

Utah State University

DigitalCommons@USU

All Graduate Plan B and other Reports

Graduate Studies

5-2014

Learning Two Languages: Maze Behaviors in Narrative Discourse for Spanish-English Bilinguals

Joel Hopper
Utah State University

Follow this and additional works at: <https://digitalcommons.usu.edu/gradreports>



Part of the [Speech and Rhetorical Studies Commons](#)

Recommended Citation

Hopper, Joel, "Learning Two Languages: Maze Behaviors in Narrative Discourse for Spanish-English Bilinguals" (2014). *All Graduate Plan B and other Reports*. 440.

<https://digitalcommons.usu.edu/gradreports/440>

This Thesis is brought to you for free and open access by the Graduate Studies at DigitalCommons@USU. It has been accepted for inclusion in All Graduate Plan B and other Reports by an authorized administrator of DigitalCommons@USU. For more information, please contact digitalcommons@usu.edu.



Learning Two Languages: Maze Behaviors in Narrative Discourse for Spanish-English
Bilinguals

By

Joel Hopper

A publication submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Speech-Language Pathology

Approved:

Sandra Gillam Ph.D., CCC-SLP
Major Professor

Jamison Fargo Ph.D.
Committee Member

Ron Gillam Ph.D., CCC-SLP
Committee Member

UTAH STATE UNIVERSITY
Logan, Utah

2014

Copyright © Joel Hopper, 2014

All Rights Reserved

ABSTRACT

Learning Two Languages: Maze Behaviors in Narrative Discourse for Spanish-English bilinguals

by

Joel Hopper

Utah State University, 2014

Major Professor: Dr. Sandra Gillam
Department: Communicative Disorders and Deaf Education

The purpose of this current study was to examine the differences in mazing behavior longitudinally between English and Spanish narratives in school-age bilingual children learning English. Narrative retells in English and Spanish from 216 children between the ages of 5 and 9 were analyzed for average number of mazes, percent mazes, NSS scores, and specific mazing behaviors in simple and complex utterances using Systematic Analysis of Language Transcription (SALT) software. Data for students who produced at least one complex sentence at the beginning and ending of each school year (K, 1st, 2nd) were analyzed in this study. Results obtained from a Rectangular Data Analysis were used in multilevel modeling to determine any changes over time, and a post-hoc analysis was conducted to determine any further changes in regards to the specific maze behaviors of repetitions, revisions, and filled pauses. Results indicated a statistically significant difference (main effect) in average maze scores between sentence types, with higher scores for complex as compared to simple sentence types. Additionally, a statistically significant interaction between wave (time) and language indicated that the trend over time for average maze scores significantly differed between English and Spanish, with English maze scores remaining level over time and Spanish maze scores

decreasing slightly. These results present implications for explaining the causes of bilingual disadvantage and discrepancies of maze behaviors in English and Spanish.

PUBLIC ABSTRACT

Learning Two Languages: A longitudinal investigation of discourse skills for Spanish-English bilinguals

This study was designed to determine whether mazing behavior in narrative retells of bilingual English and Spanish speaking children was affected by language, complexity of sentence structure, and language development over time. Language transcripts were analyzed from English and Spanish narrative retells elicited from 216 children between the ages of 5 and 9 at the beginning and end of kindergarten, first, and second grade years. Findings revealed a statistically significant difference in the average maze scores and sentence types. A significant interaction was also found between time and language, with maze scores remaining nearly level over time in English and decreasing slightly in Spanish.

Joel Hopper

CONTENTS

ABSTRACT.....	iii
ACKNOWLEDGEMENTS.....	v
LIST OF TABLES.....	viii
INTRODUCTION.....	1
METHODOLOGY.....	9
Participants.....	9
General Procedures.....	12
Reliability.....	15
Data Analysis.....	16
RESULTS.....	16
DISCUSSION.....	17
Limitations.....	21
Summary and Implications.....	22
REFERENCES.....	24
APPENDICES.....	29
Appendix A Demographic Variables.....	30
Appendix B Average Maze Scores	32
Appendix C Multilevel Modeling Results.....	37
Appendix D Narrative Scoring Scheme Results.....	38

ACKNOWLEDGEMENTS

I would like to thank Dr. Sandra Gillam for her support, assistance, and encouragement throughout the course of this project and my graduate studies. I extend my heartfelt gratitude to Nate Salisbury, Woodrow Laing, Sam Stone, and Ari Watt for their assistance in separating transcripts. I would also like to thank Sarah Clement for all of her assistance in completing this project. I would also like to thank Dr. Ronald Gillam and Jamison Fargo for their dedication to making this project a success. A special thank you goes to Aquiles Iglesias for sharing this data with us.

Ultimamente, me gustaría agradecer a mi familia y a Dios para todo el apoyo que he necesitado para cumplir este gran proyecto. Los quiero muchísimo.

Joel Hopper

LIST OF TABLES

Table		Page
1	Demographic Variables.....	30
2	Participant Data Files.....	31
3	Average Maze Scores per Utterance	32
4	Percent Mazed Words per Transcript.....	36
5	Multilevel Model Results	37
4	Narrative Scoring Scheme Results	38

Introduction

The number of school-age children (ages 5–17) who speak a language other than English at home increased from 9 to 20% between the years 1979 and 2005 (National Center for Education Statistics, 2004). 76.9% or 2,963,256 students were Spanish speakers. Well over 80% of schools in the United States are currently serving English Language Learning (ELL) children who speak Spanish as their primary language (National Clearinghouse for English Language for Acquisition, 2007). The National Literacy Panel on Language-Minority Children and Youth reported that the nature of the relationship between English oral language proficiency and reading comprehension is of crucial concern for ELL children (August & Shanahan, 2006). One area of oral language ability that is particularly problematic for ELL children is narration (August, Carlo, Dressler, & Snow, 2005).

There is a strong relationship between narrative proficiency and academic performance (Wellman, Lewis, Freebairn, Avrich, Hansen, & Stein, 2011). For example, narrative skill has been shown to predict reading and academic success (Griffin, Hemphill, Camp, & Wolf, 2004) and to differentiate between students performing within the average or below average range in terms of language ability (Bishop & Donlan, 2005). Deficiencies in narrative skills has been reported in a number of child populations including those with learning disabilities (Fletcher, Lyon, Fuchs, & Barnes, 2006, p. 244), hearing loss (Soares, Goulart, & Chiari, 2010), intellectual deficits (Roberts, Martin, Moskowitz & Harris, 2007) and autism (Eigsti, de Marchena, Schuh, & Kelley, 2001).

Narrative language proficiency is both related to and predictive of difficulties in academic achievement for monolingual children developing typically and those with language learning impairments (Bishop & Edmundson, 1987; Fazio, Naremore, & Connell, 1996). Recent

research suggests that narrative language proficiency is also related to reading comprehension in Spanish-speaking bilingual students (Miller, Heilmann, Nockerts, Iglesias, Fabiano, & Francis, 2006). Miller et al., (2006) studied whether traditional measures of oral language proficiency (vocabulary, syntax, fluency, and discourse) elicited using narration would predict reading achievement within and across languages for children who were bilingual. The narrative retells of over 1,000 Hispanic/Latino Spanish-speaking English language learners were elicited using a *Frog* story narrative at 6 time points (pre K, post K, pre 1st, post 1st, pre 2nd, post 2nd) in both English and Spanish to measure language proficiency (in both languages) over time. In addition, the study was conducted to evaluate if particular oral language features were associated with reading proficiency skills in English and Spanish. Results indicated English oral language measures predicted Spanish reading scores and Spanish oral language measures predicted English reading scores. Thus, narrative proficiency is an important language skill for children learning two languages to master.

There have been a limited number of studies that have examined narrative development in ELL children (Gutiérrez- Clellen, 2002; Muñoz, Gillam, Peña, & Gulley-Faehnle, 2003; Uccelli & Paez, 2007), and there have been even fewer that have examined factors related to increasing narrative proficiency over time. One of these factors is maze behaviors. Mazes were first described by Loban (1976) and referred to as "...a series of words (or initial parts of words), or unattached fragments which do not constitute a communication unit and are not necessary to the communication unit" (p. 22). In other words, mazes do not contribute meaning to the ongoing flow of language. Levelt (1989) categorized mazes as revisions, filled pauses, or repetitions that occur as a result of uncertainty. According to Levelt (1989) production of mazes reflect the speaker's reaction to the demands of language. Many studies provide varying information concerning developmental change in children's maze behavior while focusing on using mazing

behavior as an indicator of language learning difficulty (Bedore, Fiestas, Pena & Nagy, 2006).

Research has indicated that increased maze use may be reflective of language learning difficulty (Levelt, 1989; Levelt, 1999) and may often be considered a red flag for language impairment. (Loban, 1963; Nippold, 1993; Bedore et al., 2006). It has been noted that bilingual speakers seem to produce more mazes in their second language than in their first (Lennon, 1990; Poulisse, 1999; Rieger, 2003; Wiese, 1984). Research has also suggested that a disadvantage in language production exists for those who speak multiple languages (Sandoval, Gollan, Ferreira, & Salmon, 2010). Some studies have provided support for the notion of a bilingual disadvantage by analyzing speech disfluencies (Gollan & Acenas, 2004; Gollan, Montoya, & Bonanni, 2005; Gollan, Montoya, & Werner, 2002; Gollan & Silverberg, 2001). Other studies have indicated little to no differences in verbal or speech fluency between bilingual and monolingual speakers (Bedore et al., 2006). For example, Bedore et al. (2006) examined language status and mazes produced during story telling for 66, 4-7 year olds, some were bilingual English and others bilingual Spanish. Results suggested that maze types were similar among all of the children whether speaking in English or Spanish. Findings from this study also suggested that after a year or so of school, there was no indication that bilinguals were experiencing more difficulty in their language acquisition than their functionally monolingual peers (Collier, 1989; Hakuta, 2000; Jacobsen and Schwartz, 2005).

The purpose of the current study was to examine potential changes in maze behaviors occurring in narrative discourse over a 3-year period for 216 Spanish-English Bilingual children. In this study, narratives were elicited at the beginning and the end of the school year in Spanish and again in English (2 weeks later). This study adds to the literature by allowing maze behavior to be examined in cohorts of children as they were becoming more proficient in their narrative skills in both languages.

Literature Review

It has been proposed that learning and using two languages simultaneously may pose a heavier cognitive load (ie., cognitive load hypothesis) on learners than acquiring only one language (Edmunds, 2006; Gutiérrez-Clellen & Kreiter, 2003; Silva-Corvalan, 1994). In a study of two variables (choice of copula and verb form used in apodoses) among speakers of Mexican-American Spanish, Gutiérrez-Clellen and Kreiter (2003) reported evidence of accelerated language change leading to a more simplified linguistic system. This simplification was equated with a speaker strategy aimed at lightening the cognitive load experienced by those speakers, who must manage two languages.

A study by Paul Edmunds (2006) found nine examples of repair (speech disfluencies) preceding English-origin words in New Mexican Spanish. In their study, responses from six participants in the New Mexico-Colorado Spanish Survey (NMCOS; Bills & Vigil, 1999) were coded and analyzed for repairs (revisions or mazes) of content and function words and instances of English-origin word use occurring immediately after an instance of repair of a speech disfluency. Edmunds (2006) concluded that the semantic load, the length of words, and the level of grammaticization appeared to have an effect on the rate and type of repair, and that these types of disfluencies most likely resulted from speaker attempts to “ease the cognitive load of expression” (p. 212). It is possible that bilingual speakers may experience a heavier cognitive load as they must manage progressively more complex language structures over time. If so, they may experience a higher rate of repair, and/or increasing rates of repair over time.

A cohort study was conducted by Sandoval, Gollan, Ferreira, & Salmon (2010) to determine the possible causes of a bilingual disadvantage seen in bilingual speakers. This study explored the theories of the dual-task analogy, the weaker links account, and the category size analogy as three possible models for explaining the bilingual disadvantage. The theory of

retrieval slowing with interference between languages, or the dual-task analogy, assumes that bilingual speakers have to manage interference from their non-target language when using their target language. Retrieval slowing without interference, or the weaker links account assumes that bilingual disadvantages are a direct result of the frequency of language use. The reduced vocabulary hypothesis, or the category size analogy, assumes that bilingual speakers also have a smaller vocabulary in either language, but a larger vocabulary overall (Bialystok, Craik, & Luk, 2008). Using the mean response latency as a system of measurement, two different experiments were implemented to determine which of the three models outlined is correct in predicting the bilingual disadvantage.

Experiment 1 was designed to test these three models by manipulating the type of task and the semantic category size. This experiment consisted of a total of 24 participants and compared Spanish-English bilinguals to English-speaking monolinguals. Experiment 2 tested 45 English dominant bilingual speakers to examine the interference model on non-dominant language production. The participants were tested in 60-second trials, and word retrieval times were recorded. Half of the twelve semantic categories were completed in each language. Experiment 1 supported the dual-task analogy of the bilingual disadvantage. This was verified by the observation of an impaired production of high-frequency words as compared to low-frequency words, since high-frequency words have more possible interference targets from the second language. Experiment 2 indicated more of an interference effect on non-dominant language production, suggesting that the dominant language in bilinguals is not affected by the non-dominant language. The overall results of this study confirmed that bilinguals produced fewer correct responses than monolinguals, and that bilinguals produced fewer correct responses in their non-dominant language. These findings were consistent with previous studies that have indicated the same effects (Gollan, Montoya, Cera, & Sandoval 2008; Ivanova and Costa, 2008).

Studies of Mazes in children

Studies of maze production suggest that excessive maze behavior may be an indicator of a language-learning problem (Thordardottir & Weismer, 2002). This notion was tested in a cohort study conducted by Elin T. Thordardottir and Susan Ellis Weismer that compared the frequency of filled pauses and content mazes in narratives of school-age children with NL (normal language) and with SLI (specific language impairment). The participants ranged from ages 5:5 to 9:8 including 50 children with SLI and 50 children with NL matched by age with subgroups of 25 children with NL and 25 children with SLI matched on MLU. Thordardottir and Weismer (2002) used narrative language samples that were previously collected as part of previous studies. These samples were transcribed and coded, and content mazes and filled pauses were separated using SALT conventions (Miller & Chapman, 1993). Their maze measures included: the number of mazes, the number of filled pauses, and the number of content mazes.

Thordardottir and Weismer (2002) reported that the interaction for maze type and group interaction, for age-matched, was significant, and content mazes were more frequent for both groups than filled pauses. However, the difference between maze types was much larger and more statistically significant for the children with SLI than the children with NL. A significant difference was also noted in the main effect of maze type, suggesting that the children with SLI used fewer mazes overall than the children with NL.

In the subgroups that matched the children according to MLU it was reported that the children with SLI had a significantly large difference between the two types of mazes, whereas the children with NL did not. Thordardottir and Weismer (2002) also suggested that the “content mazes were overall more frequent than filler mazes,” (p. 590) since the main effect of maze type was significant. It was expected that the children with SLI would produce more mazes than their MLU-matched peers. In this regard only content mazes were produced in higher frequency.

Additionally, this study suggested that the two type of mazes, filled pauses and content mazes, were not affected by the same factors, since the children with SLI used fewer filled pauses than children with NL in both matched groups.

Bedore, Fiestas, Peña, & Nagy (2006), as mentioned before, examined the relationship between maze behaviors and bilingual status and characteristics of the bilingual's language. In their study, 66 participants ranging from 4;3 to 7;3 years in age (mean age 5;9) were selected from a pool of Mexican-American children already participating in a study of semantic development. These children were recruited from school districts in central Texas and consisted of typically developing bilingual children that were paired with functionally monolingual Spanish-speaking and English-speaking children closely matched the bilingual child's age. The classification for being bilingual was a minimum of 20% output in both Spanish and English as well as additional ratings provided by teachers and the Idea Proficiency Test (IPT-1; Dalton, Amori, Ballard and Tighe, 1991). This study addressed the following questions: 1) Is the frequency of mazes higher in bilingual children than their monolingual peers? 2) Do patterns of maze use differ as a function of bilingual status? 3) Do patterns of maze use differ in Spanish and English? 4) Do patterns of maze use relate to indices of productivity in Spanish and English? (Bedore et al., 2006).

Two narratives, one in English and one in Spanish, were randomly elicited on separate days within a four-week time window from the bilingual English and bilingual Spanish groups using one of the four wordless picture books by Mercer Mayer and transcribed using SALT. Spanish and English narratives were elicited from the FME and FMS groups, according to their primary language. The maze content for filled pauses, repetitions, connectors, grammatical revisions, lexical revisions, and phonological revisions were coded in the narrative transcriptions

according to SALT conventions.

Results from the authors' analysis indicated that the functionally monolingual groups produced narratives of comparable length to those of the bilingual group. The MLU of the bilingual children did not differ from those of the children in the monolingual groups, and the number of different words, or NDW, was also comparable across groups. Moreover, the monolingual children produced more grammatical utterances than the bilingual children (grammatical utterances included code switched elements). According to Bedore et al. (2006), no significant difference was noted between the percentage of utterances with mazes used by the bilingual and functionally monolingual group.

The patterns of maze types were reported as similar in both groups with a general non-significant trend of higher maze use in bilingual children in most instances. Additional data was obtained by combining the English and Spanish groups and comparing the types of patterns used by language. Overall, the authors only found grammatical revisions to differ in Spanish and English, and significantly more grammatical revisions were produced in Spanish with monolingual and bilingual children. An additional analysis to classify grammatical revisions by type was conducted, such as distinguishing revisions to noun or verb morphology. This analysis reported similar kinds of elements were revised in English and Spanish, with some variation to the characteristics of the revised element.

In the article by Bedore et al. (2006) several important results were highlighted including the suggestion that different representations of language do not cause any negative impact in communicative outcomes, as shown by the lack of difference between the monolingual and bilingual groups. This evidence converges with other studies suggesting children develop communicative competence within a year or two of exposure to a second language in a school

environment and that children rapidly develop productive morphosyntactic skills in their second language (Collier, 1989; Hakuta, 2000; Jacobsen and Schwartz, 2005).

The present study was designed to extend the research of Bedore et al. (2006) by examining maze behaviors occurring in narrative discourse over a 3-year period for 3 cross-sectional samples obtained from a pool of 216 Spanish-English Bilingual children.

Research Hypotheses

It was predicted that children would produce more mazes 1) in English than Spanish, and 2) in complex versus simple sentence structures having less proficiency in English and in complex sentence use. It was hypothesized that the average mazes per utterance would increase over time proportional to the general complexity of language being used (Edmunds, 2006), or would remain stable as is consistent in monolingual English speaking children (Evans, 1985; Leadholm & Miller 1995; Loban, 1976). The types of mazes made in English and Spanish over time were examined as part of a descriptive analysis.

Methodology

Participants

This study employed a longitudinal, cohort study design with Spanish-English bilingual children ranging in age from 5 to 9. Narrative skills were measured in English and in Spanish before and after Kindergarten, 1st and 2nd grades by asking them to retell one of four Frog Stories: *Frog, Where are you?* (Mayer, 1969); *Frog on his own* (Mayer, 1973); *Frog goes to dinner* (Mayer, 1974); and *One frog too many* (Mayer, 1975). Participants were drawn from a pool of 1,723 children who participated in a larger study [the Biological and Behavioral Variation in the Language Development of Spanish-Speaking Children (BBVLDSC; R305U01001, Francis, Carlson, Fletcher, Foorman, Goldenberg, Vaughn, Miller, Iglesias, & Papanicolaou, 2005), awarded by the U.S. Department of Education's Development of English Literacy in Spanish-

Speaking Children Research Program and the Institute of Education Sciences in 2002. This multi-level study was designed to identify factors and conditions that contribute to the development of language and literacy skills of Spanish-speaking children learning English. The BBVLDSC project collected both cross-sectional and longitudinal data over a five-year period.

Larger Study Sampling Strategies

The sampling procedures for the larger study from which the data for this study were drawn were instituted in an effort to represent educational experiences for the majority of Spanish-speaking children in the United States. First, schools were selected for inclusion in the larger study if 40% or more of the school's population was Latino. Second, no less than 30% of the Kindergarten students at the school were identified to be limited in their English language skills. Third, the students in selected schools were required to be performing in the average range on major state assessments. Fourth, the schools that participated in the project needed to be actively engaged in providing educational programs to improve language skills for bilingual students (e.g., Structured English Immersion, Early and Late Transitional Bilingual Education, and Dual-Language and Maintenance Programs; Branum-Martin, Foorman, Francis, & Mehta, 2010; Francis et al., 2005; Vaughn, 2005).

General Procedures for the larger study

The larger study was designed to provide data for a measurement grant, a skill development study, a classroom language and instruction project, a family, community and schools project and finally, an intervention project. The data for this study came from the skill development project that was designed to examine English and Spanish language and literacy acquisition over the course of three years (K-3rd grade).

Narrative language samples were elicited from all of the students using wordless picture

books by Mercer Mayer and included *Frog, Where are You?* (Mayer, 1969); (b) *Frog Goes to Dinner* (Mayer, 1974); (c) *Frog on His Own* (Mayer & Mayer 1975a); and (d) *One Frog Too Many* (Mayer, 1975b). The stories were selected because they have a longstanding reputation in the literature for being an appropriate method for obtaining narrative language samples from children from many different culturally and linguistically diverse backgrounds (Heilmann, Miller, Nockerts, & Dunaway, 2010; Massey, & Gleitman, 2002; Pavlenko, 2009).

The examiner and child sat across from each other during the narrative elicitation task and the examiner presented the story orally to the child, while the participant looked at the pictures in the book. The examiner then gave the book to the participant and asked him or her to tell the story again in English and in Spanish. The English and Spanish retells were elicited with at least one intervening week.

The stories told by the children were transcribed orthographically and then coded using Systematic Analysis of Language Transcription Conventions (SALT; Miller & Iglesias, 2003-2004). The narrative samples were also digitally recorded. All utterances were segmented into Modified Communication Units (SALT; Miller & Iglesias, 2003-2004; Gutiérrez-Clellen and Hofstetter, 1994) to ensure that the differences between Spanish and English syntax could be taken into account, specifically with reference to pronouns.

Current Study

A subset of children from the larger data set ($n = 216$) were asked to tell stories at the beginning and end of Kindergarten, 1st and 2nd grades in English and in Spanish. Of the 216 participants in this sample, 80 were in schools using Structured English Immersion, 120 were receiving either Dual Language or Transitional Bilingual Instruction, and 10 were receiving Maintenance services (Figure A.1). This study included the data for only those children who produced both simple and complex sentences at the beginning of the school year for each of the

3 years (K, 1st and 2nd). Of the 216 participants from the larger data set who were included in the longitudinal sample, there were 110 participants in K, 156 in 1st grade and 184 in 2nd grade that met the following inclusionary criteria: (a) produced narrative retells in both languages and (b) produced at least one complex utterance in both English and Spanish retells at the beginning and end of the year. The total number of participants that met this criteria for every year (across all six time points) was 88. A Subordination Index score (SI) of 2 or higher, as coded in the original transcripts, was used to determine which transcripts contained complex utterances (Loban, 1963). Similarly, Subordination Index scores of 1 were considered simple utterances. Original transcripts at each time point were examined individually by undergraduate students under the direction of graduate level students and a professor to determine which transcripts met the inclusionary criteria for each year.

General Procedures

As previously stated, the transcripts from the larger data set were originally transcribed using the Systematic Analysis of Language Transcription (SALT) program (Miller and Chapman, 1984-2004) and coded for several factors such as: code switching, utterance-level error, word-level error, non-standard English word order, and overgeneralization errors. As part of the current study, the original transcripts that met the inclusion criteria were divided into two separate transcripts that consisted of either “complex” or “simple” utterances in English and Spanish. Therefore, each participant had an original file in English and in Spanish that contained the complete narrative sample, a file in both English and Spanish that contained only simple sentences, and a file in both English and Spanish that contained only complex sentences, equaling a total of 6 files for each participant. Subordination Index scoring (SI) is “...a measure of syntactic complexity which produces a ratio of the total number of clauses to the total number of C-units” (Miller & Chapman, 1984-2004, p. 1). This score was used to code each utterance as

to the number of clauses it contained. A clause is defined as any statement that contains both a subject and a predicate (Miller & Chapman, 1984-2004), and a C-unit is defined as “an independent clause with its modifiers” (Loban 1976). A C-unit includes one main clause with all subordinate clauses attached to it, and cannot be further divided without the disappearance of its essential meaning. A complete guide for determining SI scores is outlined by Loban (1963). “Simple” transcripts contained only utterances that were coded with a subordination index score of 1 (SI-1), and “complex” transcripts contained only utterances with an SI score of 2 or higher. In order to be coded as SI-1, the utterance had to contain only one main clause. In order to receive a SI score of >1, the utterance must contain 2 or more clauses. For example, an utterance with an SI score of 1 may say, “They were both really sad.” An utterance with an SI score of 2 would include an additional clause that may say, “They were both really sad *because they couldn’t find their friend.*” An utterance with an SI score of 3 would include additional clauses and may something like this, “When they were sitting down looking at the menus the frog jumped into a saxophone.”

After the original transcripts were divided into transcripts that contained only simple or complex utterances, a Rectangular Data Analysis was conducted using the SALT program (Miller and Chapman, 1984-2004). A Rectangular Data Analysis, or Rectangular Data File (RDF), is used to organize and categorize language samples according to specific attributes that were originally coded in the transcripts that are of interest. The RDF for the current study included the following parameters: Number of Different Words (NDW), Number of Total Words (NTW), Average Maze Per Utterance, Total Number of Mazes, Percent of Maze Words, Maze Revisions of Part-Words, Words, and Phrases, Maze Repetitions of Part-Words, Words, and Phrases, Maze Filled Pauses of Single Words, Maze Filled Pauses of Multiple Words, and Narrative Scoring Scheme (NSS).

NSS is a method used for scoring narratives that is included in the SALT software program. It incorporates a 0 – 5 point scale ranging from minimal (0) to proficient use (5) of story grammar elements with ratings of 0-1 indicating low or minimal use of an element; 2-3 indicating a medium or emerging use; and 4-5 indicating a high or proficient use of elements. The story grammar elements assessed in NSS consist of; Introduction, Character Development, Mental States, Referencing, Conflict/Resolution, Cohesion, and Conclusion (Miller and Chapman, 1984-2004). These story grammar elements are defined as such: Introduction scores were determined by the inclusion and qualitative depiction of character and setting components. Character Development scores are based on how well characters and their significance are acknowledged throughout the story. Mental States scores were determined by the vocabulary used in conveying the emotions and thoughts of characters as well as the diversity of mental state words. Referencing scores are based on the accurate and consistent use of pronouns, proper names, antecedents, and clarifiers. Conflict/Resolution scores are determined by the inclusion and thorough use of conflicts and resolutions throughout the story. Cohesion scores are given according to the details given to sequencing and transitions between each event. Conclusion scores are based how well the final event is concluded and how well the entire story is wrapped up (Miller and Chapman, 1984-2004).

Additional variables that were included in the RDF were the age of the child at the time of testing (in months), gender (male, female), language (English, Spanish), story (one of 4 Frog Stories), complexity (simple, complex), year (K, 1st, 2nd grade), phase (Fall, Spring) and wave (year + phase). As mentioned earlier, complexity was the type of utterance contained in the transcripts that were cut from the original transcript (simple or complex). Each student had one complete file in English, one file that contained only simple utterances ($SI \leq 1$) in English, and one file that contained only complex utterances ($SI \geq 1$) in English. In addition, each student had

one complete file in Spanish, one file that contained only simple utterances ($SI \leq 1$) in Spanish, and one file that contained only complex utterances ($SI \geq 1$) in Spanish. An illustration of the files that were analyzed for each participant is shown in Table A.1. Wave was determined as a sequential order of the narrative retells: Wave 1) Beginning of Kindergarten year, Wave 2) End of Kindergarten year, Wave 3) Beginning of 1st grade year, Wave 4) End of 1st grade year, Wave 5) Beginning of 2nd grade year, and Wave 6) End of 2nd grade year. The phase also indicated the sequential order of the narrative retells and was labeled as Fall or Spring.

Reliability

Reliability for orthographic transcription use of SALT transcriptions, word, morpheme and MC unit segmentation was >98% as reported in the larger study. Approximately, 40 (½ English, ½ Spanish) of the narratives collected during the larger study and used in the current study were randomly selected for calculating inter-rater reliability for transcript cutting.

Complete transcripts were randomly assigned to research assistant for cutting into simple and complex files. After research assistants cut the files, another research assistant cut the same files and they were compared for accuracy. Discrepancies between any transcript cuts resulted in the division and comparison of all additional transcripts in an assignment set by a different student. The accuracy with which transcripts were originally divided was determined to be greater than 95%.

Data Analysis

Data were compiled using the Rectangular Data Analysis function of the SALT software. Average maze Per Utterance scores were analyzed using multilevel modeling to test for significant changes over time (phase nested within wave) as a function of sentence type (simple vs. complex), language (English vs. Spanish), narrative scores (NSS), gender, and age (in months). Two-way interactions were examined between wave and 1) phase, 2) sentence type, 3)

language, 4) NSS score, 5) gender, and 6) age. Each model had 4 nested levels: 1) sentence type and language nested within phase, 2) phase (Fall, Spring) and NSS nested within wave, 3) wave nested within individuals, and 4) individuals. The analysis allowed intercepts and slopes for wave, phase, sentence type, language, and NSS to vary randomly in the models.

Results

Descriptive statistics for Average Mazes Per Utterance for Waves 1-6 for complex and simple sentences in English and Spanish are shown in Table B.1. Results of the multilevel model showed a statistically significant difference (main effect) in average maze scores between sentence types overall, with higher scores for complex as compared to simple sentence types ($p < .0001$). Additionally, a statistically significant interaction between wave and language ($p < .001$) indicated that the trend over time for average maze scores significantly differed between English and Spanish, with a slight downward trend for Spanish and a more flat trend for English (see Figure B.2 for an illustration). Post hoc analyses (paired t-tests) suggested that, as is shown in Figure B.1, there were significantly higher mazes in English than in Spanish at every time point ($p < .05$), and a slight decrease in mazes during the Fall of 2nd grade in Spanish that was significantly lower than mazes measured during the Fall of Kindergarten and 4th Grades.

No other main effects or interactions were statistically significant in either model. Table C.1 presents the results of the multilevel model. Because of the small range of percentages observed for the maze subtypes, as well as many zero values, individual maze scores were not analyzed.

Discussion

The purpose of this study was to examine potential differences in mazing behavior between English and Spanish narratives in school-age bilingual children learning English. It was hypothesized that the children would demonstrate more maze behaviors 1) in English than in

Spanish and 2) in complex versus simple utterances. It was also predicted that children's use of mazes in both languages may increase or remain stable over time as a function of language proficiency. The results obtained from this study showed higher average mazes per utterance in complex sentence types as compared to simple sentence types. This is consistent with previous findings that implied an increase in the frequency of mazes proportional to the general complexity of language being used (Leadholm & Miller, 1992; Maclachlan & Chapman, 1988). Moreover, the higher maze scores in complex sentence types may be unique to syntactic complexity rather than narrative complexity since the relationship between sentence complexity and narrative complexity was not explored in the current study. Details of significant findings regarding narrative measurements are highlighted later in this discussion.

Results from the multi-level modeling also determined a statistically significant interaction between wave (time) and language, indicating that maze use significantly differed over time between English and Spanish. Maze scores in English displayed more of a flat trend overall, and Spanish maze scores exhibited a slight downward trend through time. These findings were similar to previous findings which indicate higher productions of mazes in L2 (second language) than in L1 (first language) (Lennon, 1990; Poullisse, 1999; Rieger, 2003; Wiese, 1984). Several studies have found that productions of mazes typically remained stable or slightly increased over time in monolingual English speaking children, which is consistent with the flat trend exhibited with maze scores in English (Evans, 1985; Leadholm & Miller 1995; Loban, 1976). In their study of maze productions, Leadholm and Miller (1995) noted that individual variability in the production of mazes occurred when children tried to express complex ideas. The downward trend of maze scores in Spanish contradicts the findings by Evans (1985), Leadholm & Miller (1995), and Loban (1976) since they do not remain constant over time and do not increase with general language complexity in complex or simple sentences.

There are several previous studies that provide possible explanations for the difference in maze scores between English and Spanish as well as the decreasing trend of maze scores in Spanish. Considering the study by Edmunds (2006), it may be possible that less repair strategies (mazes) may be needed as English proficiency increases. A speaker may experience a lightened cognitive load as he or she becomes more proficient and familiar with the English language, resulting in a decrease in maze behaviors in the dominant language, Spanish (Edmunds, 2006). It may be possible that maze scores in Spanish were originally lower before the children were enrolled in kindergarten and increased shortly after they began learning English. The downward trend in maze scores may have simply returned to the original frequency of production after the children became more proficient in English. Evidence for this could be obtained by an analysis of maze scores before enrollment into an English instructional program and continued evaluation past the second grade.

The greater difference in maze scores between simple and complex utterances in English than in Spanish may be explained by reduced language interference. As previously mentioned, Sandoval et al. (2010) indicated that language interference tends to occur more with non-dominant language production than with dominant language production. More interference would have occurred in English since it was the non-dominant language, causing more frequency of mazes. This theory, coupled with findings by Evans (1985), Leadholm & Miller (1995), and Loban (1976) would explain the higher frequency of mazes in English as well the greater difference between sentence types, with complex sentences in English being the most affected by complexity and interference.

Although no additional statistically significant interactions were determined from the multi-level modeling, notable patterns were found throughout other measures such as NSS scores and specific types of maze behaviors. It should be noted that NSS scores could not be accurately

examined for interactions with sentence complexity due to the method used for separating language transcripts into simple and complex components. To clarify, one original transcript was separated into two distinct transcripts containing either simple or complex utterances only. Both transcripts represent the same NSS score, as both simple and complex utterances are used in the original completed narrative. For this reason, a single NSS score determined by the original transcript was attributed to a separate transcript with simple utterances and a separate transcript with complex utterances, and the relationship between sentence complexity and narrative performance could not be examined.

NSS scores were tested for interactions between the other variables in the multi-level modeling including wave, phase, complexity, and language, and no statistically significant interactions were determined. Data showed a general increase of NSS scores over time (Figure D.1), which is consistent with typical developmental patterns and indicates language development (Heilmann et al., 2010). Even though no statistical difference was measured between English and Spanish NSS scores, Spanish NSS scores appeared to be higher than English NSS scores at all time points, especially after the kindergarten year (see Figure D.1). This may be a reflection of the complexity of the narrative, language uncertainty, and/or generally undeveloped English language knowledge and skills as compared to Spanish.

As part of the analysis, an additional post-hoc analysis was conducted to determine the existence of statistically significant trends in the specific types of maze behaviors of 1) Filled Pauses, 2) Part-word, whole-word, and phrase repetitions, and 3) Part-word, whole-word, and phrase revisions. This analysis was originally completed by obtaining the average number of these specific maze types on a per-transcript basis. Results from this analysis indicated that more part-word revisions were produced in English than in Spanish in simple sentences overall. More part word revisions were produced in both languages for complex sentences overall excluding

the last time point when there were more part-word revisions in English than Spanish. Additionally, a general increase in whole-word revisions in simple sentences in English was noted while a general decrease was observed in waves 3 through 6 in Spanish. These results presented contradictory evidence compared to some previous studies and interesting implications when compared to other studies.

Previous studies have indicated that specific maze behaviors may be the result of grammatical differences between the two languages being spoken (Choi, 1997; Ford, Fox, & Thompson, 2003; Rieger, 2003). Bedore et al. (2006) presented findings that were, for the most part, contradictory to results from the post-hoc analysis of the current study. Bedore et al. (2006) reported similar patterns of maze use in different maze types between English and Spanish with the exception of grammatical revisions, with grammatical revisions being higher in Spanish in both monolingual and bilingual groups. Results from the post-hoc analysis in the current study suggested that lexical and grammatical revisions (part-word, whole-word, and phrase) were generally higher in English simple sentences. The frequency of all revisions in complex sentences was similar in both languages. Bedore et al. (2006) cited previous studies suggesting that revisions were more common in Spanish due to more advanced language skills in that language, allowing the speaker to more readily revise any differences attributed to the language (Evans, 1985; Kormos, 1999). Findings from the current study suggest that the opposite may be true.

Results from the initial post-hoc analysis should be viewed descriptively and with caution, since the overall length of each individual transcript and the amount of utterances produced by each participant were not taken into account. Regarding this, an additional post-hoc analysis was conducted using the calculated differences between the total number of words (NTW) and each specific maze behavior. Results from this secondary analysis were inconclusive, since there

weren't enough instances of any one type of maze behavior to indicate any significant findings.

Limitations

There are several limitations in the study related to methodology, analysis, and participants. One of the biggest limitations to the study was the lack of participant representation in every wave (time point) from the beginning of kindergarten until the end of second grade. As part of the methodology for the study, transcripts were selected according to the presence of at least one complex utterance (Subordinating Index (SI) score of two or higher) in both languages at the beginning and end of the school year. This was necessary for the analysis in order to compare simple and complex utterances from the same narratives produced by the same participants. Accordingly, fewer transcripts were used for analysis from the kindergarten year than from the first and second grade years, and any participant included in the sample in any given wave may not be represented in the successive or previous waves.

A significant limitation to the study is in the specific type of instructional method that each of the participants were receiving at the time that the narratives were elicited. Though this information was provided to indicate which instructional methods the participants were receiving as a whole, this information was not accounted for in the multi-level modeling concerning individual participants or groups. It is possible that varying results may be obtained if this information were to be considered in an additional analysis.

Summary and Implications

This study examined the difference in the frequency of maze behaviors between simple and complex utterances, English and Spanish, and time in narrative retells produced by English and Spanish speaking bilingual school-age children. The findings indicate that bilingual children enrolled in various instruction programs, mostly provided in English, produce more mazes overall in complex sentences than in simple sentences in both English and Spanish. The findings

also indicate that the frequency of maze productions tends to remain constant in English from kindergarten through second grade while slightly decreasing in Spanish. No other significant interactions were obtained. Larger differences in maze scores between complex and simple sentences were noted in English than in Spanish.

These findings may contribute to further research regarding the discrimination between maze scores that are a result of bilingual language learning and maze behaviors that indicate language impairment. These findings might help to determine the extent that bilingual children may produce mazes in their second, or non-dominant, language, improving the accuracy with which language impairments are classified in the bilingual population. Additional analysis examining the frequency that specific maze types that are produced by bilingual children with both typical language development language impairments may also provide insight in this regard.

References

- August, D., Shanahan, T. (2006). *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth*. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- August, D., Carlo, M., Dressler, C., & Snow, C. (2005). The critical role of vocabulary development for English language learners. *Learning Disabilities Research & Practice*, 20 (1), 50-57.
- Bedore, L. M., Fiestas, C. E., Peña, E. D., & Nagy, V. J., (2006). Cross-language comparisons of maze use in Spanish and English in functionally monolingual and bilingual children. *Bilingualism: Language and Cognition*, 9 (3), 233-247.
- Bishop, D. V. M., & Donlan, C. (2005). The role of syntax in encoding and recall of pictorial narratives: Evidence from specific language impairment. *British Journal of Developmental Psychology*, 23, 25-46.
- Bialystok, E., Craik, F. I. M. & Luk, G. (2008). Lexical access in bilinguals: Effects of vocabulary size and executive control. *Journal of Neurolinguistics*, 21, 522-528.
- Bishop, D. V., & Edmundson, A. (1987). Specific language impairment as a maturational lag: Evidence from longitudinal data on language and motor development. *Developmental Medicine & Child Neurology*, 29 (4), 442-459.
- Bills, G. D., & Vigil, N. A. (1999). Ashes to ashes: the historical basis for dialect variation in New Mexican Spanish. *Romance Philology*, 53, 43-66.
- Branum-Martin, L., Foorman, B. R., Francis, D. J., & Mehta, P. D. (2010). Contextual effects of bilingual programs on beginning reading. *Journal of Educational Psychology*, 102 (2), 341-355.
- Choi, S. (1997). Language-specific input and early semantic development: Evidence from children learning Korean. In D. Slobin (ed.), *The crosslinguistic study of language acquisition*, vol. 5: *Expanding contexts*, pp. 41-133, Mahwah, NJ: Lawrence Erlbaum Associates.
- Collier, V. P. (1989). How long? A synthesis of research on academic achievement in a second language, *Tesol Quarterly*, 23 (3), 509-531.
- Dalton, E. F., Amori, B. A., Ballard, W. S., & Tighe, P. L. (1991). *The IDEA oral language proficiency tests*. Brea, CA: Ballard & Tighe.
- Edmunds, Paul. (2006). Buen- -- Buena Gente: Repair in the Spanish of the southwest. Selected proceedings of *The 8th Hispanic Linguistics Symposium*, ed. Timothy L. Face and Carol A. Klee, 204-213. Somerville, MA: Cascadilla Proceedings Project.

- Eigsti, Inge-Marie., de Marchena, A. B., Schuh, J. M., Kelley, E. (2011). Language acquisition in autism spectrum disorders: A developmental review. *Research in Autism Spectrum Disorders*, 5 (2), 681-691.
- Evans, J. L. (1985). Self-initiated repairs: A reflection of communicative monitoring in young children. *Developmental Psychology*, 21 (2), 365-371.
- Fazio, B. B., Naremore, R. C., & Connell, P. J. (1996). Tracking children from poverty at risk for specific language impairment. *Journal of Speech, Language, and Hearing Research*, 39 (1), 611-624.
- Fletcher, J., Lyon, G., Fuchs, L., & Barnes, M. (2007). Written expression disabilities. In Fletcher et al, (Eds.), *Learning disabilities: From identification to intervention*. New York, NY: Guilford Publications Inc.
- Ford, C. E., Fox, B. A, & Thompson, S. A. (2003). Social interaction and grammar. In M. Tomasello (ed.), *The new psychology of language: Cognitive and functional approaches to language structure* (vol. 2), pp. 119-144. Mahway, NJ: Lawrence Erlbaum Associates.
- Francis, D., Carlson, C.D., Fletcher, J.M., Foorman, B., Goldenberg, C., Vaughn, S., Miller, J., Iglesias, A., Papanicolaou, A. (2005). Oracy/Literacy development of Spanish speaking children: A Multi-level program of research on language minority children and the instruction, school and community contexts, and interventions that influence their academic outcomes. *Perspectives* (The International Dyslexia Association) 31 (2), 8-12.
- Gollan, T. H., & Acenas, L. A. (2004). What is a TOT? Cognate and translation effects on tip-of-the-tongue states in Spanish-English and Tagalog-English bilinguals. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 30 (1), 246-269.
- Gollan, T.H., Montoya, R. I. & Bonanni, M.P. (2005). Proper names get stuck on bilingual and monolingual speakers' tip of the tongue equally often. *Neuropsychology*, 19 (3), 278-287.
- Gollan, T.H., Montoya, R.I., Cera, C.M. & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, 58, 787– 814.
- Gollan, T.H., Montoya, R.I. & Werner, G.A. (2002). Semantic and letter fluency in Spanish-English bilinguals. *Neuropsychology*, 16, 562–576.
- Gollan, T. H. & Silverberg, N. B. (2001). Tip-of-the-tongue states in Hebrew-English bilinguals. *Bilingualism: Language and Cognition*, 4 (1), 63-84.
- Griffin, T., Hemphill, L., Camp, L., & Wolf, D. (2004). Oral discourse in the preschool years and later literacy skills. *First Language*, 24, 123–147.
- Gutiérrez-Clellen, V. F. (2002). Narratives in two languages: Assessing performance of bilingual children. *Linguistics and Education*, 13(2), 175–197.

- Gutiérrez-Clellen, V. & Kreiter, J. (2003), Understanding child bilingual acquisition using parent and teacher reports. *Applied Psycholinguistics*, 24 (2), 267-288.
- Gutierrez-Clellen, V. F., & Hofstetter, R. (1994). Syntactic complexity in Spanish narratives: A developmental study. *Journal of Speech and Hearing Research*, 37, 645-654.
- Hakuta, K., Goto Butler, Y. & Witt, D. (2000). *How long does it take English learners to attain proficiency?* Santa Barbara, CA: University of California Linguistic Minority Research Institute Policy Report.
- Heilmann J., Miller J. F., Nockerts A., Dunaway C. (2010). Properties of the narrative scoring scheme using narrative retells in young school-age children. *American Journal of Speech-Language Pathology*, 19, 154-166.
- Ivanova, I. & Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychologica*, 127, 277– 288.
- Jacobsen, P. F. & Schwartz, R. G. (2005). English past tense use in bilingual children with language impairment. *American Journal of Speech Language Pathology*, 14 (1), 313-323.
- Kormos, J. (1999). Monitoring and self-repair in L2. *Language Learning*, 49 (2), 303-342.
- Leadholm, B., & Miller, J. F. (1992). *Language sample analysis: The Wisconsin guide*. Madison, WI: Department of Public Instruction.
- Lennon, P. (1990). Investigating fluency in EFL; A quantitative approach. *Language Learning*, 40 (3), 387-417.
- Levelt, W. J. M., Roelofs, A. & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22 (1), 1-75.
- Levelt, W. (1989). *Speaking; From intention to articulation*. Cambridge, MA; Bradford Books.
- Loban, W. (1976). *Language development: Kindergarten through grade twelve*. Urbana, IL: National Council of Teachers of English.
- Loban, W (1963). The language of elementary school children: A study of the use and control of language and the relations among speaking, reading, writing and listening. Urbana, Illinois: National Council of Teachers of English.
- Maclachlan, B. G. & Chapman, R. S. (1988). Communication breakdowns in normal and language learning-disabled children's conversation and narration. *Journal of Speech and Hearing Disorders*, 53 (1), 2-7.
- Mayer, M. (1974). *Frog goes to dinner*. New York: Dial Press.

- Mayer, M. (1973). *Frog on his own*. New York: Dial Press.
- Mayer, M. (1969). *Frog, where are you?* New York: Dial Press.
- Mayer, M. (1967). *A boy, a dog, and a frog*. New York: Dial Press.
- Miller, J. & Chapman, R. (1984-2004). *Systematic analysis of language transcripts (SALT)* (Research version 8.0) [Computer software]. Madison, WI: University of Wisconsin Madison.
- Miller, J., & Iglesias, A. (2003–2004). *Systematic analysis of English and Spanish language transcripts*. Language Analysis Laboratory, Waisman Center, University of Wisconsin-Madison, Madison, WI.
- Miller, J. F., & Iglesias, A. (2008). Systematic Analysis of Language Transcripts (SALT), English & Spanish (Version 9) [Computer software]. Madison, WI: University of Wisconsin-Madison, Waisman Center, Language Analysis Laboratory.
- Miller, J., Heilmann, J., Nockerts, A., Iglesias, A., Fabiano, L., & Francis, D. (2006). Oral language and reading in bilingual children. *Learning Disabilities Research & Practice*, 21, 30–43.
- Muñoz, M.L., Gillam, R. B., Peña, E. D., & Gulley-Faehnle, A. (2003). Measures of language development in fictional narratives of Latino children. *Language, Speech, and Hearing Services in Schools*, 34 (1), 332-342.
- Nippold, M. (1993). Developmental markers in adolescent language: Syntax, semantics, and pragmatics. *Language, Speech, and Hearing Services in Schools*, 24, 21-28.
- National Clearinghouse for English Language Acquisition. (2007). The Growing Numbers of Limited English Proficient Students [Data file]. Retrieved January 16, 2008 from http://www.ncela.gwu.edu/policy/states/reports/statedata/2005LEP/GrowingLEP_0506.pdf.
- Papafragou, A., Massey, C. & Gleitman, L. (2002). Shake, rattle, ‘n’ roll: the representation of motion in language and cognition. *Cognition*, 84 (2), 189-219.
- Pavlenko, Aneta. (Eds.). (2009). *The bilingual mental lexicon: Interdisciplinary approaches*. Bristol, BS1 2AW, UK: Multilingual Matters.
- Pouliisse, N. (1999). Language production in bilinguals. In A. deGroot & J. Kroll (eds.). *Tutorials in bilingualism: Psycholinguistic perspectives*, pp. 201-224. Mahwah, NJ: Lawrence Erlbaum Associates.
- Rieger, C. L. (2003). Repetitions as self-repair in English and German conversations, *Journal of Pragmatics*, 35 (1), 47-69.
- Roberts, J., Martin, G. E., Moskowitz, L., Harris, A. A., Foreman, J., & Nelson, L. (2007). Discourse skills of boys with fragile X syndrome in comparison to boys with Down syndrome. *Journal of Speech, Language, and Hearing Research*, 50, 475-492.

- Sandoval T. C., Gollan T. H., Ferreira V. S., & Salmon D. P. (2010). What causes the bilingual disadvantage in verbal fluency? The dual-task analogy. *Bilingualism: Language and Cognition*, 13 (2), 231-252.
- Silva-Corvalan, C. (1994). *Language Contact and Change: Spanish in Los Angeles*. New York, NY; Oxford University Press.
- Soares, A., Goulart, B., & Chiari, B. (2010). Narrative competence among hearing-impaired and normal-hearing children: Analytical cross-sectional study. *Sao Paulo Medical Journal*, 128, 284-288.
- Thordardottir, E. & Ellis Weismer, S. (2002). Content mazes and filled pauses on narrative language samples of children with specific language impairment. *Brain and Cognition*, 48 (2-3), 587-592.
- Uccelli, P., & Paez, M., (2007). Narrative and vocabulary development of bilingual children from kindergarten to first grade: Developmental changes and associations among English and Spanish skills. *Language, Speech, and Hearing Services in Schools*, 38, 225-236.
- U.S. Department of Education, National Center for Education Statistics. (2004). *The condition of education 2004* (NCES 2004-077). Washington, DC: U.S. Government Printing Office.
- Vaughn, S., Mathes, P. G., Linan-Thompson, S., & Francis, D. J. (2005). Teaching English language learners at risk for reading disabilities to read: Putting research into practice. *Learning Disabilities Research & Practice*, 20 (1), 58-67.
- Wiese, R. (1984). Language production in foreign and native languages: Same or different? In H. Dechert, D. Möhle & M. Raupach (eds.), *Second language productions*, pp. 11-25. Tübingen: Narr.
- Wellman, R. L., Lewis, B. A., Freebairn, L. A., Avrich, A. A., Hansen, A. J., & Stein, C. M. (2011). Narrative ability of children with speech sound disorders and the prediction of later literacy skills. *Language, Speech, and Hearing Services in Schools*, 42 (1), 561-579.

APPENDICES

APPENDIX A. Demographic Variables

Figure A.1. Instructional Programs of Participants

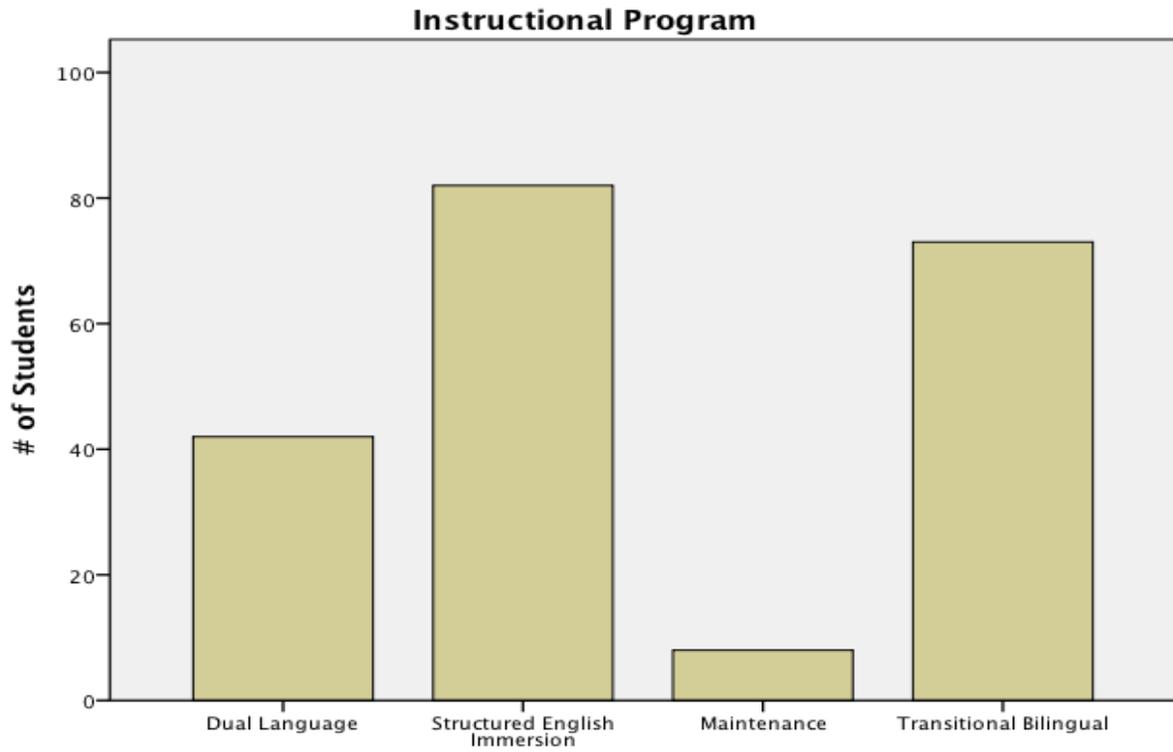


Table A.1. Illustration of the data files analyzed for each participant at each of the 6 time points

data were collected over the course of the study

Fall of K (Frog Story A)

Spanish complete file		English complete file	
Simple utterances only	Complex utterances only	Simple utterances only	Complex utterances only

Spring of K (Frog Story A)

Spanish complete file		English complete file	
Simple utterances only	Complex utterances only	Simple utterances only	Complex utterances only

APPENDIX B. Average Maze Scores

Table B.1. Means and SDs for Average Mazes Per Utterance for Waves 1-6, for complex and simple sentences in English and Spanish

Wave	Phase	Sentence Type	Language	Average Mazes per utterance Means(SD)	N
1	Fall	complex	English	0.7513 (0.71)	110
1	Fall	complex	Spanish	0.341 (0.52)	110
1	Fall	simple	English	0.5024 (0.25)	110
1	Fall	simple	Spanish	0.1863 (0.26)	110
2	Spring	complex	English	0.7455 (0.47)	110
2	Spring	complex	Spanish	0.3283 (0.40)	110
2	Spring	simple	English	0.5131 (0.28)	110
2	Spring	simple	Spanish	0.2351 (0.22)	110
3	Fall	complex	English	0.7372 (0.63)	156
3	Fall	complex	Spanish	0.2771 (0.40)	156
3	Fall	simple	English	0.5115 (0.25)	156
3	Fall	simple	Spanish	0.2083 (0.25)	156
4	Spring	complex	English	0.7319 (0.52)	156
4	Spring	complex	Spanish	0.2593 (0.35)	156
4	Spring	simple	English	0.502 (0.25)	156
4	Spring	simple	Spanish	0.1856 (0.23)	156
5	Fall	complex	English	0.7122 (0.54)	185
5	Fall	complex	Spanish	0.1826 (0.34)	185
5	Fall	simple	English	0.4856 (0.23)	185
5	Fall	simple	Spanish	0.1076 (0.17)	185
6	Spring	complex	English	0.8697 (0.45)	185

6	Spring	complex	Spanish	0.3189 (0.33)	185
6	Spring	simple	English	0.4985 (0.26)	185
6	Spring	simple	Spanish	0.2162 (0.22)	185

Figure B.1 Average Mazes Per Utterance for Waves 1-6, in English and Spanish

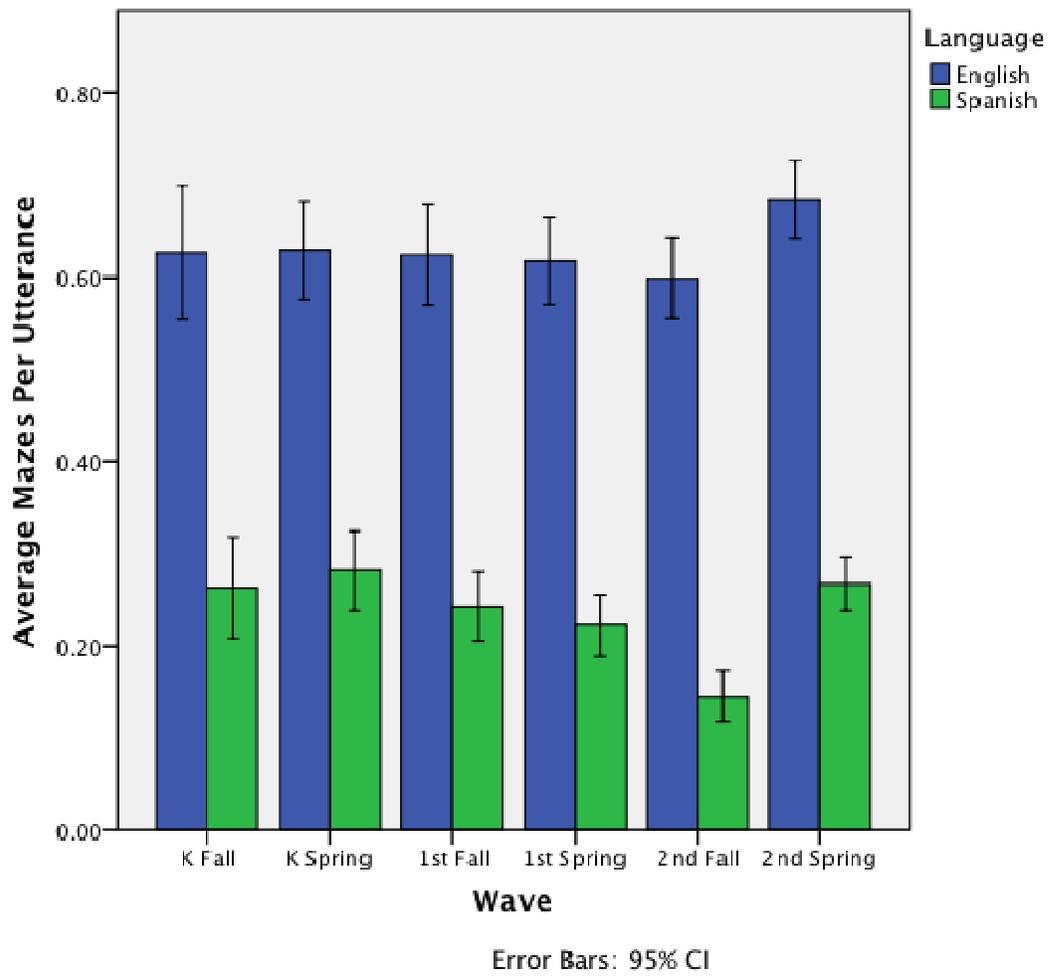


Figure B.3. Main Effect: Average Maze Score per Utterance

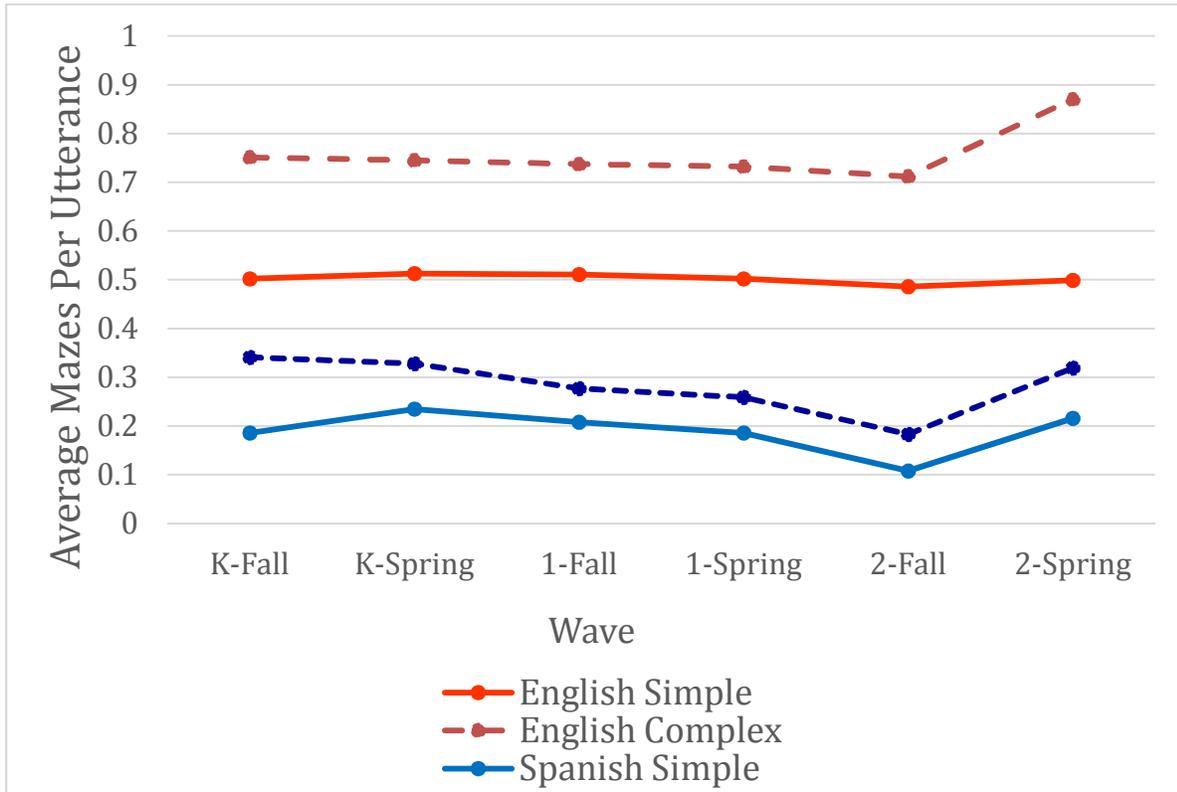


Table B.2. Means and SDs for Percent of Mazed Words per transcript for Waves 1-6, for complex and simple sentences in English and Spanish

Wave	Phase	Sentence Type	Language	Percent Mazed Words per transcript Mean (SD)	N
1	F	complex	English	0.1266 (0.12)	110
1	F	complex	Spanish	0.0648 (0.10)	110
1	F	simple	English	0.1425 (0.07)	110
1	F	simple	Spanish	0.0588 (0.08)	110
2	S	complex	English	0.1395 (0.1)	110
2	S	complex	Spanish	0.0694 (0.08)	110
2	S	simple	English	0.1353 (0.07)	110
2	S	simple	Spanish	0.0796 (0.07)	110
3	F	complex	English	0.1233 (0.11)	156
3	F	complex	Spanish	0.0528 (0.08)	156
3	F	simple	English	0.1328 (0.06)	156
3	F	simple	Spanish	0.0709 (0.08)	156
4	S	complex	English	0.1194 (0.08)	156
4	S	complex	Spanish	0.0508 (0.07)	156
4	S	simple	English	0.1292 (0.06)	156
4	S	simple	Spanish	0.0608 (0.07)	156
5	F	complex	English	0.1044 (0.07)	185
5	F	complex	Spanish	0.0335 (0.06)	185
5	F	simple	English	0.1243 (0.06)	185
5	F	simple	Spanish	0.0365 (0.06)	185
6	S	complex	English	0.1221 (0.06)	185
6	S	complex	Spanish	0.0566 (0.06)	185
6	S	simple	English	0.1258 (0.06)	185
6	S	simple	Spanish	0.0652 (0.07)	185

APPENDIX C. Multilevel Model Results

Table C.1. Results of Multilevel Model for Average Maze Scores.

Fixed Effects	Estimate	Standard Error	t-statistic
(Intercept)	1.48	0.36	4.13
NSS	0.00	0.01	0.60
Sentence Type	-0.18	0.02	-8.80
Language	-0.34	0.34	-1.00
Phase	-0.28	0.26	-1.05
Wave	0.03	0.02	1.52
Total Months	-0.01	0.00	-1.85
Gender	0.03	0.02	1.38
Language x Wave	-0.02	0.01	-2.29

Group	Random Effects	Standard Deviation	Correlation
ID	Intercept	0.23	
	Wave	0.04	-0.81
Wave (K, 1, 2)	Intercept	0.30	
	Phase	0.23	0.99
Phase (Fall, Spring)	Intercept	0.36	
	Sentence Type	0.02	-1.00
	Language	0.48	0.72

Note. Bold indicates $p < .05$.

APPENDIX D. Narrative Scoring Scheme Results

Table D.1. Average Narrative Scoring Scheme (NSS) Scores per Transcript

Wave	Language	Mean (SD)	95% Confidence Interval	
1	English	16.18 (0.28)	15.63	16.74
2	English	20.11 (0.28)	19.55	20.67
3	English	20.35 (0.24)	19.89	20.82
4	English	21.26 (0.24)	20.80	21.73
5	English	19.27 (0.22)	18.84	19.70
6	English	21.35 (0.22)	20.92	21.78
1	Spanish	16.90 (0.28)	16.34	17.46
2	Spanish	21.11 (0.28)	20.55	21.67
3	Spanish	23.68 (0.24)	23.21	24.15
4	Spanish	23.96 (0.24)	23.49	24.43
5	Spanish	21.61 (0.22)	21.18	22.04
6	Spanish	26.32 (0.22)	25.90	26.75

Figure D.1. Average Narrative Scoring Scheme (NSS) Scores per Transcript

