements when complement formation is a different process (possibly order) in the two languages (Blas Arroyo and Tricker, 2000; Sankoff, Poplack and Vanniarijan, 1990), morphological case (Ghafar and Meechan, 1998; Sankoff, Poplack and Vanniarijan, 1990), gender assignment/agreement (Aaron 2014; Blas Arroyo and Tricker, 2000; Budzhak-Jones, 1998; Poplack and Dion, 2012;), and plural marking (Poplack and Dion, 2012; Sankoff, Poplack and Vanniarijan, 1990).

With the exception of a few factors that solely affect the noun (plural marking, gender, and case), the majority of the aforementioned variables are not ideal predictors of integration if we consider that it is possible to codeswitch at the N level and not solely at the NP level. Testing only the use of determiner, contractions, and word order makes the *a priori* assumption that

codeswitching cannot occur at the individual word level. For example, the use of determiners in the recipient language would not differentiate from the use of determiners in this language if the speaker is codeswitching at the N level from the donor language. Therefore, analyzing only the use of the determiner assumes that codeswitching occurs at the NP level and does not take into account the possibility that codeswitching can occur at the level of an individual word. The results that determiner use of LOLIs patterns after the recipient language confirms that bilinguals can incorporate LOLIs, but does not shed light on the status of the item as borrowing or code-switching.

In subsequent research, Torres Cacoullos and Aaron (2003a, 2003b) devised a way to test morphosyntactic integration at the NP and N levels. They tested the probability of whether a determiner or a bare noun would be used based on several linguistic factors that condition the use of determiners: specificity (specific, non-specific, generic), syntax (subject, object, oblique, predicate nominal/existential), use of a modifier (prenominal, post nominal, none), string position (first noun or only noun, second noun), polarity (negated, non-negated), and semantic class (occupation/status, institution, coincidence sites). Using a multivariate regression analysis, the authors compared the constraint hierarchies of the factors that condition determiner usage of nouns from Spanish (recipient language), nouns from English (donor language, found in multi-word codeswitching), established borrowings from English in Spanish discourse, and ambiguous LOLIs in Spanish discourse. Following research from Poplack and associates, they predicted that the constraint rankings should be similar for nouns from Spanish, established borrowings from English in Spanish discourse, and ambiguous LOLIs, but different for nouns from English. Their hypothesis was borne out in the findings. The factor of specificity was selected by the statistical model for all groups, and ambiguous LOLIs indeed patterned like nouns from Spanish and

established borrowings in that nonspecific referents favored bare nouns and generic referents

favored the use of a determiner. Additionally, the pattern of effects of ambiguous LOLIs for the

syntax and semantic factors, while not significant, also patterned after nouns from Spanish and

established borrowings. The syntax factor revealed that the groups of nouns from Spanish,

established borrowings, and ambiguous LOLIs favored bare nouns in predicate

nominal/existential constructions, while the group of nouns from English favored the use of a

determiner. The semantics factor revealed that occupation/status nouns from Spanish, established

borrowings and ambiguous LOLIs favored the absence of the determiner more than institution

nouns, but the opposite pattern was found for the nouns from English. This allowed the authors

to conclude that in terms of the factors that condition determiner usage, ambiguous LOLIs do

show linguistic integration and therefore should be considered borrowings and not codeswitches.

While this method can only attest to the integration of determiner usage, it does make two very

important advancements to testing whether or not LOLIs are codeswitches or borrowings. First,

it is the first study to our knowledge that controls the factors in order to test morphosyntactic

integration at the N level. Secondly, it is also the first study to incorporate a multivariate analysis

of the factors being examined. The later advancement is important due to the fact that it has been

attested that multivariate analysis is better equipped as a tool to make comparisons between

groups. The use of overall percentages can be biased if participants of one group have more

tokens in contexts that favor one variant and thus show a higher overall rate even though they

have the same constraint hierarchy (see work on subject pronoun expression, for example, Torres

Cacoullos and Travis, 2010; Travis, 2007; among others).

1.2 Description of current study and research questions

Having summarized the previous research, we now turn to the present study on the

morphosyntactic integration of LOLIs. Due to the fact that the method introduced by Torres

Cacoullos and Aaron (2003a, 2003b) has brought about advancements to research on LOLIs, the

present investigation follows their method by examining determiner realization of English LOLIs

produced during Spanish discourse by bilinguals from Arizona. Furthermore, since monolingual

comparisons have yet to be included in such studies, ensuring that patterns found in patrimonial

words among bilinguals are comparable to monolingual patterns, the present study also includes

monolingual control groups for both Spanish and English. In order to avoid the issue of false

comparisons brought up by Bley-Vroman (1983) and Aaron (2014), this study creates a strong

control in the bilingual-monolingual comparisons that it measures. The bilingual participants are

all from Southern Arizona (Tucson/Nogales) and from families from Sonora, and comparing

their speech to Spanish monolinguals from Sonora and English monolinguals from Tucson does

in fact adhere to the warnings of both Bley-Vroman (1983) and Poplack and Levey (2010) who

call for comparing bilinguals to relevant monolingual communities. In doing so, the current

study expects to make five contributions to our understanding of determiner realization and

LOLI morphosyntactic integration in Spanish. First, we will document the overall frequency of

type of determiner and overall frequency of bare nouns between the groups. Second, we will

determine if bilinguals do in fact compare to monolinguals in terms of determiner realization.

Third, we will explore possible continuities of the Spanish spoken in the Southwest United States

by comparing the results from Arizona bilinguals to those of New Mexico bilinguals in Torres

Cacoullos and Aaron (2003a, 2003b) to determine if there are similarities in determiner

realization between the two groups. Fourth, we will establish the differences of determiner realization between English and Spanish. And finally, we will document the extent to which LOLIs are morphosyntactically integrated into the Spanish discourse of Arizona bilinguals. To accomplish these aims, the discussion of the results will be separated by analyses that will answer the following research questions:

RQ1: What is the distribution of determiner types and what is the overall frequency of bare nouns among the several comparison groups? (Explored in section 3.1.)

RQ2: What is the overall frequency of bare nouns among the several comparison groups? (Detailed in sections 3.2.)

RQ3: How does the conditioning of determiner realization compare between the Spanish of the bilinguals and the Spanish monolinguals, and between the English of the bilinguals and the English monolinguals, and how do these results compare to those of New Mexico bilinguals in Torres Cacoullos and Aaron (2003a)? (Reported in section 3.3.)

RQ4: What are the key conflict sites between Spanish and English determiner realization? (Outlined in section 3.4.)

RQ5: To what extent do established and nonce LOLIs show morphosyntactic integration into Spanish and what does the amount of integration suggest in terms of the status of LOLIs as codeswitches or borrowings? (Examined in section 3.5.)

Before discussing the results at length and answering the research questions outlined

above in section 3, we will first describe the participants, data, codification, and data analysis for

the study in section 2.

2. Methodology

2.1 Participants

A total of 40 participants are included in the study. All participants are young adults (no

older than 32 years of age) and are divided evenly between men and women. First, 8 of the

participants are from the monolingual Spanish speech community of Hermosillo, the capital of

the Mexican state of Sonora, which shares a border with the state of Arizona in the United States

(see Map 1).

Map 1: Arizona and Sonora border
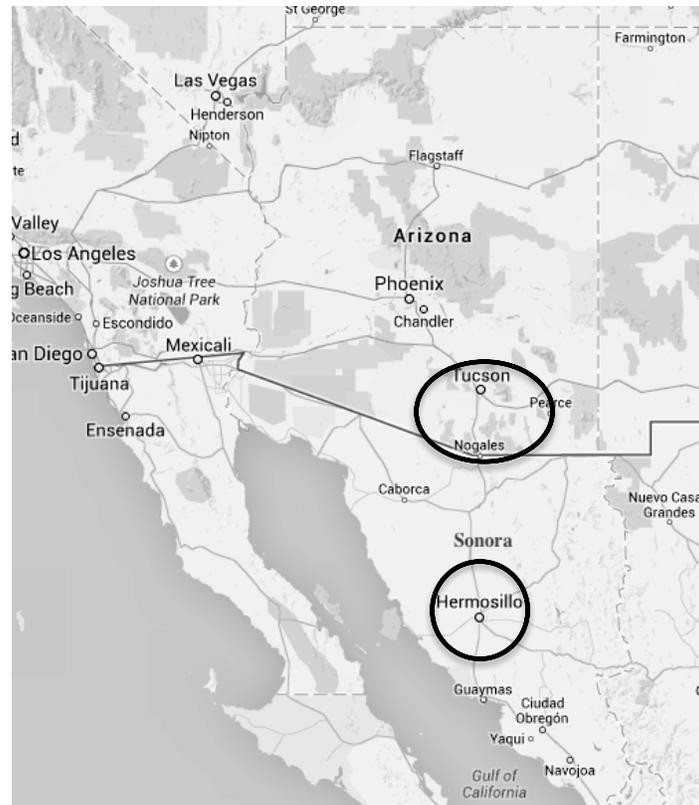
Map courtesy of Google Maps

Hermosillo is home to 784,342 inhabitants as of the 2010 census (Instituto Nacional de Estadística y Geografía). These participants were interviewed by the author and were friends of the author or friends of other participants. A second group of 8 speakers are English monolinguals and come from Arizona, the US state that borders Sonora.

The remaining 16 participants reside in Arizona, either in Tucson, located sixty miles north of the border with Sonora, or Nogales, located at the border with Sonora. According to the 2010 census, Tucson has 520,116 inhabitants, 42% of whom reported being Hispanic or Latino, and Nogales is home to 20,837 residents, 95% of whom are Hispanic or Latino (United States

Census Bureau). In Southern Arizona there is evidence of both language maintenance (due to

constant immigration of monolingual Spanish speakers) and language shift (from Spanish

monolingual, to bilingual, to English monolingual), as reported by Jaramillo (1995). The

participants of both communities (Arizona and Sonora) also describe crossing the border on a

frequent basis. Those from Sonora travel to Tucson to visit family and to go shopping, and those

from Arizona visit Sonora to be with family. These trans-frontier movements create a bridge

between the two communities, where the monolingual Spanish of Sonora is in contact with the

bilingual Spanish of Arizona. It is important to note that this contact is not sufficient for dialect

assimilation to occur given that constant face-to-face interaction is required for such a change

(Otheguy and Zentella, 2012, p. 106).

The participants are part of four different sociolinguistic corpora and can be divided into

five groups. A summary of the participants can be found in Table 1.

Table 1: List of participants by gender and by corpus

	Group 1 Spanish Monolinguals (Bessett 2012)	Group 2 Spanish/English Bilinguals (Bessett 2012)	Group 3 Spanish/English Bilinguals (Carvalho 2012-)	Group 4 Spanish/English Bilinguals (Carvalho 2012-)	Group 5 English Monolinguals (Kern 2017)
Males	4	4	4	4	4
Females	4	4	4	4	4
Total	**8**	**8**	**8**	**8**	**8**
Total participants in the study: **40**					

Groups 1 and 5 are both formed by monolinguals and serve as control groups. Group 1 consists

of 8 monolingual Spanish speakers from Hermosillo, Sonora, Mexico, who come from the

Bessett (2012) corpus, and are interviews between a participant and the interviewer. Group 5

contains 8 monolingual English participants and is part of the Kern (2017) corpus based on interviews between pairs of participants who are familiar with each other.

The remaining participants are all Spanish-English bilinguals and are part of three separate corpora. Group 2 (N=8) was interviewed in Spanish by the author (Bessett, 2012) using the sociolinguistic interview style. Group 3 comes from the *Corpus del Español en el sur de Arizona* (Carvalho, 2012-) and consists of sociolinguistic interviews between an interviewer and a participant. The final group, Group 4, is a subset of the CESA corpus which partners two bilingual speakers and records their conversation. The Spanish spoken during the interviews with participants in these groups is considered "Spanish bilingual speech in monolingual Spanish mode" due to the fact that the interviews were conducted in Spanish.

The grand majority of the bilingual participants were born in the United States or arrived before they were ten, with the exception of one participant who arrived at age fifteen. They are all part of or descendants of families who emigrated from Sonora, Mexico, thus allowing for a comparison of bilinguals (Groups 2-4) to monolinguals (Group 1) from the same dialectal area. In order to ensure that the participants in the bilingual groups (Groups 2-4) were bilingual, the participants were asked to evaluate their proficiency in both English and Spanish on a 0-10 scale, where zero meant very low proficiency and ten meant very high proficiency (in the case of the CESA corpus, Groups 3 and 4, the scale was 0-6). This assessment was adapted from the Bilingual Language Profile created by Birdsong, Gertken, and Amengual (2012). All of the participants reported a proficiency of at least six out of ten (or 4.5 of 6 in the case of participants from CESA) for both English and Spanish. The participants demonstrated their Spanish proficiency in that they were able to carry out a conversation in Spanish for the duration of the

hour-long interview. Additionally, all participants attended school in the United States and live

and work in a bilingual setting.

2.2 Data collection

In order to assess the extent to which the bilingual speakers are morphosyntactically

integrating LOLIs, we will make the following comparisons: Spanish monolingual speech,

Spanish bilingual speech in monolingual Spanish mode, established borrowings in Spanish

bilingual speech, nonce LOLIs in Spanish bilingual speech, English bilingual speech in

monolingual English mode, and finally English monolingual speech. Table 2 summarizes the

comparison groups and number of tokens in each group:

Table 2: Comparison groups and expected number of noun tokens

Comparison group	Spanish monolingual speech	Spanish bilingual speech	Established borrowings	Nonce LOLIs	English bilingual speech	English monolingual speech
Expected number of tokens	400	400	368	388	400	400
Corpus where data will come from	Spanish Monolinguals (Group 1)	Spanish/English Bilinguals (Group 4)	Spanish/English Bilinguals (Groups 2-4)	Spanish/English Bilinguals (Groups 2-4)	Spanish/English Bilinguals (Group 4)	English Monolinguals (Group 5)

Since nouns in all categories (except the established and nonce LOLIs) are quite common, a

proportionate number of patrimonial words from each speaker was collected in order to reach

similar numbers of nouns in each category. Since 368 of the LOLIs identified were established

and 388 were nonce (see section 2.3 for a description of how nouns were classified by

establishedness), 400 tokens from each of the other groups were collected. Since there are 8

speakers in each, 50 tokens per speaker were collected. In the cases of the monolingual samples,

the first 50 nouns starting after the first 20 minutes of the interview were extracted. For the

Spanish bilingual speech in monolingual Spanish mode, the first 50 nouns in the longest

stretches of monolingual mode discourse were included. Lastly, the tokens for the English-

Spanish bilingual speech were collected from English interviews with the bilingual participants

and the first 50 tokens after the 20-minute mark were also included for this group.

The analysis of the 24 interviews with bilingual speakers in Groups 2-4, revealed a total

of 756[9] LOLIs. Not included in the tokens extracted were names of cities and or people (e.g.

Tucson, Phoenix, Flagstaff, Ronald Regan). It is noteworthy to mention here that while we

concur with Torres Cacoullos and Aaron (2003a, 2003b) in removing city names and people/pet

names from the analysis, we are in favor of leaving in the analysis nouns that refer to stores (e.g.

Walmart, McDonald's) as well as street names (e.g. Grant, Oracle). While names of cities and

people[10] in both Spanish and English highly disfavor the use of an article, store names favor

them in Spanish, but not English. Consider example 1 below showing the same noun in both

Spanish and English:

---

[9] Originally 799 tokens were identified. Of the 24 participants, 6 produced nouns of sport names
without a determiner (for example, *fútbol "football"*), even though these are very established
nouns and the Spanish pattern is to use a determiner when speaking of them generally. All sport
nouns were then removed for these 6 participants (a total of 43 tokens, leaving 756 in the
analysis).

[10] In Sonoran Spanish the use of a definite article is actually common for well-known referents,
but this is also dependent on the listener's familiarity with the referent as well. Additionally,
many of the bilinguals do not adhere to this norm, possibly due to interaction with other Spanish
dialects in Arizona, which makes the usage stigmatized in Arizona Spanish. To avoid a more
complex analysis, we will consider Arizona Spanish to be lacking this use.

(1) a. Vamos al McDonald's.

b. Let's go to McDonald's.

Additionally, there is a large number of two-word other language items including: *hookah lounge*, *ultimate Frisbee*, *student union*, *summer school*, *turning lane*, *little league*, *buzz kill*, *extra credit*, *border patrol*, *social security*, *full time*, *part time*. While Torres Cacoullos and Aaron (2003a, 2003b) initially separate these words from the other LOLIs, Torres Cacoullos and Aaron (2003a, p. 319) make an argument that such words act as compound borrowings and not as constituent insertions as Poplack and Meechan (1995, p. 224) suggest. Since in English these items act as compound nouns and separating them into their parts changes their meaning, we consider them as LOLIs in the analysis.

2.3 Coding practices

The data from the corpora described in section 2.1 allowed for the classification of LOLIs by establishedness: established and nonce LOLIs. Since established loanwords are clearly borrowings, well integrated into the recipient language grammar, they are a key piece in determining the status of LOLIs as borrowings or codeswitches. If established loanwords do not act like patrimonial words, despite lengthy presence in the language, one should not expect sporadic (nonce) LOLI borrowings to act as such. It is for this reason that the present study also marks a difference between the established and nonce LOLIs. The classification of a borrowing as established or nonce (ambiguous) has been made based on the existence of the word in dictionaries (Aaron, 2014; Torres Cacoullos, 2003a, 2003b), by their frequency (Aaron, 2014; Poplack and Dion, 2012; Torres Cacoullos, 2003a, 2003b), and by the number of participants

that also use the LOLI at least once (Torres Cacoullos, 2003a, 2003b). Establishing the

classification of LOLIs based on the frequency or the number of participants who also use the

same LOLI is problematic here due to the small size of the corpus (only 24 bilingual speakers).

This is easily seen with the following examples from the corpus:

(2) "Pues no, casi no porque, los <u>coaches</u> [kó.tʃes] no saben enseñar, pues…" -T17[11]

'Well no, almost never because the <u>coaches</u> don't know how to teach…'

(3) "Sí, el año pasado, cuando hice mis <u>taxes</u> , no agarré mucho dinero…" -CESA049

'Yes, last year, when I did my <u>taxes</u>, I didn't get much money back…'

In examples (2) and (3), the LOLIs *coaches* and *taxes* are actually representative of well-

established borrowings in the community (through the researcher's personal experience, through

a small inquiry of native Tucsonan bilinguals, and also their incorporation in documented lists of

words used in Spanish by bilinguals in the Southwest United States). In fact, many signs in

Spanish in South Tucson that advertise help with taxes use the word *taxes* rather than the Spanish

impuestos. However, in both cases, each word is only used once in the entire corpus. For this

reason, the current study, differently from Torres Cacoullos and Aaron (2003a, 2003b), used

several other measures to discern established borrowings from nonce LOLIs. First, the words

were looked up in official dictionaries (Academia Mexicana de la Lengua, 2015, 2010;

Corominas, 1967; García Icazbalceta, 1899; Lara Ramos, 1996; Lugones, 1932; Real Academia

de la Lengua Sonorense, 2006; Real Academia Española, 2016). Second, the words were

searched in local dictionaries and reported lists of borrowings in Southwest Spanish (Cobos,

---

[11] These numbers refer to the way the participants are organized in the corpus: H = Sonora, T = Arizona, both from the (Bessett, 2012) corpora, and CESA from the Carvalho (2012-) corpus.

1983, 2003; Cerda, Cabaza, and Farias, 1970, 1953; Galván and Teschner, 1985, 1975; García, 1935; Gross, 1935; Hamel, 1994; Reyna, 1980). Words found in any of the aforementioned dictionaries were considered established borrowings. Finally, the monolingual corpus of 24 Sonoran Spanish speakers in Hermosillo (Bessett, 2012) was consulted. If a word appears in the monolingual corpus collected in Hermosillo, we can assume that the word is an established borrowing, expected to be fully adapted to the local Spanish grammar (monolinguals, by definition, cannot codeswitch). The remaining words that were not found in dictionaries or the Hermosillo corpus were classified as nonce LOLIs. Of the 756 LOLIs identified in the bilingual corpora, about half were identified as established (368/756, 49%) and the other half as nonce (388/756, 51%).

After categorizing the data into its respective comparison groups (see Table 2) each noun was coded for several linguistic factors, following the methodology in Torres Cacoullos and Aaron (2003a, 2003b). The use of determiner versus bare noun was used as the dependent variable. In order to determine the linguistic conditioning of determiner usage, the following linguistic factors groups were analyzed (examples are given for each factor):

- Specificity:
  - Specific (specific nouns referring to specific people or thing, they are not interchangeable)

    (4) "Mi **mamá** nació en Hermosillo." -T02

      'My **Mom** was born in Hermosillo.'

    (5) "Pues, tengo un hermano mayor…" -T02

      'Well, I have an older **brother**…'

  - Non-specific (any member of a class)

(6) "Jugábamos con lo que tuviéramos…si encontrábamos un **palo**…" -T02

    'We would play with what we had…If we found a **bat**…"

(7) "…él dijo que yo había dicho una mala **palabra**…" -T02

    '…He said that I had said a bad **word**…'

- o Generic (refers to an entire class)

(8) "…y yo creo que **los videojuegos** era lo que más nos unía." -T02

    '…and I think that **videogames** are what bonded us.'

(9) "…no me gustan mucho las **películas** entonces no me gusta ir ahí…" -T02

    '…I don't like **movies**, so I don't like going there…'

- Syntax:
  - o Subject

(10) "Mi **nana** hace tortillas grandes." -T11

    'My **grandma** makes big tortillas.' -T11

  - o Object

(11) "…no tenía casi **problemas**..." -T11

    '…I didn't have almost any **problems**…'

  - o Oblique (object of a preposition)

(12) "Con los **clients**, todo estaba bien." -T11

    'With the **clients**, everything was good.'

  - o Predicate nominal/existential

(13) "Pues, había **gente** grosera." -T11

    'Well, there were some rude **people**.'

(14) "Pues, ni modo, son los **clientes**." -T11

'Well, it doesn't matter, they are the **clients**.'

- Presence of a modifying element

  - Prenominal

    (15) "Sí, está a la misma **distancia**." -T12

    'Yes, it's the same **distance**.'

  - Post-nominal

    (16) "Mis papás vienen de un pueblito **chiquito**." -T12

    'My parents come from a small **town**.'

  - None

    (17) "…pero luego me metí en un **accidente**." -T12

    '…but later I got myself into an **accident**.'

- Polarity

  - Affirmative

    (18) "…I saw one of my **friends**…" -1EM

  - Negative

    (19) "I have no **idea**." -1EM

- String Position

  - First or only noun in a series

    (20) "Casi todos los **sábados** vamos…" -T17

    'Almost every **Saturday** we go…"

  - Second (or subsequent) noun in a series

    (21) "Pues nomás, tengo cuadernos y **lápices** y todo lo que necesitamos." -T17

    'Well just, I have notebooks and **pencils** and everything we need.'

- Semantic Class[12]

    o Occupation nouns

    (22) "Soy **recepcionista**…" -T18

    'I am a **receptionist**…" -T18

    o Institution nouns

    (23) "Pues, ya terminé el **colegio**." -T18

    'Well, I already finished **high school**.'

    o Other

    (24) "…hablo mucho con la gente que entra." -T18

    '…I talk a lot with the people that come in.'

    (25) "…y así estamos, toda la **conversación**…" -T18

    '…and that's how we go, the whole **conversation**…'

2.4 Data analysis

Once the data was coded, it was run using the statistical program GoldVarb to determine

how the linguistic factors condition the realization of determiners among the six comparison

groups. An analysis of the results will be discussed at length in section 3. In order to answer our

research questions presented in the introduction (section 1.2), we will first discuss the

distribution of determiners among the several groups and compare them to previously studied

communities (section 3.1). We will then explore the overall rate of bare nouns in the groups and

---

[12] The semantic class category was limited to occupational and institutional nouns based on the results of Torres Cacoullos and Aaron (2003b), which found that the aforementioned groups were the only conflict sites in the semantic class category between English and Spanish.

among previously studied varieties of Spanish (section 3.2). Next, in section 3.3 we will compare

the Spanish noun data in the monolingual Spanish and bilingual Spanish groups and determine if

their distribution is similar (section 3.3.1), before comparing the English nouns of the

monolingual English and bilingual English groups to measure the same (3.3.1). We will then

summarize the differences between the constraint ranking in Spanish and English (section 3.4),

before finally being able to incorporate the LOLI data (section 3.5), separating them into

established (section 3.5.1) and nonce (section 3.5.2) groups for a more detailed analysis.

3. Results

3.1 Distribution of determiners by type

To better understand the type of data we have collected from our various speakers it is

important to consider the distribution of the determiners by type, thus answering RQ1.

RQ1: What is the distribution of determiner types and what is the overall frequency of bare

nouns among the several comparison groups?

While differences in the types of determiners most prevalent may differ between English and

Spanish, it would be ideal that the distribution between the same language monolingual and

bilingual pairs be similar. Table 3 shows the distribution of determiners amongst the varying

groups.

Table 3: The distribution of determiners by type in NPs extracted from Spanish monolinguals
from Sonora, Spanish/English bilinguals from Arizona, and English monolinguals from Arizona

Type	Spanish Monolingual		Spanish bilingual		Established LOLIs		Nonce LOLIs		Bilingual English		Monolingual English	
Definite	179	44.8%	99	24.8%	185	50.3%	116	29.9%	55	13.8%	61	15.2%
Indefinite	59	14.8%	120	30.0%	44	12.0%	60	15.5%	93	23.3%	77	19.3%
Possessive	28	7.0%	76	19.0%	23	6.2%	23	5.9%	77	19.2%	81	20.2%
Quantitative	7	1.8%	14	3.5%	1	0.3%	0	0.0%	25	6.2%	14	3.5%
Demonstrative	6	1.5%	22	5.5%	2	0.5%	2	0.5%	18	4.5%	16	4.0%
Complex	3	0.8%	15	3.8%	0	0.0%	0	0.0%	11	2.8%	8	2.0%
Zero	118	29.5%	54	13.5%	113	30.7%	187	48.2%	121	30.2%	143	35.8%
Total	400		400		368		388		400		400	

At first glance, Table 3 demonstrates that there are differences between English and Spanish.

While in Spanish there are more definite articles than indefinite (by almost triple in the Spanish

monolingual group), in English the difference between definite and indefinite is not very

pronounced. This pattern follows the results of Torres Cacoullos and Aaron (2003a), which

showed that there was a higher percentage of definite articles (38%) as opposed to indefinite

(9%) in Spanish and an almost equal distribution of definite and indefinite articles in English

(22% definite, 15% indefinite).

In terms of the distribution of determiners by monolingual and bilingual pairs, the

English group is very cohesive. For both monolingual and bilingual English data sets, the first

three determiner types are the same: indefinite, followed by definite, and then possessive. On the

other hand, there does seem to be a discrepancy between the monolingual and bilingual Spanish

groups. In the monolingual Spanish data set, there are more definite articles than indefinite articles, but the bilingual data set shows the opposite. One possible explanation for this is the higher tokens with possessive determiners in bilingual Spanish (N = 76) versus in monolingual Spanish (N = 28). A second, and likely, possibility is that the sample of 400 tokens is simply not as representative of the speech in the community as it is for the monolingual speakers (meaning that in these 400 tokens there just happened to have more indefinite determiners). The Spanish data provided in Torres Cacoullos and Aaron (2003a) was made up of 1,386 tokens, a much larger data set. The difference with the bilingual Spanish data here is not, by itself, telling that the bilinguals are behaving differently from the monolinguals.

Lastly, when focusing on the two LOLI groups, established and nonce, it is clear that both groups follow the Spanish pattern (the recipient language pattern) as previous studies have also reported (Blas Arroyo and Tricker, 2000; Torres Cacoullos and Aaron, 2003a, 2003b). On the other hand, Aaron (2014) in terms of determiner distribution, found that while the established LOLIs followed the Spanish (recipient language) pattern, the nonce LOLIs followed the English (donor language) pattern. While the frequency of determiner types in the data sets is not the clearest indication of morphosyntactic integration of the LOLIs (take for example the contradictory distributions of the Spanish monolingual versus bilingual data sets), the fact that the LOLIs do pattern after the Spanish (recipient language) data, is one indication that the LOLIs may be morphosyntactically integrated into the recipient language, thus aiding in their classification as borrowings, rather than codeswitches. However, this idea must be explored in greater detail, an undertaking that can be found in the next subsections of this chapter.

3.2 Overall frequencies

After establishing the distribution patterns of determiner type among the various

comparison groups, we will discuss the overall frequency of bare nouns in the data sets, thus

allowing us to answer RQ2:

RQ2: What is the overall frequency of bare nouns among the several comparison groups?

Table 4 shows the overall frequency of determiner realization in the 6 comparison groups for the

current study:

Table 4: The overall frequency of determiner realization in Spanish monolinguals from Sonora, Spanish bilinguals from Arizona, established LOLIs, nonce LOLIs, English bilinguals, and English monolinguals from Arizona

Overall %	Spanish Monolingual		Spanish Bilingual		Established LOLIs		Nonce LOLIs		English Bilingual		English Monolingual	
Bare	118/400	29.8%	54/400	13.5%	113/368	30.7%	187/388	48.2%	121/400	30.2%	143/400	35.8%
Determiner	282/400	70.5%	346/400	86.5%	255/368	69.3%	201/388	51.8%	279/400	69.8%	257/400	64.2%

Monolingual Spanish, monolingual and bilingual English, and established LOLIs follow a very

similar pattern of about 30% bare nouns. The greatest difference is that the bilingual Spanish

group has a much lower rate of bare nouns, by about half that of their monolingual Spanish

counterparts. This could be due to the higher rate of possessive determiners found in the

bilingual Spanish data set, which was discussed in the previous subsection. It just happened that

in the small segments of speech gathered from each of the 8 bilingual Spanish speakers, they

were often describing family members, and thus had more reason to use possessive nouns. Future

studies may consider not including possessives in the analysis to avoid such discrepancies between groups. Lastly, the nonce LOLIs have a higher overall frequency of bare nouns (almost half, N = 48.2%).

These data sets compare well to previously studied communities. First, it is generally found that the overall frequency of bare nouns does not differ between Spanish and English (Aaron, 2014; Torres Cacoullos and Aaron, 2003a, 2003b). Second, it also has been commonly found that nonce LOLIs have a higher rate of bare nouns as compared to patrimonial words and established LOLIs in various language pairings: Catalan words in Spanish discourse (Blas Arroyo and Tricker, 2000), English words in Spanish discourse (Aaron, 2014; Torres Cacoullos and Aaron, 2003a, 2003b;), and English words in French discourse (Poplack and Dion, 2012).

Due to the fact that frequencies cannot explain if the presence or absence of a determiner is conditioned by other factors, this measure cannot be used to shed light on whether the LOLIs are borrowings or codeswitches. We must then refer to the constraint hierarchies for the probability of a bare noun among the various comparison groups. We will first compare the monolingual and bilingual pairs, and then discuss the main differences between Spanish and English, before finally describing the LOLI pattern.

3.3 Patrimonial word patterns for Spanish and English

In this section, we will first explore the patterns of Spanish (monolingual versus bilingual) and then English (monolingual versus bilingual), discussing the similarities or differences between each monolingual-bilingual pair in this study as well as that of Torres Cacoullos and Aaron (2003a). This section will allow us to answer RQ3:

RQ3: How does the conditioning of determiner realization compare between the Spanish of the bilinguals and the Spanish monolinguals, and between the English of the bilinguals and the English monolinguals, and how do these results compare to those of New Mexico bilinguals in Torres Cacoullos and Aaron (2003a)?

3.3.1 Monolingual Spanish vs Bilingual Spanish

The data sets for both Spanish groups (monolingual and bilingual) consist of 400 tokens each. It is important to consider the conditioning of bare nouns in the two groups to determine whether or not the distributional differences reported in sections 3.1 and 3.2 are indeed important, whether they change the conditioning of the variable. If we find the two groups are similar in their conditioning, we can conclude that the distributional differences are rather superficial and not representative of the grammar involved in constructing noun phrases in the two groups. Table 5 shows the constraint hierarchies for bare nouns in the Spanish monolingual and Spanish bilingual groups.

Table 5: The conditioning of the probability that a noun will be bare in the Monolingual Spanish
and Bilingual Spanish groups

Spanish Monolingual (N = 400)				Spanish Bilingual (N = 400)			
Specificity	Factor Weight	Number	Percent	Specificity	Factor Weight	Number	Percent
Non-Specific	0.65	52/123	42.3%	Non-Specific	0.81	46/161	28.6%
Generic	0.59	58/173	33.5%	Generic	0.420	3/49	6.1%
Specific	0.21	8/104	7.7%	Specific	0.25	5/190	2.6%
	Range = 44				Range = 56		
Syntax				Syntax			
Object	0.59	72/183	39.3%	Pred. nom/Exist	0.64	7/32	21.9%
Pred. nom/Exist	0.52	16/47	34.0%	Object of Prep	0.60	18/123	14.6%
Object of Prep	0.41	28/156	17.9%	Object	0.58	28/180	15.6%
Subject	0.26	2/14	14.3%	Subject	0.13	1/65	1.5%
	Range = 33				Range = 51		
Modifier				Modifier			
Prenominal	0.71	29/54	53.7%	Prenominal	0.80	11/23	47.8%
Post-nominal	0.49	22/73	30.1%	Post-nominal	0.53	4/29	13.8%
None	0.46	67/273	24.5%	None	0.48	39/348	11.2%
	Range = 25				Range = 32		
String Position				Semantic Class	(NOT SIG)		
Second or more	0.68	27/53	50.9%	Occupation	[0.64]	4/16	25.0%
First or only	0.47	91/347	26.2%	Other	[0.50]	49/373	13.1%
	Range = 21			Institution	[0.47]	1/11	9.1%
					Range = [17]		
Polarity	(NOT SIG)						
Affirmative	[0.50]	114/384	29.7%	Polarity	(NOT SIG)		
Negative	[0.44]	4/16	25.0%	Negative	[0.648]	1/4	25.0%
	Range = [6]			Affirmative	[0.498]	53/396	13.4%
					Range = [15]		
Semantic class	(Knock out)						
Occupation	KO	2/4	50.0%	String Position	(NOT SIG)		
Other	KO	116/375	30.9%	Second or more	[0.41]	1/14	7.1%
Institution	KO	0/21	0.0%	First or only	[0.50]	53/386	13.7%
	Range = NA				Range = [9]		
Log Like = -206.378		Sign = p < 0.005		Log Like = -121.165		Sign = p < 0.05	

The results offered in Table 5 demonstrate that the two groups are strikingly similar in how bare nouns are conditioned. The first three factor groups for both communities are specificity, followed by syntax, and then modifier. Within these three factor groups the order is also very similar. The biggest discrepancy between the two is that while for Spanish monolinguals string position is significant, it is not for Spanish bilinguals.

Specificity is the most robust conditioner of determiner relation in NP constraining in the two groups. In both groups, non-specific nouns highly favor bare determiners (FW = 0.65 for monolinguals, FW = 0.81 for bilinguals), while specific nouns highly favor the use of a determiner (FW = 0.21 for monolinguals, FW = 0.25 for bilinguals). In both communities, generic nouns are between non-specific and specific, they slightly favor zero determiner in the monolingual group (FW = 0.59) and slightly favor determiners in the bilingual group (FW = 0.42). These results compare to those of bilingual speakers in New Mexico (Torres Cacoullos and Aaron, 2003a) where specificity was also the highest ranking factor group and the order was also non-specific, generic, specific. The main difference between the communities presented here and that in Torres Cacoullos and Aaron (2003a) is that in New Mexico generic highly disfavored bare nouns (FW = 0.27).

The second factor group in both communities is syntax. Here we notice that noun phrases in subject position strongly disfavor bare determiners (FW = 0.26 for monolinguals, FW = 0.13 for bilinguals). For the monolingual group, NPs that are objects or predicate nominal/existential favor bare nouns, while NPs that are objects of a preposition favor the realization of a determiner. The pattern is similar for the bilingual group, with the exception that NPs that are objects of a preposition also favor no determiner. In New Mexico Spanish, this factor group acted very similar to both the monolinguals and bilinguals in Arizonan Spanish in the current

study, where NPs that were predicate nominal/existential and objects favored bare determiners, objects of prepositions slightly disfavored bare determiners, and subjects highly disfavored bare determiners (Torres Cacoullos and Aaron, 2003b).

The presence of a modifier is another factor group that shows cohesion between the monolingual and bilingual Spanish data sets. In both groups, NPs with pre-nominal modifiers strongly favor (FW = 0.71 for monolinguals, FW = 0.80 for bilinguals) being realized as bare nouns, followed by post-nominal, and lastly no modifier. There is a difference between the two communities in that post-nominal slightly disfavor bare nouns in the monolingual group (FW = 0.49) and favor bare nouns in the bilingual group (FW = 0.53), however, the order of the factor rankings is the same in the two groups. Again, these findings are on par with those of bilinguals in New Mexico who showed that NPs with pre-nominal modifiers favored bare nouns and no modifier and post-nominal modifiers favored determiners (Torres Cacoullos and Aaron, 2003a).

The last factor group that is significant in either community is string position, which is significant for the monolingual group but not the bilingual group. For monolinguals, second and subsequent nouns in a series favor being realized as bare nouns, while first or only nouns favor determiners. This finding is identical to that of bilinguals in New Mexico (Torres Cacoullos and Aaron, 2003a).

Polarity is not a significant factor for either group, and the distribution of the tokens is very uneven, the vast majority of tokens fall into the affirmative category in both groups (96% for monolinguals, 99% for bilinguals), leaving less than 5% of the data in the negative category. This factor group was significant in the bilingual Spanish of New Mexican speakers, and negative NPs favored zero determiner, while affirmative favored a realized determiner.

The last factor group is semantic class. In both communities, this factor was either non-significant or resulted in a knock out (100% of the tokens behaved in the same way). The pattern of the effect is the same for both communities where occupation nouns favor zero determiner (50% for monolinguals, 25% for bilinguals), and other nouns (30.9% for monolinguals, 13.1% for bilinguals) and institutional nouns (0% for monolinguals, 9.1% for bilinguals) favor determiner realization. The same pattern was also found in the New Mexico bilinguals (Torres Cacoullos and Aaron, 2003a).

Based on the results shown in Table 5, it is abundantly clear that the Spanish monolingual and bilingual groups do indeed show very similar conditioning of bare nouns in their respective grammars. The same three factors with almost identical distributions are found in both data sets: specificity, syntax, and modifier. Additionally, the factor group of semantic class demonstrates the same pattern of occupation nouns favoring zero determiner and institutional nouns favoring a realized determiner. We can conclude, then, that when bilinguals use Spanish patrimonial words in Spanish discourse, they are following the pattern that Spanish monolinguals use, and as such, at least for Spanish, we can assume that LOLIs, if borrowed and integrated into the language, should also follow the monolingual Spanish pattern. It is of further interest to note that the trends found in the two data sets of this study parallel those found in New Mexico by Torres Cacoullos and Aaron (2003a). For all three communities, the same factor group order appears: Specificity, syntax, modifier. Additionally, the within-factor conditions are also similar among the communities, showing the same tendencies. Table 6 summarizes the similarities between the three communities.

Table 6: Similarities (highlighted in gray) of factors that favor/disfavor bare nouns in US-Mexico border Spanish communities (Sonora, Arizona, and New Mexico)

Factor	Sonora Monolingual (Current Study) N = 400	Arizona Bilingual (Current Study) N =400	New Mexico Bilingual (Torres Cacoullos and Aaron 2003a) N = 1386
Specificity			
Non-Specific	+	+	+
Generic	+	-	-
Specific	-	-	-
Syntax			
Object	+	+	+
Prenom/Exist	+	+	+
Object of Prep	-	+	-
Subject	-	-	-
Modifier			
Prenominal	+	+	+
Post-nominal	-	+	-
None	-	-	+
String Position			
Second or more	+	-	+
First or only	-	+	-
Polarity			
Affirmative	+	-	-
Negative	-	+	+
Semantic class			
Occupation	+	+	+
Other	-	-	+
Institution	-	-	-

Table 6 compares whether each within-group factor favors or disfavors a bare noun and highlights in gray the similarities between the three border Spanish communities. Due to the amount of similarities, it appears that these factors are very strongly ingrained in the variational grammar of Spanish.

3.3.2 Monolingual English vs Bilingual English

After discussing the Spanish patterns of the conditioning of bare nouns, we now turn to the English data sets based on 400 tokens from Group 4 and 400 tokens from Group 5, as detailed in section 2.2. Table 7 shows the results from the logistic regression analysis.

Table 7: The conditioning of the probability that a noun will be bare in
the Monolingual English and Bilingual English groups

Monolingual English (N = 400)				Bilingual English (N = 400)			
Specificity	Factor Weight	Number	Percent	Specificity	Factor Weight	Number	Percent
Generic	0.96	119/131	90.8%	Generic	0.94	87/104	83.7%
Non-Specific	0.36	17/110	15.5%	Non-Specific	0.45	20/100	20.0%
Specific	0.09	3/159	1.9%	Specific	0.2	14/196	7.1%
	Range = 87				Range = 74		
Semantic Class	(NOT SIG)			Syntax	(NOT SIG)		
Institution	[0.96]	14/15	93.3%	Object	[0.60]	56/156	35.9%
Other	[0.47]	129/384	33.6%	Subject	[0.57]	13/54	24.1%
Occupation	[0.00]	0/1	0.0%	Pred. nom/Exist	[0.42]	15/73	20.5%
	Range = [96]			Object of Prep	[0.39]	37/117	31.6%
					Range = [21]		
String Position	(NOT SIG)						
Second or more	[0.78]	8/12	66.7%	String Position	(NOT SIG)		
First or only	[0.49]	135/388	35.8%	Second or more	[0.69]	11/28	39.3%
	Range = [29]			First or only	[0.49]	110/372	29.6%
					Range = [20]		
Syntax	(NOT SIG)						
Object of Prep	[0.58]	49/114	43.0%	Modifier	(NOT SIG)		
Pred. nom/Exist	[0.51]	27/76	35.5%	None	[0.51]	88/283	31.1%
Object	[0.50]	60/171	35.1%	Prenominal	[0.49]	30/101	29.7%
Subject	[0.29]	7/39	17.9%	Post-nominal	[0.35]	3/16	18.8%
	Range = [29]				Range = [16]		
Modifier	(NOT SIG)			Polarity	(NOT SIG)		
None	[0.55]	120/302	39.7%	Affirmative	[0.50]	117/385	30.4%
Post-nominal	[0.40]	4/15	26.7%	Negative	[0.46]	4/15	26.7%
Prenominal	[0.35]	19/83	22.9%		Range = [4]		
	Range = [20]						
				Semantic Class	(NOT SIG)		
Polarity	(NOT SIG)			Institution	KO	11/11	100.0%
Negative	[0.64]	3/6	50.0%	Other	KO	110/386	28.5%
Affirmative	[0.50]	140/394	35.5%	Occupation	KO	0/3	0.0%
	Range = [14]						
Log Like = -117.804		Sign = p < 0.001		Log Like = -146.793		Sign = p < 0.001	

Just like with the Spanish data sets, we see striking similarities between the English monolingual and bilingual groups.

The only significant factor group in the English data sets is that of specificity. In both the monolingual and bilingual communities, NPs with generic nouns highly favor zero determiners (FW = 0.96 for monolinguals, FW = 0.94 for bilinguals), non-specific nouns disfavor zero determiners (FW = 0.36 for monolinguals, FW = 0.45 for bilinguals), and specific nouns highly favor a realized determiner (FW = 0.09 for monolinguals, FW = 0.20 for bilinguals). A comparable pattern was found in the English of New Mexico bilinguals (Torres Cacoullos and Aaron 2003a), with the exception of non-specific nouns, which were second in the ranking but favored zero determiners instead of realized determiners.

The remaining factor groups were not significant, but analyzing the pattern of the effect for each shows more similarities between the monolingual and bilingual groups than differences. For instance, with semantic class, NPs of institutional nouns highly favor bare nouns in both communities (93.3% for monolinguals, 100% for bilinguals), "other nouns" while disfavoring bare nouns are somewhat in the middle of the three semantic classes (33.6% for monolinguals, 28.5% for bilinguals), while occupational nouns clearly favor determiner expression (100% for monolinguals, 100% for bilinguals). In New Mexico, the bilinguals also favor bare nouns for institutional nouns, very slightly favor bare nouns for other nouns, and highly favor a realized determiner for nouns of occupation (Torres Cacoullos and Aaron, 2003a).

While string position does not prove to be significant in the logistic regression, it is third in the hierarchy of factor groups for both monolingual and bilingual data sets. Furthermore, the order of effect is identical for both groups, where the second or subsequent noun in a series favors a zero determiner (FW = [0.78] for monolinguals, FW = [0.69] for bilinguals) and the first

or only noun in a series favors a realized determiner (FW = [0.490] for monolinguals, FW = [0.49] for bilinguals). This same effect was also found among bilinguals in New Mexico and to almost the same degree (FW = 0.80 for second or subsequent, FW = 0.49 for first or only) as reported in Torres Cacoullos and Aaron (2003a).

The presence or absence of a modifier is the last factor group that is cohesive between the bilingual and monolingual communities. For both communities, no modifier favors a bare noun (FW = [0.55] for monolinguals, FW = [0.51] for bilinguals) and the use of a modifier favors an expressed determiner (pre-nominal: FW = [0.35] for monolinguals, FW = [0.49] for bilinguals; post-nominal: FW = [0.40] for monolinguals, FW = [0.35] for bilinguals). For English/Spanish bilinguals in New Mexico, the same pattern was found in Torres Cacoullos and Aaron (2003a), where NPs without a modifier favored bare nouns and NPs with a modifier favored a realized determiner.

The last two factor groups, both of which were not selected as significant in the logistic regression analysis, show the greatest difference between the two data sets. First, polarity shows that while in the monolingual group NPs that have negation favor bare nouns (FW = [0.61]) and affirmative NPs favor a realized determiner (FW = [0.50]), the opposite is true for the bilingual group (affirmative: FW = [0.50], negative: FW = [0.46]). However, the total number of tokens for the negative category is extremely low for both communities (N = 6 for the monolinguals and N = 15 for the bilinguals). The bilingual speakers in New Mexico followed the same pattern as the monolinguals of this study, in that negated NPs favored bare nouns and affirmative NPs favored a determiner (Torres Cacoullos and Aaron, 2003b). Second, the categories within the factor group of syntax are completely different in the two data sets. For the monolingual group objects of a preposition (FW = [0.58]) and predicate nominal/existential (FW = [0.51]) both

favor bare nouns, while objects (FW = [0.50]), and subjects (FW = [0.29]) favor the realization

of a determiner. For the monolingual group, objects (FW = [0.60]) and subjects (FW = [0.57])

favor bare nouns, while predicate nominal/existential (FW = [0.42]) and objects of preposition

(FW = [0.39]) disfavor zero determiners. In contrast, for the bilinguals of New Mexico, syntax

was the second most significant factor group. NPs that were objects, and objects of a preposition

favored bare nouns, while predicate nominal/existential and subject NPs favored the realization

of a determiner (Torres Cacoullos and Aaron, 2003a). This may be the first sign of a regional

difference between New Mexican and Arizonan English. However, in New Mexico (the authors)

measured 772 tokens of English, while in this study only 400 were measured. This difference

could also be related to the smaller sample size.

As we noted above, there are strong ties between the monolingual and bilingual English

data sets. Specificity is the only factor group selected as significant in the logistic model, and the

pattern of the effect is the same for both groups (generic nouns favor zero determiners and non-

specific and specific nouns favor a determiner). Three of the remaining non-significant factor

groups also show parallels, semantic class, string position, and modifier. While polarity was

opposite for the two communities, there were so few tokens in the negative polarity category that

this difference is most likely not true in the community grammars themselves, but rather an

artifact of the data sets. The only factor that could possibly show a difference due to bilingualism

is syntax, where both groups differed on every category within the factor group. However, since

this is the only factor group that shows such a striking difference, and the factor group was not

significant in the logistic model, it is safe to conclude that the similarities between the two

communities are strong enough to support the idea that the bilingual group does share the same

variational grammar as the monolingual group. It is therefore safe to postulate that since

bilinguals treat patrimonial English words the same way as monolinguals do, if the LOLIs show

English patterns, it is most likely due to the fact that the bilinguals are codeswitching, rather than

borrowing them.

The next sub-section will summarize the differences between the Spanish and English

patterns for the constraints on bare nouns in the two languages.

3.4 Differences between Spanish and English patterns

Having identified the conditioning of bare nouns in both English and Spanish separately,

we now provide a synthesis of the differences that are found between the two languages in order

to answer RQ4:

RQ4: What are the key conflict sites between Spanish and English determiner realization?

The main conflict sites, the constraint rankings that differ between the two languages, are found

in the behavior of four factor groups: specificity, semantic class, modifier and (possibly) syntax.

In the factor group of specificity, we find the most robust difference between Spanish and

English. Whereas in Spanish, non-specific nouns most highly favor bare nouns, in English,

generic nouns highly favor zero determiners. An additional conflict site is found within the factor

group of semantic class. In Spanish, occupation nouns favor zero determiners while institutional

nouns favor a realized determiner. In English, the opposite pattern is true, institutional nouns are

more often bare, while occupation nouns are realized with a determiner. A third difference

between Spanish and English determiner realization can be found in the modifier factor group.

Spanish nouns are more often bare if there is a modifier in the NP, especially if it is a pre-nominal modifier. In English, there is no effect of modifier and, in fact, nouns are more often realized with a determiner if the NP contains a modifier. Lastly, while this result was not found in the communities examined in this study, in New Mexico, Torres Cacoullos and Aaron (2003a) found that NPs in the syntactic function of predicate nominal/existential favored zero determiner, while English NPs in the same syntactic category favored a realized determiner. However, the high rate of zero determiners in Spanish for the nominal/existential category is most likely due to occupational nouns. Whereas in English occupational nouns in an existential phrase of "to be" are realized with a determiner as in example (26), in Spanish nouns of the same category are realized with zero determiner as seen in example (27):

(26) I am a teacher.

(27) Soy   maestro.

    (∅I) am (∅a) teacher.

It would be better, then, to use specificity, semantic class, and modifier as the key factor groups in comparing LOLIs to Spanish and English patrimonial words. Table 8 summarizes the key differences between Spanish and English determiner realization:

Table 8: Conflict sites for determiner realization
in Sonoran Spanish and Arizonan English

Factor	Spanish	English
Specificity	1. Non-specific 2. Generic 3. Specific	1. Generic 2. Non-specific 3. Specific
Semantic Class	Occupation + Institution -	Institution + Occupation -
Modifier	Pre-nominal + Post-nominal (+/-)	No effect (Pre-nominal -)

The factors outlined above as conflict sites are crucial in categorizing LOLIs as either borrowings or codeswitches. In the section 3.5 we will explore the conditioning of bare nouns in established and nonce LOLIs, paying close attention to conflict sites between English and Spanish in order to determine if the LOLI groups are integrated (Spanish like, borrowings) or not integrated (English like, codeswitches).

3.5 LOLI patterns

In this section we will analyze the constraint hierarchies for established and nonce LOLIs to determine if they pattern like Spanish nouns (thus indicating their status as borrowings), or if they pattern like English nouns (thus demonstrating that they would be more accurately classified as codeswitches) in order to answer RQ5:

RQ5: To what extent do established and nonce LOLIs show morphosyntactic integration into Spanish and what does the amount of integration suggest in terms of the status of LOLIs as codeswitches or borrowings?

3.5.1 Established LOLIs

In this subsection, we will discuss the conditioning of bare nouns in the established LOLIs found in Spanish discourse by the bilingual speakers in this study and compare it to the Spanish and English conditioning outlined in section 3.4. Table 9 summarizes the results.

Table 9: The conditioning of the probability that a noun will be bare among Established LOLIs in Arizona bilingual Spanish

Established LOLIs (N = 368)			
Specificity	Factor Weight	Number	Percent
Non-Specific	0.65	34/84	40.5%
Generic	0.59	72/198	36.4%
Specific	0.19	7/86	8.1%
	Range = 46		
Semantic Class			
Occupation	0.88	5/6	83.3%
Institution	0.62	33/83	39.8%
Other	0.45	75/279	26.9%
	Range = 43		
Syntax	(NOT SIG)		
Object of Prep	[0.54]	51/160	31.9%
Object	[0.51]	51/169	30.2%
Pred. Nom./Exist	[0.44]	10/25	40.0%
Subject	[0.17]	1/14	7.1%
	Range = [37]		
Modifier	(NOT SIG)		
Pre-nominal	[0.64]	2/6	30.7%
None	[0.50]	110/358	30.7%
Post-nominal	[0.42]	1/4	25.0%
	Range = [22]		
String Position	(NOT SIG)		
Second or more	[0.60]	12/28	42.9%
First or only	[0.49]	101/340	29.7%
	Range = [11]		
Polarity	(NOT SIG)		
Negative	KO	0/1	0.0%
Affirmative	NA	113/367	30.8%
	Range = [NA]		
Log Like = -205.301		Sign = $p < 0.01$	

As witnessed in Table 9, the established LOLIs seem to pattern very similarly to Spanish patrimonial nouns based on the factors presented in Table 8 in section 3.4. In terms of specificity, established nouns strongly favor zero determiners when the noun is non-specific (FW = 0.62), slightly favor zero determiners for generic nouns (FW = 0.59), and highly disfavor bare determiners among specific nouns (FW = 0.19). This same pattern was found in the Spanish monolingual group of this study, and among bilinguals in New Mexico (Torres Cacoullos and Aaron, 2003a). While the order is the same for the established LOLIs and Spanish bilingual groups, in the bilingual group, generic nouns disfavored zero determiners.

A second set of evidence of incorporation of the established LOLIs can be found in the semantic class factor group. Occupation nouns favor zero determiner for established LOLIs, while institutional nouns favor a realized determiner. When comparing this to Spanish patrimonial nouns, the same effect is observed for the Spanish monolinguals and bilinguals of this study, as well as the bilinguals in Torres Cacoullos and Aaron (2003a). This is a second strong parallel between Spanish patrimonial nouns and established LOLIs and a further continuity between Arizona and New Mexico Spanish.

The last conflict site between the conditioning of bare nouns in Spanish as compared to English that was outlined in section 3.3.3 was the factor group of modifier. For established LOLIs, pre-nominal modifiers favor bare nouns, while no modifier or post-nominal favor the use of a determiner. This result is again identical to the conditioning of Spanish patrimonial nouns seen in Table 8 in section 3.4. In all three studied communities, Spanish monolinguals and bilinguals in this study and New Mexican bilinguals in Torres Cacoullos and Aaron (2003a), Spanish NPs with pre-nominal modifiers is the category that most favors zero determiners.

Additionally, in New Mexican bilingual Spanish (Torres Cacoullos and Aaron, 2003a), the order

of the factors within the modifier factor group is the same (pre-nominal, none, post-nominal).

Established LOLIs pattern like Spanish patrimonial nouns in the three most

distinguishable conflict sites between Spanish and English determiner realization, specificity,

semantic class, and modifier, as shown in Table 10.

Table 10: Morphosyntactic integration of established LOLIs in Arizona bilingual Spanish

Factor	Spanish	Established LOLIs	English
Specificity	1. Non-specific 2. Generic 3. Specific	1. Non-specific 2. Generic 3. Specific	1. Generic 2. Non-specific 3. Specific
Semantic Class	Occupation + Institution -	Occupation + Institution -	Institution + Occupation -
Modifier	Pre-nominal + Post-nominal (+/-)	Pre-nominal + Post-nominal -	No effect (Pre-nominal -)

Based on the gray shading in Table 10 which represents factors that are conditioned the same

between Spanish patrimonial nouns and established LOLIs, it is clearly seen that established

LOLIs uttered in Spanish discourse by the bilinguals in this study are integrated into the Spanish

variational grammar and therefore should be classified as borrowings.

3.5.2 Nonce LOLIs

After having classified established LOLIs as borrowings, we now turn to a discussion of

the conditioning of the variable of bare nouns among the nonce LOLIs. Table 11 shows the

constraint ranking of bare nouns on nonce LOLIs.

Table 11: The conditioning of the probability that a noun will be
bare among nonce LOLIs in Arizona bilingual Spanish

Nonce LOLIs (N = 388)			
Specificity	Factor Weight	Number	Percent
Generic	0.82	135/171	78.9%
Non-Specific	0.331	36/102	35.3%
Specific	0.16	16/115	13.9%
	Range = 66		
Semantic Class			
Occupation	0.90	15/19	78.9%
Other	0.48	146/325	44.9%
Institution	0.45	26/44	59.0%
	Range = 45		
Modifier	(NOT SIG)		
Pre-nominal	[0.75]	10/16	62.5%
None	[0.49]	177/371	47.7%
Post-nominal	[KO]	0/1	0.0%
	Range = [75]		
String Position	(NOT SIG)		
Second or more	[0.68]	20/31	64.5%
First or only	[0.48]	167/357	46.8%
	Range = [20]		
Syntax	(NOT SIG)		
Pred. Nom./Exist	[0.53]	33/64	51.6%
Object	[0.52]	97/197	49.2%
Object of Prep	[0.47]	52/113	46.0%
Subject	[0.37]	5/14	35.7%
	Range = [16]		
Polarity	(NOT SIG)		
Negative	[0.56]	2/4	50.0%
Affirmative	[0.50]	185/384	48.2%
	Range = [6]		
Log Like = -192.780		Sign = $p < 0.01$	

As seen in Table 11, nonce LOLIs do show signs of morphosyntactic integration into Spanish. First, nonce LOLIs follow the Spanish pattern for semantic class where occupation nouns favor zero determiners (FW = 0.89) and institutional nouns favor a realized determiner (FW = 0.45). This pattern is the opposite in English (institutional nouns favor zero determiners while occupation nouns favor realized determiners). While New Mexican bilinguals did not follow the same pattern when using nonce LOLIs in that occupation nouns disfavored zero determiners, the order was similar in that institutional nouns disfavored bare nouns more (FW = 0.36) than occupation nouns (FW = 0.43) as reported in Torres Cacoullos and Aaron (2003a).

A second sign of integration into Spanish is found in the modifier factor group. Nonce LOLIs favor bare nouns when the NP includes a pre-nominal modifier (FW = 0.75). This finding is in line with the Spanish monolingual and Spanish bilingual groups of this study, as well as the New Mexican bilinguals in Torres Cacoullos and Aaron (2003a). In English, NPs that have pre-nominal modifiers disfavor zero determiners. The pattern of the nonce LOLIs, then, is clearly the Spanish pattern.

In addition to the evidence that nonce LOLIs show signs of morphosyntactic integration into the Spanish grammar for the bilinguals in the present study, there is one factor group, specificity, that suggest that the nonce LOLIs do not show the distribution seen in the use of established LOLIs or patrimonial words. In terms of specificity among the nonce LOLIs, generic nouns highly favor a zero determiner (FW = 0.82), while non-specific nouns highly disfavor a zero determiner (FW = 0.33). In Spanish, non-specific nouns most favor bare determiners in all three data sets available (Spanish monolinguals and bilinguals in this study, as well as bilinguals in New Mexico from Torres Cacoullos and Aaron, 2003a).

Based on the results outlined above, the nonce LOLIs in this data set do not perfectly match up with the recipient language, but instead show signs of integration while keeping some of the conditioners found to affect bare nouns in the donor language. The integration is summarized in Table 12 below.

Table 12: Morphosyntactic integration of nonce LOLIs in Arizona bilingual Spanish

Factor	Spanish	Nonce LOLIs	English
Specificity	1. Non-specific 2. Generic 3. Specific	1. Generic 2. Non-specific 3. Specific	1. Generic 2. Non-specific 3. Specific
Semantic Class	Occupation + Institution -	Occupation + Institution -	Institution + Occupation -
Modifier	Pre-nominal + Post-nominal (+/-)	Pre-nominal + Post-nominal -	No effect (Pre-nominal -)

From Table 12, it is clear that while nonce LOLIs do not show integration in terms of specificity, they do show integration in semantic class and the presence of a modifier. This leads to the conclusion that nonce LOLIs should also be classified as borrowings, following the results of every previous study on the integration of LOLIs (Aaron, 2014; Blas Arroyo and Tricker, 2000; Ghafar and Meechan, 1998; Poplack and Dion, 2012; Poplack and Meechan, 1995; Sankoff, Poplack and Vannianajan, 1990; Torres Cacoullos and Aaron, 2003a, 2003b; Torres Cacoullos and Virgil, 2002, 2003; among others).

The fact that the nonce LOLIs in this study are not completely integrated is also in line with previous accounts of LOLIs in Spanish discourse which have shown nonce LOLIs to be integrated into the recipient grammar, but not as fully integrated as established LOLIs (Aaron,

2014; Torres Cacoullos and Aaron, 2003a, 2003b). For example, in New Mexico Spanish, nonce LOLIs patterned like patrimonial Spanish nouns in specificity and syntax, but not in semantic class (Torres Cacoullos and Aaron, 2003a, 2003b). These findings also match well to the typology of lexical transfer upheld by Hougan (1950), Poplack (1993), and Weinreich (1953), which considers other language lexical elements to be on a continuum from completely integrated borrowed material to completely unintegrated codeswitching. While established LOLIs borrowings that are completely integrated into the grammar of the recipient language, nonce LOLIs are somewhere between established LOLIs and codeswitches, but clearly do act as borrowings.

4. Conclusion

In this chapter we have discussed in detail the conditioning of bare nouns in both English and Spanish, including among monolingual and bilingual speakers, and LOLIs in Spanish discourse. In general, we found that bilingual speakers share the same variable grammar as their monolingual counterparts for both Spanish and English determiner realization. It was also found that Spanish and English differ on the conditioning of three key factor groups: specificity, semantic class, and modifier. When comparing the results from the current study to New Mexico Spanish/English bilinguals (Torres Cacoullos and Aaron, 2003a), we also found overwhelming connections between the Spanish, English, and LOLI integration of New Mexico and Arizona bilinguals. These connections allowed us to conclude that the Spanish and English of Arizonan

and New Mexico bilinguals share a variational grammar, at least for the conditioning of determiner realization, showing strong ties between the two communities.

When examining the LOLIs, we discovered that established LOLIs are very integrated into Spanish, they follow a Spanish-like conditioning of determiners and diverge from an English-like conditioning, and therefore should be considered borrowings, in line with all previous accounts of LOLIs (Aaron, 2014; Blas Arroyo and Tricker, 2000; Ghafar and Meechan, 1998; Poplack and Dion, 2012; Poplack and Meechan, 1995; Sankoff, Poplack and Vannianajan, 1990; Torres Cacoullos and Aaron, 2003a, 2003b; Torres Cacoullos and Virgil, 2002, 2003; among others). Nonce LOLIs also follow Spanish-like patterns and thus should also be classified as borrowings, which again has been seen in all previous accounts (Aaron, 2014; Blas Arroyo and Tricker, 2000; Ghafar and Meechan, 1998; Poplack and Dion, 2012; Poplack and Meechan, 1995; Sankoff, Poplack and Vannianajan, 1990; Torres Cacoullos and Aaron, 2003a, 2003b; Torres Cacoullos and Virgil, 2002, 2003; among others). However, in this data set, the incorporation is not absolute, as some previous studies have found (Aaron, 2014; Torres Cacoullos and Aaron, 2003a, 2003b). While nonce LOLIs were incorporated for the factor groups of semantic class, modifier, and syntax, they were more English-like with specificity. These results indicated that nonce LOLIs are partially incorporated into Spanish and, while they should also be considered borrowings, they are most likely somewhere along the spectrum between established borrowings (fully integrated) and codeswitches (not at all integrated), in line with the typology of lexical transfer (Hougan, 1950; Poplack, 1993; and Weinreich, 1953).

The current study has furthered our understanding of LOLI behavior in several ways. First, by introducing relevant monolingual comparisons, we have ensured that the bilingual patterns that are used as a basis to determine if LOLIs are integrated into the recipient language

are similar to patrimonial words produced by monolinguals. This is an important advancement in

research on LOLIs because previous studies have always assumed that bilinguals would behave

like monolinguals, but never quantified the similarities. Second, we have provided a detailed

account of determiner realization in Arizonan Spanish and English, Sonoran Spanish, and among

LOLIs borrowed into Arizonan Spanish. Third, we have shown that not only does the method

laid out in Torres Cacoullos and Aaron (2003a, 2003b) work for additional data sets (i.e. the

Sonoran and Arizonan communities), but that there are strong continuities between Arizona/New

Mexico Spanish and English in terms of determiner realization. Fourth, we have provided more

evidence in favor of the nonce loan hypothesis (Sankoff, Poplack and Vannianajan, 1990) that

suggests that LOLIs in general show a trend towards morphosyntactic integration. The results

from this study are also crucial to the next chapters in this dissertation, which will explore the

correlation between phonological and morphosyntactic integration.

Chapter 3:

Phonological integration of LOLI nouns

1. Introduction

The study of lone other language items (LOLIs[13]) in bilingual discourse is the subject of a great breadth of research in variationist sociolinguistics (Aaron, 2014; Blas Arroyo and Tricker, 2000; Ghafar and Meechan, 1998; Myers-Scotton, 1993, 2006, 2009; Poplack, 2012; Poplack and Meechan, 1995; Poplack, Sankoff, and Miller, 1988; Sankoff, Poplack and Vanniarijan, 1990; Torres Cacoullos and Aaron, 2003a, 2003b; Torres Cacoullos and Vigil, 2002, 2003; among countless others). The main goal of this research has been to determine how such elements are incorporated, or not, into the recipient language. It is widely agreed upon that LOLIs show morphosyntactic integration and that the longer the element is in the language, the more consistent the integration takes place (Haugen, 1950). This integration brings the interpretation that LOLIs are more like borrowings and not like codeswitches, which are unintegrated segments of speech. Previous studies have focused on a wide range of morphosyntactic variables including case marking (Ghafar and Meechan, 1998; Sankoff, Poplack and Vanniarijan, 1990;), word order (Ghafar and Meechan, 1998; Poplack and Dion, 2012; Sankoff, Poplack and Vanniarijan, 1990,1990), plural marking (Ghafar and Meechan, 1998; Poplack and Dion, 2012; Sankoff, Poplack and Vanniarijan, 1990,1990), and determiner

---

[13] This acronym was coined by Poplack (2012).

realization (Aaron, 2014; Blas Arroyo and Tricker, 2000; Ghafar and Meechan, 1998; Poplack and Dion, 2012; Poplack and Meechan, 1995; Torres Cacoullos and Aaron, 2003b; Torres Cacoullos and Virgil, 2002, 2003; among many others). While morphosyntactic integration is widely studied, phonological integration has been left out. Since the seminal work of Poplack and Sankoff (1984) and Poplack, Sankoff, and Miller (1988), it has been accepted that phonological integration is too variable to be a good measure of integration (Aaron 2014; Blas Arroyo and Tricker 2000; Poplack and Dion 2012; Poplack, Sankoff, and Vanniarijan 1990; Torres Cacoullos and Aaron 2003a, 2003b; among many others). Poplack and Dion (2012) further assert, "recipient-language phonology often colors donor-language items even in the absence of mixing ("foreign accent"), so phonological criteria are not reliable indicators of loanword integration" (p. 284). The main purpose of the present study is to examine the phonological integration of LOLIs in bilingual speech in order to determine if a case can be made that phonological and morphosyntactic integration are correlated, and if phonological integration can be used as an additional measure when discerning borrowings from codeswitches. To this end, the present study examines LOLIs produced by bilingual (Spanish-English) speakers in Arizona, testing if the morphosyntactic variables that conditioned determiner realization of English LOLIs in Spanish discourse discussed in the previous chapter are correlated with phonological integration. Before exploring in detail the methodology that will be used in this study, we will first provide a more in-depth review of previous related literature (in Section 1.1), followed by a breakdown of the main and sub research questions that will allow us to examine the role of phonology in LOLI implementation (in Section 1.2).

1.1 Theoretical background and previous research

In the field of contact sociolinguistics, while morphosyntactic adaptation is well agreed upon as a diagnosis for the status of LOLIs as a borrowing, phonological integration is usually discarded for being too variable and therefore not a reliable factor in discerning one-item codeswitches from borrowings. This claim finds support in, or is mentioned in relation to, the work of Poplack and Sankoff (1984) and Poplack, Sankoff, and Miller (1988). In both studies, phonological integration is measured using impressionistic data to decide if each feature in a single token is from the donor language (English for both studies) or the recipient language (Spanish for Poplack and Sankoff (1984) and French for Poplack, Sankoff, and Miller (1988)). The results are reported in terms of the phonological integration per lemma and indicate that the frequency with which phonological integration takes place for a given lemma is variable. This is to say, that among the total tokens speakers have for a given lemma, sometimes the word is phonologically adapted and sometimes it is not. Specifically, in Poplack, Sankoff, and Miller (1988), established borrowings show phonological integration 79% of the time, while nonce words show 49% integration and codeswitching 27% (p. 73). Similar findings are reported in Poplack and Sankoff (1984). Additionally, some extralinguistic factors have been associated to the frequency with which LOLI lemmas are integrated phonologically in the community including the following: length of time the borrowing has been in the community (Poplack and Sankoff, 1984; Poplack, Sankoff and Miller, 1988), bilingual ability (Haugen, 1950, 1969; Mougeon, Beniak and Valois, 1985; Poplack and Sankoff, 1984; Poplack, Sankoff, and Miller, 1988), and the community itself (Poplack, Sankoff and Miller, 1988). In an attempt to theorize why phonological adaptation of LOLIs is variable among bilingual speakers, Poplack and

Meechan (1998) state that it is impossible to know whether a given token failed to be phonologically integrated because it is a codeswitch or whether the speaker "has borrowed both the word *and* its phonemes, in much the same way as speakers sometimes treat polymorphemic words as single units (e.g., French-origin infinitives in a variety of language pairs)" (p. 134-135). Subsequent research refers back to the results of Poplack and Sankoff (1984) and Poplack, Sankoff and Miller (1988) to reaffirm the idea that phonological integration is variable (Poplack, 2000; Poplack and Dion, 2012; Poplack and Meechan, 1998; Sankoff, Poplack and Vanniarajan, 1990; Torres Cacoullos and Aaron, 2003a) and therefore exclude it from the analysis. Other studies simply do not report on phonological integration and provide an analysis of morphosyntactic factors only (Aaron, 2014; Poplack and Meechan, 1995; Torres Cacoullos and Aaron, 2003b; Torres Cacoullos and Virgil, 2002, 2003).

Phonologists, on the other hand, based on lexical items borrowed in monolingual communities and under Optimality Theory, claim that any word borrowed in the discourse should go through the phonological constraints of the recipient language, and therefore should show phonological adaptation (Holden, 1976; Jacobs, Haike, and Gussenhoven, 2000; Kenstowicz and Suchato, 2006; Repetti, 2006; Sayahi, 2005; Yip, 1993, 2006). To the best of our knowledge the only formal phonology study to have examined phonological adaptation of borrowing among bilinguals, in the context of bilingual discourse, is MacSwan and Colina (2014) who analyze codeswitched segments.

MacSwan and Colina (2014) test the productions of intervocalic voiced stops in Spanish and English codeswitching. In Spanish, voiced stops are realized as approximants in intervocalic position. In their optimality theoretic analysis, the constraint ranking proposed is:

Agree(stricture)[14] >> Ident-IO (continuant)[15], Ident-IO(sonorant)[16]. On the other hand, in

English, voiced stops are realized as stops and the constraint hierarchy is: Identity-

IO(continuant), Ident-IO(sonorant) >> Agree(stricture). Since the two rankings are contradictory,

the authors predict that when faced with an intervocalic stop at the point of codeswitching, the

result will be a change in the grammar at the point of switch and when the change is to English, a

stop will be produced. Five simultaneous bilinguals (Spanish/English) were asked to produce 27

sentences involving a multi-segment switch from Spanish to English at the point of an

intervocalic stop, for instance "Hablamos de mi ghost yesterday" 'We talked about my ghost

yesterday'. The overwhelming majority of the cases for all five participants were realized as

stops, rather than approximants, supporting the idea that codeswitching is indeed abrupt and

total. The authors then tested whether a voiced segment in English could trigger voicing at the

end of a preceding Spanish word (in Spanish /s/ is voiced when followed by a voiced consonant:

Ident-IO Onset(voice)[17] >> Agree (voice)[18] >> Ident-IO Coda(voice)[19]). The prediction was that

the English voiced segment should trigger voicing in the Spanish /s/ since there is no

contradiction in the constraint rankings between the two languages (e.g. "Hablamos de mis

ghosts yesterday" 'We talked about my ghosts yesterday') and the [+voice] feature trigger is not-

language specific, and this result was borne out.

Further research by Bullock (2009) finds that Spanish-English bilinguals show English

voice onset time (VOT) rates that begin to converge with Spanish VOT rates at the site of

codeswitching, while the Spanish VOT rates remain unaffected by English. She uses these

---

[14] Adjacent segments must agree in manner of articulation.
[15] The input and output must agree in terms of the [+/− continuant] feature.
[16] The input and output must agree in terms of the [+/− sonorant] feature.
[17] The input and output must agree in terms of the [+/− voice] feature in the onset.
[18] Adjacent segments must agree in the [+/− voice] feature.
[19] The input and output must agree in terms of the [+/− voice] feature in the coda.

results to conclude that phonetics does show a cross-linguistic effect of codeswitching.

Additionally, it was found that when /l/ in coda position of a Spanish word was at the boundary

of a switch it was often produced as the English [ɫ] (dark l). These results lead Bullock to

conclude that there also may be phonological cross-linguistic effects along the codeswitching

border, in direct contradiction with MacSwan and Colina (2014). However, it is important to

consider the type of bilinguals that are being studied. Bullock (2009) incorporates bilinguals who

are dominant in one language over the other (some participants are Spanish dominant and others

are English dominant) and most likely began speaking the less dominant language at a later age

(she considers that bilinguals may have an "accent" in one or both of the languages they speak)

(p. 167). We should be careful to separate research on balanced bilinguals versus bilinguals who

are dominant in one language. On the other hand, MacSwan and Colina (2014) examine

balanced bilinguals who grew up speaking both English and Spanish. The clear phonological

separation at the point of codeswitch found in MacSwan and Colina (2014), provides further

evidence to incorporate phonological integration as another predictor, alongside morphosyntactic

integration, in discerning LOLI borrowings from codeswitches.

Having established a lack of investigation into the phonological integration of LOLIs and

the theoretical relevance to doing so, we now turn to a summary of how morphosyntactic

integration has been analyzed and the findings of the studies that undertake such analyses, since

we will also need to measure morphosyntactic incorporation in order to establish how

morphology and phonology are related in terms of the incorporation of LOLIs into a recipient

language. Two main procedures have been undertaken towards measuring degree of

morphosyntactic integration of LOLIs. First, Poplack and Meechan (1995, 1998) began applying

the sociolinguistic comparative method to loanwords. In doing so, the authors compare overall

frequencies of morphosyntactic variables that are realized differently between the two language pairings (donor and recipient). Comparisons are made between patrimonial words of the recipient language, established borrowings, nonce LOLIs, and words from the donor language (in the form of multi-word codeswitches). While many factors have been used as predictors, one of the most frequently included factors is the use of determiners (Aaron, 2014; Blas Arroyo and Tricker, 2000; Ghafar and Meechan, 1998; Poplack and Dion, 2012; Poplack and Meechan, 1995). These studies quantify the percentage of usage of definite articles, indefinite articles, demonstratives, possessives, and zero determiners among the various groups (donor language, recipient language, established borrowings, nonce LOLIs) and find that nonce LOLIs pattern after the recipient language, thus showing morphosyntactic integration.

The second procedure, from Torres Cacoullos and Aaron (2003a, 2003b), involves a multivariate analysis of the probability a NP will have a determiner or be realized as a bare noun based on several linguistics variables: specificity (specific, non-specific, generic), syntax (subject, object, oblique, predicate nominal/existential), use of a modifier (prenominal, post nominal, none), string position (first noun or only noun, second noun), polarity (negated, non-negated), and semantic class (occupation/status, institution, coincidence sites). Just like in the first procedure outlined above, comparisons are made between nouns of the recipient language, the donor language (in the form of codeswitched segments), established borrowings, and nonce borrowings. This method is more appropriate than the first since it is well attested that overall percentages can be misleading, because overall rates can be higher even when the same constraint hierarchy conditions the variable (see work on subject pronoun expression, for example, Torres Cacoullos and Travis, 2010; Travis, 2007; among others). The current study will follow the methodology offered in Torres Cacoullos and Aaron (2003a, 2003b).

Nevertheless, in both cases, measuring overall frequency or constraint ranking of variables within a multivariate analysis, the comparative sociolinguistic method allows investigators to compare LOLI behavior to patrimonial nouns in both the recipient and donor language. It has been hypothesized that if LOLIs pattern after nouns in the recipient language, they show signs of morphosyntactic integration, and therefore should be considered borrowings. However, if LOLIs pattern after nouns from the donor language, they show a lack of morphosyntactic integration, and therefore should be classified as codeswitches. To date, all previous research shows that LOLIs behave more like the recipient language, showing integration, and therefore should be considered borrowings.

More specifically, when assessing the factors that condition determiner realization in the language pairing of English and Spanish, four factor groups show divergent behavior in terms of the within-group factors, thus providing conflict sites that can be used to analyze LOLI conditioning (Chapter 2; Torres Cacoullos and Aaron, 2003a, 2003b). First, specificity (whether a noun represents a specific referent, a non-specific referent, or a generic/entire class referent), has shown to differ between nouns in English and Spanish. In English, generic nouns highly favor bare determiners, while in Spanish, non-specific nouns most favor zero determiners (Chapter 2; Torres Cacoullos and Aaron, 2003a). A second conflict site is found in the semantic class factor group. Whereas in English, occupational nouns favor a realized determiner and institutional (i.e. 'high school') nouns favor bare nouns, in Spanish institutional nouns favor a realized determiner and occupational nouns favor the bare form (Chapter 2; Torres Cacoullos and Aaron, 2003a). Additionally, the presence of a modifier affects Spanish and English determiner realization differently. Namely, in Spanish, the presence of a modifier favors zero determiner, and to the largest degree when the modifier is pre-nominal. In English, there is no

effect of the presence of a modifier (Chapter 2; Torres Cacoullos and Aaron, 2003a). Lastly, in New Mexico, Torres Cacoullos and Aaron found that in Spanish, nouns in nominal/existential syntactic position highly favor zero determiners, while in English, the same category favor the realization of a determiner (2003a). This same finding was not shared by the Arizona community (see Chapter 2), and therefore may not be a cross-dialectal pattern like the other three conflict sites that do share continuities between New Mexico and Arizona bilingual Spanish and English. Therefore, specificity, semantic class, and modifier are the morphosyntactic factors that will be correlated with phonological integration rates in the current chapter.

1.2 Research questions

The present study seeks to determine whether or not phonological and morphosyntactic integration are correlated, through the exploration of English LOLIs in Spanish discourse, as produced by Spanish/English bilinguals in Arizona. To that end, we must first explore the phonological integration that is taking place within the LOLIs and establish the type of variation that exists. If there is significant intra-word variation of Spanish/English phonology, we cannot use the method outlined above to compare morphosyntactic and phonological variation. However, if the variation is inter-lemma, meaning a particular lemma is sometimes produced in Spanish and sometimes in English, this would give more merit to testing how morphosyntactic and phonological integration are correlated. If established lemmas are not always produced in Spanish (which we would expect since they have been in the local grammar the longest), we need to explore how phonological integration is linked to morphosyntactic integration.

In order to organize the arguments provided in the results section, we will identify several research questions. First, our main research question for the experiment:

MRQ1: Are phonological and morphosyntactic integration correlated, and if so, what does the correlation mean to current methods that disambiguate one-item CS from borrowings?

However, before being able to answer MRQ1, we must first define the phonological variation that exists in the data set, which will be done categorically between non-integration (English phonology) and integration (Spanish phonology). In doing so, we will be able to answer sub research question 1 (SRQ1):

SRQ1: Can LOLIs be classified as integrated or unintegrated by phonological integration? If so, what ratio of Spanish and English phonology exists within the data set?

Once we have clearly established the level of phonological integration, we can then move on to testing the correlation between phonological and morphosyntactic integration by answering SRQ2:

SRQ2: Is there a difference between the constraint hierarchies of determiner realization between LOLIs with Spanish phonology and LOLIs with English phonology? If so, how do the differences compare to previously documented English and Spanish

morphosyntactic patterns (as represented in Torres Cacoullos and Aaron 2003a and

Chapter 2)?

Only after answering SRQs 1 and 2 can we then finally discuss MRQ1. The next section details

the methodology needed to fully explore MRQ1 and SRQs 1 and 2.

2. Methodology

2.1 The participants and the community in which they live

A total of 24 participants were included in the study. All participants were young adults

(no older than 30 years of age) and are divided evenly between men and women. The

participants reside in Arizona, either in Tucson, located sixty miles north of the border with

Sonora, or Nogales, located at the border with Sonora. According to the 2010 census, Tucson has

520,116 inhabitants, 42% of whom reported being Hispanic or Latino, and Nogales is home to

20,837 residents, 95% of whom are Hispanic or Latino (United States Census Bureau). In

Southern Arizona there is evidence of both language maintenance (due to constant immigration

of monolingual Spanish speakers) and language shift (from Spanish monolingual, to bilingual, to

English monolingual), as reported by Jaramillo (1995). The participants also describe crossing

the border on a frequent basis to visit Sonora and to be with family.

The participants are part of three different sociolinguistic corpora. A summary of the participants can be found in Table 1.

Table 1: List of Arizona participants by gender and by corpus

	Group 1 (Bessett 2012)	Group 2 (Carvalho 2012-)	Group 3 (Carvalho 2012-)
Males	4	4	4
Females	4	4	4
Total	**8**	**8**	**8**
	Total participants in the study: **24**		

The participants are all Spanish-English bilinguals and are part of three separate groups. Group 1 (N=8) was interviewed in Spanish by the investigator (Bessett, 2012) and maintained monolingual mode for virtually the entirety of the corpus (only 3 total cases of codeswitching are present in the corpus of approximately 24 hours). Group 2 (N=8) comes from the *Corpus del español en el sur de Arizona* (CESA) and the interviews were conducted by various interviewers (Carvalho 2012-). The final group, Group 3 is a subset of the CESA corpus, which partners two bilingual speakers and records their conversation.

The grand majority of the bilingual participants were born in the United States or arrived before they were ten, with the exception of one participant who arrived at age fifteen. They are all part of, or descendants of, families who emigrated from Sonora, Mexico, thus allowing for a cohesive sample from the same dialectal area. In order to ensure that the participants were bilingual, the participants were asked to evaluate their proficiency in both English and Spanish on a 0-10 scale where zero meant very low proficiency and ten meant very high proficiency (in

the case of the CESA corpus, Groups 4-5, the scale was 0-6). This assessment was adapted from the Bilingual Language Profile created by Birdsong, Gertken, and Amengual (2012). All of the participants reported a proficiency of at least six out of ten (or 4.5 of 6 in the case of participants from CESA) for both English and Spanish. The participants demonstrated their Spanish proficiency in that they were able to carry out a conversation in Spanish for the duration of the hour-long interview. Additionally, all participants attended school in the United States and live and work in a bilingual setting.

2.2 Measuring the phonological integration of LOLIs

First, all LOLIs from the bilingual participants were analyzed impressionistically by the investigator, classifying LOLIs as incorporated (Spanish phonology), non-incorporated (English phonology), or mixed (part of the LOLI contains English phonology and part contains Spanish phonology). After this global exploration was performed, integration ratios for the most common lemmas (having 3 or more tokens) were tabled.

In order to verify the investigator's judgments, a subset of the data (N = 179 words, approximately 24% of the total number of LOLIs) was used to create a perception experiment in which 19 English monolingual participants heard a LOLI (in isolation, separated from the Spanish sentence in which it was produced) and were asked how foreign or native the word was pronounced. Additionally, 20 LOLI lemmas were chosen randomly to create stimuli from a monolingual English speaker and a monolingual Spanish speaker as a control group making the total number of stimuli 219 words. Participants were given the following prompt at the start of the experiment:

A company is looking to hire workers to deal with customer phone calls. They have selected candidates and need to know if these candidates have Native-like English pronunciation. For each word you hear, please indicate how foreign or native the pronunciation sounds.

For each word, the participant was asked to move a slider from the left side of the screen (foreign pronunciation) to the right side of the screen (native pronunciation) with a continuous scale (from 0 to 1, giving 100 possible response values). Each participant heard each stimuli three times. An average of the responses from the 19 monolingual English participants (native judges) was calculated for each stimulus and was compared to the average of the responses of the investigator for each stimulus. A Pearson product-moment correlation coefficient was computed to assess the relationship between the responses of the investigator and the responses of the native judges. There was a positive correlation between the two variables (r=0.878, t=27.03, df=217, $p < 0.001$). Chart 1 shows the directionality of these results.

Chart 1: Correlation between investigator's response and native judge's response of perceived nativeness of LOLI pronunciation
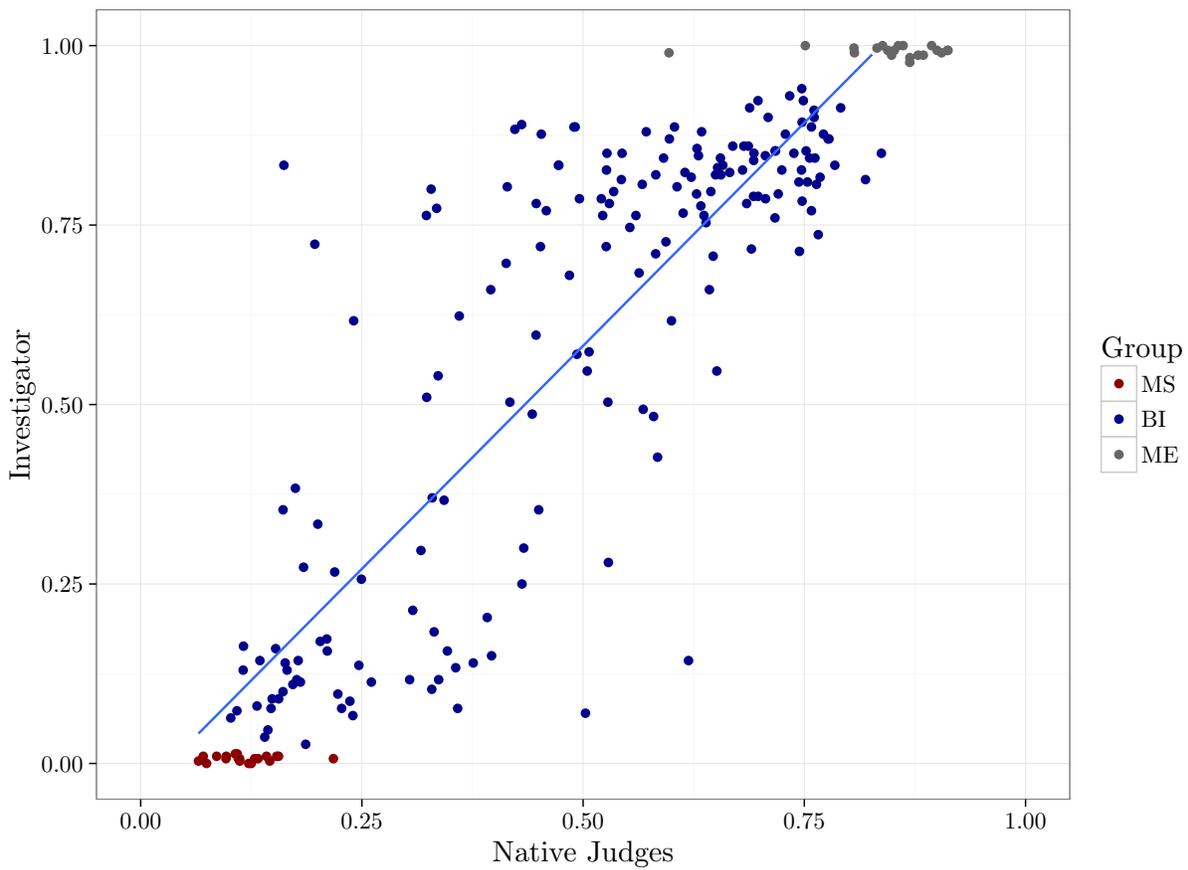
These results show that when the investigator perceives the word to be more native (approaching 1.00), the native judges also perceive the word to be more native. Conversely, as the investigator perceives the words to be more foreign (approaching 0.00), the native judges also perceive the words to be more foreign. We can see this relationship more closely if we consider only 16 words that were included and were produced by the MS (monolingual Spanish), ME (monolingual English) and BI (bilingual) speakers. Chart 2 shows this distribution.

Chart 2: Mean response of native judges compared to investigator's response of perceived nativeness of LOLI pronunciation
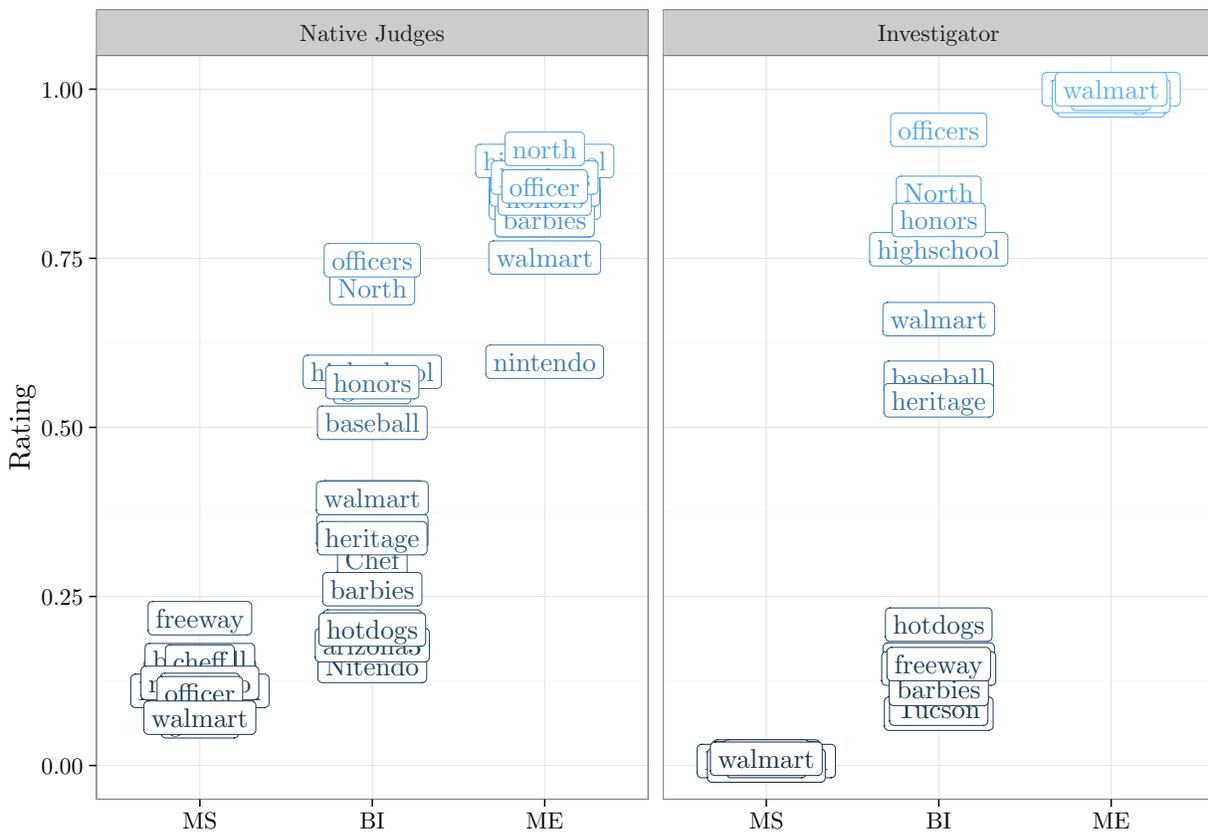
The largest difference that can be seen between the native judge responses and the investigator is that the investigator had less variation for the MS and ME groups. This was expected as the investigator knew the MS speaker and ME speaker and was able to recognize the voices. It is however, more important to notice that the native judges did respond in the expected manner, the MS speaker was perceived as being foreign (approaching 0.00), while the ME speaker was perceived as being native (approaching 1.00). Most importantly, the results show that the same words that were perceived as foreign by the investigator were also seen as foreign by the native judges. Interestingly, the native judges were fairly harsh in their assessment, most noticeably the

judgment of "Nintendo" by the ME speaker was about 0.60. Nevertheless, due to the high

correlation found (0.87) between the responses of the investigator and the native judges, it is

evident that the two groups perceive the words in a very similar manner.

As a further measure of phonological integration, two[20] commonly integrated elements

were analyzed with a more detailed analysis in PRAAT: the realization of rhotics and syllable

final laterals. First, the realization of rhotics was analyzed in 352 tokens (46.6% of the data), and

compared with the results of the investigator's classification of 287 tokens that were pronounced

with English phonology (produced with the retroflex [ɹ]) and 65 with Spanish phonology

(produced as the flap [ɾ] or the trill [r]). In Spanish, the flap [ɾ] is characterized by a single

occlusion at the end of the sound and the trill [r] is characterized by multiple occlusions (Willis,

2007). Figure 1 shows an example of "r" in Spanish and English.

---

[20]The decision to use rhotics and the lateral in coda position was based on previous research.
First, we wanted to only include categorical features as criteria based on the discrepancy in the
results seen in Bullock (2009) and McSwan and Colina (2014), where categorical features show
distinctive change at the boundary of a codeswitch, while gradient features (like VOT) show a
more gradual change. Other categorical elements were considered but excluded for various
reasons. The /b/-/v/ distinction was not included because in Southwest Spanish it has been
reported that some bilinguals do use /v/ (Torres Cacoullos and Ferreira, 2000). Additionally, the
assimilation of /s/ to [z] before voiced consonants was not included because this assimilation
does not always occur in Mexican dialects (Schwegler, Kempff, and Ameal-Guerra, 2010, p.
311).

Figure 1: Sample of Spanish and English "r" spectrums from LOLIs in Arizona Spanish

Spanish [ɾ] in "traila" 'trailer'   English [ɹ] in "freshman"
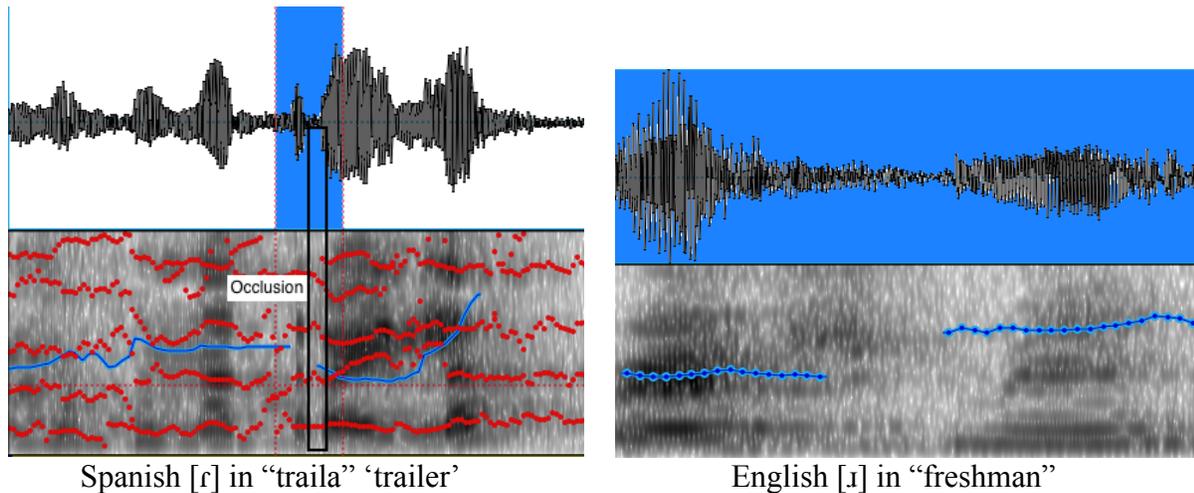
An analysis in PRAAT, by looking at the spectrogram to visually confirm the impressionistic data, revealed that 279 of the 287 tokens identified as having English phonology could be identified as showing signs of the retroflex English [ɹ], while 7 tokens had too much interference to properly read the spectrogram. Of the 65 tokens that had Spanish phonology, 64 were confirmed to have spectrograms matching what was expected for the Spanish flap [ɾ] or trill [r], while only 1 token had too much interference to be analyzed. Thus, the PRAAT analysis confirms the previous impressionistic classification.

Second, the use of the lateral in coda position was analyzed in PRAAT by viewing the spectrogram, again to confirm the impressionistic data. The key difference between "light l" (found in coda position in Spanish) and "dark l" (found in coda position in English) is that "dark l" has a much lower F2 value (Recasens and Espinosa, 2005). Figure 2 shows an example of "light l" and "dark l" from the LOLIs examined in this data set.

Figure 2: Sample of "light l" and "dark l" spectrograms from the LOLIs in Arizona Spanish

"Light l" from Spanish "fútbol" 'footbal'   "Dark l" from English "legal"
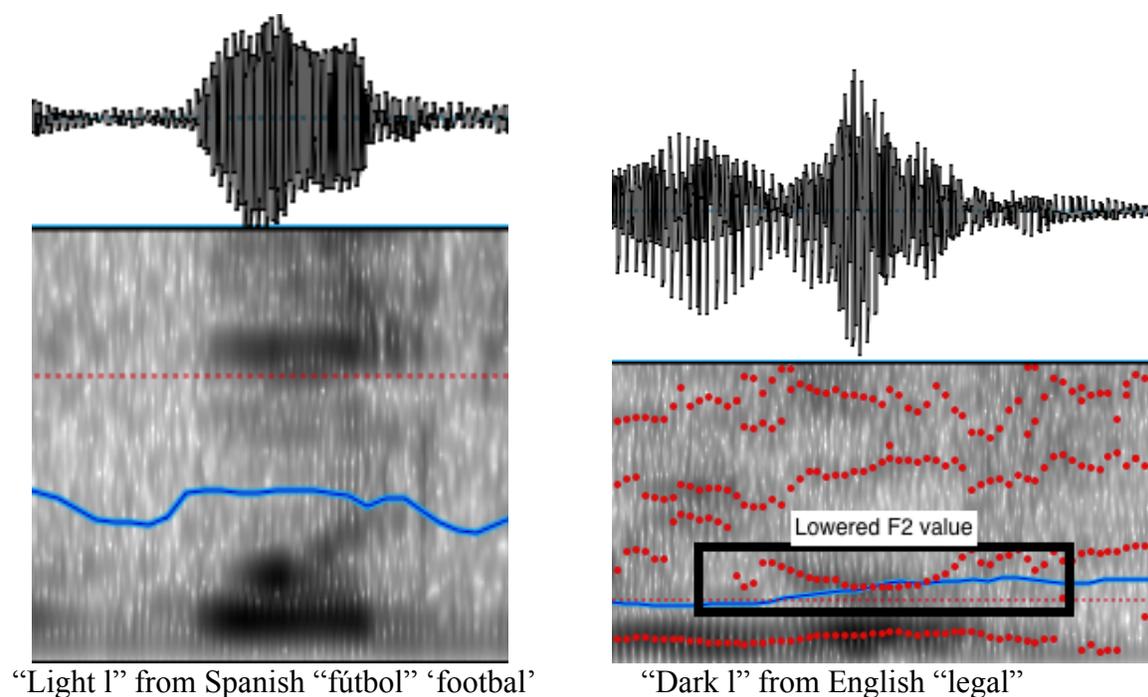
Among all 756 LOLIs, there were 193 tokens that contained a lateral in coda position (25.5% of the data). Spectrogram analyses revealed that among the 119 LOLIs that impressionistically appeared to have the English "dark l" (61.7% of all the /l/s produced in the corpus), 116 did, in fact, show acoustic properties of dark /l/, while 3 included too much interference for proper verification. Meanwhile, of the 74 tokens that impressionistically had the Spanish alveolar /l/, 70 were confirmed by the spectrogram and 4 were unable to be read due to interference.

The fact that the investigator's judgments correlate strongly with the 19 English monolingual participants in the perception study and that the spectrum analyses of /r/ and /l/ confirm those judgments, suggests a high saliency of Spanish-English phonology of the LOLIs present in this data set. These results demonstrate that the classification of the LOLIs by

phonology by the investigator is accurate and therefore will be used in the subsequent analyses in the results section.

The purpose of this descriptive analysis is to fill a gap in the literature in order to better understand what types of phonological variation exist in LOLIs produced by bilingual speakers, specifically for the Spanish-English bilinguals found in Arizona. The preliminary, two factor, detailed analysis in PRAAT is a starting point. Future studies will hopefully continue to analyze other phonological features using more sophisticated techniques than the impressionistic data presented in this section. However, the analysis of these two factors should serve as a starting point for future research and provide preliminary results that can contribute to the current lack of information about phonological adaptation of LOLIs in bilingual speech.

2.3 Correlation between morphosyntactic and phonological integration of LOLIs

For this experiment, phonological integration will be determined by first using the results of the impressionistic analysis described above. After separating the LOLIs with Spanish phonology and LOLIs with English phonology, we will compare the linguistic factors that condition determiner realization of the LOLIs with Spanish phonology to the LOLIs with English phonology, and afterward compare both groups to previous studies that analyze determiner realization of English nouns to Spanish nouns (namely, Chapter 2 and Torres Cacoullos and Aaron, 2003a). Each LOLI will be coded for several linguistic factors, following the methodology in Torres Cacoullos and Aaron (2003a, 2003b). The use of determiner versus bare noun will be used as the dependent variable. In order to determine the linguistic

conditioning of determiner usage, the following linguistic factors groups will be analyzed (examples are given for each factor):

- Specificity:

  - Specific (specific nouns referring to specific people or thing, they are not interchangeable)

    (1) "Pues, como, yo siempre estaba en el **board** preparando la comida." –T11

    'Well, like, I was always at the **board** preparing food.'

    (2) "Pues no, casi no, porque los **coaches** no saben enseñar…" –T17

    'Well no, almost never, because the **coaches** don't know how to teach'

  - Non-specific (any member of a class)

    (3) "Para mí es el peor día del año…porque es un **buzzkill**." –T12

    'For me, (it) is the worst day of the year…because (it) is a **buzzkill**."

    (4) "Pues no más fuimos a un, era un **gay bar**…" –T18

    'Well, (we) just went to a, (it) was a **gay bar**…'

  - Generic (refers to an entire class)

    (5) "Vamos al parque y jugamos **béisbol**." –T11

    '(We) go to the park and play **baseball**.'

    (6) "…los **customs** sí son groseros." –T18

    '…the **customs** (officers) really are rude.'

- Syntactic position:

  - Subject

    (7) "Los **border patrol** vienen de allá de Texas la mayoría..." –T18

    'The **border patrol** come from there, from Texas, the majority..."

- o Object

    (8) "Nos ponemos a hacer **popcorn**" –T18

    '(We) go about making **popcorn**.'

- o Oblique (object of a preposition)

    (9) "Tenía un carrito de esos del **power wheels**." -T18

    '(I) had a little car one of those of **Power Wheels**.'

- o Predicate nominal/existential

    (10) "…y ya es **senior**." –T02

    '…and (she) is (a) **senior**.'

- Presence of a modifier

    - o Prenominal

        (11) "Tenías que hacerlo fuera de la escuela, como un cierto **club**." -T18

        '(You) had to do it outside of school, like a certain **club**.'

    - o Post-nominal

        (12) "Estoy haciendo un **minor** en African American Studies." -6BS

        '(I) am doing a **minor** in African American Studies.'

    - o None

        (13) "De niña me gustaba mucho jugar con las **Barbies**." –T02

        'As a girl I really liked to play with **Barbies**.'

- Polarity

    - o Affirmative

        (14) "Me acuerdo mucho de jugar a **Hot Wheels**." –T02

        '(I) really remember playing **Hot Wheels**.'

- o Negative

    (15) "No tiene ni un **GED**." -8BS

    '(He) doesn´t even have a **GED**."

- String Position

    - o First or only noun in a series

    (16) "Está en el **Student Union**." –T02

    '(It) is in the **Student Union**'

    - o Second (or subsequent) noun in a series

    (17) "O sea, tienen un IHOP y un **Applebee's** y un **Peter Piper** y **Chili's**." –T02

    'So, (they) have an IHOP and an **Applebee's** and a **Peter Piper** and **Chili's**.'

- Semantic Class

    - o Occupation nouns

    (18) "…era **copy editor**…" –T02

    '…(I) was (a) **copy editor**.'

    - o Institution nouns

    (19) "Apenas este verano fue cuando ya pues, terminé la **high school**." –T23

    'Just this summer was when finally, well, (I) finished (the) **high school**.'

    - o Other (mass nouns, countable things, humans, locations, time expressions)

    (20) "…pero, pues, como **hobbies,** me gusta escuchar música…." –T23

    '…but, like, for **hobbies**, I like to listen to music...'

    (21) "…nos dejó dar una **U-turn**…" –T23

    '…(he) let us make a **U-turn**…'

2.4 Data analysis

The following logic will be applied to the interpretation of the results of this analysis.
First, let us assume that it is found that the LOLIs impressionistically analyzed in Section 2.2
show there to be no, or very infrequent, variation within a single token (tokens that show both
Spanish and English phonological elements). We do expect, as reported in the literature (Poplack
and Sankoff, 1984; Poplack, Sankoff, and Miller, 1988), that there will be variability within
lemmas (specific word types that sometimes are pronounced with Spanish phonology and
sometimes with English phonology), but we are referring to partially integrated tokens here. If
the conditioning of the factors of determiner use for the English phonological LOLIs differs from
the factors conditioning determiner use for the Spanish phonological LOLIs, and the differences
match the nouns from the respective language pairing (i.e. Spanish phonological LOLIs pattern
like Spanish nouns), we can conclude that phonological integration is to a certain extent
correlated to morphosyntactic integration, or the use of determiners. These results would then
show that phonological integration can be used as a factor in predicting the difference between
codeswitches and borrowings. If both morphosyntactic and phonological factors are found to be
correlated, we can conclude that examining phonological integration is fruitful when trying to
disambiguate a borrowing from a one-item codeswitch, countering the well-held view in
sociolinguistic research that phonology does not play a role in this distinction.

3. Results

A total of 799 LOLI nouns were extracted from the 24 interviews of Arizona Spanish with the Spanish-English bilingual speakers. However, 6 of the 24 participants demonstrated an unexpected use of nouns of sport names (*fútbol* 'football', *béisbol* 'baseball', ect.) in that they appeared without determiners on a(n almost) categorical bases (sport nouns in Spanish highly favor the use of a determiner). Upon further review of sports nouns, it became clear that this is true for all sports nouns from these speakers, including nouns like *baloncesto* 'basketball', which is not a borrowing, since the stem comes from the Spanish *balón* 'ball'. Due to this divergent behavior, the 43 tokens of sports nouns from the 6 participants were excluded from the data set, leaving a total of 756 nouns to be further analyzed.

In the next section we will fully explore the phonological integration that is found among the 756 LOLIs. We will then discuss the relationship between phonological integration and LOLI lemma. Finally, we will separate LOLIs with Spanish phonology from LOLIs with English phonology and examine the conditioning of determiner realization (morphosyntactic integration) between the two groups, thus exploring the correlation between phonological and morphosyntactic integration of LOLIs.

3.1 Phonological integration

3.1.1 Overall Frequency of Spanish, English, and mixed phonology

The overall analysis of phonological integration shows that the overwhelming majority of

LOLIs in the three data sets are produced with English phonology (72.2%), while only 26.7% are

produced with Spanish phonology, as seen in Table 2.

Table 2: Frequency of phonology type (Spanish, English, mixed)
of LOLI produced in Spanish discourse by Arizona bilinguals

Phonology Type	Number	%
Spanish	202/756	26.7%
English	546/756	72.2%
Spanish/English separate words	5/756	0.7%
Spanish/English same word	3/756	0.4%

There is only a total of 8 tokens (of the 756 total, 1% of the data) that do not maintain all one

phonology (Spanish or English). Of these, 5 show mixed phonology at the word level (as part of

a multi-word LOLI). Only three tokens show Spanish and English phonology in the same word

(out of 756 total tokens). Table 3 shows the list of words, the phonological pattern, and the

participant who uttered the LOLI that presented mixed phonology.

Table 3: Phonology pattern of multi-word LOLIs with
mixed phonology (Spanish-English) in Arizona Spanish

LOLI	Phonology pattern (Spanish or English)	Participant
Ford Raptor	Ford (English); Raptor (Spanish)	CESA048
Pontiac GTO	Pontiac (Spanish); GTO (English)	CESA049
Circle K's	Circle (Spanish); K's (English)	T01
Latin American Studies	Latin (Spanish); American Studies (English)	CESAGi5
Latin American Studies	Latin (Spanish); American Studies (English)	CESAGi5

The tendency, 4 out of 5 of the cases, is that the phonology goes from Spanish to English, a trend

that is expected since the utterances all happen in contexts where Spanish words precede the

LOLI utterance. The three remaining LOLIs with mixed phonology are within-word mixes and

can be seen in Table 4 along with a description of where the phonology changes.

Table 4: Phonology pattern of within-word
mixed phonology (Spanish-English) in Arizona Spanish

LOLI	Phonology pattern (Spanish or English)	Participant
Funtasticks	fun [fan] (Spanish); tastics (English)	T01
Hotdogs	Spanish phonology on the word except that the /s/ is voiced ([z]) after the voiced [ɣ] which is the English voicing pattern (progressive assimilation) and not the Spanish (regressive assimilation) since the following sound is a vowel.	CESA049
English	First Vowel Spanish, rest of the word English	T12

For two of the three tokens ("Funtasticks" and "hotdogs"), the change in phonology comes at a

morpheme boundary. First, "Funtasticks" is a compound composed of the word "fun" and the

second half of the word "fantastic". Second, the change in phonology in "hotdogs" happens on

the plural morpheme "-s". There is only one clear token in which the mixed phonology happens within a single morpheme, "English", leaving 1 of 756 tokens (0.1% of the data) with a truly mixed phonology within the word. For this reason, we can conclude that mixing phonology is not a normal part of the process by which bilinguals incorporate LOLIs from a donor language into a recipient language in these data sets. We can consider this one token a performance error or a bilingual slip-of-the-tongue.

Overall, LOLIs are produced with either Spanish or English phonology (in 99% of the data), there seems to be little evidence that phonological integration happens within a word. This is to say, it seems that phonological integration is an all or nothing process, whereby words are either integrated into Spanish phonology or maintained with English phonology. These results support the theoretical accounts to the phonology of codeswitching, for instance MacSwan and Colina (2014) state "there are therefore no opportunities for switching from one phonological system to another within words" (p. 190). Furthermore, the results of the present study coincide with previous phonological accounts of loanword adaptations discussed in the introduction, in that adaptation occurs when a word goes through the constraints of the recipient language's grammar (Holden, 1976; Jacobs, Haike, and Gussenhoven, 2000; Kenstowicz and Suchato, 2006; Repetti, 2006; Sayahi, 2005; Yip, 1993, 2006). Bilinguals, do act as monolinguals in the sense that they use a word-by-word implementation of the constraints. This idea also provides evidence from spontaneous speech to support formal phonology approaches to loanword adaptation and also to the hypothesis that phonological processes happen at both the lexical and post-lexical levels (Ito and Mester 2003a, 2003b; among others). The key difference is that while monolinguals must always use the recipient language's grammar, bilinguals, who have access to both the recipient and donor languages' grammars, are able to use either set of constraints when

incorporating a LOLI from the donor language to the discourse of the recipient language. The question, then, is whether we can explain when bilinguals will be more likely to choose one grammar over the other. Sections 3.1.2 and 3.1.3 will begin to answer this question and Section 3.2 will further the inquiry through the exploration of the correlation between phonological and morphosyntactic integration.

3.1.2 Results of phonology by lemma

In order to further understand the phonological integration that is found in the LOLIs in Arizona Spanish, we will now focus on phonological integration by lemma. Table 5 shows the phonology (Spanish or English) of each LOLI lemma that had three or more tokens.

Table 5: Phonological integration (Spanish versus English phonology)
by most frequent lemmas (3 or more tokens) in Arizona Spanish

Lemma	Spanish	English
high school	19/39	20/39
football (fútbol)	19/19	0/19
Truck (troque)	14/14	0/14
Telephone (teléfono)	11/11	0/11
Facebook	1/10	9/10
minor	0/9	9/9
Ford	1/8	7/8
Freshman	0/8	8/8
cellular (celular)	7/7	0/7
freeway	5/6	1/6
hotdogs*	2/6	3/6
mall	4/6	2/6
McDonald's	5/6	1/6
ticket	2/6	4/6
break	0/5	5/5
club	4/5	1/5
experience	0/5	5/5

Park Place	0/5	5/5
sushi	4/5	1/5
trolley	0/5	5/5
U of A	0/5	5/5
antifreeze	4/4	0/4
border patrol	0/4	4/4
Instagram	0/4	4/4
Kinder	4/4	0/4
Nintendo	4/4	0/4
percent	0/4	4/4
Pima	4/4	0/4
Spanglish	3/4	1/4
AutoZone	0/3	3/3
business	0/3	3/3
carwash	0/3	3/3
Chick-fil-A	0/3	3/3
cowboy	0/3	3/3
customer service	0/3	3/3
customs	0/3	3/3
extra credit	0/3	3/3
gasoline (gasolina)	3/3	0/3
Little Caesar's	0/3	3/3
master's	0/3	3/3
Muppet	0/3	3/3
muscle car	0/3	3/3
mustang	1/3	2/3
NAU	0/3	3/3
Pontiac*	2/3	0/3
record	3/3	0/3
social security	0/3	3/3
Toyota	3/3	0/3
UMC	0/3	3/3
Walmart	2/3	1/3
warning	0/3	3/3
Wings Over Broadway	0/3	3/3

*1 mixed

It is clear that some high-frequency and well-established borrowings show full integration into

Spanish phonology in that all tokens are produced with Spanish phonology (those that are

highlighted in dark grey, for example 'football', 'truck', and 'antifreeze'). At the same time, there are some high-frequency but non-attested borrowings that only have English phonology (those that are in light grey, for instance 'warning', 'muscle car', and 'carwash'). There are also established lemmas that are sometimes produced with Spanish phonology and sometimes produced with English phonology (e.g. 'high school', 'McDonald's', and 'hotdogs'). Whereas there is very little within-word variation (as seen in the previous section), there does seem to be a great deal of within-lemma variation (14/52 of the lemmas that have 3 or more tokens show variation). This variation does not always take place among unattested nouns, which is where we would expect to find it due to the fact that it is believed that the longer an element is in the recipient language, the more integrated it will be (Haugen 1950). The variation of the phonological integration previously reported must have been within-lemma variation and not within-word variation (see work by Poplack and associates). As stated in the introduction, within-lemma variation is worth examining to see if morphosyntactic integration is correlated to phonological variation.

When further analyzing the within-lemma variation, of particular interest are high-frequency and well established borrowings that are sometimes produced with Spanish phonology and sometimes produced with English phonology. These nouns often act morphosyntactically like Spanish when produced with Spanish phonology and morphosyntactically like English when produced with English phonology. Take for example lemmas 'high school' and 'Walmart'. Both of these nouns are institutional nouns, a category that highly favors a determiner in Spanish but not in English (see Chapter 2; Torres Cacoullos and Aaron, 2003a). Of the 19 tokens of 'high school' that are produced with Spanish phonology, all 19 were realized with a determiner. On the other hand, of the 20 tokens of 'high school' produced with English phonology, 15 were bare and

only 5 had a realized determiner. Additionally, for 'Walmart', the two Spanish phonology tokens had a realized determiner while the 1 English token was bare. These findings suggest that morphosyntactic integration may indeed be correlated to phonological integration, a notion we will test in greater detail in the Section (3.2).

Another way to explore the relationship between phonological integration and lemma, is through the establishedness of the lemma. In a separate experiment (see Chapter 2), the same LOLIs from the current study were separated into well-established borrowings (those that could be found in official Spanish dictionaries (Academia Mexicana de la Lengua, 2015, 2010; Corominas, 1967; García Icazbalceta, 1899; Lara Ramos, 1996; Lugones, 1932; Real Academia de la Lengua Sonorense, 2006; Real Academia Española, 2016), regional borrowings (found in local dictionaries or lists of borrowings in the Spanish from the Southwest United States: Cerda, Cabaza, and Farias, 1970, 1953; Cobos, 1983, 2003; Galván and Teschner, 1985, 1975; García, 1935; Gross, 1935; Hamel, 1994; Reyna, 1980; or the Bessett, 2012 corpus of monolingual Hermosillo, Sonora, Mexico Speakers), and unattested borrowings. Chart 3 shows the results of percent of phonological integration (Spanish phonology) of LOLIs in the three separate establishedness categories.

Chart 3: Phonological integration (Spanish phonology) of LOLIs in Arizona Spanish by establishedness of the LOLI lemma
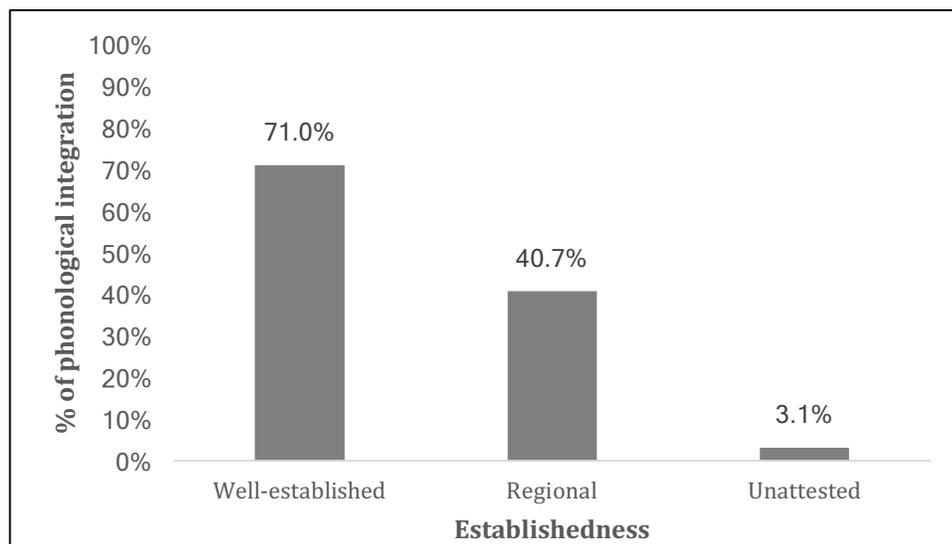
As seen in Chart 3, lemmas that have been in the local grammar for the longest period of time (well-established) are more often phonologically integrated into Spanish (98/138, 71%), followed by regional lemmas (92/226, 40.7%), and less often integrated are unattested borrowings (12/384, 3.1%). These results match those of Poplack, Sankoff, and Miller (1988) who also found that established borrowings were most often integrated, followed by nonce borrowings, and lastly codeswitches. Furthermore, this pattern upholds to the widely held notion that the longer a lemma is in the language, the more stable the phonological integration will be (Haugen, 1950; Ortigosa and Otheguy, 2007; Poplack, 2012; Poplack, Sankoff, and Miller, 1988).

3.1.3 Summary of phonological integration

In Section 3.1 we explored the way in which phonological integration is undertaken when Arizona bilinguals borrow LOLIs from English and implement them into Spanish discourse. In doing so, we are able to answer SRQ1 which asks:

SRQ1: Can LOLIs be classified as integrated or unintegrated by phonological integration? If so, what ratio of Spanish and English phonology exists within the data set?

The overwhelming answer is, yes, LOLIs can be classified by phonological integration. Of the 756 LOLIs in the current study, 99% showed either Spanish or English phonology. The majority of the 1% of LOLIs produced with mixed phonology involved mixing at the word-level (0.7% of

the data) or at a morpheme boundary (0.4% of the data), while only 1 token was produced with mixed phonology within a word (0.1% of the data). In relation to the second part of SRQ1, 72.22% of the LOLIs were produced with English phonology, while 26.72% were produced with Spanish phonology.

A closer look at phonological integration by lemma revealed that, as previously reported (Poplack, Sankoff, and Miller, 1988), established borrowings are most often integrated, followed by regional, and lastly unattested borrowings are most often unintegrated. Additionally, while some high frequency and well-established lemmas are categorically produced with Spanish phonology, this is not the case for all. Interestingly, there seems to be a correlation between when a word is phonologically integrated and when it is morphosyntactically integrated, a point emphasized viewing the relationship between determiner realization and phonological integration in institutional nouns (for instance, "high school and "Walmart"). These findings warrant a more thorough investigation into the relationship between phonological and morphosyntactic integration of LOLIs. Section 3.2 endeavors to explore said relationship.

3.2 Correlation between phonological and morphosyntactic integration

As detailed in the methodology section, the LOLIs, once classified into Spanish or English phonology, were then coded for the factors that condition determiner realization in NPs, following the methodology of Torres Cacoullos and Aaron (2003a, 2003b). By comparing determiner realization between LOLIs with Spanish phonology and LOLIs with English phonology, we expect to be able to further detail the connection between phonological and morphosyntactic integration of LOLIs. If the two groups differ in the conditioning of

determiners, and LOLIs with Spanish phonology follow the pattern of Spanish patrimonial nouns, while LOLIs with English phonology follow the pattern of English patrimonial nouns, we can then make the claim that phonology and morphology are indeed correlated in terms of LOLI integration. In the next sub-sections, we will perform such an analysis, determining if there is a correlation between morphosyntactic and phonological integration. We will start by examining the overall frequency of bare nouns and after that we will determine the factors that condition determiner realization.

3.2.1 Overall Frequency

While overall frequency rates are not always the best means of measuring differences between communities in general (Torres Cacoullos and Travis, 2010; Travis, 2007; among others), and rates are relatively similar in previous studies on determiner realization between the language pairings of Spanish and English, there are slight differences that can be observed. Table 6 provides a list of the percentage of a bare noun in previously studied communities of English patrimonial words, Spanish patrimonial words, and the LOLIs from the current study (separated by phonology).

Table 6: Use of a bare
noun among nouns in previously reported Spanish/English
Southwest US/Northern Mexico communities

Bare nouns in Spanish		
Community (Study)	Frequency	Overall %
Arizona Spanish Bilinguals (Chapter 2)	54/400	13.5%
Spanish Phonology (Current Study)	**29/202**	**14.4%**
New Mexico Spanish Bilinguals (Aaron 20114)	253/865	29.2%
Sonora Spanish Monolinguals (Chapter 2)	118/400	29.5%
New Mexico Spanish Bilinguals (Torres Cacoullos and Aaron 2003a)	413/1386	30.0%
Bare nouns in English		
Community (Study)	Frequency	Overall %
New Mexico English Bilinguals (Aaron 20114)	176/608	28.9%
New Mexico English Bilinguals (Torres Cacoullos and Aaron 2003a)	233/772	30.0%
Arizona English Bilinguals (Chapter 2)	121/400	30.2%
Arizona English Monolinguals (Chapter 2)	143/400	35.8%
English Phonology (Current Study)	**266/546**	**48.7%**

The overall frequency for LOLIs with Spanish phonology (14.4%) appears at the low end of the

spectrum of previous reported bare noun rates for Spanish in the US Southwest and Northern

Mexico. The rate is also very close to the rate reported for Spanish nouns among Arizona

bilingual speakers (see Chapter 2). LOLIs with English phonology show a very high rate of bare

nouns (48.7%) compared to previously reported rates for English in the US Southwest, however

the rate is comparable to Arizona monolingual English speakers who also showed a higher

(35.8%) rate (Chapter 2). While both English and Spanish are very close in the rate in which

bare nouns occur, there is a slight tendency that English has more bare nouns (with a range of

28.9%-35.8%) than Spanish (with a range of 13.5%-30.0%). The fact that the LOLIs with

Spanish phonology are low (14.4%), while the ones with English phonology are so high (48.7%),

could be an indication that the Spanish LOLIs act as Spanish patrimonial nouns while the

English LOLIs act as English patrimonial nouns. However, it is important to consider the

constraints that condition the probability of a bare noun, which will be examined in Section

3.2.2.

3.2.2 Constraint ranking of determiner realization by phonological integration of LOLIs

After comparing the overall frequency of bare nouns between LOLIs with Spanish and

English phonology, it is important to see how bare nouns are conditioned between the two

groups and note if there are differences. Table 7 shows the results of the regression analysis.

Table 7: The conditioning of the probability of a bare noun
among LOLIs in the Spanish discourse of Arizona Spanish-English bilinguals,
separated by phonological integration

LOLIs with Spanish phonology (N = 202)				LOLIs with English phonology (N = 546)			
Specificity	Factor Weight	Number	Percent	Specificity	Factor Weight	Number	Percent
Non-Specific	0.86	20/48	41.7%	Generic	0.80	196/254	77.2%
Generic	0.44	8/103	7.2%	Non-Specific	0.37	49/135	36.3%
Specific	0.19	1/42	2.3%	Specific	0.15	21/136	13.4%
	Range = 67				Range = 65		
Semantic Class	(NOT SIG)			Semantic Class			
Occupation	[1.00]	2/2	100.0%	Occupation	0.87	18/23	78.3%
Other	[0.56]	25/138	15.3%	Institution	0.60	56/87	64.4%
Institution	[0.27]	2/37	5.4%	Other	0.46	192/436	44.0%
	Range = [73]				Range = 41		
Syntax	(NOT SIG)			Modifier	(NOT SIG)		
Pred. nom/Exist	[0.67]	5/9	35.7%	Prenominal	[0.78]	11/16	68.8%
Object	[0.53]	16/105	15.2%	None	[0.49]	253/527	48.2%
Object of Prep	[0.43]	8/73	11.0%	Post-nominal	[0.42]	1/3	33.3%
Subject	[0.00]	0/10	0.0%		Range = [36]		
	Range = [67]						
				Syntax	(NOT SIG)		
String Position	(NOT SIG)			Pred. nom/Exist	[0.56]	37/72	51.4%
Second or more	[0.70]	3/11	27.3%	Object	[0.55]	130/258	50.4%
First or only	[0.49]	26/191	13.6%	Object of Prep	[0.42]	93/198	47.0%
	Range = [21]			Subject	[0.40]	6/18	33.3%
					Range = [16]		
Modifier	(NOT SIG)						
None	[NA]	29/195	14.9%	String Position	(NOT SIG)		
Prenominal	[NA]	0/5	0.0%	Second or more	[0.61]	28/46	60.9%
Post-nominal	[NA]	0/2	0.0%	First or only	[0.49]	238/500	47.6%
	Range = [NA]				Range = [12]		
Polarity	(NOT SIG)			Polarity	(NOT SIG)		
Affirmative	[NA]	29/201	14.4%	Negative	[0.58]	2/2	50.0%
Negative	[NA]	0/1	0.0%	Affirmative	[0.50]	264/278	48.7%
	Range = [NA]				Range = [8]		
Log Like = -63.614		Sign = p < 0.001		Log Like = -276.709		Sign = p < 0.001	

The highest ranking, significant, factor group for both the LOLIs with Spanish phonology and those with English phonology is specificity. Previous studies have found this factor group to consistently be the highest (or second ranked) factor group in both English and Spanish (Chapter 2; Torres Cacoullos and Aaron, 2003a). The second similarity is that semantic class, while not significant for LOLIs with Spanish phonology, is the second highest ranking factor group. The differences between the two groups, however, are more telling in terms of determining a correlation between phonological and morphosyntactic integration.

The first major discrepancy between the two groups is found with the way specificity affects determiner realization. While among LOLIs with Spanish phonology, non-specific nouns highly favor bare nouns (FW = 0.86), generic nouns are the highest ranking category that favor zero determiners for LOLIs with English phonology (FW = 0.80). These patterns are consistent with patrimonial nouns from the respective languages where Spanish nouns favor bare nouns for non-specific nouns and English favors bare nouns in generic nouns (Chapter 2; Torres Cacoullos and Aaron, 2003a).

Secondly, while semantic class is not significant for LOLIs with Spanish phonology, it is the second highest ranking factor group for the LOLIs with English phonology, which was the same result as found in English for New Mexican bilinguals (Torres Cacoullos and Aaron 2003a). Additionally, LOLIs with Spanish phonology follow the Spanish patrimonial noun pattern (see Chapter 2; Torres Cacoullos and Aaron, 2003a) in that occupational nouns highly favor a zero determiner (FW = 1.00) and institutional nouns highly favor the realization of a determiner (FW = 0.27). On the other hand, LOLIs with English phonology show signs of similarity between both English and Spanish patrimonial noun norms. While these LOLIs favor bare determiners for institutional nouns (FW = 0.60), following the English pattern (see Chapter

2; Torres Cacoullos and Aaron, 2003a), they also favor bare determiners for occupational nouns (FW = 0.87), following the Spanish pattern (see Chapter 2; Torres Cacoullos and Aaron, 2003a).

Lastly, as identified in the introduction, the presence (or absence) of a modifier can also be used to measure differences between Spanish and English determiner realization. LOLIs with English phonology do pattern after English nouns in that the category was not significant (see Chapter 2; Torres Cacoullos and Aaron, 2003a); however, the pattern of effect does, in part, follow the Spanish pattern in that pre-nominal modifiers most favor bare nouns (see Chapter 2; Torres Cacoullos and Aaron, 2003a). Among LOLIS with Spanish phonology the pattern of effect does not seem consistent with Spanish nouns, the presence of a modifier does not favor zero determiner as it does for the previously studied Spanish nouns (Chapter 2; Torres Cacoullos and Aaron, 2003a).

Overall, LOLIs with Spanish phonology do seem to act like Spanish patrimonial nouns in terms of the factors that condition determiner realization. This is true in the behavior of 2 of the 3 important factor groups that present clear conflict sites between Spanish and English, specificity and semantic class, both of which are the highest ranking factor groups. On the other hand, LOLIs with English phonology appear to be more similar to English patrimonial nouns in terms of specificity (generic nouns favoring zero determiners) and the institutional noun category of semantic class. There are also similarities between LOLIs with English phonology and Spanish patrimonial nouns, specifically with the occupational noun category in the semantic class factor group and with the presence of a modifier factor group. The fact that LOLIs with English phonology show partial integration is not that unsurprising since nonce LOLIs are often found to pattern similarly to patrimonial words of the recipient language (Aaron, 2014; Blas Arroyo and Tricker, 2000; Ghafar and Meechan, 1998; Poplack and Dion, 2012; Poplack and Meechan,

1995; Sankoff, Poplack and Vannianajan, 1990; Torres Cacoullos and Aaron, 2003a, 2003b;

Torres Cacoullos and Virgil, 2002, 2003).

3.2.3 Summary of the relationship between phonological and morphosyntactic integration

 The results from section 3.2 help to answer sub research question 2:

  SRQ2: Is there a difference between the constraint hierarchies of determiner realization

  between LOLIs with Spanish phonology and LOLIs with English phonology? If so, how

  do the differences compare to previously documented English and Spanish

  morphosyntactic patterns (as represented in Torres Cacoullos and Aaron 2003a and

  Chapter 2)?

Overall, yes, there is a clear difference in the constraint hierarchies of determiner realization

between LOLIs of Spanish phonology and LOLIs with English phonology. Furthermore, the

differences seem to coincide with the expected patterns of each language pairing (LOLIs with

Spanish phonology to Spanish patrimonial nouns and LOLIs with English phonology to English

patrimonial nouns). LOLIs with Spanish phonology do seem to act like Spanish patrimonial

nouns in terms of the factors that condition determiner realization. This is true in the behavior of

2 of the 3 important factor groups that present clear conflict sites between Spanish and English,

specificity and semantic class, both of which are the highest-ranking factor groups. On the other

hand, LOLIs with English phonology appear to be more similar to English patrimonial nouns in

terms of specificity (generic nouns favoring zero determiners) and the institutional noun category

of semantic class. There are also similarities between LOLIs with English phonology and

Spanish patrimonial nouns, specifically with the occupational noun category in the semantic

class factor group and with the presence of a modifier factor group. The fact that LOLIs with

English phonology show partial integration is not that unsurprising since nonce LOLIs are often

found to pattern similarly to patrimonial words of the recipient language (Aaron, 2014; Blas

Arroyo and Tricker, 2000; Ghafar and Meechan, 1998; Poplack and Dion, 2012; Poplack and

Meechan, 1995; Sankoff, Poplack and Vannianajan, 1990; Torres Cacoullos and Aaron, 2003a,

2003b; Torres Cacoullos and Virgil, 2002, 2003). These results also help to answer our main

research question, which was:

> MRQ1: Are phonological and morphosyntactic integration correlated, and if so, what
>
> does the correlation mean to current methods that disambiguate one-item CS from
>
> borrowings?

The answer is yes, LOLIs are more morphosyntactically integrated into Spanish when they have Spanish

phonology than LOLIs that have English phonology. What is interesting is that LOLIs with English

phonology are not as well integrated morphosyntactically into Spanish as nonce LOLIs are (see Chapter

2). While nonce LOLIs share the main characteristics of Spanish nouns (occupational nouns favor zero

determiners, while institutional nouns disfavor zero determiners, and pre-nominal modifiers highly favor

bare nouns), they only differ from Spanish nouns in that generic NPs favor zero determiners (see

Chapter 2). On the other hand, LOLIs with English phonology show more parallels to English than

Spanish. This is something that should be explored further in future studies. With a larger sample, one

could create 4 groups to be tested, well-established LOLIs with Spanish phonology, nonce LOLIs with

Spanish phonology, well-established LOLIs with English phonology, and nonce LOLIs with English

phonology. Based on the results of this study and Chapter 2, it would be expected that well-established LOLIs with Spanish phonology show the most morphosyntactic integration, while nonce LOLIs with English phonology show the least, and nonce LOLIs with Spanish phonology should find themselves somewhere along the continuum between the other two. If this outcome is realized, it would give further merit to the integration continua presented by Haugen (1950) and Poplack (1993), who propose that on one end of the spectrum there is codeswitching (in this case, nonce LOLIs with English phonology) and on the other end of the spectrum there are borrowings that are completely integrated (which would be the well-established LOLIs with Spanish phonology group). At any rate, the results from this section indicate that phonological integration does play a role in the incorporation of LOLIs into Spanish on the part of Arizona bilinguals.

4. Conclusion

The purpose of this study was to determine if there is a correlation between phonological integration and morphosyntactic integration of LOLIs. We found that LOLIs can overwhelmingly be categorized by phonological integration as either presenting Spanish or English phonology (this was true for 99% of the data). An analysis of phonological integration by lemma showed that while LOLIs become more established in the local grammar, they are more often phonologically integrated, there was still variation among even the well-established groups. Further examination of morphosyntactic, in the form of the conditioning of determiner realization, revealed that LOLIs produced with Spanish phonology are more morphosyntactically

similar to Spanish patrimonial nouns and LOLIs produced with English phonology are more morphosyntactically similar to English patrimonial nouns. These findings lead us to conclude that phonological integration is indeed correlated with morphosyntactic integration, and thus should be included in the analysis when discern LOLIs as borrowings or codeswitches.

The methodology of this study is novel in several ways. First, it is the first study on discerning LOLIs as borrowings from codeswitches since the seminal work of Poplack and Sankoff (1984) and Poplack, Sankoff, and Miller (1988) to incorporate phonology into the analysis. Furthermore, it is the first study to report the phonological integration in its full detail. The way in which the correlation between phonology and morphosyntax is conducted is also a novel approach. Lastly, the perception experiment that was used to help determine whether a LOLI was phonologically integrated (produced with Spanish phonology) or unintegrated (produced with English phonology), in addition to the PRAAT analysis, proved a high rate of reliability in relation to the analyst's impressionistic method of classification.

However, the fact that there were so few tokens of LOLIs with Spanish phonology (N = 202, only 27% of the data) as opposed to LOLIs with English phonology (N = 546, 72% of the data), poses a possible issue Type I error for the correlation between phonology and morphosyntax, that while the result was found in the limited data of the study, it is not true of the local grammar as a whole.

The results of the present study vary drastically from the findings in Poplack and Sankoff (1984) and Poplack, Sankoff, and Milller (1988). One possible explanation for this is the type of bilingual that is being analyzed. In both previous studies, there was a mix of balanced bilinguals and bilinguals that were dominant in their L1. After viewing a presentation by Poplack, Robillard, and Dion, and at the 2016 New Ways of Analyzing Variation (NWAV) conference, a

subsequent conversation between Natalie Dion and myself revealed that the French-English bilinguals in the Poplack, Robillard, and Dion presentation showed evidence of French phonological features on their English, even during codeswitched segments that lasted as long as three minutes. If the bilinguals have French features on their English, this is most likely due to the fact that they have "accented" English, that they have not fully acquired the English phonological grammar. If these participants have not fully acquired the English phonological grammar, we would expect to see phonological variation within a word, but this should be interpreted as interference from their L1, and not as a choice that the bilinguals make when incorporating other language nouns into their French. On the other hand, the bilinguals in the present study are all balanced bilinguals and do maintain separate phonological grammars in both English and Spanish. Future studies will need to explore the difference in phonological integration of LOLIs in both balanced bilinguals and L1 dominant bilinguals in order to further test the results and interpretations of the present study.

Whether or not the correlation between phonological and morphosyntactic integration holds up in future research with larger data sets, the data from this study clearly shows that phonology is not being integrated in the expected manner. We would expect that the longer a lemma from an English noun is in the local Spanish grammar, the more integration would take place in accordance with Haugen (1950). The lack of total phonological integration for well-established borrowings like 'high school' can be explained in two ways. First, it may simply be that phonology does not follow the same path as morphology. This has certainly been the way previous research deals with variation in phonological integration and the reason why phonology has been removed from analysis since the early 1980s. However, a second option exists, one that is in line with the results of the current study. When assessing the production of well-established

borrowing lemmas, it may be that when a bilingual produces the LOLI with phonology of the recipient grammar, the word was accessed from the lexicon of the well-established lemma from the grammar of the recipient language. On the other hand, when a bilingual produces the same lemma with phonology from the donor language, it is possible that the bilingual is re-borrowing the word from the donor language grammar and it is, then, a nonce borrowing (or a codeswitch). Since bilinguals have access to both English and Spanish grammars, this borrowing behavior is possible. The results from the comparison of determiner realization between Spanish LOLIs and English LOLIs also support this hypothesis. If the nouns were always part of the integrated lemma, we would not expect to see a difference in the morphosyntactic integration (or at least not as marked of a difference) between LOLIs produced Spanish versus English phonology. Even the comparison of the factors that condition determiner realization between established and nonce borrowings (see Chapter 2; Torres Cacoullos and Aaron, 2003a) were not as distinguished as the groups when LOLIs were separated by phonological integration. This may be due to the fact that using lemma to determine if a LOLI is established or nonce is not the only criteria that needs to be considered. Future studies should repeat the methodology provided here to determine if the pattern holds to more communities and larger data sets, which would give more evidence in favor of the idea that phonology helps to determine established versus nonce borrowings.

To further explore the hypothesis that when a LOLI is phonologically integrated, it comes from a well-established lemma within the lexicon of the recipient language, and when it is not integrated, it is a codeswitch (or at least a nonce borrowing) from the donor language (even if a well-established borrowing already exists within the lexicon of the recipient language), the next chapter will examine the role of pauses and false starts between LOLIs produced with Spanish phonology (integrated) and English phonology (not integrated). Since some studies have

suggested that pauses and false starts are signs of cognitive strain and codeswitching involves higher online processing (Moreno, Federmeier and Kutas, 2002), a higher amount of pausing or false starts in one group (expectedly the English phonology group) may be a sign that the LOLIs in the group are codeswitches, rather than borrowings.

Chapter 4:

Flagging of LOLI nouns

1. Introduction

As we thoroughly explored in Chapters 2 and 3, lone other language items (LOLIs[21]) have been the focus of multiple variationist studies (Aaron, 2014; Blas Arroyo and Tricker, 2000; Chapter 2; Chapter 3; Ghafar and Meechan, 1998; Myers-Scotton, 1993, 2006, 2009; Poplack, 2012; Poplack and Meechan, 1995; Poplack, Sankoff, and Miller, 1988; Sankoff, Poplack and Vanniarijan, 1990; Torres Cacoullos and Aaron, 2003a, 2003b; Torres Cacoullos and Vigil, 2002, 2003; among countless others). In each case, researchers work to discern LOLIs' status as either borrowings (morphosyntactically incorporated material) or codeswitches (morphosyntactically unincorporated material). The LOLIs are separated into established (lemmas that have been in the grammar for a period of time) and nonce (lemmas that have just been introduced by a speaker) groups, to determine if nonce LOLIs show integration or lack thereof. While the extent to which nonce LOLIs are morphosyntactically incorporated has varied (namely in Aaron, 2014; Chapter 2), it is unanimously agreed upon that LOLIs, in general, are borrowings (Aaron, 2014; Blas Arroyo and Tricker, 2000; Chapter 2; Chapter 3; Ghafar and Meechan, 1998; Myers-Scotton, 1993, 2006, 2009; Poplack, 2012; Poplack and Meechan, 1995; Poplack, Sankoff, and Miller, 1988; Sankoff, Poplack and Vanniarijan, 1990; Torres Cacoullos

---

[21] Acronym from Poplack (2012)

and Aaron, 2003a, 2003b; Torres Cacoullos and Vigil, 2002, 2003; among countless others).

Phonological integration has been largely removed from the discussion since the work of

Poplack and Sankoff (1984) and Poplack, Sankoff, and Miller (1988) who found that phonology

is too variable to be considered when determining if LOLIs are borrowings or codeswitches.

However, results from Chapter 3 made a strong case that the variation in the phonological

integration of LOLIs can not only be explained, but also may be seen as a key component in

furthering our understanding of LOLIs as either borrowings or codeswitches.

In Chapter 3, we found a correlation between morphosyntactic and phonological

integration of LOLIs. Specifically, it was found that LOLIs that are produced with Spanish

phonology are more morphosyntactically similar to Spanish patrimonial nouns and LOLIs that

are produced with English phonology are more morphosyntactically similar to English

patrimonial nouns. Table 1 below summarizes the morphosyntax integration of LOLIs based on

their phonological integration found in Chapter 3.

Table 1:
Morphosyntactic integration of LOLIs with Spanish vs English phonology
in Sonoran bilingual Spanish

Factor	Spanish (N = 400)	Spanish Phon. LOLIs (N = 202)	English Phon. LOLIs (N = 546)	English (N = 400)
Specificity	1. Non-specific 2. Generic 3. Specific	1. Non-specific 2. Generic 3. Specific	1. Generic 2. Non-specific 3. Specific	1. Generic 2. Non-specific 3. Specific
Semantic Class	Occupation + Institution -	Occupation + Institution -	Occupation + Institution +	Institution + Occupation -
Modifier	Pre-nominal + Post-nominal (+/-)	Pre-nominal - (?) Post-nominal - (?)	Pre-nominal + Post-nominal -	No effect (Pre-nominal -)

As shown in Table 1, LOLIs that are produced with Spanish phonology match Spanish patrimonial words for specificity, semantic class, and post-nominal modifiers. On the other hand, LOLIs produced with English phonology pattern like English patrimonial words for specificity and institutional nouns. It is clear that in this sample, LOLIs with Spanish phonology are well-integrated morphosyntactically into Spanish discourse, while LOLIs with English phonology are unintegrated. Based on these results, we concluded with the tentative hypothesis that LOLIs that are produced with English phonology behave morphosyntactically more like codeswitches, while LOLIs with Spanish phonology seem to tend to be incorporated also morphosyntactically, and therefore, behave more like borrowings. In this chapter we will provide further evidence of this hypothesis by examining the pauses and false starts that are present before LOLIs with Spanish versus English phonology.

1.1 Theoretical background and previous research

Beyond the simple measure of morphosyntactic integration, another feature of LOLIs that may be useful to further the discussion of LOLI status as a borrowing or a codeswitch deals with what has been called "flagging" (Dumont, 2010; Torres Cacoullos and Aaron, 2003a), which includes features like pausing and false starts.

The observation that pausing is found around other language material has been noticed within research on borrowed discourse markers who often attribute the feature as a way for bilinguals to highlight the discourse marker (Dajko and Carmichael, 2014; de Rooij, 2000; Maschler, 1994, 1997, 2009). In their work on disambiguating LOLIs as borrowings or codeswitches, Torres Cacoullos and Aaron (2003a) found several flagging techniques before

both patrimonial words and LOLIs: hesitations, pauses, fillers, parentheticals, and metalinguistic

comments (p. 317). The authors conclude that these flaggings do not influence determiner

realization because they account for a small percentage of the data (flaggings are present only

before 9% of nonce LOLIs tokens and before 2% of patrimonial noun tokens) and that the

distribution of bare nouns (in terms of percentages) mostly remains the same between flagged

and unflagged tokens (p. 317). In a subsequent study based on 10 highly proficient bilingual

speakers, Dumont (2010) also found that false starts do not happen more often before

codeswitches, as one might predict; however, it is important to note that pauses were not

included in this analysis (p. 347).

Unlike sociolinguistics findings (Dumont, 2010; Torres Cacoullos and Aaron, 2003a),

research in psycholinguistics has found contradictory evidence to whether or not there is a cost

associated with codeswitching where studies using spontaneous speech find no cost (Gardner-

Chloros, McEntee-Atalianis, and Parasekva, 2013; Hennecke, 2013) and studies based on

laboratory speech do find a cost to codeswitching (Meuter and Allport, 1999; Moreno,

Federmeier and Kutas, 2002; Navracsics, 2004; Thomas and Allport, 2000). For instance,

Moreno, Federmeier and Kutas (2002) measured electrophysiological responses to expected,

unexpected, and codeswtiched responses to idiomatic expressions. Participants were presented

with idiomatic expressions that sometimes finished in the expected way ('out of sight out of

mind'), an unexpected way ('out of sight out of ***brain***'), or with a codeswitch ('out of sight out

of ***mente***'). They found that the participants' electrophysiological responses to idiomatic

expressions that included a codeswitch patterned as unexpected or improbable events, but the

responses were different from when the idiomatic expression ended in the expected way. While

codeswitches were found to be costlier than expected responses, unexpected responses were costlier than codeswitches.

Flagged material does deserve more attention because it may be a product of codeswitching. Given that some psycholinguistic research suggests that pausing can indicate a higher level of cognitive processing (Schilperoord, 2002) and that some research has shown a higher cost in codeswitching (Meuter and Allport, 1999; Moreno, Federmeier and Kutas, 2002; Navracsics, 2004; Thomas and Allport, 2000), it stands to reason that flaggings (specifically, pausing and false starts) could be signs of difficult processing, and as such also signs of codeswitching. This hypothesis will be explored in more detail in this chapter.

1.2 Research questions

Having discussed some of the relevant literature that centers on discerning LOLI borrowings from codeswitches, we will now turn to the research questions that will be addressed in the dissertation and the methodology that will be used to answer them. This final part of the dissertation explores more thoroughly the role that pauses and false starts play on LOLIs in order to provide more evidence that phonological integration is a good predictor of LOLIs' status as borrowings or codeswitches. Specifically, we seek to answer Main Research Questions 1 and 2 through the analyses outlined below.

MRQ 1: Does the way in which flagging is distributed by phonological integration provide further evidence that phonology can be used in discerning LOLI borrowing from codeswitching?

MRQ2: When comparing flagging by establishedness and by phonological integration, which better explains the variation in the data?

However, before we are able to answer MRQ1, we must first answer several sub-research questions:

SRQ1: What is the distribution of pausing by phonological integration of LOLIs?

SRQ2: What is the distribution of false starts by phonological integration of LOLIs?

SRQ3: What is the distribution of flagging by phonological integration and by establishedness?

To accomplish this task, we will examine the number of LOLIs with pauses and false starts and their level of phonological integration and establishedness, extracted from a pool of LOLI nouns found in spoken bilingual Sonoran Spanish in Southern Arizona. The methodology section that follows will explore in more detail the participants, data, and analyses that will be used to answer the research questions provided above.

2. Methodology

2.1 The participants

      The participants of this study are 24 young adults (no older than 30 years of age). 12 participants are men and 12 are women. The participants reside in Southern Arizona, Tucson or Nogales, two cities with close proximity to the border between Sonora, Mexico and Arizona, US. The participants belong to three different sociolinguistic corpora. A summary of the participants can be found in Table 2.

Table 2:

List of Arizona participants by gender and by corpus

	Group 1 (Bessett 2012)	Group 2 (Carvalho 2012-)	Group 3 (Carvalho 2012-)
Males	4	4	4
Females	4	4	4
Total	**8**	**8**	**8**
Total participants in the study: **24**			

      The participants are all Spanish-English bilinguals and are part of three separate groups. Group 1 (N=8) was interviewed in Spanish by the investigator (Bessett, 2012) and maintained monolingual mode for virtually the entirety of the corpus (only 3 total cases of codeswitching are present in the corpus of approximately 24 hours). Group 2 (N=8) comes from the *Corpus del*

español en el sur de Arizona (CESA) and the interviews were conducted by various interviewers, some with ties to the community and others that are not members of the community (Carvalho 2012-). The final group, Group 3 is a subset of the CESA corpus, which partners two bilingual speakers and records their conversation.

The grand majority of the bilingual participants were born in the United States or arrived before they were ten, with the exception of one participant who arrived at age fifteen. They are all part of, or descendants of, families who emigrated from Sonora, Mexico, thus allowing for a cohesive sample from the same dialectal area. In order to ensure that the participants were bilingual, the participants were asked to evaluate their proficiency in both English and Spanish on a 0-10 scale where zero meant very low proficiency and ten meant very high proficiency (in the case of the CESA corpus, Groups 4-5, the scale was 0-6). This assessment was adapted from the Bilingual Language Profile created by Birdsong, Gertken, and Amengual (2012). All of the participants reported a proficiency of at least six out of ten (or 4.5 of 6 in the case of participants from CESA) for both English and Spanish. The participants demonstrated their Spanish proficiency in that they were able to carry out a conversation in Spanish for the duration of the hour-long interview. Additionally, all participants attended school in the United States and live and work in a bilingual setting.

2.2 The data

The participants described in section 2.1 took part in sociolinguistic interviews or casual conversations, conducted in Spanish, which last for approximately an hour. From the almost 24 hours of participant speech, 756 LOLIs were identified. In order to answer the research questions

presented in section 1.2, the LOLIs were further classified by phonological integration and establishedness.

2.2.1 Phonological integration of LOLIs

Phonological integration was coded for Spanish (integrated), English (non-integrated) and Spanish/English (mixed) and was established based on the impressionistic observations of the investigator, a proportion of which were verified using a perception study involving 19 English monolingual participants, and spectral analyses of /r/ and /l/ in PRAAT. The perception experiment was based on 24% of the total number of LOLIs present among the 24 Spanish-English bilingual participants described in section 2.1. The results showed that when the English monolingual participants heard a LOLI and thought it sounded more English-like, the investigator also reported that the LOLI was more English-like. Additionally, the spectrogram analyses revealed that for both /r/ (found in 38% of LOLIs) and /l/ (found in 26% of LOLIs), the investigator's impressionistic judgments were confirmed in the vast majority of the cases (97% for /r/ and 96% for /l/). For a more detailed explanation of the perception experiment and spectrogram analyses, please see Chapter 3. Due to the fact that the results of the perception experiment and spectrogram analyses confirmed the impressionistic observations of the investigator in such a robust manner, the investigator's impressionistic observations were used for the final determination of LOLI integration. Table 3 shows the frequency of the three phonology types.

Table 3:

Frequency of phonology type (Spanish, English, mixed)
of LOLI produced in Spanish discourse by Arizona bilinguals

Phonology Type	Number	%
Spanish	202/756	27%
English	546/756	72%
Spanish/English mixed	8/756	1%

Since the comparison that needs to be made, in order to answer our research questions, is the

difference between LOLIs with Spanish phonology and LOLIs with English phonology, the 8

LOLIs with mixed phonology were removed from the rest of the analyses, leaving 748 LOLIs to

be further analyzed.

2.2.2 Establishedness of LOLIs

LOLIs were also separated into categories based on whether they were established, or

nonce. Established LOLIs were lemmas that were found in official dictionaries (Academia

Mexicana de la Lengua, 2015, 2010; Corominas, 1967; García Icazbalceta, 1899; Lara Ramos,

1996; Lugones, 1932; Real Academia de la Lengua Sonorense, 2006; Real Academia Española,

2016), local dictionaries and reported lists of borrowings in Southwest Spanish (Cerda, Cabaza,

and Farias, 1970, 1953; Cobos, 2003; Galván and Teschner, 1985, 1975; García, 1935; Gross,

1935; Hamel, 1994; Reyna, 1980), or monolingual corpus of 24 Sonoran Spanish speakers in

Hermosillo (Bessett, 2012). Nonce LOLIs were any lemmas that did not appear in these three

contexts. Table 4 shows the distribution of the LOLIs of this study by establishedness type.

Table 4:

Frequency of established and nonce LOLIs
produced in Spanish discourse by Arizona bilinguals

Phonology Type	Number	%
Established	384/748	51%
Nonce	364/748	49%

2.2.3 Flagging

The last factor to be coded was flagging. All 748 LOLIs were coded for if there was a

false start immediately before the LOLI, or not. Additionally, the length in seconds (recorded to

the 7th decimal place) between each LOLI and the previous word was calculated in PRAAT by

measuring the last zero crossing of the previous word and the zero crossing right before the first

band of sound from the LOLI. When there was no space between the previous word and the

LOLI, the token was labeled as 0 seconds.

2.3 Data Analysis

In order to answer Sub-Research Questions 1-3, the overall frequencies of pausing and

false starts were tallied for LOLIs produced with Spanish phonology and LOLIs produced with

English phonology, as well as flagging (pausing and false starts) for phonological integration and

establishedness. Once the descriptive comparisons were made, a generalized linear mixed

effects model with a logit link function was calculated in R in order to determine which predictor

(phonological integration or establishedness) better accounted for the variation in the data.

3. Results

The results are separated into sub-sections in order to better answer the research

questions outlined section 1.2 of the introduction. In section 3.1, we will discuss the distribution

of pausing by phonological integration, thus answering SRQ1. We will then turn to the

distribution of false starts by phonological integration in section 3.2 in order to answer SRQ2.

Then, in section 3.3, we will present the distribution of flagging by phonological integration and

establishedness, allowing for us to answer SRQ3. Finally, in section 3.4, we will present the

results of the linear model that will answer MRQ 1 and 2.

3.1 Pausing by phonological integration

Of the 546 total LOLIs with English phonology, 150 (or 27.47%) contain pauses (when

considering a pause to be anything longer than $0.00s^{22}$). Meanwhile, of the 202 LOLIs with

Spanish phonology, only 10 (4.95%) have pauses. Chart 1 shows the distribution of these results.

---

[22] A pause was considered anything more than 0 seconds because there were so few tokens (160
total) that contained any pause between the LOLI and the previous word.

Chart 1:

Percentage of Pauses (longer than 0.00s) by phonology type in
LOLIs spoken during Spanish discourse from Spanish/English bilinguals in Arizona
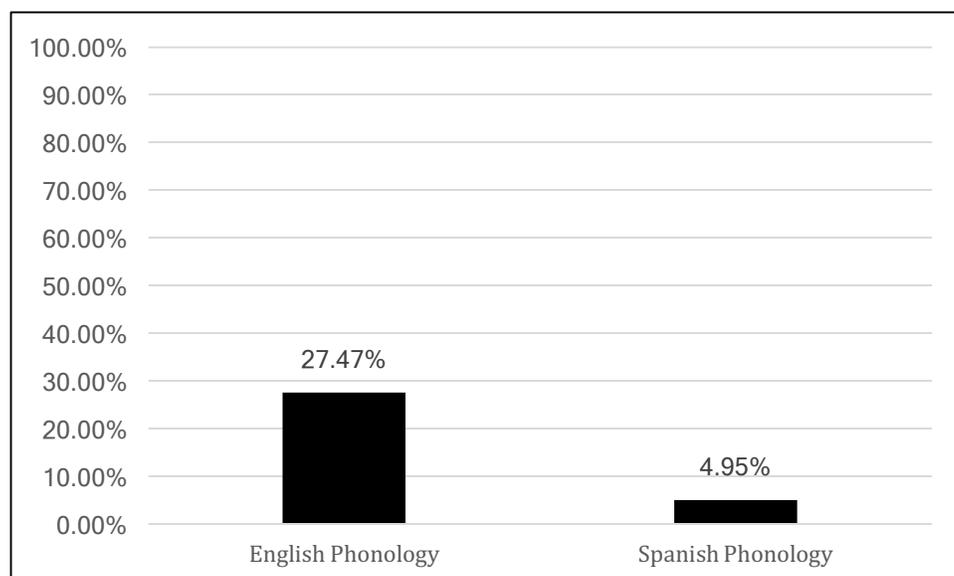

Not only do we find more pauses among LOLIs with English phonology, but the average pause

length[23] (including values in 0.00s) is larger for LOLIs with English phonology (0.14s) than for

LOLIs with Spanish phonology (0.02s). Chart 2 shows a box plot of the pausing values for

LOLIs separated by phonological integration.

---

[23] The length in seconds (recorded to the 7th decimal place) between each LOLI and the previous
word was calculated in PRAAT by measuring the last zero crossing of the previous word and the
zero crossing right before the first band of sound from the LOLI. When there was no space
between the previous word and the LOLI, 0 seconds was recorded.

Chart 2:

Boxplot of pause length by phonology type in
LOLIs spoken during Spanish discourse from Spanish/English bilinguals in Arizona
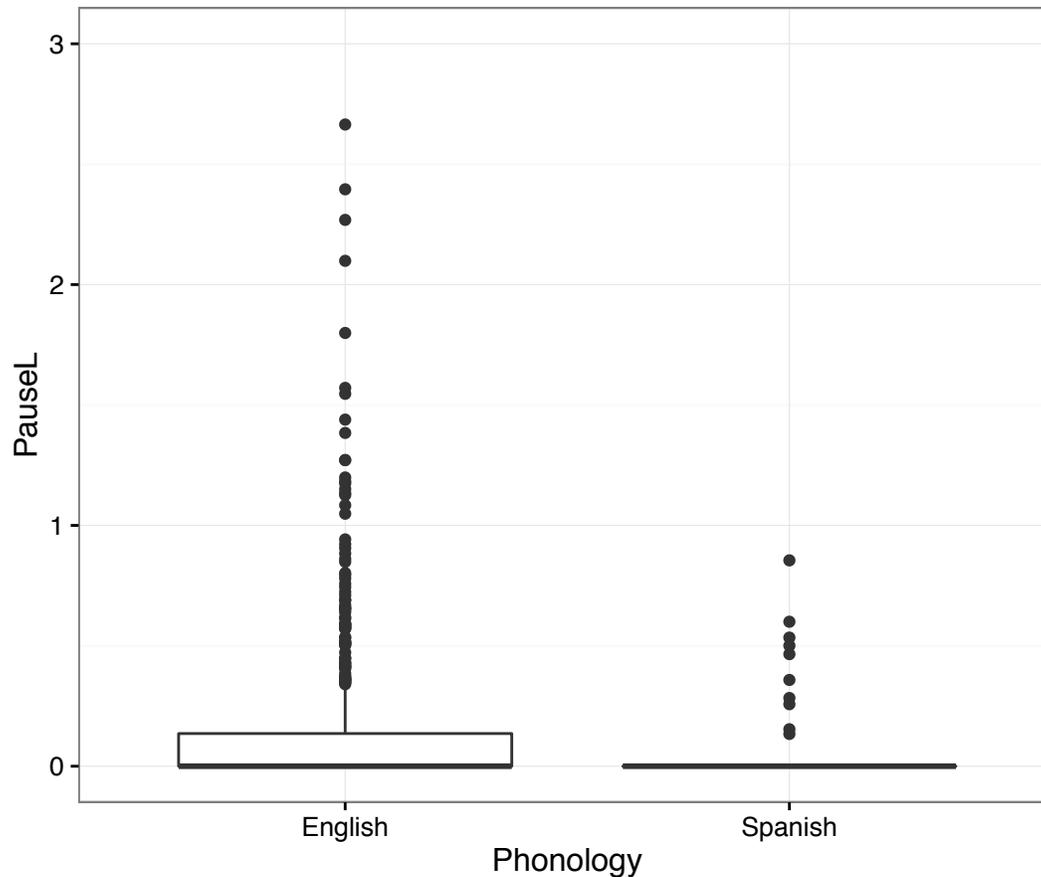
The results from this section answer SRQ1, repeated here below:

SRQ1: What is the distribution of pausing by phonological integration of LOLIs?

In response, there is more pausing among LOLIs with English phonology, than LOLIs with
Spanish phonology. Also, the length of the pauses tends to be longer among LOLIs with English
phonology than LOLIs with Spanish phonology.

3.2 Distribution of False Starts

Examples (1)-(7) below demonstrate the variety of false start types that are found among LOLIs with English phonology in the data set. In (1), (2), (5), and (7) the material directly before the LOLI is repeated. Another false start strategy can be found in (4), where the participant starts the phrase with "fui a:" '(I) went to' but then changes the direction of the sentence by restarting it with "de high school" 'for high school'. A third false start type can be seen in (1), (2), (3), and (6), which all involve the insertion of a repair word such as "uh" and "um". Interestingly, the only repair words found among the LOLIs with English phonology are also English, not traditionally associated with Spanish.

(1) "Es una clase enorme, quinientos XXX yo creo, **es en-** es en social sciences …" –T01

'(It) is a huge class, five-hundred XXX I think, **(it) is in-** (it) is in social sciences…"

(2) "Ella está estudiando historia **y- y ya es uh-** ya es senior." –T02

'She is studying history **and- and (she) already is uh-** (she) already is (a) senior."

(3) "Y fui **a:l um** college." –T11

'And (I) went **to the um** college."

(4) "…que **fui a:-** de high school pues fui a XY." –T11

'…that **(I) went to:-** for high school well, (I) went to XY."

(5) "El que pegué era… un **ex-** ex-policeman." –T12

'The (one) that (I) hit was… an **ex-** ex-policeman."

(6) "Um… nos ponemos a- a hacer… **um** popcorn, cuando vamos a ver peli." –T18

'Um…we set to- to make… **um** popcorn, when (we) are going to watch a movie."

(7) "Usan mucho **el-** el slang aquí." –T18

'(They) use **the-** the slang a lot here."

Among LOLIs with English phonology, 91 of the 546 (16.67%) of the tokens had false starts. On the other hand, LOLIs with Spanish phonology had false starts in only 16 of the 202 tokens (7.92%). Chart 3 shows the distribution of the results.

Chart 3:

Percentage of False Starts by phonology type in LOLIs spoken during
Spanish discourse from Spanish/English bilinguals in Arizona
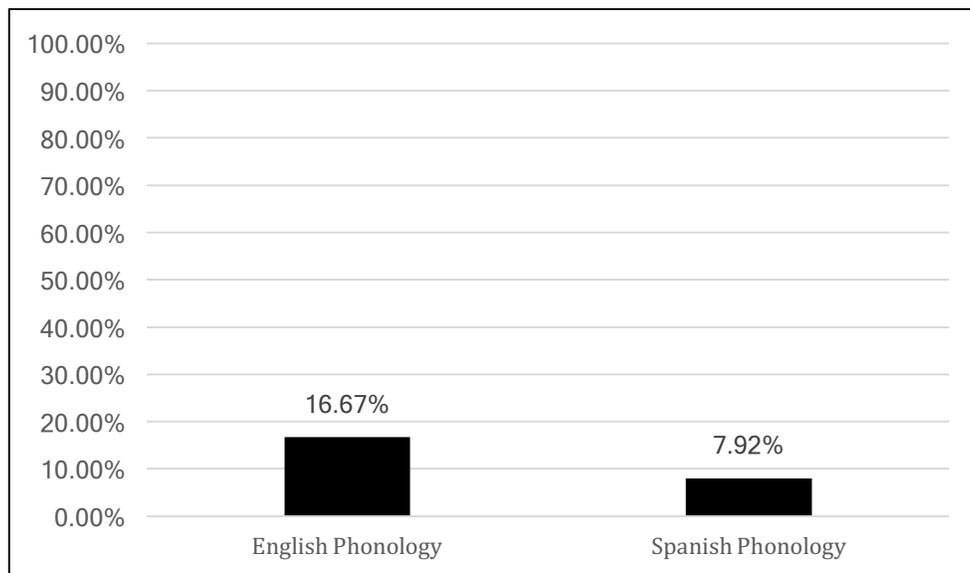
The results in Chart 3 help to answer SRQ2:

SRQ2: What is the distribution of false starts by phonological integration of LOLIs?

We can again, see a clear difference in the distribution of flagging material by phonological integration, where unintegrated material is more often flagged than integrated material.

3.3 Flagging (pausing and false starts together)

Since there are so few data points in general, the pausing and false starts were combined into a larger category, flagging, which was then used for statistical analysis. For LOLIs with English phonology, there are a total of 212/546 tokens that contain flagging material (38.82%). LOLIs with Spanish phonology are far less often flagged, 24/202 (11.88%). Chart 4 shows the distribution of these results.

Chart 4:
Percentage of Flagging by phonology type in LOLIs spoken during
Spanish discourse from Spanish/English bilinguals in Arizona
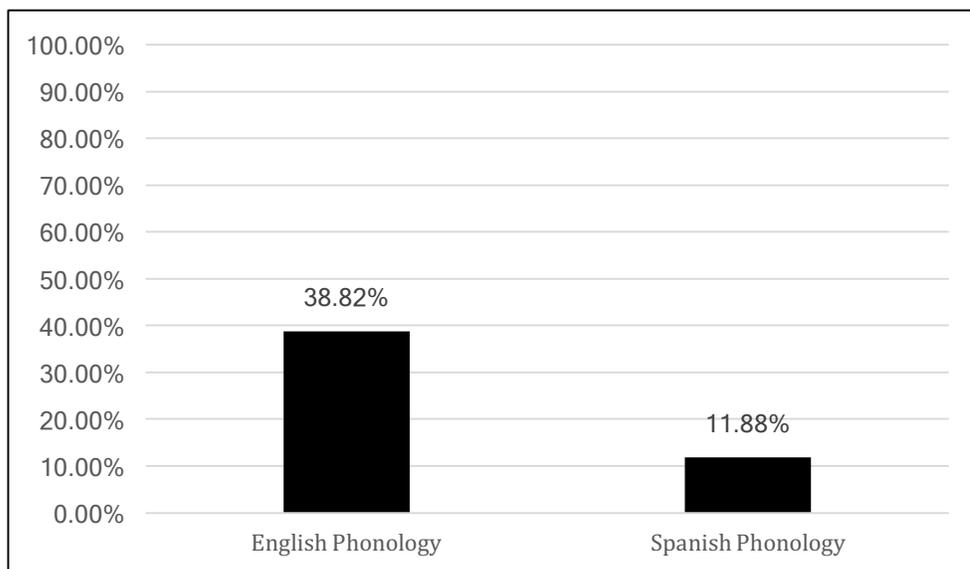
In addition to reporting solely whether or not phonology correlates to flagging, we also tested for the correlation between establishedness and flagging. This was done to compare how

well establishedness accounts for the variation in the data as compared to phonology. Since established borrowings have been in the recipient language grammar (the Spanish grammar of the bilinguals in this study), they are highly incorporated both morphosyntactically (see Chapter 2) and phonologically (see Chapter 3) into the recipient language (Spanish). For this reason, if flaggings are a sign of higher cognitive processing, we would expect established borrowings to have less flaggings than nonce borrowings (which have considerably less time in the recipient language grammar). These results are born out, as attested in Chart 5, which shows the distribution of flagging by establishedness.

Chart 5:
Percentage of Flagging by establishedness type in LOLIs spoken during
Spanish discourse from Spanish/English bilinguals in Arizona
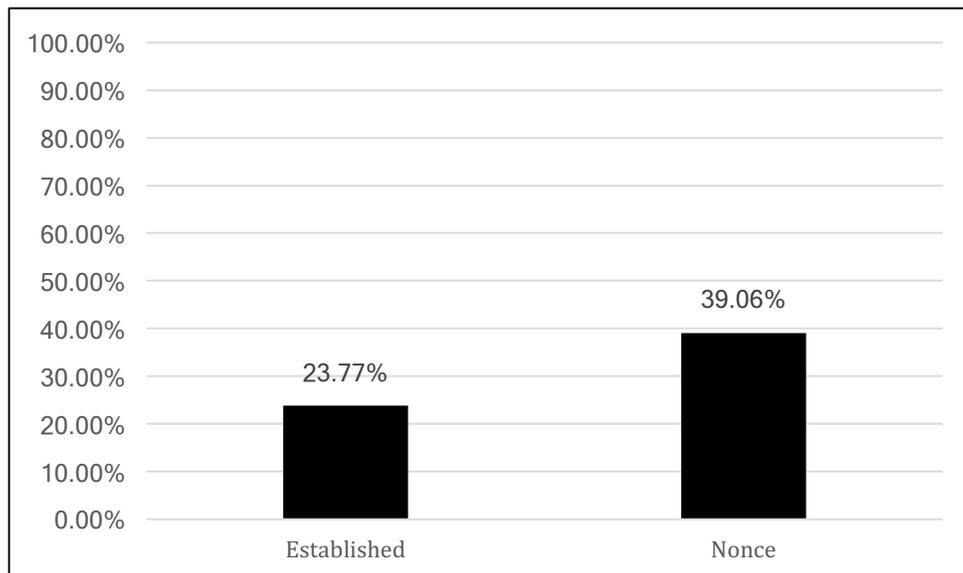
Charts 4 and 5 serve to answer SRQ3:

SRQ3: What is the distribution of flagging by phonological integration and by establishedness?

The results from this sub-section indicate that both phonological integration and establishedness show differences in the presence or absence of flagging. LOLIs that are phonologically integrated are less often flagged than LOLIs that are not phonologically integrated. Similarly, established LOLIs are less often flagged than nonce LOLIs, but this seems to be at a lesser extent than the difference between LOLIs with Spanish versus English phonology. However, these observations are based solely on distributional frequencies and need to be tested for statistical significance before we can answer MRQ 1 and 2.

3.4 Testing the significance of flagging by phonological integration and establishedness

In order to test whether the distributions described in 3.3 were statistically significant, the data was analyzed using a generalized linear mixed effects model with a logit link function in order to assess the probability of Flagging as a function of phonology (English, Spanish) and establishedness (Established, Nonce). Participants were given random intercepts. Main effects were found for phonology ($X^2(1) = 54.75$; $p < 0.001$), and establishedness ($X^2 (1) = 19.298$; $p < 0.001$). An additive model including both predictors did not improve the fit above and beyond that of either predictor alone, and the two predictors did not interact ($X^2 (1) = 0.12$; $p > 0.05$). The model fit was assessed using conditional R^2 (Nakagawa and Schielzeth, 2013). Phonology accounted for 13.4% of the variance within the data, whereas establishedness accounted for

3.9%. The likelihood of flagging was increased by 17.5% when the phonology was English (β = 1.55 ± 0.23 standard errors; z = -6.58; p < 0.001).

To better understand the results of the linear model, we will provide a discussion of how the results pertain to the two main research questions we presented in the introduction. First, we will explore MRQ1:

MRQ 1: Does the way in which flagging is distributed by phonological integration provide further evidence that phonology can be used in discerning LOLI borrowing from codeswitching?

In section 3.3, we noted that flagging occurred more before LOLIs that were produced with English phonology, LOLIs that were not phonologically adapted. The results from the linear model confirm that this distribution is significant. If we take into account the fact that that LOLIs with English phonology show lack of integration into Spanish (as demonstrated in Chapter 3), and we consider flagging to be a sign of higher cognitive strain as suggested by Schilperoord (2002), and we then in turn consider the higher strain to be a sign of codeswitching, the fact that LOLIs with English phonology are more often flagged is further indication that phonology can be used as a means of discerning borrowings from codeswitches. The findings of Chapter 3 and this current chapter indicate that LOLIs that are phonologically integrated (in this case, LOLIs with Spanish phonology) should be considered borrowings, while LOLIs that are not phonological integrated (LOLIs with English phonology) should be considered codeswitches.

The results of this study agree with previous psycholinguistic accounts based on laboratory speech (Meuter and Allport, 1999; Moreno, Federmeier and Kutas, 2002; Navracsics,

2004; Thomas and Allport, 2000), but differ from findings based on spontaneous speech (Dumont, 2010; Gardner-Chloros, McEntee-Atalianis, and Parasekva, 2013; Hennecke, 2013; Torres Cacoullos and Aaron, 2003a). This may be due to the small sample size in the current study, or may be due to differing codeswitching practices at the community level. It could be that codeswitching is found less in Southern Arizona than in the Spanish of New Mexico (Dumont, 2010; Torres Cacoullos and Aaron, 2003a). At any rate, regardless of the reason for the differing results, if we take into account (Schilperoord, 2002) that pauses are a sign of higher cognitive processing, for the bilinguals of this study, LOLIs with English phonology do come at a higher cost. This could, then, be an indication that LOLIs of this type are codeswitches. The fact that nonce borrowings also show a higher percentage of flagging may be further evidence. If we assume that the less-established a word is, the less defined the entry for that word is in the lexicon, then higher flagging rates in nonce borrowing would be expected to come at a higher cost, much like codeswitches.

As we discussed in the introduction, since the 1980s, studies exploring LOLIs' status as borrowings or codeswitches have used establishedness as the factor for separating LOLIs into categories. However, since this study is the first to introduce phonology into the analysis in such a prominent way, we also introduced our second main research question to determine if phonology may be a better predictor than establishedness:

MRQ2: When comparing flagging by establishedness and by phonological integration, which better explains the variation in the data?

The results from the linear model presented above seem to indicate that phonology accounts for more of the variation (13.4%) than establishedness (3.9%). Chart 6 below shows how adding establishedness to the model does not account for more variation.

Chart 6:

Lines of best fit for flagging by establishedness and phonology type
in LOLIs spoken during Spanish discourse from Spanish/English bilinguals in Arizona
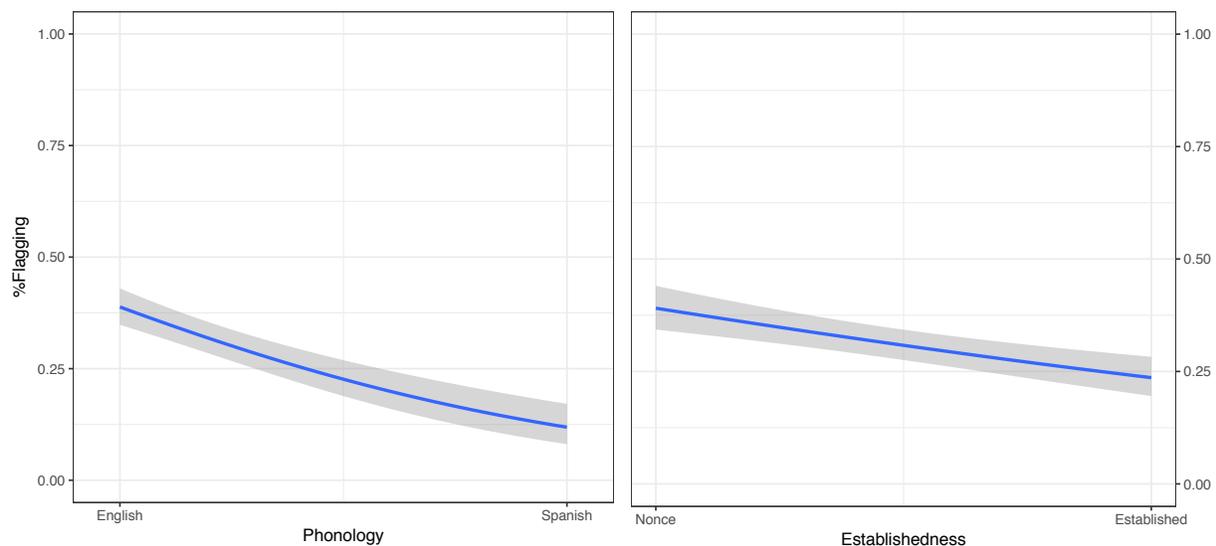
While phonological integration has been so long criticized and discarded from analyses on discerning LOLIs as borrowings or codeswitches, the results from this study present clear evidence that phonology does not only help predict whether a LOLI is a borrowing or a codeswitch, but quite possibly that it does so to a greater extent than the traditional division of LOLIs by establishedness.

4. Conclusion

In this chapter, we explored how pauses and false starts correlate to the phonological integration of LOLIs. The results indicate that LOLIs with English phonology are more often preceded with pauses and false starts than LOLIs with Spanish phonology. We determined that this finding most likely attests to the fact that LOLIs with English phonology, LOLIs that are unincorporated phonologically, are codeswitches. This conclusion is further evidenced by the fact that in Chapter 3 we saw that LOLIS with English phonology are also much less morphosyntactically integrated than LOLIs with Spanish phonology.

We also compared the correlation between flagging and phonological integration to the correlation between flagging and establishedness and found that phonological integration explains more of the variation than establishedness. This provides even more evidence that phonological integration is a key factor when discerning LOLIs' status as a borrowing or a codeswitch.

Based on the investigator's observations with the LOLIs in this data set, there are other phonological factors that may prove useful tools in determining if LOLIs that are not phonologically adapted into the recipient language are codeswitches. Many LOLIs with English phonology seem to not follow the rest of the Spanish phrase in the way LOLIs with Spanish phonology do. For instance, LOLIs with English phonology are sometimes produced at a higher volume, with a larger intensity, with a different pitch (on the suprasegmental, phrase, level). Additionally, for many LOLIs with English phonology, the preceding word is produced with a lengthened vowel, especially when the preceding word is a determiner. Future studies may

consider these factors when analyzing how LOLIs are inserted into the discourse of a recipient language.

While the findings of this chapter may not hold true in the community in general, for other communities, or for larger data sets, there is enough evidence here to warrant further investigation into how phonology affects the incorporation of LOLIs into a recipient language. Future studies should include phonological integration into their analyses and further explore its correlation with the presence of preceding flaggings.

Chapter 5:

Conclusion

This dissertation investigated lone other language nouns in the Spanish discourse of 24 bilinguals in Southern, Arizona. Three separate studies were carried out to explore the morphosyntactic and phonological integration of LOLIs among the bilingual speakers. Chapter 2 examined the morphosyntactic integration of LOLIs, while Chapter 3 analyzed the phonological integration of LOLIs. Finally, Chapter 4 examined flagging (pauses and false starts) before LOLIs.

The first study, Chapter 2, extensively documented the morphosyntactic integration of LOLIs among Sonoran Spanish bilingual speakers in Southern Arizona. To accomplish this task, we adopted the methodology of Torres Cacoullos and Aaron (2003a, 2003b) by testing the linguistic factors that condition the probability that an NP would contain a realized determiner, or consist of a bare noun. We first compared monolingual and bilingual Sonoran Spanish and found that bare nouns are most favored with a non-specific referent, nouns relating to occupations, and in NPs with a prenominal modifier. Determiner realization was conditioned in an almost identical way not only between the Sonoran monolingual speakers and the Arizona bilingual speakers of this study, but also among the New Mexican bilingual speakers in Torres Cacoullos and Aaron (2003a, 2003b), showing continuities of Southwest Spanish along the US-Mexican border. A second comparison was calculated for monolingual and bilingual English, which showed similarities in the conditioning of determiner realization between the two groups. Specifically, in English, bare nouns are favored in noun types that are generic and institutional.

Lastly, we compared established and nonce LOLIs to the patterns we found for both English and Spanish. We found that LOLIs in general show morphosyntactic adaptation into the recipient language grammar, in this case Spanish. While established LOLIs are very well integrated, nonce LOLIs are not fully integrated in that they share the English pattern for specificity. This pattern is to be expected, since material that is transferred into a recipient language falls along a continuum, from well-established and highly-integrated borrowings to unintegrated codeswitches (Hougan, 1950; Poplack, 1993; Weinreich, 1953).

Chapter 2 led to several contributions to our understanding of how bilinguals integrate LOLIs into the discourse of a recipient language. First, this was the first study to introduce a monolingual comparison and prove that bilinguals do treat patrimonial nouns the same way as monolinguals, at least in terms of determiner realization. While previous research had assumed that bilinguals would not differ from monolinguals in this regard, it was important to test this assumption empirically. The continuities found in determiner realization in the Spanish spoken along the US-Mexico border between Sonora (Chapter 2), Southern Arizona (Chapter 2), and New Mexico (Torres Cacoullos and Aaron, 2003a) not only provide evidence that these dialects are related, but also that the conditioning of determiner realization is a valid test when discerning LOLI nouns as either borrowings or codeswitches. Finally, the result that nonce LOLIs in this study are not fully adapted into the recipient language gives further evidence to the notion that the material that bilinguals introduce into a recipient language can be found in a continuum from fully integrated borrowings to fully unintegrated codeswitches.

After determining the patterns of morphosyntactic integration of LOLIs in bilingual Spanish discourse by establishing the patterns for Sonoran Spanish, Arizonan English, and LOLIs, we turned our attention to the phonological integration of LOLIs in Chapter 3, since

phonology has been removed from most analyses of LOLIs since the 1980s. By categorizing

LOLIs as having Spanish phonology, English phonology, or mixed Spanish/English phonology,

we discovered that overall (99% of the tokens) LOLIs are either integrated (Spanish phonology)

or unintegrated (English phonology). This observation was a crucial first step in the argument

that phonological integration is a feature that can be used to determine LOLIs' status as a

borrowing or a codeswitch. Since phonological integration does not vary within a word, but

instead by LOLI lemma, we were able to test the correlation between phonological and

morphosyntactic integration. By comparing the conditioning of determiner realization between

Spanish, LOLIs with Spanish phonology, LOLIs with English phonology, and English, we

discovered that unlike when we separated LOLIs by establishedness in Chapter 2, the two LOLI

groups did not both pattern after Spanish nouns. Specifically, LOLIs with Spanish phonology

were morphosyntactically integrated, showing the same conditioning of determiner realization as

Spanish nouns. However, LOLIs with English phonology patterned more closely to English

nouns. Additionally, we looked at the correlation of morphosyntactic and phonological

integration by lemma. *High school*, the most common LOLI lemma in the data set, was produced

with Spanish phonology in about half of the tokens and with English phonology about half of the

tokens. When *high school*, which is an institutional noun, was produced with Spanish phonology

it was always produced with a determiner, matching the Spanish pattern that institutional nouns

highly favor the realization of a determiner. On the other hand, when *high school* was produced

with English phonology, it highly favored a bare realization, following the English pattern. The

fact that in general LOLIs with Spanish phonology were morphosyntactically integrated into

Spanish and LOLIs with English phonology were morphosyntactically unintegrated and that this

trend held true for even well-established LOLI lemmas, led us to conclude that phonological

integration is an important factor to be considered when studying LOLI integration patterns. It may well be that when a LOLI, even from a well-established lemma, is produced with English phonology, the LOLI is accessed from the lemma of the English grammar rather than from the well-established lemma in the Spanish grammar. We further hypothesized that LOLIs with Spanish phonology were borrowings, while LOLIs with English phonology were codeswitches (or at the least nonce borrowings). The results of this chapter strikingly contradict those of previous work by Poplack and associates (i.e. Poplack and Dion, 2012; Poplack and Meechan, 1995, 1998). Further research on a wider range of communities and language pairings will be necessary to better understand the significance of phonological integration on the incorporation of LOLIs in bilingual discourse. One particular contention point may be found in the type of bilinguals included in the current study, namely the bilinguals are highly proficient in the two languages they speak (Spanish and English), including the phonological level. The bilinguals of this study have developed native Spanish and English phonological grammars, they do not have an accent in either language. Previous work by Poplack and Sankoff (1984) and Poplack, Sankoff, and Miller (1988) included L1 dominant bilinguals in the analysis. The higher amount of variation found in those studies may be more to do with phonological transfer or interference than with a choice that bilinguals make when phonologically integrating (or not) words from a donor language. Future studies should report in detail the linguistic competency of their bilingual participants in order to compare results from balanced bilinguals and L1 dominant bilinguals.

The work laid out in Chapter 3 has advanced the field in multiple ways. This study is among very few on LOLIs that engage in the treatment of phonological integration at all. It is also the first study to provide, in detail, an analysis of the phonological adaptations (or lack thereof) that occur among LOLIs. The perception experiment involving native English judges

that was designed to verify the impressionistic observations of the investigator was novel and proved to be very effective in the classification of the phonological integration of LOLIs. Future studies should consider using the judgments of native judges as the sole classification of phonological integration in order to remove the investigator's impressionistic observations and increase the reliability of the results. Another contribution of this dissertation is that it demonstrated that while there is considerable variation in the phonological integration of LOLIs, this variation is meaningful and can be explained. Future studies need to not only include phonology in their analysis, but move beyond reporting the frequency with which LOLIs are phonologically integrated. The next step, as illustrated in Chapter 3, is exploring the correlation between morphosyntactic and phonological integration. The hypothesis formulated based on the results in Chapter 3 suggests that LOLIs with the phonology of the recipient language are well established borrowings and are accessed from the lemma in the lexicon of the recipient language, while LOLIs with the phonology of the donor language are accessed from the lemma in the lexicon of the donor language and therefore should be considered codeswitches, or at the very least, nonce borrowings.

In order to further explore this hypothesis, we analyzed pauses and false starts in Chapter 4. In this chapter, we compared the amount of flagging (pauses and false starts) that was present before LOLIs with Spanish phonology and LOLIs with English phonology. Based on data from studies in psycholinguistics (Meuter and Allport, 1999; Moreno, Federmeier and Kutas, 2002; Navracsics, 2004; Schilperoord, 2002; Thomas and Allport, 2000;), we suggested that a higher amount of flagging among LOLIs with English phonology could be a sign of higher cognitive processing, and in turn a sign of codeswitching. The results indeed showed that flagging occurred at a much higher rate among LOLIs with English phonology than LOLIs with Spanish

phonology. These findings, combined with those in Chapter 3, provide strong evidence in support of the hypothesis that LOLIs with English phonology should be seen as codeswitches, while LOLIs with Spanish phonology should be considered borrowings. Also tested was the amount of the variation in the data that could be explained by separating LOLIs by establishedness versus by phonology. We found that while both categorizations were significant predictors, separating LOLIs by phonology explained a greater amount of variation than separating LOLIs by establishedness. The results in this chapter demonstrate that phonology should be used in the study of LOLIs' status as borrowings or codeswitches, since it is a stronger predictor than the conventional classification by establishedness.

Chapter 4 provided insights into the classification of LOLIs as borrowings or codeswitching by discussing flaggings, which have not been well addressed in previous variationist studies on LOLIs. This analysis further evidenced our interpretation of the results in Chapter 3, which stated that LOLIs with phonology of the recipient language are established borrowings and LOLIs with phonology of the donor language are codeswitches.

Whether or not the results found in Chapters 3 and 4 hold true in other communities or with larger data sets, the results are robust enough to warrant more research on the relevance of phonological integration in discerning LOLIs as borrowings or codeswitches. We hope future studies of LOLIs will consider phonology in their analysis in order to further test the hypothesis that LOLIs that are phonologically adapted into the recipient language are borrowings, while LOLIs that are not phonologically adapted into the recipient language are codeswitches. We also suggest, with a larger data set, a more thorough analysis of the interface between phonological integration and establishedness by further dividing LOLIs into four groups: 1) established and phonologically integrated LOLIs, 2) established but not phonologically integrated LOLIs, 3)

nonce but phonologically integrated LOLIs, and 4) nonce and phonologically unintegrated LOLIs. The prediction is that these groups would fall along the integration continuum (Hougan, 1950; Poplack, 1993; Weinreich, 1953), where (1) would be very well established and most like patrimonial words, (2) and (3) would be somewhere between borrowings and codeswitches, and (4) would prove to be codeswitches.

The results of this dissertation have furthered our understanding of how bilingual speakers incorporate donor language material into a recipient language. Specifically, the results of Chapter 2 established that bilinguals follow the same constraints that condition determiner realization for patrimonial nouns in both of the languages they speak. Additionally, it was found that LOLIs, both established and nonce, are morphosyntactically integrated into the recipient language, in this case Spanish. Chapter 3 brought phonology back into the discussion, which had been removed since the mid 1980s, and found that the variation observed between LOLIs that are phonologically adapted and LOLIs that are not phonologically adapted can be explained through the correlation between morphosyntactic and phonological integration. LOLIs that are produced with the phonology of the recipient language are also morphosyntactically integrated into the recipient language. On the other hand, LOLIs that are produced with the phonology of the donor language act more morphosyntactically like nouns in the donor language. Lastly, Chapter 4 measured the amount of flagging (pauses and false starts) among LOLIs that are phonologically integrated and those that are not. The results showed that flagging occurs more often before LOLIs that are not phonologically integrated, which was taken as a sign that LOLIs of this type are codeswitches. The main contribution of this dissertation is that it demonstrated that phonological integration is a factor that should be considered in any analysis that seeks to discern LOLIs as borrowings or codeswitches.

References:

Aaron, J. E. (2014). Lone English-origin nouns in Spanish: The precedence of community

   norms. In C. Travis & R. Torres Cacoullos (Eds.), *International Journal of Bilingualism*,

   Special issue, doi:10.1177/1367006913516021.

Academia Mexicana de la Lengua. (2015). http://www.academia.org.mx.

Academia Mexicana de la Lengua. (2010). *Diccionario de mexicanismos*.

   http://www.academia.org.mx/DiccionarioDeMexicanismos.

Auer, P. (1999). From code-switching via language mixing to fused lects: toward a dynamic

   typology of bilingual speech. *International Journal of Bilingualism 3(4),* 309-332.

Bessett, R. M. (In progress). Testing English influence on subject pronoun expression in Sonoran

   Spanish.

Bessett, R. M. (2015). The extension of *estar* across the Mexico-US border: Evidence against

   contact induced acceleration. *Sociolinguistic Studies, 9(3)*, 421-444.

Bessett, R. M. (2012). *La extensión de estar a través del tiempo y el espacio: Un estudio

   diacrónico de los siglos XIII-XX y sincrónico de Hermosillo, Sonora y el Sur de Arizona*.

   (Master's Thesis). University of Arizona, Tucson.

Birdsong, D., Gertken, L. M., & Amengual, M. (2012). *Bilingual Language Profile: An Easy-to-Use Instrument to AssessBilingualism.* COERLL, University of Texas at Austin. <https://sites.la.utexas.edu/bilingual/>.

Blas Arroyo, J. L., & Tricker, D. (2000). Principles of variationism for disambiguating language contact phenomena: The case of lone Spanish nouns in Catalan discourse. *Language Variation and Change, 12*, 103–140.

Bley-Vroman, R. (1983). The comparative fallacy in interlanguage studies: The case of systematicity. *Language Learning*, *33*, 1–17.

Budzhak-Jones, S. (1998). Developing diagnostics: Word-internal code-switching versus borrowing. *Cahiers linguistiques d'Ottawa,* 26, 1-14.

Bullock, B. (2009). Phonetic reflexes of code-switching. In B. Bullock & A. J. Toribio (Eds.), *The Cambridge handbook of linguistics code-switching*, 163-181. Cambridge: Cambridge University Press.

Carvalho, A. M. (2012-). *Corpus del Español en el Sur de Arizona (CESA)*. University of Arizona. cesa.arizona.edu.

Cerda, G., Cabaza, B. & Farias, J. (1970). *Vocabulario español de Texas*. Austin/London: The University of Texas Press.

Cerda, G., Cabaza, B. & Farias, J. (1953). *Vocabulario español de Texas*. Austin/London: The
University of Texas Press.

Cobos, R. (2003). *A Dictionary of New Mexico & Southern Colorado Spanish.* Santa Fe:
Museum of New Mexico Press.

Cobos, R. (1983). *A Dictionary of New Mexico and Southern Colorado Spanish*. Santa Fe:
Museum of New Mexico Press.

Corominas, J. (1967). *Breve diccionario etimológico de la lengua castellana.* Madrid: Editorial
Gredos, S.A.

Dajko, N., & Carmichael, K. (2014). But qui c'est la difference? Discourse Markers in Louisiana
French: The Case of but vs. mais. *Language in Society*, 43(2), 159-183.

de Rooij, V. A. (2000). French Discourse Markers in Shaba Swahili Conversations. *International
Journal of Bilingualism*, 4(4), 447-469.

Galván, R. A. & Teschner, R. V. (1985). *El diccionario del español chicano/The Dictionary of
Chicano Spanish.* Lincolnwood, IL: National Textbook Company.

Galván, R. A. & Teschner, R. V. (1975). *El diccionario del español de Tejas/The Dictionary of*

the Spanish of Texas (Spanish-English). Maryland: Institute of Modern Languages, Publishers in the Languages of the World.

García Icazbalceta, J. (1899). *Vocabulario de mexicanismos: comprobado con ejemplos y comparado con los de otros países hispano-americanos.* México: imp. y lit. "La Europea" de J. Aguilar Vera y Ca.

García, T. (1935). "A Vocabulary of New Mexican Spanish Words of English Origin from Southwestern New Mexico." M.A. Thesis, Stanford University.

Gardner-Chloros, P., McEntee-Atalianis & Paraskeva, M. (2013). Code-switching and pausing: an interdisciplinary study. *International Journal of Multilingualism, 10(1)*, 1-26.

Ghafar Samar, R. & Meechan, M. (1998). The null theory of code-switching versus the nonce borrowing hypothesis: Testing the fit in Persian-English bilingual discourse. *International Journal of Bilingualism, 2(2)*, 202-219.

Gross, S. M. (1935). "A Vocabulary of New Mexican Spanish." M.A. Thesis, Leland Stanford Junior University.

Hamel, B. H. (1994). *Hamel's Bilingual Dictionary of Mexican Spanish/Diccionario bilingüe Hamel de Mexicanismos.* Los Angeles: Bilingual Book Press.

Haugen, E. (1950). The analysis of linguistic borrowing. *Langauge, 26(2)*, 210-231.

Haugen, E. (1969). *The Norwegian language in America*. Philadelphia: University of

Pennsylvania Press.

Hennecke, I. (2013). Self-repair and language selection in bilingual speech processing. *Discours:*

Revue de linguistique, psycholinguistique et informatique. A journal of linguistics,

psycholinguistics and computational linguistics, 12, 1-20.

Holden, K. (1976). Assimilation rates of borrowings and phonological productivity. *Language,*

52(1), 131-147.

Ito, Junko, and Armin Mester. (2003a). On the sources of opacity in OT: coda processes in

German. In Féry, Caroline and van de Vijver, Ruben (Eds.), Structure and Typology of

the Syllable. Cambridge, UK: Cambridge University Press.

Ito, Junko, and Armin Mester. (2003b). Weak Parallelisn: Allophonic Interactions in German.

In Lombardi, Linda (Ed.), Segmental Phonology in Optimality Theory, 271-303.

Cambridge, UK: Cambridge University Press.

Jacobs, Haike & Gussenhoven, Carlos. (2000). Loan phonology: Perception, salience, the

lexicon and OT. In Dekkers, Joost, van der Leeuw, Frank, & van de Weijer, Jeroen

(Eds.), *Optimality Theory: Phonology, syntax, and acquisition*, 193-210. Oxford: Oxford University Press.

Jaramillo, J. A. (1995). The passive legitimization of Spanish: A macrosociolinguistic study of a quasi-border: Tucson, Arizona. *International Journal of the Sociology of Language* 114, 67–91. Doi: http://dx.doi.org/10.1515/ijsl.1995.114.67.

Kern, J. (Forthcoming). Corpus of bilingual (Spanish-English) speakers and English monolingual Speakers (official name: TBD).

Lara Ramos, L. F. (1996). *Diccionario del español usual en México*. Alicante : Biblioteca Virtual Miguel de Cervantes.

Liberman, M. (2011). /l/ variation in American English: A corpus approach. *Journal of Speech Sciences,* 1(2), 35-46.

Lugones, Leopoldo. (1932). *Diccionario etimológico del castellano usual*. Buenos Aires: Academia Argentina de Letras.

MacSwan, J. & Colina, S. (2014). Some consequences of language design: Codeswitching and the PF interface. In EDITOR(S) (Ed.), *Grammatical Theory and Bilingual Codeswitching*, 185-210. MA: MIT Press.

Maschler, Y. (2009). *Metalanguage in Interaction: Hebrew Discourse Markers*. Amsterdam; Philadelphia: John Benjamins Pub. Co.

Maschler, Y. (1997). Emergent Bilingual Grammar: The Case of Contrast. *Journal of Pragmatics*, 28(3), 279-313.

Maschler, Y. (1994). Metalanguaging and Discourse Markers in Bilingual Conversation. *Language in Society*, 23(3), 325-366.

Meuter, R. F. I. & Allport, A. (1999). Bilingual language switching in Naming: Asymmetrical costs of language selection. *Journal of Memory and Language, 40*, 25-40.

Moreno, E. M., Federmeier, K. D., & Kutas, M. (2002). Switching languages, switching *palabras* (words): An electrophysiological study of code switching. *Brain and Language, 80*(2), 188-207.

Mougeon, R., Beniak, E., & Valois, D. (1985). *Variation in the Phonological Integration of Loanwords in a Bilingual Speech Community.* Toronto: Centre for Franco-Ontarian Studies.

Muysken, P. (2005). Two linguistic systems in contact: Grammar, phonology and lexicon. In T. K. Bhatia & W. C. Ritchie (Eds.), *The Handbook of Bilingualism,* 148-168. Maiden/Oxford: Blackwell.

Muysken, P. (2000). *Bilingual Speech: A Typology of Code-mixing*. Cambridge University Press.

Myers-Scotton, C. (2009). Code-switching. In N. Coupland & A. Jaworski (Eds.), *The new sociolinguistics reader,* 473-489. New York: St. Martin's Press.

Myers-Scotton, C. (2006). *Multiple Voices. An Introduction to Bilingualism*. Blackwell Publishing.

Myers-Scotton, C. (1993). *Duelling languages*. Oxford: Clarendon Press.

Navracsics, J. (2004). The question of control in bilingual speech production in different language modes. *Grazer Linguistische Stu*dien, 62, 95-110.

Otheguy, R. & A. C. Zentella. (2012). *Spanish in New York: Language Contact, Dialect Leveling, and Structural Continuity*. Oxford: Oxford University Press.

Poplack, S. (2012). What does the Nonce Borrowing Hypothesis hypothesize? *Bilingualism: Language and Cognition, 15(3)*, 644-648.

Poplack, S. (2004). Code-switching. In U. Ammon, N. Dittmar, K. J. Mattheier, & P. Trudgill, *Soziolinguistik: an international handbook of the science of language*, *2nd edition*, 589-596. Berlin: Walter de Gruyter.

Poplack, S. (1993). Variation theory and language contact. In D. R. Preston (Ed.), *American dialect research*, 251-286. Amsterdam: John Benjamins.

Poplack, S. & Levey, S. (2010). Contact-induced grammatical change: A cautionary tale. In P. Auer & J. E. Schmidt (Eds.), *Language and Space: An international handbook of linguistic variation*, *vol. 1: Theories and methods*, 391-419. Berlin: Mouton de Gruyter.

Poplack, S, & Meechan, M. (1998). Introduction: How languages fit together in codemixing. *International Journal of Bilingualism, 2(2),* 127-138.

Poplack, S., & Meechan, M. (1995). Patterns of language mixture: Nominal structure in Wolof-French and Fongbe-French bilingual discourse. In P. Muysken & L. Milroy (Eds.), *One speaker, two languages*, 199–232. Cambridge: Cambridge University Press.

Poplack, S & Dion, N. (2012). Myths and facts about loanword development. *Language Variation and Change, 24(3),* 279-315.

Poplack, S., Zentz, L., & Dion, N. (2012). What counts as (contact-induced) change. *Bilingualism: Language and Cognition, 15(2)*, 247-254.

Poplack, S. & Sankoff, D. (1984). Borrowing: the synchrony of integration. *Linguistics, 22,* 99-135.

Poplack, S., Sankoff, D., & Miller, C. (1988). The social correlates and linguistic processes of

    lexical borrowing and assimilation. *Linguistics, 26*, 47-104.

Potowski, K., & Carreira, M. (2010). Spanish in the USA. In K. Potowski (Ed.), *Language*

    diversity in the USA, 66-80. Cambridge: Cambridge University Press.

Real Academia Española. (2016). *Diccionario de la lengua española.* http://www.rae.es. Madrid.

Real Academia de la Lengua Sonorense. (2006). *Diccionario: ¡¡¡Arriba Sonora!!!.*

Repetti, L. (2006). The emergence of marked structures in the integration of loans in Italian. In

    Gess, Randall & Arteaga, Deborah (Eds.), *Historical Romance linguistics: Retrospective*

    and perspectives, 209-235. Amsterdam& Philadelphia: John Benjamins.

Rescansens, D. & Espinosa, A. (2005). Articulatory, positional and coarticulatory characteristics

    for clear /l/ and dark /l/: evidence from two Catalan dialects. *Journal of the International*

    Phonetic Association, 35(1), 1-25.

Reyna, J. R. (1980). *Modismos de Tejas*. Texas: Penca Books.

Sankoff, D. (2002). Linguistic outcomes of language contact. In J. K. Chambers, P. Trudgill, &

    N. Schilling-Estes (Eds.), *The handbook of language variation and change*, 638-668.

    Oxford: Blackwell.

Sankoff, D., Poplack, S., & Vanniarajan, S. (1990). The case of the nonce loan in Tamil. *Language Variation and Change, 2(1),* 71-101.

Sayahi, L. (2005). Phonological adaptation of Spanish loanwords in northern Moroccan Arabic. *University of Pennsylvania Working Papers in Linguistics, 11*(1), 253-263.

Schilperoord, J. (2002). On the cognitive status of pauses in discourse production. In Olive, Thierry and Levy, C. Michael (Eds.), Contemporary tools and techniques for studying writing (Vol. 10), 61-88. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Schwegler, A., Kempff, J., & Ameal-Guerra, A. (2010) *Fonética y fonología esapñolas, 4th Edition*. New Jersey: John Wiley & Sons, Inc.

Thomas, M. S. C. & Allport, A. (2000). Language switching costs in bilingual visual word recognition. *Journal of Memory and Language, 43*, 44-66.

Thomason, S. G. (2001). *Language contact*. Edinburgh: Edinburgh University Press.

Thomason, S. G., & Kaufman, T. (1988). *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.

Torres Cacoullos, R., & Aaron, J. E. (2003a). Bare English-origin nouns in Spanish: Rates, constraints, and discourse function. *Language Variation and Change, 15*, 289-328.

Torres Cacoullos, R., & Aaron, J. E. (2003b). Determiner variation with English origin nouns in New Mexican Spanish: Borrowing bare forms. *Working Papers in Linguistics*, *9*, 159-172.

Torres Cacoullos, R., & Ferreira, F. (2000). Lexical frequency and voiced labiodental-bilabial variation in New Mexican Spanish. *Southwest Journal of Linguistics,* 19(2), 1-18.

Torres Cacoullos, R., & C. Travis. (2010a). Variable yo expression in New Mexico: English influence? In S. Rivera-Mills & D. Villa Crésap (Eds.), *Spanish of the U.S. southwest: A language in transition*, 189-210. Madrid: Iberoamericana/Vervuert.

Torres Cacoullos, R., & C. Travis. (2010b). Testing convergence via code-switching: priming and the structure of variable subject expression. *International Journal of Bilingualism,* XX(X), 1-27.

Torres Cacoullos, R., & Vigil, N.A. (2003). Sustantivos de origen ingles en discurso español: ¿cuál es su gramática? *A Journal of the Borderlands 2(1),* 69-84.

Torres Cacoullos, R., & Vigil, N. A. (2002). 'Mi *dad* habla mexicano': ¿Préstamos o alternancias de códigos? In V. M. I. Sánchez Corrales (Ed.), A*ctas XIII Congreso Internacional Asociación de Lingüística y Filología de América Latina (ALFAL) 2002,* 455-462. San Juan, Costa Rica.

Travis, C. E. (2007). Genre effects on subject expression in Spanish: Priming in narrative and conversation. *Language Variation and Change,* 19(2), 101-135.

Willis, E. W. (2007). An acoustic study of the 'pre-aspirated trill' in narrative Cibaeño

    Dominican Spanish. *Journal of the International Phonetic Association, 37(1),* 33-49.

Yip, M. (1993). Cantonese loanword phonology and optimality theory. *Journal of East Asian*

    Linguistics, 2(3), 261-291.

Yip, M. (2005). The symbiosis between perception and grammar in loanword phonology.

    Lingua, 116(7), 950–975.