Undergraduate Research   La Salle Scholar

Spring 2017

Code Mixing in Narratives of Spanish-English Dual Language Learners

Janine Peca
pecaj1@student.lasalle.edu

Follow this and additional works at: http://digitalcommons.lasalle.edu/undergraduateresearch

Part of the Speech and Hearing Science Commons, and the Speech Pathology and Audiology Commons

Recommended Citation

Peca, Janine, "Code Mixing in Narratives of Spanish-English Dual Language Learners" (2017). *Undergraduate Research*. 24.
http://digitalcommons.lasalle.edu/undergraduateresearch/24

Code Mixing in Narratives of Spanish-English Dual Language Learners

Janine Peca A.A.

Dana Bitetti Ph.D., CCC-SLP (Faculty Sponsor)

La Salle University

Abstract

As the U.S. diversifies, the number of Spanish-English dual language learners (DLLs) has grown significantly and is projected to continue to do so through 2060 (Colby & Ortman, 2015). Although research has shown that DLLs implement strategies, like code mixing, to compensate for varying language proficiencies (Vu, Bailey, & Howes, 2010), what elements of code mixing are typical has not been widely researched. This study aims to add to the body of work on typical development of DLL children by examining code mixing in narratives of 212 Spanish-English dual language learners at the beginning of preschool, to determine if differences in language exposure correspond to differences in the types and frequency of code mixing. Close inspection of English and Spanish narratives revealed that the participants preserved the mean length of utterance (MLU) in their code mixes, the majority of which were nouns, and that they code mixed more frequently in their Spanish narratives than their English narratives.

The U.S. population is becoming more diverse. In fact, the Census Bureau projects that between 2014 and 2060 the number of foreign-born residents is expected to grow by 84.7% or 36 million individuals (Colby & Ortman, 2015). Notably, the increase in foreign-born residents is made more significant by the fact that the native population is only expected to grow by 22% or 62 million individuals (Colby & Ortman, 2015). In other words, the rate of growth for foreign-born residents is expected to overtake the rate of growth of native residents by a margin of 62.7%. Consequently, if the Census Bureau's predications are realized, it will result in approximately one in every five U.S. residents being foreign-born by 2060 (Colby & Ortman, 2015). Partially as a result of the rising immigrant population, by 2044 it is expected that the majority of U.S. residents will identify with a minority group (Colby & Ortman, 2015). Overall, America is becoming increasing more diverse.

Out of all the minority groups to which many foreign-born residents identify, the Latino population in particular is projected to undergo a significant growth spurt. Already between 2000 and 2010, the population underwent a 43% growth (Ennis, Rios-Vargus, & Albert, 2011). In terms of percent of the population, the number of for foreign-born individuals identifying as having Latino origin will hold at over 40% through 2060 (Colby & Ortman, 2015). In addition, the Latino population is expected to more than double what it was in 2004 to 119,044,000 in 2060 (Colby & Ortman, 2015). All in all, the dynamic demographics are increasing the diversity in America.

The growth of the Latino population is especially significant in terms of language use. There are over 40 million residents that speak Spanish at home (United States Census Bureau). In the Latino population specifically, over 72.9% of the population uses Spanish

as a home language (United States Census Bureau). In addition to Spanish, 57.4% of the Latino population is reported to hold English fluency (United States Census Bureau). Considering both the growth in foreign-born Latinos and their command of two languages, the U.S. is becoming increasing more linguistically diverse and the preschool population in particular is no exception. Head Start reports that 405,989 (37.8%) of its participants had Latino origin and 266,005 (24.7%) of participants used Spanish as the home language (Administration for Children and Families, 2014). Consistent with the data on the Latino population as a whole, the population of Spanish-English dual language learners (DLLs) in preschools specifically is on the rise.

Research has shown that these dual language learners are likely to implement effective strategies to compensate for varying language proficiencies (Vu, Bailey, & Howes, 2010). To note, for the purposes of this study, a DLL is an individual who is learning the home language of Spanish while gaining experience with English at school. For example, DLLs might alternate between languages for whole utterances do to the sociopragmatic nature of the situation (Vu, Bailey, & Howes, 2010). For instance, a DLL might alternate the language they are speaking based on the language preferences of the listeners they encounter over the course of the day. For the purposes of this study the swapping languages for entire utterances will be hereafter referred to as code switching.

Additionally, and pertinent to the present study, DLLs often swap the language of individual words and/or morphemes within an utterance (Vu, Bailey, & Howes, 2010). This strategy, which will hereafter be referred to as code mixing, occurs twice in the English utterance "A fish y a muchacho," as taken directly from one of the narratives reviewed in the present study. Although research has shown DLLs use code mixing as an

effective strategy, what elements of code mixing are typical has not yet been widely researched. The existing body of research does, however, give credit to the Usage-Based theory of language acquisition in DLLs that suggests these individuals, because of their nature as DLLs, receive less exposure to individual languages (Paradis, Nicoladis, Crago, & Genesee, 2011). According to the theory and its supporting research, it can be expected that DLLs make fewer code mixes in their language of greater exposure as they have greater knowledge of that particular language (Paradis et al., 2011). To build off exisiting research and fill in knowledge gaps, this study examines code mixing in narratives of Spanish-English dual language learners at the beginning of preschool, to determine if differences in language exposure correspond to differences in the types and frequency of code mixing.

Methods

Participants

As part of a larger longitudinal language, self-regulation, and literacy project, children were recruited from Head Start programs and school district early childhood centers that serve children from low socioeconomic backgrounds in New York and Florida (Hammer, 2009). About 95% of the children were born in the mainland United States. To be included in the larger study, participants had to meet the following criteria:

1. Have at least one parent of Latino decent.

2. Be exposed to Spanish by a family member.

3. Be exposure to English prior to entering the study.

4. Have no parent or teacher concerns about development.

To participate in the present study, participants had to have a usable narrative in both English and Spanish as defined by three qualifications:

1. The child had to produce a minimum of 3 utterances in the target language.

2. The child had to produce more than half of their narrative in the target language.

3. The child needed to have language exposure and usage information on the consent from.

Of all the participants for whom the larger study collected data, 212 children met the above criteria to be included in the present study. Children ranged in age from 37 to 65 months (M = 54.56, SD = 4.36). Information on home language exposure and usage was obtained from the mother via a written consent form. Participants fell into three groups reflecting their language exposure:

Language Mother Speaks to Child		
Reported Language Exposure Group	Frequency	Percent
Mostly to All Spanish	118	55.7%
Mostly Spanish to Some English	66	31.1%
Equal Amounts of Spanish and English	28	13.2%
Total	212	100.0%

A subset of 10 participants was randomly selected from each language exposure group to create a sample set of 30 participants.

Procedures

In the larger study, one English and one Spanish narrative sample was elicited from each participant. Narratives were elicited using age-appropriate wordless picture books. *A Boy, a Dog, a Frog, and a Friend* was used to obtain the English narrative sample and *One Frog too Many* was used to obtain the Spanish narrative sample (Mayer & Mayer, 1971;

Mayer, 1975). Both books, written by Mercer Mayer, were the same length, contained the same main characters, and an early conflict from which the book's events unfolded. Narratives were elicited about one week apart and the order of the language of elicitation was counterbalanced. All narrative samples were elicited by trained data collectors who were only allowed to restate the participant's utterances and ask open-ended questions. Data collectors were fluent in the language of the narrative sample they elicited and all narrative samples were recorded for later transcription.

All narratives were transcribed using the Spanish-English version of the Systematic Analysis of Language Transcripts (SALT) (Miller & Iglesias, 2016). Transcription was performed by graduate and undergraduate students who were fluent in the language of the narrative samples they coded. All of the transcribers were trained and supervised by two speech-language pathologists at the doctoral level. Transcripts were checked three times for reliability:

1. All transcripts were checked by a second trained transcriber.

2. All transcripts were checked by one of the supervising speech-language pathologists at the doctoral level.

3. All transcripts were checked by the speech-language pathologist principal investigator.

Additionally, 20% of the narrative samples were chosen at random and re-transcribed a second time by a transcriber who had not already assisted with that narrative sample's transcription. Reliability was 82% for English transcripts and 80% for Spanish transcripts at the word level.

Specifically relevant to the present study, transcribers classified instances of code mixing into two types based on their relation to the mean length of utterance (MLU). MLU is a measure of linguistic complexity in which the total number of morphemes is divided by the total number of utterances. Code mixes that did not preserve MLU (changed level of linguistic complexity) were coded as mixed code mixes. Code mixes that preserved MLU (maintained the same level of linguistic complexity), were coded as mixed lexical code mixed.
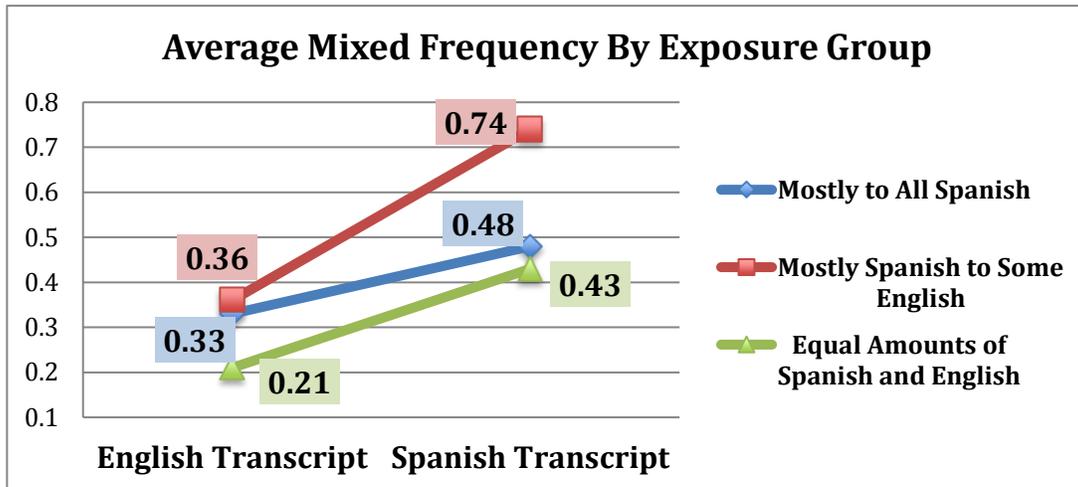
Data Analysis

The Statistical Package for Social Sciences software program (SPSS) Version 24, 2016 for Mac was used to complete statistical analyses (IBM Corp., 2015). To determine if differences in language exposure resulted in differences in code mixing frequency and frequency by transcript language, two separate 3 (language exposure group) x 2 (language of transcript) mixed ANOVAs were completed with mixed and mixed lexical frequencies as the dependent variables inclusive of Spanish and English transcripts of all 212 participants. Additionally, the Spanish and English transcripts from the randomly selected sample set of 10 participants from each language exposure group were analyzed for type of code mixing and part of speech code mixed to determine which of each occurred most frequently.

<center>Results</center>
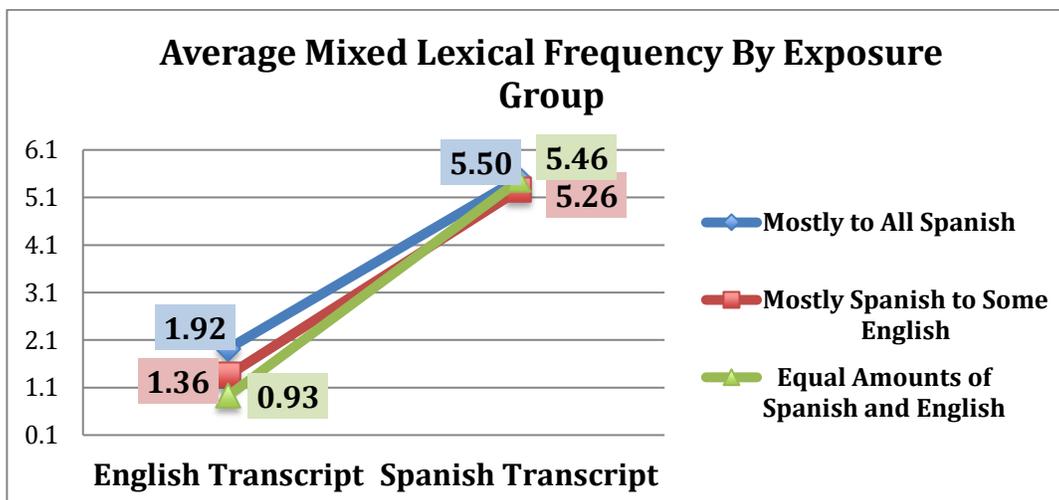
Frequency of Code Mixing (Mixed Type)

After statistical analysis, the language exposure group proved not to be significant as the three groups did not differ significantly in mixed code mix frequency in English and

Spanish (*p* = 0.669). Language exposure groups also did not differ in the frequency of mixed code mixes by language of transcripts. In other words, participants from all language exposure groups had similar English and Spanish mixed code mixing frequency.

Average Mixed Frequency By Exposure Group

- Mostly to All Spanish: English Transcript 0.33, Spanish Transcript 0.48
- Mostly Spanish to Some English: English Transcript 0.36, Spanish Transcript 0.74
- Equal Amounts of Spanish and English: English Transcript 0.21, Spanish Transcript 0.43

Frequency of Code Mixing (Mixed Lexical Type)

After statistical analysis, the language exposure group proved not to be statistically significant as the three groups did not differ significantly in mixed lexical code mix frequency in English and Spanish (*p* = 0.839). Participants from all language exposure groups, however, made significantly more mixed lexical code mixes in Spanish transcripts (*M* = 5.41) than in English transcripts (*M* = 1.41) (*p* < 0.001).

Average Mixed Lexical Frequency By Exposure Group

- Mostly to All Spanish: English Transcript 1.92, Spanish Transcript 5.50
- Mostly Spanish to Some English: English Transcript 1.36, Spanish Transcript 5.26
- Equal Amounts of Spanish and English: English Transcript 0.93, Spanish Transcript 5.46

Frequency of Code Mixing (Sample Set)

Of the 60 transcripts from the sample set that were reviewed, 177 utterances contained instances of code mixing. There were 14 instances of the mixed code mix type and 163 instances of the mixed lexical code mix type. Of the 163 instances of mixed lexical code mixes, 128 were from the Spanish narrative transcripts, which is consistent with the findings of the larger dataset. It is important to note, however, that 20 of the 60 transcripts reviewed did not contain any code mixes of either type.   Interestingly, 18 of these transcripts were of English narratives and only 2 were of Spanish narratives.

Total Number of Utterances Containing Code Mixes							
	Mostly to All Spanish		Mostly Spanish to Some English		Equal Amounts		
Type of Code Mix	English	Spanish	English	Spanish	English	Spanish	Totals
Mixed	4	3	1	2	0	4	14
Mixed Lexical	25	36	8	36	2	56	163
No Mixes	4	1	6	1	8	0	20

Part of Speech of Code Mixing (Sample Set)

In the 60 sample set transcripts reviewed there were 240 code mixed words. Of these 62 code mixes were identified in English transcripts and 178 mode mixes were identified in Spanish transcripts. Participants from all language exposure groups code mixed nouns in both language transcripts more frequently than any other part of speech. In fact, nouns accounted form more code mixes than all other part of speech combined.

Parts of Speech Code Mixed							
Part of Speech	Mostly to All Spanish		Mostly Spanish to Some English		Equal Amounts		Totals
	English	Spanish	English	Spanish	English	Spanish	
Noun	12	39	1	28	1	42	123
Conjunction	15	1	0	4	0	11	31
Article	6	4	9	2	1	5	27
Verb	6	3	1	2	0	7	19
Adverb	3	0	0	2	0	8	13
Adjective	1	4	0	4	0	4	13
Pronoun	3	1	1	1	0	3	9
Preposition	2	0	0	0	0	2	4
Interjection	0	1	0	0	0	0	1

Discussion

In terms of the frequency of code mixes by code mix type, the averages of code mix type across all language exposure groups demonstrate a clear preponderance of the mixed lexical code mixes over the mixed code mixes. The prevalence of utterances containing mixed lexical code mixes could have resulted from the fact that MLU is preserved. It is possible that DLLs attempt to maintain the linguistic complexity of an utterance when they code mix. The amount of mixed lexical utterances, however, could also be a coincidence, as the preschool-aged participants are not expected to have great linguistic complexity.

Overall, there was a greater abundancy of mixes lexical code mixes in Spanish narrative transcripts than in English narrative transcripts. This finding may be due, in part, to the fact that participants are more familiar with telling narratives in the predominantly English school environment. So, although the majority of the participants have a home environment in which Spanish is the primary language of exposure, the children still associate some aspect of the preschool environment or act of sharing a narrative with the

English language and therefore code mix English words into their Spanish narratives frequently.

In terms of the part of speech most often code mixed, nouns were found to be code mixed in the overwhelming majority of cases. The pervasiveness of nouns could be due in part to the preschool age of the participants. This hypothesis is supported by the fact that more complex parts of speech, like prepositions and modifiers, were not as commonly code mixed. It is important to note also that the researcher noticed many of these nouns consisted of content words from the pictures books, like "froggie," which could account for the large amount of nouns code mixed.

Clinically speaking, based on the results of the present study, speech-language pathologists and other professionals working in the early childhood environment can expect their DLLs to preserve MLU when code mixing. Additionally, professionals can expect to see their DLLs code mix nouns more frequently than any other part of speech.

Future Directions

Analysis of the transcripts of all 212 participants would lead to more reliable results than the sample size of 30 participants used in the present study. Additionally, comparative case studied might reveal if some DLLS tend to favor certain parts of speech for their code mixing. This might reveal if code mixing is an idiosyncratic phenomenon or if it does indeed follow a predictable pattern. To see if code mixing frequency changed over time, the Fall narrative samples used in the present study should be compared with the Spring narrative samples collected as part of the larger study.

Acknowledgements

For more information contact Janine Peca:

pecaj1@student.lasalle.edu

References

Administration for Children and Families. (2014). *Office of Head Start – Services Snapshot.*

[PDF file]. Retrieved from

https://eclkc.ohs.acf.hhs.gov/hslc/data/psr/2014/NATIONAL_SNAPSHOT_ALL_PR

OGRAMS.pdf

Colby, S. L., & Ortman J. M. (2015, March). *Projection of the size and composition of the U.S.*

Population: 2014-2060 [PDF file]. Retrieved from

http://www.census.gov/content/dam/Census/library/publications/2015/demo/p

25-1143.pdf?

Ennis, S. R., Rios-Vargas, M., & Albert, N.G. (2011, May). *The Hispanic Population: 2010* [PDF

file.].   Retrieved   from   https://www.census.gov/prod/cen2010/briefs/c2010br-

04.pdf

Hammer, C.S. (2009) Tools of the Mind: Promoting ELLs' Language, Self-Regulation &

School-Readiness (1U01HD060296-01).

IBM Corp. (2016). IBM SPSS Statistics for Mac, Version 24.0. Armonk, NY: IBM Corp.

Mayer, M., & Mayer, M. (1971). *A boy, a dog, a frog and a friend*. New York, NY: Dial Books

for Young Readers.

Mayer, M. (1975). *One frog too many*. New York, NY: Puffin.

Miller, J., & Iglesias, A. (2016). Systematic Analysis of Language Transcripts (SALT)

Research Version 2016. Middleton, WI: SALT Software LLC.

Paradis, J., Nicoladis, E., Crago, M., & Genesee F. (2011). Bilingual children's acquisition of

the past tense: a usage-based approach. *Journal of Child Language*, 38, 554-578

doi:10.1017/ S0305000910000218

United States Census Bureau. (2016, October). *Profile America facts for features: Hispanic heritage month 2016* [PDF file]. Retrieved from http://www.census.gov/content/dam/Census/newsroom/facts-for-features/2016/cb16-ff16.pdf

Vu, J. A., Bailey, A. L. & Howes C. (2010). Early cases of code-switching in Mexican-heritage children: Linguistic and sociopragmatic considerations. *Bilingual Research Journal: The Journal of the National Association for Bilingual Education, 33,* 200-219. doi:10.1080/15235882.2010.502798