Exploring the Phonological Integration of Lone Other-Language Nouns in the Spanish of Southern Arizona

Ryan M. Besset
University of Arizona

This paper is posted at ScholarlyCommons. http://repository.upenn.edu/pwpl/vol23/iss2/5
For more information, please contact repository@pobox.upenn.edu.
Exploring the Phonological Integration of Lone Other-Language Nouns in the Spanish of Southern Arizona
Exploring the Phonological Integration of Lone Other-Language Nouns in the Spanish of Southern Arizona

Ryan M. Bessett

1 Introduction

It is widely attested that the lexicon is the part of grammar that is most permeable to the effects of contact (Muysken, 2000, 2005; Sankoff, 2002; Thomason, 2001; Thomason and Kaufman, 1988; among many others). Within this category, language contact usually results in the use of one-item lexical elements from a donor language into a recipient language, which have been referred to as "lone other language items" or "LOLIs" (Poplack, Sankoff, and Miller, 1988; Poplack and Meechan, 1995; and Poplack, 2012). LOLIs have been the subject of a great breadth of variationist research which has centered around determining if LOLIs should be classified as borrowings (incorporated into the recipient language) or unincorporated codeswitches (Poplack, Sankoff, and Miller, 1988; Sankoff, Poplack and Vanniarajan, 1990; Poplack and Meechan, 1995; Myers-Scotton, 1993, 2006; Myers-Scotton, 2009; Ghafar and Meechan, 1998; Blas Arroyo and Tricker, 2000; Torres Cacoullos and Vigil, 2002, 2003; Torres Cacoullos and Aaron, 2003a, 2003b; Poplack, 2012; Aaron, 2014; among countless others). In order to make the determination, variationist studies have used a variety of morphological and syntactic features in their analyses and have unanimously determined LOLIs demonstrate morphosyntactic integration and therefore should be classified as borrowings (Sankoff, Poplack and Vannianajan, 1990; Poplack and Meechan, 1995; Ghafar and Meechan, 1998; Torres Cacoullos and Aaron, 2003a, 2003b; Blas Arroyo and Tricker, 2000; Poplack and Dion, 2012; Aaron, 2014). Based on evidence from two studies (Poplack and Sankoff, 1984; Poplack, Sankoff and Miller, 1988), phonological integration has been rejected as a method of differentiating LOLIs as borrowings or codeswitches due to its variability (Poplack and Meechan, 1995; Torres Cacoullos and Aaron, 2003b; Aaron, 2014). The present study, using data from highly proficient Spanish-English bilinguals in Southern Arizona, seeks to test the variability of phonological integration of LOLI nouns (LOLNs) in the Spanish of this border community. The way in which the variability manifests itself, either within a word or at the lemma level, will be fully explored. If we find variability at the lemma level, but not within words, we can then test the correlation between morphosyntactic integration and phonological integration in order to formulate a fuller picture of LOLIs’ status as borrowings or codeswitches.

The seminal work of Poplack and Meechan (1995, 1998) applied the sociolinguistic comparative method to loanwords. Under this framework, the frequencies of morphosyntactic factors are compared to the frequencies in patrimonial words of the recipient language, patrimonial words of the donor language (usually gathered from multi-word codeswitches), and established borrowings. The idea is that established borrowings should pattern after patrimonial words from the recipient language, given that they are well integrated into the grammar of the recipient language. If ambiguous LOLIs pattern after words from the recipient language, they too show incorporation and therefore should be considered borrowings. However, if LOLIs pattern after words from the donor language, they should be considered codeswitches as they are not incorporated into the recipient language grammar. Several morphosyntactic variables have been considered in previous analyses: the realization of determiners (Poplack and Meechan, 1995; Ghafar and Meechan, 1998; Blas Arroyo and Tricker, 2000, Poplack and Dion, 2012; Aaron, 2014), word order (Sankoff, Poplack and Vanniarajan, 1990; Ghafar and Meechan, 1998; Poplack and Dion, 2012), the formation of contractions (Blas Arroyo and Tricker, 2000), complements when complement formation is a different process (possibly order) in the two languages (Sankoff, Poplack and Vanniarajan, 1990; Blas Arroyo and Tricker, 2000), morphological case (Sankoff, Poplack and Vanniarajan, 1990; Ghafar and Meechan, 1998), gender assignment/agreement (Blas Arroyo and Tricker, 2000; Budzhak-Jones, 1998; Poplack and Dion, 2012; Aaron, 2014) and plural marking (Sankoff, Poplack and Vanniarajan, 1990; Poplack and Dion, 2012).

Guided by the notion that overall frequency is not always the best method of testing differences between two communities (see Travis, 2007), Torres Cacoullos and Aaron (2003a, 2003b) advance the technique used to discern LOLIs as borrowings or codeswitches by testing the conditioning of
the morphosyntactic integration. In their study, the researchers examine the probability of whether a determiner is realized based on several linguistic factors that condition the use of determiners: specificity (specific, non-specific, generic), syntax (subject, object, oblique, predicate nominal/existential), the use of a modifier (prenominal, post nominal, none), string position (first or only noun, second or subsequent noun), polarity (negated, non-negated), and semantic class (occupation/status, institution, coincidence sites).

In terms of phonological integration, very little research has been conducted. Poplack and Sankoff (1984) and Poplack, Sankoff and Miller (1988) measured phonological integration based on impressionistic observations of the data by coding each feature in a token as having the recipient language phonology or the donor language phonology. The researchers reported that sometimes a particular lemma was produced with the recipient language phonology (demonstrating integration) and sometimes the same lemma was produced with the donor language phonology (demonstrating it was not integrated). The variation was found among ambiguous LOLIs (49% integration), established borrowings (79% integration), and codeswitches (27% integration). These findings led the authors to conclude that phonological variation is too variable to be used as a measure of LOLI integration. Under their interpretation, LOLIs that are not integrated phonologically to the recipient language can still be classified as borrowings. This interpretation is at odds with phonological theory, however. In phonology, under Optimality Theory, any foreign word borrowed into the discourse of the recipient language must go through the phonological constraints of the recipient language and show phonological adaptation (Holden, 1976; Yip, 1993, 2006; Jacobs, Haeke and Gussenhoven, 2000; Sayahi, 2005; Kenstowicz and Suchato, 2006; Repetti, 2006). In order to reconcile the discrepancy between sociolinguistic and phonological accounts, it is necessary to not only determine the amount of integration by lemma, but also the correlation between phonological and morphosyntactic integration.

The present study explores in detail the phonological integration of LOLNs in the Spanish discourse of Spanish-English bilinguals in Southern Arizona. First, we provide an in-depth analysis of the phonological variation both within-word and at the lemma level and provide a discussion of the implications of the type of variation found in this community. We then provide an exploratory examination of the correlation between the morphosyntactic and phonological integration of the LOLNs found in this data set. Finally, we provide implications for the results and avenues for future research.

2 Methodology

2.1 The Participants

A total of 24 participants took part in this study. All 24 were between the ages of 18 and 30 and were divided evenly between men and women. They are residents of Southern Arizona, from either Tucson or Nogales, located along the US border with the Mexican state of Sonora and their families originate from Sonora. The participants come from three different corpora: 8 took part in sociolinguistic interviews with the investigator, 8 took part in sociolinguistic interviews with various interviewers as part of the Corpus del Español en el Sur de Arizona (CESA) (Carvalho 2012-), and the remaining 8 participants are from a sub-section of CESA and involve two bilingual speakers conversing without the presence of an investigator\(^1\).

The participants were born in the United States or arrived before the age of ten, except for one participant who arrived at the age of fifteen. The participants auto evaluated their proficiency for both English and Spanish as at least a 6/10 (or 4.5/6 for the two CESA groups). The speakers also demonstrated their high proficiency in both languages by holding a conversation in Spanish for the duration of the hour-long interview, attending school in the United States (in English), and living and working in a bilingual setting.

\(^1\)The three groups were formed with the hypothesis that each group would present varying degrees of codeswitching. It was believed that the 8 interviewed by the investigator (and L2 speaker of Spanish) would have few codeswitches and the 8 participants who conversed with another bilingual speaker and without an interviewer would codeswitch quite often. These results did not come to fruition, all three groups presented very few codeswitches.
2.2 Phonological Integration

Every LOLN of English origin was extracted from the nearly 24 hours of interviews. Each noun was coded via impressionistic observations of the investigator as having English (unincorporated), Spanish (incorporated), or mixed phonology. In order to test the validity of the investigator’s impressionistic judgements, three further measures were undertaken. First, 352 tokens contained a rhotic and these were analyzed in PRAAT to determine if the Spanish flap [ɾ], characterized by a single occlusion at the end of the sound, or the Spanish trill [ɾ], characterized by multiple occlusions (Willis 2007), or the English retroflex [ɹ] could be identified. Of the 65 tokens that were perceived by the interviewer as having a Spanish rhotic, 64 were confirmed by the presence of an occlusion in the spectrogram reading in PRAAT. Second, 193 tokens contained a lateral in coda position, where in Spanish we would expect to find a “light l” and in English we would find a “dark l”. In the spectrogram, a “dark l” has a much lower F2 value (Recasens and Espinosa 2005). Of the 119 perceived “dark l” tokens, 116 were confirmed in the spectrogram reading in PRAAT. Lastly, a subset of the LOLNs (N = 179) was extracted in isolation and was incorporated into a perception experiment with 19 monolingual English participants that were recruited from a first semester Spanish course. As a control, a native Spanish speaker and a native English speaker each produced 20 lemmas found in the subset of the LOLNs. The participants heard the words and were asked to rate them as having foreign (left side of the screen) or native (right side of the screen) pronunciation. A total of 100 response values was possible. Each participant heard each word 3 times. The average response values for the participants was compared to the average response values of the investigator for each stimulus. A Pearson product-moment correlation revealed that there was a strong positive correlation between the responses of the participants and the responses of the investigator (r=0.878, t=27.03, df=217, p < 0.001). Due to the fact that the impressionistic judgments of the investigator could be confirmed in PRAAT for the rhotics and the laterals in coda position and the strong correlation between the judgments made by the investigator and the 19 English monolingual judges, it was reasonable to conclude that the judgements by the investigator were accurate, and therefore those judgments were used to code for phonological integration.

2.3 Morphological Integration

All the LOLNs were coded for having a realized determiner or as bare nouns. Following the methodology of Torres Cacoullos and Aaron (2003a), each token was also coded for the linguistic variables that condition determiner realization: specificity (specific, non-specific, generic), syntax (subject, object, oblique, predicate nominal/existential), the use of a modifier (prenominal, post nominal, none), string position (first or only noun, second or subsequent noun), polarity (negated, non-negated), and semantic class (occupation/status, institution, coincidence sites). Additionally, the tokens were classified as based on establishedness: well-established (found in official dictionaries), regional (found in local dictionaries, lists of borrowings in Southwest Spanish, or in the AUTHOR (2012) corpus of monolingual Sonoran speakers), or unattested.

2.4 Data Analysis

In section 3, we report the phonological integration of the LOLNs found in this data set. We first report the frequency of integrated (Spanish phonology), unintegrated (English phonology), and mixed tokens. We then provide an analysis of integration by lemma. We expect to find variability within lemma, as reported in Poplack and Sankoff (1984) and Poplack, Sankoff and Miller (1988). This is to say that within a particular lemma, some tokens will be produced with Spanish phonology and others with English phonology. However, in accordance with phonological theory, we do not expect to find variation within a word. If these hypotheses are borne out, we will then provide a preliminary analysis of the correlation between morphological and phonological integration.
3 Results

A total of 756 LOLNs were identified in the Spanish discourse of the 24 bilingual participants. In section 3.1 we discuss the overall frequency of Spanish, English and mixed phonology. In section 3.2 we provide the results of phonology by lemma. Section 3.3 examines phonological integration by establishedness. Lastly, in section 3.4 we conduct a preliminary analysis of the correlation between phonological and morphosyntactic integration.

3.1 Overall Frequency of Phonology Types

LOLNs are more often produced with English phonology (72%, 546/756) than Spanish phonology (202/756). Only 8 (1%) of the tokens were produced with mixed phonology and of them, the majority (5) were found at a word boundary and only 3 were within the same word. Chart 1 provides a detailed account of the phonological integration types.

![Chart 1](image)

Figure 1: Frequency of phonology type (Spanish, English, mixed) of LOLI produced in Spanish discourse by Arizona bilinguals.

Overall, LOLNs 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 (only present in 0.4% of the data). For this data set, then, phonological integration is an all or nothing process, words are either integrated into Spanish phonology or maintained with English phonology.

3.2 Phonology by Lemma

In order to further understand the phonological integration that is found in the LOLNs in Arizona Spanish, we will now focus on phonological integration by lemma. Table 1 shows the phonology (Spanish or English) of each LOLN lemma that had three or more tokens.

<table>
<thead>
<tr>
<th>Lemma</th>
<th>Spanish</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>high school</td>
<td>19/39</td>
<td>20/39</td>
</tr>
<tr>
<td>football (futból)</td>
<td>19/19</td>
<td>0/19</td>
</tr>
<tr>
<td>Truck (troque)</td>
<td>14/14</td>
<td>0/14</td>
</tr>
<tr>
<td>Telephone (teléfono)</td>
<td>11/11</td>
<td>0/11</td>
</tr>
<tr>
<td>Facebook</td>
<td>1/10</td>
<td>9/10</td>
</tr>
<tr>
<td>minor</td>
<td>0/9</td>
<td>9/9</td>
</tr>
<tr>
<td>Item</td>
<td>Ford</td>
<td>Freshman</td>
</tr>
<tr>
<td>--------------------</td>
<td>------</td>
<td>----------</td>
</tr>
</tbody>
</table>
Table 1: Phonological integration (Spanish versus English phonology) by most frequent lemmas (3 or more tokens) in Arizona Spanish.

<table>
<thead>
<tr>
<th>Lema</th>
<th>Spanish phonology</th>
<th>English phonology</th>
</tr>
</thead>
<tbody>
<tr>
<td>social security</td>
<td>0/3</td>
<td>3/3</td>
</tr>
<tr>
<td>Toyota</td>
<td>3/3</td>
<td>0/3</td>
</tr>
<tr>
<td>UMC</td>
<td>0/3</td>
<td>3/3</td>
</tr>
<tr>
<td>Walmart</td>
<td>2/3</td>
<td>1/3</td>
</tr>
<tr>
<td>warning</td>
<td>0/3</td>
<td>3/3</td>
</tr>
<tr>
<td>Wings Over Broadway</td>
<td>0/3</td>
<td>3/3</td>
</tr>
</tbody>
</table>

*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 well 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.

3.3 Phonology Integration by Establishedness

Another way to explore the relationship between phonological integration and lemma, is through the establishedness of the lemma. Chart 2 shows the results of percent of phonological integration (Spanish phonology) of LOLIs in the tree separate establishedness categories (well-established, regional, and unattested).

Figure 2: Phonological integration (Spanish phonology) of LOLIs in Arizona Spanish by establishedness of the LOLI lemma.
As seen in Chart 2, 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 untested 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; Poplack, Sankoff and Miller, 1988; Poplack, 2012).

3.4 Correlation between Phonological and Morphosyntactic Integration

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 Torres Cacoullos and Aaron, 2003a). Table 2 shows the phonological integration and morphological integration for both ‘Walmart’ and ‘high school’.

<table>
<thead>
<tr>
<th></th>
<th>Phonology</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Spanish</td>
<td>English</td>
</tr>
<tr>
<td>Morphosyntax</td>
<td>Bare</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Determiner</td>
<td>19</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Phonology</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Spanish</td>
<td>English</td>
</tr>
<tr>
<td>Morphosyntax</td>
<td>Bare</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Determiner</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 2: Phonological and morphosyntactic integration of ‘Walmart’ and ‘high school’.

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 one 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 subsequent publications of this working paper.

4 Conclusion

In this paper we examined the phonological integration of LOLNs in Southern Arizona Spanish. We determined that LOLNs are overwhelmingly produced with either Spanish or English phonology (99%) and not produced with mixed phonology (only 1%). In agreement with Poplack and Sankoff (1984) and Poplack, Sankoff and Miller (1988), we did find variation at the lemma level. Well-established lemmas are the most phonologically integrated (71%), followed by regional (41%), and least integrated are untested LOLNs (3%). Of the high frequency lemmas, some are fully integrated, some never integrated, and others are partially integrated. When comparing this phonological integration to the morphological integration, our preliminary results suggested that the two correlate. For both ‘high school’ and ‘Walmart’, LOLNs that were phonologically integrated were also morphosyntactically integrated. These results are promising for the argument that phonological integration should be included in analyses of LOLNs.

The next step in this research is to further explore the correlation between phonological and
morphological integration by completing the analysis of the factors that condition determiner realization for all 756 LOLNs. If LOLNs that are phonologically integrated (Spanish phonology) are morphosyntactically conditioned like Spanish nouns and LOLNs that are not phonologically integrated (English phonology) are morphosyntactically conditioned like English nouns, then we can conclude that there is a strong correlation between phonological and morphosyntactic integration. Future studies should consider phonological integration a valid instrument which, when combined with morphosyntactic integration, provides a more detailed measure in determining if LOLNs are borrowings or codeswitches.

References


Carvalho, Ana M. (2012-). Corpus del Español en el Sur de Arizona (CESA). University of Arizona. cesa.arizona.edu.


Department of Writing & Language Studies
University of Texas Rio Grande Valley
Edinburg, TX 78539-2999
bessetr@outlook.com