ABSTRACT

Title of Thesis: THE RELATIONSHIP BETWEEN
                                             LANGUAGE EXPERIENCE AND
                                             PERFORMANCE ON LANGUAGE
                                             ASSESSMENT MEASURES IN TYPICALLY-
                                             DEVELOPING SPANISH-ENGLISH
                                             BILINGUAL CHILDREN

                                             Daniella Otarola-Seravalli, Master of Arts, 2021

Thesis Directed By: Dr. Nan Bernstein Ratner, Department of Hearing
                                             and Speech Sciences

This study aimed to better understand the factors that affect bilingual children's assessment

performance and compare the effects of language experience on different types of

measures. English language sample measures (i.e., Index of Productive Syntax, Mean

Length of Utterance in morphemes, number of Brown's morphemes, and Vocabulary

Diversity) and English/Spanish nonword repetition (NWR) from 29 children with varying

degrees of English and Spanish language experience were analyzed. Language experience,

age, and baseline language abilities were identified as factors that influence and predict

performance on language samples. Additionally, it was determined that NWR ability was

not influenced by language-specific knowledge, due to the lack of significant correlation

between nonword repetition accuracy and language experience. These preliminary findings

suggest that NWR, even in a child's second language, is a relatively unbiased tool. Future

studies should compare the role of language experience on different measures in other

languages.

THE RELATIONSHIP BETWEEN LANGUAGE EXPERIENCE AND
PERFORMANCE ON LANGUAGE ASSESSMENT MEASURES IN
TYPICALLY-DEVELOPING SPANISH-ENGLISH BILINGUAL CHILDREN

by

Daniella Otarola-Seravalli

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Master of Arts
2021

Advisory Committee:
 Dr. Nan Bernstein Ratner, Chair
 Dr. Yasmeen Faroqi-Shah
 Dr. Min Wang
 Mr. José Ortiz

Acknowledgements

I would like to sincerely thank my advisor, Dr. Nan Ratner, for her guidance, extensive knowledge, and patience throughout the thesis process. I am also grateful for the other members of my committee, Dr. Yasmeen Faroqi-Shah, Dr. Min Wang, and Mr. José Ortiz, whose support and expertise were instrumental in the development of this project. Special thanks to Dr. Ning Hsu for her help with statistical analysis; Einat Korman for her help setting up coding procedures for repetition tasks; research assistant, Camila Lopez, for her help coding the data; and all others member of the Language Fluency Lab who participated in data collection. I very much appreciate Grandpa Lee VanderVelden for proofreading my work and inspiring me to do research. Finally, thank you to my family, fiancé, and friends for their endless encouragement.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

DLD: Developmental Language Disorder

IPSyn: Index of Productive Syntax

L1: First Language

L2: Second Language

LSA: Language Sample Analysis

MLU: Mean-Length of Utterance

NBM: Number of Brown's Morphemes

NWR: Nonword Repetition

NWRA: Corrected Nonword Repetition accuracy

RWR: Real Word Repetition

VocD: Vocabulary Diversity

Y1: Year 1

Y2: Year 2

Chapter 1: Introduction

Bilingual Language Experience

A bilingual child is one "who receives regular input in two or more languages during the most dynamic period of communication development" (Kohnert, 2010). Children who begin receiving input in both languages before three years of age are referred to as simultaneous bilinguals, as both languages develop at the same time (Paradis, 2007). On the other hand, children who begin receiving input in the second language (L2) after age three are referred to as sequential bilinguals, as they have partially developed their first language (L1) prior to learning their L2 (Paradis, 2007).

Even within simultaneous and sequential bilingual subgroups, there are variations in language experience (e.g., Years of exposure, current hours of exposure, language of communication partners, and settings of exposure) (Bedore et al., 2018; Thordardottir & Brandeker, 2011). Relative levels of experience with each language will determine a child's language proficiency (how developed each language is) and dominance (which language is stronger). The language of communication varies across partners, topics, locations, and time; thus, proficiency and dominance may change throughout the lifespan (Kohnert, 2010). Proficiency increases when the child has more opportunities to use a language in a variety of settings, and thus develops stronger skills in that language. If one language is more often used in a particular setting, context, or topic, that will be the dominant language in those cases. However, the other language may be dominant in other situations (Kohnert, 2010). For example, Cognitive Academic Language Proficiency (CALP) might be stronger in the child's L2, or the language used at school, while Basic

Interpersonal Communication Skills might be stronger in the L1, or the language used at home (Cummins, 1981).

Issues in Bilingual Assessment

Disproportionality

Accurately diagnosing bilingual children who may or may not have language delay or disorder has become increasingly critical, given the changing demographics of U.S. public schools student body. The number of English language learners (ELLs) in U.S. public schools has grown from 3.8 million (8.1% of total students) in 2000 to 5 million (10.1%) in 2017 (NCES , 2020a). Among these students, 75% were Spanish-speakers (NCES, 2020a). In 2017, 14% (7 million) of students received special education services due to disabilities (speech or language impairment being the second most common diagnosis [19% of students]) (NCES, 2020b). That same year, 14.3% (718, 400) of ELLs were identified as having a disability (NCES, 2020a).

Bilingual children are often over- or under-diagnosed with developmental language disorder (DLD), as language errors may either be attributed to a language difference or to an impairment (Boerma et al., 2015). This determination appears to be influenced by the child's language proficiency. Yamasaki and Luk (2018) found that bilingual children who were proficient in English were statistically likely to be under-identified as having DLD throughout elementary school. On the other hand, bilinguals with less English experience were first under-identified in 3rd grade, but then over-identified (relative to frequencies seen in monolingual students) by 5th grade (Yamasaki & Luk, 2018).

Further, Valenzuela et al. (2006) noted that ELLs are disproportionately placed in segregated special educational settings. Misdiagnosing bilinguals and incorrectly placing

them in special education can result in violations of Section 1412(a)(5) of the Individuals with Disabilities Education Act (2004). This states that children should receive an education in the "least restrictive environment," meaning as close to a regular educational environment with typically developing peers as possible. Thus, errored productions in bilinguals should not automatically be assumed to be caused by a language impairment. However, if there are concerns of DLD, assessment should include special considerations for bilinguals in order to obtain an accurate language profile.

Challenges in Bilingual Assessment

First, bilingual children's knowledge of one language may influence their expressive and receptive abilities in their other language (Kohnert, 2010). For example, Russian-Hebrew bilingual children are more likely to correctly produce features shared across the two languages (e.g., subject-verb agreement) than those that are not shared (e.g., definite articles) (Meir et al., 2017). Thus, what might appear to be an error in one language caused by an underlying language impairment might actually be cross-linguistic transfer.

Secondly, there is sometimes an overlap between typically-developing bilingual children's performance in the non-dominant language and that of monolingual children with DLD (Boerma et al., 2015). Blom et al. (2013) found that verb inflection error patterns were similar between typically-developing sequential Turkish-Dutch school-age bilinguals and Dutch monolinguals with DLD. However, Scheidnes (2018) found that, at least in the case of English-French sequential bilinguals, this overlap in error targets was only temporary. As the typically-developing bilingual group gained more experience in French, they produced more object clitics, but the DLD group continued to omit them (Scheidnes, 2018). Additionally, language attrition (stagnant development or decreasing ability in one

language due to increased use of the other language) may further complicate this matter, as markers of language attrition include use of more general terms and errored grammatical productions (Paradis, 2007). Since these language characteristics are also common in DLD, bilingual children's abilities in the previously dominant language might be mistakenly identified as impaired.

Third, bilinguals are a heterogenous group. Variability in language experience has been shown to influence performance on different tasks and measures. For instance, Spanish-English school-aged bilinguals needed at least 40% of English language input on a daily basis to be correctly identified as having DLD when using English grammatical forms (Bedore et al., 2018). Similarly, in word association tasks, bilingual children performed better in the language they had more experience with (Sheng et al., 2012).

Assessment Options

Standardized Language Assessment

Standardized tests are what Campbell (1997) refers to as "knowledge-based" measures, as they rely on the child's experience and knowledge. However, given the diversity of bilinguals' experience and knowledge, norms required by these tests are difficult to develop (Kohnert, 2010; Restrepo, 2001). While it is relatively easy to develop expectations of language knowledge when children experience input in a single language, it is difficult to establish norms when English or other language input vary considerably from child to child.

Further, typically developing bilingual children have been shown to perform more poorly on standardized tests than their typically developing monolingual peers (Boerma et al., 2015; Laing, 2003; Restrepo, 2001). This discrepancy has been attributed to test bias

(Laing, 2003). Two common types of bias found in standardized tests include content and linguistic bias. Content bias is defined as "the assumption that all children have been exposed to the same concepts and vocabulary or have had similar life experiences" (Laing, 2003). Thus, if a test includes stimuli (e.g., pictures, terms) that might be familiar to one group but unfamiliar to others (e.g., "robin"), children who are unfamiliar with the items would be more likely to perform poorly on them. Similarly, linguistic bias is the expectation that a child will respond in the clinician's language or dialect when the child uses different languages or dialects (Laing, 2003). Linguistic bias underlies the misidentification of language-transfer as an impairment, rather than influence from the child's other languages or dialects.

Parental Report

Parent questionnaires are a valid, cost- and time-effective way of obtaining crucial information about a child's language experience (Gutiérrez-Clellen & Kreiter, 2003; Marchman & Martínez-Sussman, 2002). It is recommended that they be included in assessment batteries to improve diagnostic accuracy (Paradis et al., 2010; Paradis et al., 2013). Parent questionnaires have been used for almost 40 years as screening and diagnostic tools in monolinguals, and as language history and language status tools for bilinguals (Gutiérrez-Clellen & Kreiter, 2003).

Language experience questionnaires vary from broad devices, such as asking parents to estimate a general percentage of exposure for each language, to specific, such as asking about hours of exposure to each language across the lifespan, conversational partners, and changes over time (DeAnda et al., 2016). Parental report of language exposure has been shown to be a valid and reliable measure (DeAnda et al., 2016). It has

been correlated with parental reports of vocabulary (either specific items or total counts), which are another type of parental report that is often used as a proxy for children's language abilities and exposure (DeAnda et al., 2016; Libertus et al., 2015).

However, language experience questionnaires might be preferred over vocabulary inventories, as they provide more detailed information about language exposure and bypass some of the limitations of vocabulary inventories. For example, as with other subtypes of standardized tests, if not carefully selected, vocabulary inventories may be subject to issues with norming procedures (Marchman & Martínez-Sussman, 2002). Also, given that some inventories in other languages are direct translations of the English version, they might not contain culturally relevant items; thus, introducing potential content and linguistic bias (Marchman & Martínez-Sussman, 2002).

Alternative Assessment

Alternatives to standardized tests include language sampling, processing-dependent measures (e.g., nonword repetition [NWR]), and dynamic assessment. Language Sample Analysis (LSA) offers an ecologically valid picture of the child's language abilities, as the child engages in spontaneous conversations or narratives. On the other hand, NWR tasks tap into cognitive-linguistic skills necessary for language development, such as phonological memory, through children's repetition of pseudo words (Boerma et al., 2015). Finally, dynamic assessment measures a child's current level and learning potential through a series of strategies (e.g., test-teach-retest and graduated prompting) (Laing, 2003). Given that dynamic assessment is not one of the main interests of this study, it will not be discussed in detail.

These measures have been recommended when working with bilinguals as they are less biased (Laing, 2003). Language sampling and NWR are worthy of particular notice, as they have been well-researched and are commonly used. Campbell (1997) specifically advocates for the use of "processing-dependent" measures, which ensure that all children start out with the same knowledge, as the task is completely novel or it is certain that all children have the same background-knowledge needed. NWR is also of particular interest as it only relies on a few properties of the target language (i.e., phonological and phonetic), while LSA relies on more components of the target language (i.e., lexical, morphological, and syntactic properties). Dynamic Assessment, another option for appraising language skill in ELL children, often relies on NWR stimuli, thus overlapping in part with the conceptual basis for NWR. In theory, measures that are semi-independent of language-specific knowledge, such as NWR, would be less likely to misdiagnose typically developing bilinguals than measures that are dependent on language-specific knowledge, since they assess general linguistic skills as opposed to proficiency in a language.

Language Sampling

LSA has been shown to have good diagnostic accuracy in monolingual children (Gutiérrez-Clellen & Simon-Cereijido, 2012). In fact, Dunn and colleagues (1996) found that LSA more correctly identified children with DLD than did standardized tests. LSA has also been found to correctly identify bilingual children with DLD (Gutiérrez-Clellen & Simon-Cereijido, 2012; Rojas & Iglesias, 2009). Gutiérrez-Clellen (2002) noted that unlike typically developing bilinguals, who show adequate narrative proficiency in at least one language, bilinguals with DLD perform poorly in both languages.

Nonword Repetition (NWR)

NWR is considered a processing-dependent measure, as it uses novel words that no children are familiar with and taps phonological working memory skills by evaluating repetition accuracy (Campbell, 1997). Stimuli are crafted to be either language-specific (i.e., nonwords that mirror the target language's syllable structures, phonotactic constraints, and stress patterns) or quasi-universal (i.e., nonword features that are shared across many languages) (Boerma et al., 2015). Nevertheless, both types of words are produced using the phonetic patterns of the target language. NWR repetition has been successfully used with school-age bilingual speakers of many languages including Arabic, Berber, Chinese, Danish, Dari, Dutch, English, French, Frisan, German, Hebrew, Japanese, Kirundi, Moroccan, Pashto, Portuguese, Punjabi, Russian, Spanish, Sighalese, Suriyoyo, Tamil, and Turkish (Armon-Lotem & Meir, 2016; Boerma et al., 2015; de Almeida et al., 2017; Guiberson & Rodriguez, 2013; Hamann & Abed Ibrahim, 2017; Thordardottir & Brandecker, 2013; Windsor et al., 2010). While NWR research has been done with school-age children, NWR has also been successfully used with monolingual children as young as 20 months of age and Spanish-English bilingual children as young as 22 months of age (Hoff et al., 2008; Parra et al., 2011).

When used for diagnostic purposes, NWR has had relatively good discrimination in monolingual children (Boerma et al., 2015). Dollaghan & Campbell (1998) found that a cut-off score of 70% accuracy, defined as percentage of phonemes correct, accurately distinguished impairment. Results in bilingual children have been more mixed, as typically-developing bilingual children have not performed comparably to their typically-developing monolingual peers (Armon-Lotem & Meir, 2016; Boerma et al., 2015; de

Almeida et al., 2017; Hamann & Abed Ibrahim, 2017; Thordardottir & Brandeker, 2013; Windsor et al., 2010).

Finally, it appears NWR has a promising predictive value of a child's language abilities (Chiat & Roy, 2007). Hoff et al. (2008) found that NWR predicted vocabulary percentile in monolinguals. Dispaldro et al. (2011) found this predictive value of NWR to be true cross-linguistically, as English and Italian NWR predicted grammatical skills in both languages. Finally, Parra et al. (2011) found that bilinguals' performance on NWR at 22 months of age was significantly correlated with vocabulary and grammar at 25 months of age.

Challenges of Alternative Assessment

Even though both LSA and NWR are recommended as part of a bilingual assessment, it is important to consider the challenges that remain with these measures. LSA and NWR are not exempt from the issues of language transfer, overlap with monolingual impairment profiles, and impact of language experience on performance, all of which were discussed earlier. Of note, since dynamic assessment may use nonwords in fast-mapping tasks, it may be subject to the same challenges as NWR.

Language Transfer

Language transfer might affect both LSA and NWR. In the case of LSA, one might expect to see code-switching to fill in lexical gaps or morphosyntactic errors that could be explained by morphosyntactic rules of the other language. Similarly, in the case of NWR, errors might be explained by the phonetic patterns of the other language.

Profile Overlap

There is evidence of overlapping performance between typically developing bilinguals and monolinguals with DLD on both NWR and LSA. To illustrate, in a study by Gutiérrez-Clellen et al. (2008), L2 English bilinguals' performance on LSA measures mirrored that of monolinguals with DLD. Further, Armon-Lotem and Meir (2016) showed that using diagnostic cut-off points based on monolingual children for Russian-Hebrew bilinguals resulted in misdiagnosis of bilingual children. It was necessary to use a separate bilingual cut-off point to increase the diagnostic accuracy of NWR (Armon-Lotem & Meir, 2016). Compared to monolinguals, the bilingual cut-off point was higher for Russian (L1) NWR and lower for Hebrew (L2) NWR (Armon-Lotem & Meir, 2016). Finally, Kohnert et al. (2006) found that even in cases when typically-developing bilingual children perform better than monolingual children with DLD on English NWR, they still do not perform as well as typically developing monolingual children.

Impact of Language Experience

Research shows an impact of language experience on LSA measures. Gutierréz-Clellen (2002) found that second grade Spanish-English bilinguals produce higher levels of narrative proficiency in the language they have more experience in. Relatedly, Jacobson and Walden (2013) found that in early sequential Spanish-English school-age bilinguals, there is a relationship between measures of lexical diversity in narrative samples and oral language proficiency, as measured by the Woodcock-Muñoz Language Survey, which calculates a proficiency score based on relative Spanish and English input. However, there is mixed evidence about how much or if language experience influences NWR performance.

Some studies claim that language experience indeed influences NWR performance. For instance, Windsor et al. (2010) found that school-age children had higher NWR accuracy in the language they had more experience with. English monolinguals outperformed Spanish-English sequential bilinguals on English NWR, but Spanish-English bilinguals outperformed English monolinguals on Spanish NWR (Windsor et al., 2010). Likewise, Gibson et al. (2015) found that Spanish-dominant bilinguals were more accurate in Spanish NWR than English NWR. They also found that Spanish-dominant bilinguals were more accurate than English-dominant bilinguals at repeating longer nonwords, presumably because they had more practice producing longer words due to their increased exposure to Spanish, which has a higher frequency of multisyllabic words (Gibson et al., 2015). Parra et al. (2011) found that even in young 22-month-old Spanish-English bilinguals, exposure to the target language improved NWR accuracy. In fact, exposure to the target language accounted for 20% of the variance in English-NWR performance and 25% of the variance in Spanish-NWR (Parra et al., 2011). Finally, Thordardottir and Brandeker (2013) found that French-English bilinguals needed at least 35-40% exposure to English from birth to perform similarly to English monolinguals on an English NWR task.

In contrast, other studies have concluded that language experience does not influence performance or only minimally influences it. In fact, in that same study by Thordardottir and Brandeker (2013), even though English NWR performance correlated with previous exposure to the target language, French NWR did not. Similarly, de Almeida et al. (2017) concluded that there was no effect of language experience when using the French version of the Language Impairment Testing in Multilingual Settings (LITMUS)

11

tools' NWR task, as there was no correlation between measures of language experience (i.e., age of acquisition, length of exposure, quantity of language exposure, dominance) and the performance of L2 French bilinguals. Finally, Hamann and Abed Ibrahim (2017) found that the German LITMUS-NWR was only minimally influenced by language dominance of L2 German bilinguals. It must be noted that all the studies cited above included different NWR tasks. Thus, past mixed findings for the impact of language experience on NWR performance might be related to the disparate types of stimuli or differing structure of the tasks.

Summary

While alternative assessment measures are helpful in bilingual assessment, they are not perfect, especially since interpreting results can be challenging. For instance, the extent to which language experience influences bilinguals' performance is not fully understood. It remains unclear whether typically-developing bilinguals ever perform similarly to monolingual peers on language tasks, or if monolinguals always outperform bilinguals due to increased exposure to that language. Relatedly, since NWR has been more studied in school-age children, the predictive value of these tasks for younger bilingual children requires further exploration, as it might better inform decisions for early intervention. Thus, it is crucial to gain a better understanding of the value, fairness, and challenges involved in using and interpreting alternative assessment results; specifically, NWR and LSA, as they are commonly used and recommended as tools in bilingual assessment.

Current Study

Consequently, the present study investigated the relationship between language experience and performance on different types of alternative assessment measures

12

(dependent versus semi-independent of language-specific knowledge), as well as the value of these measures for bilingual assessment. We did this by studying English language sample scores and English and Spanish nonword repetition abilities of 29 typically-developing children with varying degrees of exposure to English, Spanish, and other languages over the course of 2 years. The language sample scores that we focused on were the Index of Productive Syntax (IPSyn; Scarborough, 1990), Mean Length of Utterance in morphemes (MLU; Brown, 1973), Vocabulary Diversity (VocD; Malvern & Richards, 2002), and number of Brown's morphemes (NBM; Brown, 1973).

Research Questions and Hypotheses

1. Does NWR skill have a predictive value? More specifically, does performance on NWR predict spoken English language skills? If NWR is a predictive tool that does not rely on language-specific knowledge, then English NWR accuracy at Year 1 and Year 2 will be significant predictors of English IPSyn in Year 2, even when controlling for other factors (i.e., age and baseline abilities [Year 1 IPSyn]) in the analysis.

2. Do bilingual children catch up to their monolingual peers over time, when given additional exposure to English? If cumulative input enables better skill development, then it is predicted that the monolingual group will improve more (achieve higher change scores from Year 1 to Year 2) in English NWR and English-language measures (such as IPSyn, MLU, NBM, and VocD) than will the bilingual group, as monolinguals always have higher daily and cumulative exposure to English.

3. What is the relationship between NWR performance and language experience? Specifically, if a child has more experience in a language, will they do better in that language's NWR task and worse in the other language's NWR task? If level of language experience predicts NWR performance, then children with a higher percentage of reported experience in a language will perform better in that language and worse in the other language's NWR task. In other words, NWR performance in a given language would be predicted by experience in that same language, but not experience with another language.

Chapter 2: Methods

Participants

A total of 29 typically-developing children, who were a part of a larger longitudinal study at the Language Fluency Laboratory at the University of Maryland, participated in this study. Children were recruited into this study on an on-going basis over a 4-year period. Participant characteristics are displayed in Table 1.

Table 1: Participant characteristics

	Year 1 n (n%) (total n=29)	Year 2 n (n%) (total n=23)
Age (in months)		
28-32	10 (34)	-
33-40	11 (38)	-
41-44	5 (17)	5 (22)
45-52	3 (10)	9 (39)
53-62		9 (39)
Gender		
Female	15 (52)	12 (52)
Male	14 (48)	11 (48)
Language experience (%)		
English		
0-20	4 (13)	-
21-40	2 (7)	4 (17)
41-60	5 (17)	1 (4)
61-80	2 (7)	1 (4)
81-100	10 (34)	8 (35)
Spanish		
0-20	13 (45)	8 (35)
21-40	1 (3)	2 (9)
41-60	4 (13)	1 (4)
61-80	2 (7)	4 (17)
81-100	3 (10)	-
Other		
0-20	22 (75)	-
21-40	-	-
41-60	1 (3)	-

61-80	1 (3)	-
81-100	-	-
SES (Maternal Education)		
College	8 (28)	7 (30)
Graduate Degree	20 (69)	16 (70)
Did not report	1 (3)	-

During their first year of testing, children were between 2;4-4;2 years old, while during their second year they were 3;5-5;2 years old. Prior to being recruited into the study, all children underwent an informal telephone screening using a standard set of questions asked of their parents to determine if they had any previously identified or suspected language or developmental disorders. Their status as typically-developing was confirmed during lab visits through parental inventories of vocabulary (*MacArthur-Bates Communicative Development Inventories* or the *Developmental Vocabulary Assessment for Parents*, depending on the child's age), an informal parent questionnaire (see Appendix A), and performance within normal limits on a set of standardized tests (*Peabody Picture Vocabulary Test, Clinical Evaluation of Language Fundamentals Preschool–2nd Edition* [subtests: *Concepts and Following Directions*, and *Sentence Structure]*, and *Developmental Neuropsychological Assessment–2nd edition* [word generation task]). Spanish-English bilinguals also completed standardized testing in Spanish (*Test de Vocabulario en Imágenes Peabody, Clinical Evaluation of Language Fundamentals Preschool–Spanish Edition* [*Conceptos y Siguiendo Direcciones, Estructura de Oraciones*], and *Developmental Neuropsychological Assessment–2nd edition* [word generation task in Spanish]). Children were considered typically-developing if parents reported that the child did not have any diagnoses (prior or current) of language or developmental disorders.

To be considered bilingual, children had to receive at least 20% exposure to each language, as measured by parental reports of hours of exposure (Gutiérrez-Clellen & Kreiter, 2003). On the other hand, to be considered monolingual, children could only receive minimal or inconsistent exposure to a second language other than English (i.e., less than 20% exposure). In cases of exposure to more than 2 languages, children were considered bilingual if the sum of exposure to the less common languages was at least 20%. To illustrate, it was reported that one of the children received 67% exposure to Hebrew, 17% exposure to English, and 17% exposure to Spanish. By this definition, 3 of the participants (all female) were classified as Spanish monolinguals, 10 as bilinguals (2 female, 8 male), and 10 as English monolinguals (5 female, 2 male). In addition to English and Spanish, some children had exposure to other languages during their first year, including Hebrew, Farsi, Portuguese, and Chinese.

Due to missing data (e.g., parents did not provide hours of exposure), 6 children were classified based on their status at recruitment (4 bilinguals [2 female, 2 male]), 2 monolinguals [both male]) for matching purposes. However, they were not included in analyses that involved language status or experience. Children were matched across groups for gender, SES (parent education), and age (within 12 months). Across both groups, 19 children were reported to be Caucasian (8 Hispanic Caucasian), 4 African American (3 Hispanic African American), 4 Hispanic only, and 2 biracial (Caucasian and Asian, Caucasian and Pacific Islander).

Tasks

Children participated in a battery of language assessment tasks administered by a research assistant at the Language Fluency Laboratory. However, for the purposes of this

study, only the parental questionnaire, English language sample, and the English and Spanish real and nonword repetition (RWR and NWR) will be considered. Spanish language samples were excluded since one of the main aims of this study was to compare bilinguals' performance to that of English monolinguals.

Monolingual children were tested in English, while bilingual children were tested in English and Spanish on separate days. English and Spanish sessions were, on average, within about 2.5 weeks of each other; and their order was counterbalanced across participants. Because this was a longitudinal study design, children were re-tested using the same format one year later.

Parental Questionnaire

At both first and second-year visits, parents filled out an informal questionnaire (see Appendix A and B). This questionnaire asked about parents and other caregivers' native language, primary home language, hours exposed to other languages, and relative skills in both English and Spanish compared to peers. Of particular interest was parental report of hours of language input and exposure, which was used to compute a percentage experience with each language.

Quantifying Cumulative Language Experience

Parental report has been shown to be useful in determining bilingual children's language abilities and proficiency in each language (Gutiérrez-Clellen & Kreiter, 2003; Marchman & Martínez-Sussmann, 2002). Therefore, parental report of hours of exposure to each language was used to calculate a percentage of cumulative experience in each language. The equations were as follows:

$$\% \text{ Target language } y1 = \frac{\text{Daily hours exposure to target language at year 1} * \text{age at year 1}}{12 \text{ waking hours} * \text{age at year 1}} *100$$

$$\%\text{Target language } y2 = \frac{(\text{target hours } y1 * \text{age at } y1) + (\textbf{target hours } \textbf{y2} * (\textbf{age } \textbf{y2} - \textbf{y1}))}{(12 * \text{age at } y1) + (\textbf{12} * (\textbf{age } \textbf{y2} - \textbf{y1}))} *100$$

Age was used as a multiplier to account for cumulative levels of exposure. Given that parents only reported hours of exposure for the year of testing, it was assumed that exposure for Year 1 was consistent with exposure prior to Year 1. For Year 2, the equation was modified to include hours of exposure at Year 2 multiplied by the age difference between Year 1 and 2 in order to account for changes in exposure between testing periods. The assumption that the child is awake for 12 hours per day is based on the Sleep Foundation's recommendation of hours of sleep for 2–5-year-old children (Sunni, 2020).

Language Experience by Group

Table 2 displays language experience ranges and averages for each group, according to these equations. On average, English monolingual experience remained consistently similar from Year 1 to Year 2, while Spanish-speaking children's average English experience increased from Year 1 to Year 2. Consequently, by Year 2, Spanish monolinguals had become Spanish-English bilinguals. This confirms that even after one year, there can be significant changes in language experience.

Spanish experience, on the other hand, remained relatively consistent for Spanish-English bilinguals, but decreased for Spanish monolinguals. This decrease in Spanish percentage might suggest a trend towards language attrition, or stagnant growth in Spanish at the expense of English growth. However, it could also be that these children are becoming more balanced bilinguals, as it is possible to maintain growth in Spanish while

increasing in English over time (Winsler et al., 1999). Likewise, children who were considered bilingual at Year 1, for the most part, remained balanced bilinguals at Year 2.

Nevertheless, there was wide variability in hours of exposure, particularly in the bilingual group. The persistent variability in language experience from Year 1 to Year 2, suggests that language experience should always be considered in bilingual assessments regardless of how much time has passed since the child was last assessed.

Table 2: Language experience by group and year.

	Bilingual % Range (Mean)		Spanish monolingual% Range (Mean)		English monolingual % Range (Mean)	
	Year 1 n=14	Year 2 n=4	Year 1 n=3	Year 2 n=3	Year 1 n=12	Year 2 n=8
English	17-80 (46)	31-69 (47)	13-17 (15)	21-22 (21)	92-100 (99)	94-100 (98)
Spanish	0-79 (45)	31-69 (50)	83-88 (85)	78-79 (79)	0-8 (1)	0-6 (2)
Other	0-67 (11)	n/a	n/a	n/a	n/a	0-4 (0)

Language Sample

Children engaged in naturalistic conversation during a play session with a graduate clinician or their parent, who spoke in the target language for that session. Children selected from a standard assortment of toys that they wanted to play with. Toy items included pretend food, a railroad set, dolls, and a building set. Language samples lasted for roughly 20 minutes and contained an average of 202 utterances per child. To ensure that the child produced enough utterances for future, a research assistant kept a tally of the number of utterances throughout the recording session.

Language Sample transcription and analysis

Play sessions were video recorded and later transcribed and coded by research assistants using CHAT codes and the Child Language Analysis (CLAN) program (MacWhinney, 2000). The Index of Productive Syntax (IPSyn; Scarborough, 1990), Mean

Length of Utterance in morphemes (MLU; Brown, 1973), Vocabulary Diversity (VocD; Malvern & Richards, 2002), and number of Brown's morphemes (NBM; Brown, 1973) were extracted by CLAN's Kideval utility. VocD was chosen as a measure of lexical diversity; while IPSyn, MLU, and NBM were chosen as measures of morphosyntactic development. VocD was chosen over NDW, as it considers the full length of the language sample (Jacobson & Walden, 2013). IPSyn was chosen over alternatives as it provides comprehensive information about the child's phrase structure variety (Jalilevand & Ebrahimipour, 2014). Of note, for the current study, IPSyn was calculated using 50 utterances, as opposed to the traditional 100, as some children did not produce enough utterances. Recently, Yang and colleagues (2021) found that computing IPSyn on 50 utterances yielded comparable results to when it was computed on 100 utterances. Finally, MLU and NBM were chosen as additional measures of grammatical skills since they are LSA measures commonly used in clinical practice.

Real and Nonword Repetition Task

This task was adapted from Parra, Hoff & Core (2011), who modified NWR so it could be used in children as young as 20 months. To provide some control for articulation, and given the young age of the participants, this study included repetition of real word stimuli. The lists of real and nonwords are displayed in Appendix C. Each list contained 4 one, two, and three-syllable long words, none of which contain late-developing sounds to avoid the confound of misarticulation. The 24 real words (12 English, 12 Spanish) were composed of words from the 16–30-month-old version of the *MacArthur-Bates Communicative Development Inventories* for the English stimuli and from the *Inventario del Desarrollo de Habilidades Comunicativas* for the Spanish stimuli. The 24 nonwords

(12 based on English phonology, 12 based on Spanish phonology) were developed by combining different parts of the real words of equivalent syllable lengths. The stimuli were recorded by a native speaker of the target language, who produced the words using the language's phonetic patterns.

Children were presented with a picture on a computer screen, and they were asked to repeat a word that the pre-recorded voice produced. In the case of real words, children saw an object and were asked to repeat its corresponding label (e.g., "This is a book. Can you say book?"). For the nonwords, children saw animated pictures of penguins, each with a unique symbol on its chest, and were asked to repeat the penguin's name (e.g., "Hi my name is /dʊk/. Can you say /dʊk/?"). RWR was included to distinguish between misarticulations that may occur at such a young age and phonological working memory errors (Hoff et al., 2008). For both the Spanish version and the English version of the task, RWR was always conducted before the NWR task in that language.

RWR & NWR accuracy computation

Research assistants trained in phonetic transcription transcribed participant productions live using IPA. Productions were also re-transcribed from video recordings by Spanish-English bilingual research assistants. The percentage accuracy determined through the re-transcriptions was used whenever videos were available. However, in cases of technical issues with video or audio recordings, accuracy was calculated using live transcriptions. Transcription reliability was calculated using 20% of the data by adding up the number of agreements (total=881) and disagreements (total=71) in samples that were double coded and computing a total percentage of agreements over total number of consonants in these samples (total=952). This rendered an inter-rater agreement of ~93%.

Percent consonants correct (PCC) is a commonly used measure for repetition accuracy, as it maximizes reliability (Parra et al., 2011; Hoff et al., 2008). However, to account for differences in speech production, a corrected NWR accuracy (NWRA) was calculated using PCC scores that are corrected based on real word productions. To calculate NWRA, nonword consonant targets earned 1 point if they were produced the same way in nonwords and real words (either correctly or as the same phonemic substitution), and 0 points if they were produced incorrectly or omitted. This is similar to the adjusted non-word accuracy referenced in Eaton et al. (2015). However, unlike the adjusted non-word accuracy, NWRA does not penalize errors made in the repetition of real words. NWRA was chosen over adjusted non-word accuracy because it focuses on repetition of nonwords alone, which is the main interest of this study.

Only consonants were considered when scoring because vowels are produced and perceived with more variability (Hoff et al., 2008). Non-attempts were scored as missing data, as it is difficult to determine the reason that a child is not producing a target using conventions proposed by Hoff et al. (2008).

Chapter 3: Results

Does performance on NWR predict spoken English language skills?

It was predicted that if performance on English NWR predicted changes in English language skills from Year 1 and 2, then English NWR performance would be a significant predictor of Year 2 IPSyn scores. The IPSyn was chosen over other English language measures, as it provides a more detailed measure of spoken phrase structure complexity. The relationships between NWR and Year 2 IPSyn were plotted in Figure 1.
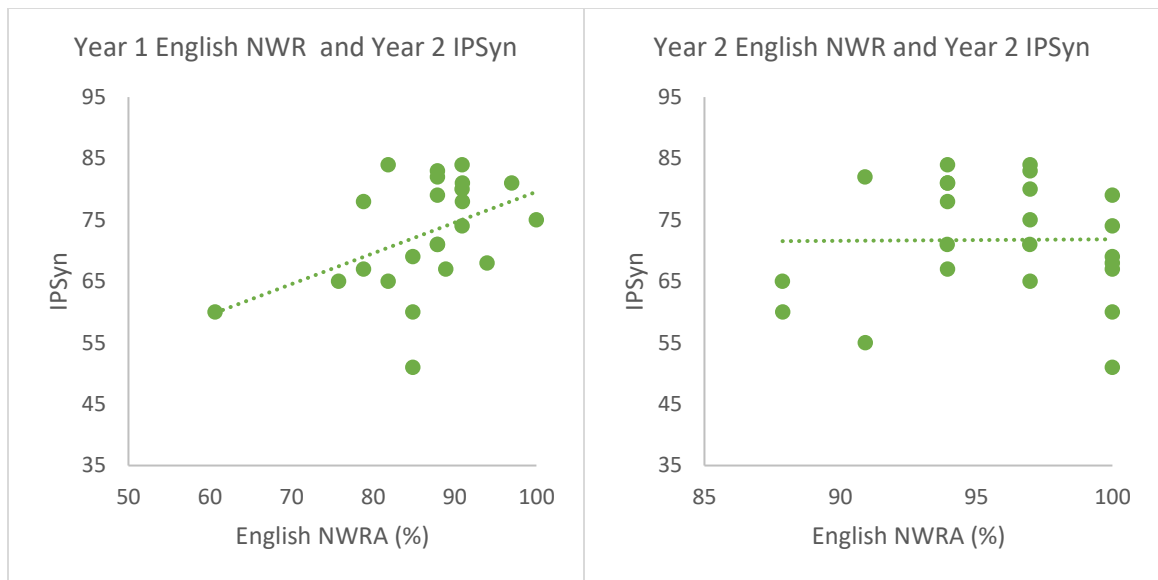
Figure 1: Plot of English NWRA and Year 2 IPSyn.

To control for other possible factors that might influence Year 2 IPSyn scores (English experience, age, and Year 1 IPSyn score), a multiple regression analysis was conducted that included these factors as covariates. Thus, Year 2 IPSyn was the dependent variable, while Year 1 English NWRA, Year 2 English NWRA, Year 1 IPSyn, and age were the independent variables. NWR for both years was included to assess both the concurrent and longitudinal predictive value of NWR. All predictors were entered into the regression

simultaneously. Data from twenty-two children were included in the analysis. One child was excluded from this analysis due to missing data from Year 1 IPSyn.

We found that the regression model was significant (R=0.839, F=10.102, p<0.001). As seen on Table 3, age (p=0.040) and Year 1 IPSyn (p=0.028) significantly predicted Year 2 IPSyn. However, Year 1 (β=0.399, p=0.066) and Year 2 English NWRA (β=0.327, p=0.056) did not significantly predict Year 2 IPSyn, even though they approached significance at the p=0.05 level.

If age and Year 1 IPSyn were removed from the regression model, Year 1 English NWRA would become a significant predictor (β=0.665, p=0.004), while Year 2 English NWRA would remain insignificant (β=-0.399, p=0.064). Further, we ran a correlation which included all 23 children and found that Year 1 English NWRA had a positive significant correlation with Year 2 IPSyn (r(21)=0.449, p=0.032). On the other hand, Year 2 English NWRA was not significantly correlated with concurrent Year 2 IPSyn (r(21)=0.010, p=0.964). Of note, there was a ceiling effect for Year 2 English NWRA, with most scores reaching (30% [n=7]) or approaching 100% accuracy. In fact, only 2 out of the 23 children in Year 2 achieved English NWRA scores lower than 90% in Year 2; thus, reducing variability and compressing the score range.

Table 3: Standardized Regression Coefficients Predicting Year 2 IPSyn

Variable	B	Standard Error	β	T
Age	0.538	0.242	0.351	2.225*
Y1 IPSyn	0.374	0.156	0.399	2.404*
Y1 NWRA	37.002	18.863	0.327	1.962
Y2 NWRA	-73.976	36.109	-0.305	-2.049

Significant if $\alpha \leq 0.05$

Do bilingual children catch up to their monolingual peers over time, when given additional exposure to English?
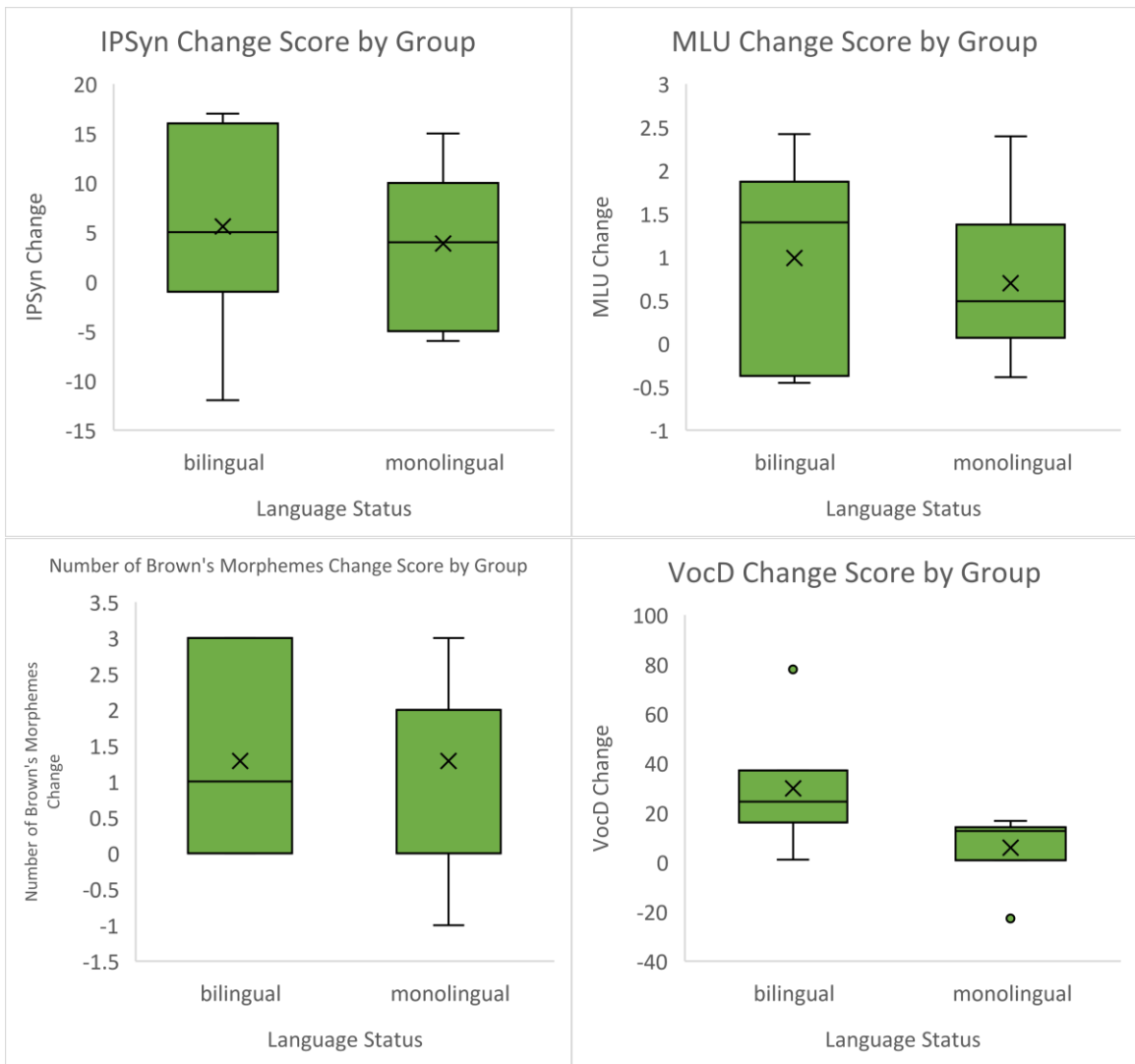
It was predicted that if cumulative exposure to a language results in better skill development in that language, then the monolingual group would improve more (i.e., achieve higher change scores) than the bilingual group due to their higher daily and cumulative English input. As this question required a comparison across groups, only children whose data and whose matches' data were available (n=14) were included in the analysis. English monolingual and bilingual groups' change scores for English measures (IPSyn, MLU, NBM, VocD, and English NWRA) are plotted in Figure 2 and further detailed in Table 4.

A Wilcoxon Rank Sum Test was used to determine if differences between groups were significant. The critical W was set to 21, as the number of children in the analysis was 14. This non-parametric test was chosen due to the high variability of scores in the bilingual group, which violated assumptions of normality of score distribution and skew according to the Shapiro-Wilk test of normality. As seen on Table 7, bilinguals' median change score for VocD (Mdn=24.5) was significantly higher (Z=33, p=0.013) than monolinguals' (Mdn=9.37). However, the difference between groups was not significant for the IPSyn (Z=47.5, p=0.522), MLU (Z=49, p=0.655), NBM (Z=49.5, p=0.693), or English NWRA (Z=42.5, p=0.343).

Table 4: Wilcoxon Rank Sum test results for group change scores.

	Bilingual Median Change Score n=7	Monolingual Median Change Score n=7	Wilcoxon Rank Sum Test W (Critical W=21)
IPSyn	5	4	47.5
MLU	1.40	0.53	49
NBM	1	1	49.5
VocD	24.48	9.37	33*
English NWRA	15.15	9.09	42.5

Significant group difference if α ≤ 0.05
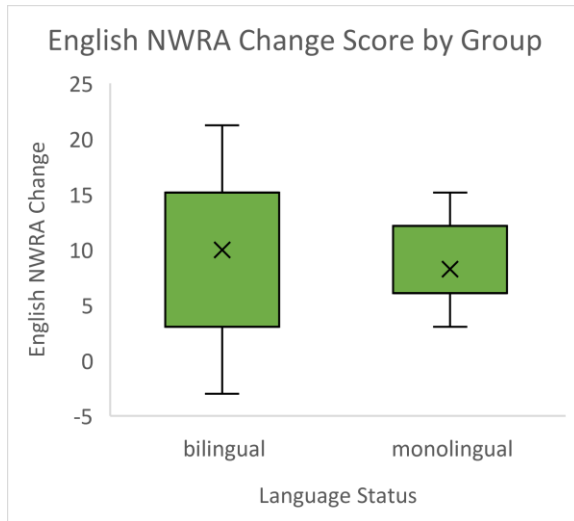
Figure 2: Box plots of English-language measures' change scores by group.

Group Scores at Year 1 and 2

In addition to comparing bilingual and monolingual change scores, we also compared their scores at Year 1 and 2 separately, and used a Wilcoxon Rank Sum Test to determine significance of differences seen between groups. As seen on Table 5, none of these measures differentiated monolinguals from bilinguals at Year 1 (IPSyn p=0.159, MLU p=0.225, NMB p=0.346, VocD p=0.085, English NWRA p=0.137) or Year 2 (IPSyn p=0.337, IPSyn p=0.159, MLU p=0.277, NMB p=0.081, VocD p=0.142, English NWRA p=0.405). However, differences between the two groups' Year 1 VocD (Z=39, p=0.085) and Year 2 NBM (Z=40.5, p=0.081) approached significance. In addition to the ceiling effect for Year 2 NWRA, there was also a ceiling effect for Year 2 NBM, as 33% (n=8) of children produced all possible 14 morphemes in their language samples.

Table 5: Wilcoxon Rank Sum test results for group scores at Year 1 and 2.

|	Bilingual Median Score n=7		Monolingual Median Score n=7		Wilcoxon Rank Sum Test W (Critical W=21)	
	Y1	Y2	Y1	Y2	Y1	Y2
IPSyn	63	69	73	74	41.5	45
MLU	3.02	4.01	3.54	4.75	43	44
NBM	12	13	13	14	45.5	40.5
VocD	33.75	50.98	41.33	47.5	39	41
English NWRA	84.85	96.47	87.87	97.83	41	42

None of the group differences were significant when α ≤ 0.05.

If a child has more experience in a language, will they do better in that language's NWR task and worse in the other language's NWR task?

It was hypothesized that, if parental reports of language exposure predict NWR performance, then higher percentages of experience in a language would result in better NWRA in that language. Conversely, it was also hypothesized that higher percentages of experience in a language would result in worse NWRA scores in the other language's NWR task. The relationships between language experience and NWRA are displayed in Figure 3.

A set of 8 correlations was computed to assess these relationships across both languages and years. Of note, only Spanish monolingual and Spanish-English bilinguals were included in the analyses involving Spanish NWR, as English monolinguals and English-Other bilinguals had not been given this task. To account for variability in the data, nonparametric correlations (Spearman's rho) were computed. Given the number of correlations run, a Bonferroni correction was used (α=0.05/8), lowering alpha to

α=0.00625. None of the relationships between language experience and NWR performance were significant or approached significance, even with a pre-correction predictability value (α=0.05). Correlation values are detailed in Table 6.
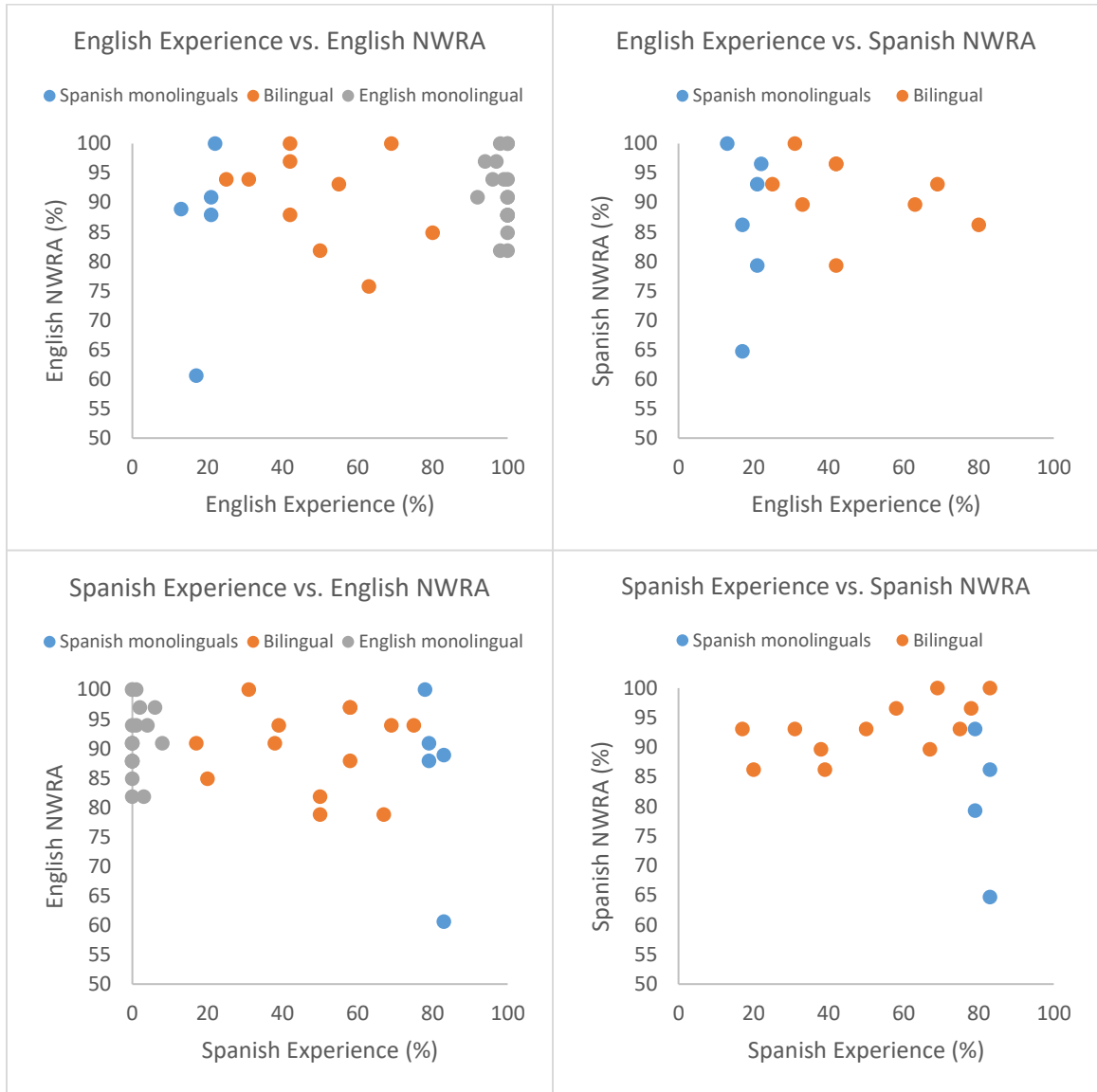
Figure 3: Scatterplots of language experience and nonword repetition performance for Year 1 and 2.

As expected, Spanish experience was negatively correlated with English NWRA at both Year 1 ($rs(20)=-0.065$, $p=0.773$) and Year 2 ($rs(14)=-0.085$, $p=0.753$). Also, English

experience was positively correlated with English NWRA (rs(14)=0.179, p=0.508) at Year 2 and negatively correlated with Spanish NWRA (rs(9)=-0.177, p=0.602) at Year 1.

However, unexpectedly, Spanish experience was negatively correlated with Spanish NWRA at both Year 1 (rs(9)=-0.002, p=0.995) and Year 2 (rs(5)=-0.119, p=0.799), although correlations were extremely weak. Also, English experience was positively correlated with Spanish NWRA (rs(5)=0.578, p=0.174) at Year 2 and negatively correlated with English NWRA (rs(9)=- (rs(20)=-0.071, p=0.753) at Year 1.

Table 6: Nonparametric correlations between language experience and NWR performance by year.

	English NWRA		Spanish NWRA (n)	
	Y1 n=22	Y2 n=16	Y1 n=11	Y2 n=7
English Experience	-0.071	0.179	-0.177	0.578
Spanish Experience	-0.065	-0.085	-0.002	-0.119

None of the correlations were significant.

English Language Experience and English Language Skill Measures

In addition to evaluating the impact of language experience on NWRA, we also evaluated its impact on performance on language sample measures for all participants. The relationship between language experience and performance on language sample measures is plotted in Figure 4. To further analyze these relationships, we ran a set of 8 Spearman's correlations, which are displayed on Table 7.

There were positive correlations between English experience and all English language sample measures considered. However, none of these correlations reached significance after Bonferroni corrections of alpha (α=0.05/4=0.0125) were made for computations at Year 1 (IPSyn p=0.037, MLU p=0.166, NMB=0.023, VocD p=0.110) or Year 2 (IPSyn p=0.185, MLU p=0.622, NBM p=0.026, VocD p=0.321). Prior to

Bonferroni corrections of alpha (α=0.05), Year 1 IPSyn (rs(17)=0.481, p=0.037) and both Year 1 (rs(20)=0.483, p=0.023) and Year 2 (rs(14)=0.554, p=0.026) NBM reached significance, but did not meet adjusted criteria for significance. As the correlation values are relatively high, lack of significance is likely due to reduced statistical power given small sample size.

Table 7: Correlations between English experience and English language measures at Year 1 and 2.

	IPSyn (n)		MLU		NBM		VocD	
	Y1 **n=19**	**Y2** **n=16**	**Y1** **n=22**	**Y2** **n=16**	**Y1** **n=19**	**Y2** **n=16**	**Y1** **n=19**	**Y2** **n=16**
English Experience	0.481*	0.3494	0.306	0.133	0.483*	0.554*	0.350	0.265

**Significant if α ≤ 0.00625, *Significant if α ≤ 0.005
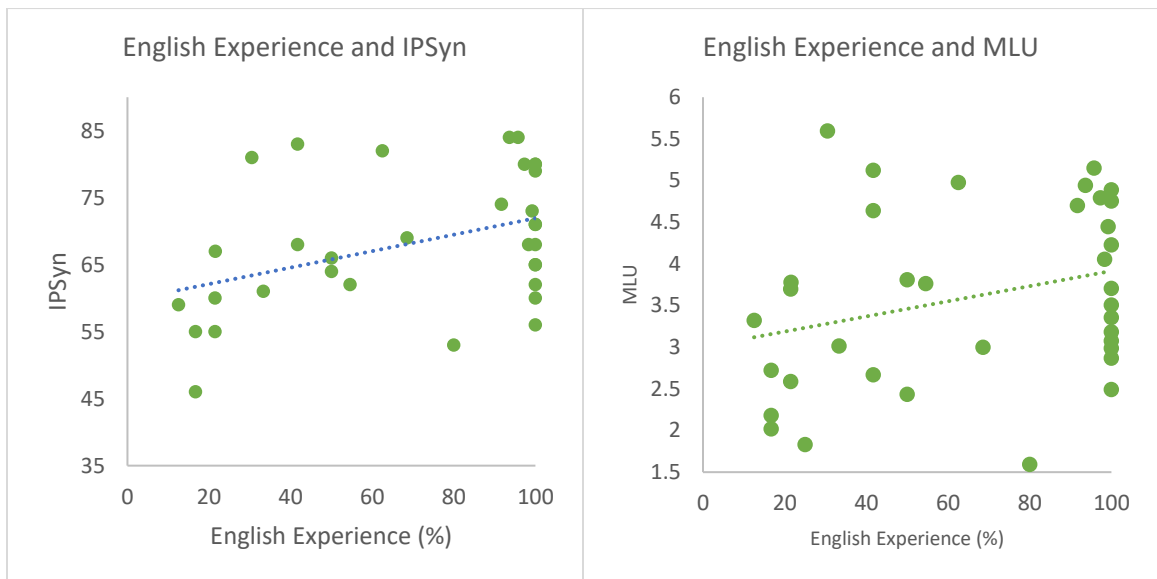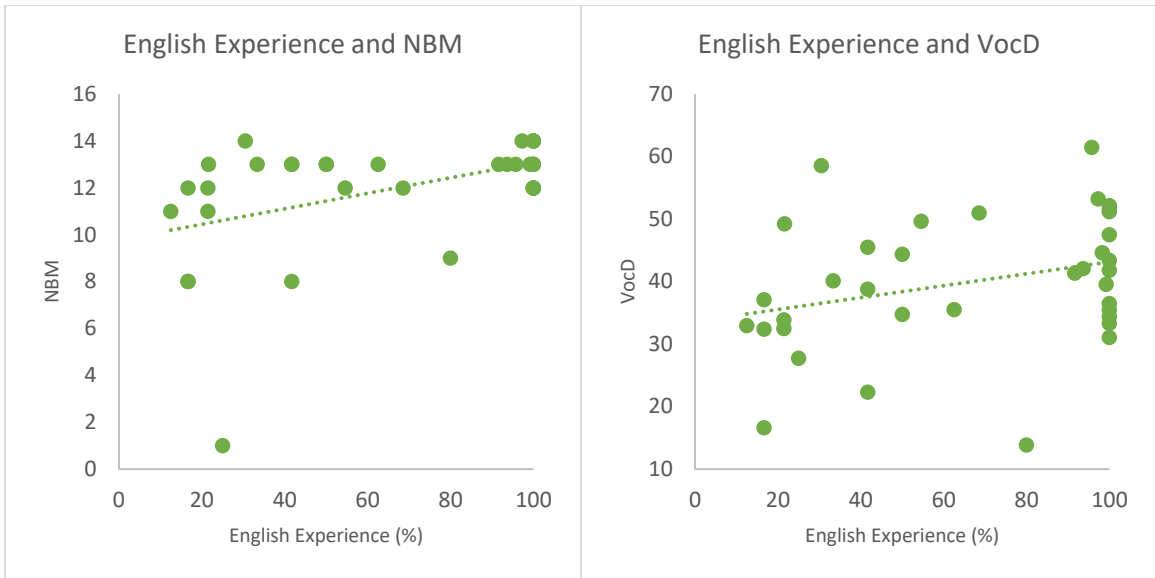
Figure 4: Relationships between English experience and English language measures.

Chapter 4: Discussion

Findings

The main aim of this study was to examine the factors that affect bilingual children's performance on language measures. The secondary aim of this study was to assess the value of NWR in assessment of bilingual children. To do this, we asked and found the following:

Does performance on NWR predict spoken English language skills?

Given that Year 1 English NWRA had a strong, positive, near significant correlation with Year 2 IPSyn, when controlling for age and baseline abilities, we conclude that performance on English NWR can predict some aspects of English language skills a year later. That correlations between Year 1 NWRA and Year 2 IPSyn did not reach adjusted levels of significance might have been due to lack of statistical power inherent in samples as small as the one we followed. However, given that age and Year 1 IPSyn were stronger predictors in the multiple regression model, we conclude that performance on English NWR is not as useful of a predictor of future language skill as are age and baseline abilities on the same measure (the IPSyn).

The fact that baseline abilities were a strong predictor was no surprise, as children who started out with better morphosyntactic abilities are likely to have been at an advantage during Year 2 because they started at a higher level. Similarly, age was expectedly an influential factor because older children have had more time to acquire both more exposure and more practice in a language, so they should have more highly developed skills. Age has been linked to changes not only in IPSyn, but also MLU and measures of lexical diversity in bilinguals (Ooi & Wong, 2012).

On the other hand, the *concurrent* predictive value of NWR was not as clear, based on our findings. This implies that using NWR to augment other diagnostic measures during a single evaluation may not solve existing problems with potential bias in current practices. The ceiling effect for Year 2 English NWRA might have affected the potential for a positive linear relationship with Year 2 IPSyn. High scores on this task were expected, as the children in this study were typically-developing (Dollaghan & Campbell, 1998). However, given that we used a NWR version that had been designed for very young children (~22 months old), this task might have been too easy once children got older (Parra et al., 2011). During Year 2, the children in our study were almost two to three times older (42-62 months old) than in the Parra et al. (2011) study.

Thus, we conclude that our findings support the existing research in which NWR performance was found to have predictive value in monolingual and bilingual children (Dispaldro et al., 2013; Hoff et al., 2008; Parra et al., 2011). In the case of Parra et al. (2011), NWR was shown to have longitudinal validity for vocabulary after 3 months. Our study extends this longitudinal period to one year, though the prediction might not be as strong. In the case of Dispaldro et al. (2013), NWR was shown to have concurrent validity with other measures of grammatical abilities. While we our findings might not support NWR's concurrent validity, they do not negate it. Additionally, since Year 1 English NWR almost reached significance as a predictor for Year 2 IPSyn, we conclude that our findings support research in which NWR performance was linked to grammatical complexity abilities.

Do bilingual children catch up to their monolingual peers over time, when given additional exposure to English?

We found that the only measure on which bilinguals significantly improved, compared to monolinguals, when given additional exposure to English over time, was VocD. In contrast, bilinguals did not significantly improve compared to monolinguals on any other English language measure (IPSyn, MLU, NBM, NWRA). However, the lack of significance, once again, may have been due to reduced statistical power as well as clear ceiling effects for both NBM and NWRA.

While this suggests that bilinguals' variety in English vocabulary improves with increased exposure to English over time, it would not be justifiable to interpret this as bilinguals catching up to monolinguals. After all, we did not find any significant differences between groups on any of the English language measures (IPSyn, MLU, NMB, VocD, NWRA) at Year 1 or Year 2. In other words, from the start of the study (Year 1), bilinguals were "caught up," and they remained this way during Year 2. These results are different from what we had expected, as we originally thought that higher cumulative exposure overall would put monolinguals at an advantage and enable them to improve more than bilinguals. However, it was not the group with higher cumulative exposure overall (monolinguals) but the group with higher cumulative increase in exposure (bilinguals) that improved on the English vocabulary measure more.

This narrow gap in performance between groups might seem to suggest that language experience does not put bilinguals at a disadvantage on English language measures. However, the large amount of variability in the bilingual group suggests that this may not be true for all children. Even with increased exposure to English, bilingual children

still might not perform like their monolingual peers. In fact, some bilingual toddlers may be 6 months to 1 year behind their monolingual peers on some aspects of English development (Hoff & Ribot, 2017).

Our findings contrast with findings from Gutiérrez-Clellen et al. (2008), in which there was a difference in performance on narrative LSA between typically-developing monolinguals and typically-developing bilinguals. This difference in findings might be attributed to wide methodological differences. For instance, the children in our study were younger (2;4-4;2 in Year 1 and 3;5-5;2 in Year 2) that those in their study (4;5-6;5) and we used different measures (IPSyn, MLU, NBM, VocD) than they did (finite verb morphology). Finally, the elicitation task differed (play versus narrative). However, we would argue that the variability in experience that we found in the bilingual group is a more influential factor, as use of language status (binary monolingual vs. bilingual categorization) neglects to consider the fullness of each child's language experience.

If a child has more experience in a language, will they do better in that language's NWR task and worse in the other language's NWR task?

The lack of significance and generally low correlations between English and Spanish experience and NWRA confirms the value of NWR for assessing underlying language abilities instead of language-specific knowledge and abilities (Boerma et al., 2015). Thus, our findings would suggest that the NWR task we used is among the types of NWR tasks that are not greatly influenced by language experience, such as the French NWR task in Thordardottir and Brandeker (2013), the French LITMUS NWR task in de Almeida et al. (2017), and the German LITMUS NWR task in Haman and Abed Ibrahim (2017). Interestingly, many of the studies in which language experience did appear to have

an impact on NWR performance were English and Spanish versions of the task (Gibson et al., 2015; Parra et al., 2011; Windsor et al., 2010). Even more surprising was the fact that Parra et al. (2011), the study from which we took the NWR stimuli for this study, observed an influence of language experience while we did not. This suggests that the mixed findings with regards to the influence of language experience on NWR performance might not just be due to the stimuli, but also the variability in participants that characterizes bilingual populations. Additionally, the children in Parra et al. were much younger and ceiling effects were not observed in that study, unlike ours.

On the other hand, the relatively high correlations between English experience and English language sample measures, especially the IPSyn and NBM, suggests that (not surprisingly) English experience does influence performance on LSA. The correlations between English experience and IPSyn scores, and between English experience and NBM presumably reached significance with a pre-corrected alpha value because both the IPSyn and NBM are measures that are more reliant on English-specific features. These findings agree with previous research, which has found that language experience influenced performance on narrative LSA measures (Gutierréz-Clellen, 2002). Nevertheless, MLU and VocD had lower, non-significant correlations with English experience, presumably because they are a more general measure of language development and can be more easily adapted for other languages. This contrasts with Jacobson and Walden (2013), who found that language experience influenced measures of lexical diversity. One of the secondary aims of this study was to evaluate the use of NWR for assessment of younger children. Given that at year 1, children were able to complete the task, we conclude that NWR is a viable tool for toddlers. Our results agree with Parra et al.'s (2011) study that found NWR

could be used with children as young as 20-22 months. Nevertheless, considering the ceiling effects on this study's NWR task, it could be argued that this NWR task was too easy for the children in this study. This might be because the children in this study were older (28-months-old and over) than the children in Parra et al.'s (2011) study (22-months-old).

Limitations

It must be noted that this study is limited by small numbers of observations. This is not only because of the small number of participants, but also because of missing data points for individual children, which could also exacerbate problems with the representative nature of the data. The small sample sizes in this study also could have affected our results, rendering some correlations insignificant simply because of inadequate power. Reasons for missing data varied from inability to locate the child's file, to video and audio issues in the file, or not enough utterances to compute the IPSyn. As previously mentioned, children were excluded from certain analyses for a variety of reasons. For instance, if they had only been tested in year 1, they were excluded from analyses involving change scores. Similarly, a common reason for exclusion from analyses concerned with language experience was the lack of parental report of hours of exposure.

As previously mentioned, the ceiling effect observed in the NWR tasks may have also impacted this study's findings. For example, since 43-47% of children achieved scores over 90% in English and Spanish NWR during Year 1, this may have compressed the data and diminished its predictive value for Year 2 IPSyn scores. Likewise, changes in English NWR between English monolinguals and bilinguals might have been significant if English

NWRA had not been limited in score range. Additionally, language experience might have predicted performance if NWRA scores had not been so condensed.

Further, many of the questions relied on parental report of exposure to be an accurate measure of language experience. However, our calculations of language experience were solely based on hours of exposure reported on the parental questionnaire and neglected to consider other factors that affect language experience, such as language of interaction with communication partners, and language experience prior to year 1. Relatedly, the classification of children as monolingual or bilingual was not always consistent. In fact, two children who had originally been recruited into the study as monolingual reported bilingual-levels of exposure to English and another language (i.e., Spanish and Persian), while three children who had been recruited as bilinguals reported Spanish-monolingual levels of exposure.

Conclusion

In summary, our findings suggest that language experience, age, and baseline abilities affect performance on LSA measures over time. This study also affirmed that language-specific NWR can be a valuable tool for assessment of underlying language abilities, instead of language-specific knowledge, as language experience was not significantly correlated with NWRA. Finally, this study found that NWR has some longitudinal predictive value for grammatical abilities observed a year later.

Clinically, the variability in the bilingual group's performance highlights the importance of clinicians thoroughly considering their clients' individual language experience level, as opposed to thinking of them in binary terms (i.e., monolingual vs. bilingual). Additionally, considering that NWR was not as strong of a predictor as other

factors, clinicians should not rely on NWR alone to gain an understanding of general language abilities. Finally, when administering NWR tasks, it may be best to use a Quasi-Universal version of the task, as Boerma et al.(2015) suggest. However, if a clinician is working with younger children, and has concerns about time or participation, administration of either the English of Spanish NWR task from Parra et al. (2011) may be acceptable. These NWR versions were designed for younger children and neither version of the task appeared to be influenced by language experience. However, both may be relatively easy tasks once children reach 3;5-5;2 (the age of our participants in Year 2), as evidenced by the ceiling effect, and thus might not tap underlying language learning aptitude or proficiency well. One way to address this issue might be to score a subset of longer nonwords (e.g., only scoring nonwords that are 3 syllables or longer), as studies have found that repetition of longer nonwords is more challenging (Boerma et al., 2015; Roy & Chiat, 2004).

This study also raises more questions for future research. For instance, what factors of language experience (e.g., hours of exposure, languages of and with communication partners, age of acquisition) are most influential for a child's language skills? These qualitative aspects of language experience merit further exploration, as they will improve quantitative measures of language experience. Additionally, it might be interesting to explore performance with regards to ranges of language experience in order to more easily distinguish effects of language status. Future research should also investigate the role of other factors (e.g., SES and gender) on performance, in addition to those considered in this study (i.e., language experience, age, and baseline abilities); as well as their impact not only on LSA measures, but also on NWRA.

Finally, this study primarily focused on the influence of language experience on English language measures. However, it would be beneficial to also assess the influence of language experience on assessment measures in the child's other languages. While this study looked at the influence of language experience on Spanish NWR, future studies could look at the effect of language experience on Spanish language sample measures. This would allow for comparison of the influence of language experience on measures that are dependent versus semi-independent of language specific knowledge, as well as inform the question of whether children with lower levels of language experience in one language (e.g., English), but high levels in another language (e.g., Spanish), perform better in the language they have higher exposure to. Hoff and Ribot (2017) found that performance on measures of vocabulary (i.e., the English and Spanish-Bilingual editions of the *Expressive One-Word Picture Vocabulary Test*) were related to parental reports of exposure at home, which suggests that children would indeed perform better in the language they have more exposure to. However, this question requires further exploration for other types of language measures, as it would better inform decisions on which language and assessment types to prioritize during the diagnostic process.

Appendices

Appendix A: Informal questionnaire year 1

Subject ID#: _____   **Date:** _____   **VISIT: 1   2   3   4   5**

Person completing form (please circle one):

Mother   Legal Guardian   Other: _____

The following questionnaire requests case history information that may be relevant to the research questions being examined in the study. This information will remain completely confidential and will only be available to the researchers conducting the study. If any of this information is used in the final research report, all identifying information will be removed.

Please fill out the following information as completely as possible.

1.1 Child's gender: M / F   **1.2 Date of Birth: ___/___/_____**

Please indicate the race/ethnicity of each parent or legal guardian and the participant. Check all that apply. These data are for reporting purposes only.

2.1 Parent / Legal guardian 1:
____ African American   ____ Hispanic   ____ Caucasian (white)
____ Asian   ____ Native American   ____ Pacific Islander
____ Other: _____
2.1.1 My Native Language: _____

2.2 Parent / Legal Guardian 2:
____ African American   ____ Hispanic   ____ Caucasian (white)
____ Asian   ____ Native American   ____ Pacific Islander
____ Other: _____
2.2.1 Native Language of Parent / Legal Guardian 2: _____

2.3Child:
____ African American   ____ Hispanic   ____ Caucasian (white)
____ Asian   ____ Native American   ____ Pacific Islander
____ Other: _____

2.4 Mother: What is your highest completed educational level? (Circle one)
     A. No formal schooling
     B. Grade school: (Some / Completed)
     C. High school: (Some / Completed)
     D. College or associate degree: (Some / Completed)
     E. Graduate degree(s):   (Masters / Doctorate)   Other:_____

1

2.5 Number of caregivers (e.g., mother, father, grandparent, nanny, others) currently in household: ____

2.6 Number of siblings: ____
    2.6.1   Ages: _____

3. Primary language spoken in the home: _____

    3.1 Is your child exposed to any other languages during the day?   Y / N
        3.1.1 If so, which one(s)? _____
        3.1.2 For approximately how many hours a day? ____

    3.2 Has your child spent one month or longer outside of the U.S. during the past year?   Y / N
        3.2.1 Where? _____
        3.2.2   For how long? _____

4. Where does your child spend most of the day Monday through Friday?

    4.1 At home with me and/or another primary caregiver (please name: _____)
        If there is an additional primary caregiver, is this person:
        4.1.1 __ a relative   __ a "nanny" or unrelated child care provider
        4.1.2 Is this person a native speaker of English?   Y / N
        4.1.3 If not, what is this person's native language? _____

    4.2 In a day care with ____ other children ages ___ to ___ (if your child is in a classroom inside a
        larger facility, estimate the number of other children in your child's "classroom")

5.   How many TVs are in the household? ____

    5.1 Please estimate how many hours per day the TV is on ____
    5.2 Please estimate how many hours of screen time (TV/movies/tablet animations) your child
        watches per day ____
    5.3 What is your child's favorite TV show(s)? _____
    5.4 What is your child's favorite DVD or video? _____

6.   Does your child play computer/tablet games?   Y / N

    6.1 If so, how many hours per day? ____
    6.2 What is your child's favorite computer/tablet game? _____

7.   Does your family own any pets?   Y / N

    7.1 If so, how many? ____
    7.2 What kind(s)? _____

8. Do you and your child ever read books together?   Y / N

    8.1 If so, on average, *how many books per week* does your child have read to him/her? _____

    8.2 Please estimate: How many books do you own? _____

    8.3 How many children's books does your child have? _____
    (do not count books that are primarily used by your child's siblings)

    8.4 Does your child ever pretend to read books?   Y / N

    8.5 Does your child have a favorite book?   Y / N
    8.5.1   If so, what is it? _____

9. On average, how much time per day do you (and other adults in your household) spend in one-to-one **conversation** with your child? _____

10. On average, how many hours per day do you (and other adults in your household) spend in one-to-one **play** with your child? _____

11. On average, how many hours per day does your child spend playing alone? _____

12. On average, how many hours per day does your child spend playing with other children? _____

13. Does your child have a history of ear infections?   Y / N

    13.1   How many? _____

    13.2   How many in the past year? _____

14 Do you have any concerns about your child's development?   Y / N

    14.1 If so, please describe: _____
    _____
    _____
    _____

15 Has your child had any major medical events?   Y / N

    15.1 If so, please describe: _____
    _____
    _____
    _____

3

15.2 At what age(s)? _____

15.3 Number of hospitalizations: _____

15.4 Length of hospital stay(s): _____

16 Has your child ever been diagnosed with a condition that might impact his/her language development or school performance?   Y / N

16.1 If so, please describe: _____

_____

_____

_____

17 Is there any history of language and/or learning disabilities in your immediate family (you or your child's other biological parent, grandparents, aunts/uncles/cousins, siblings) such as problems paying attention, reading, speaking (stuttering or articulation problems), learning, or other school problems?
Y / N

17.1 If so, please describe: _____

_____

_____

_____

18 Have any of your child's siblings had delays in meeting developmental milestones?   Y / N

18.1 If so, please describe: _____

_____

_____

_____

SLAS1.   Please rate your child's language skills in English relative to those of other native-English-speaking children his or her own age:

1	2	3	4	5	6	7
very low			normal for age			very high

SLAS2.   If your child is learning Spanish
Please rate your child's language skills in Spanish relative to those of other Spanish-speaking children his or her own age:

1	2	3	4	5	6	7
very low			normal for age			very high

SLAS3.   Please rate your child's social skills relative to those of other children his or her own age:

1   2   3   4   5   6   7

very low	normal for age	very high

.

SLAS4.   Please rate your child's ability to say sentences clearly enough to be understood by strangers:

1   2   3   4   5   6   7

very low	normal for age	very high

SLAS5.   Please rate the number of words your child knows, compared to other children his or her age:

1   2   3   4   5   6   7

very low	normal for age	very high

SLAS6.   Compared to other children, please rate how well your child responds when you call his or her name:

1   2   3   4   5   6   7

very low	normal for age	very high

SLASF.   Compared to other children his/her age, how fluent is your child's speech (free of repetitions, hesitations, and use of fillers such as "um" or "uh")

1   2   3   4   5   6   7

very low	normal for age	very high

19. Is there anything else you'd like us to know about your child, so that we can appreciate his/her development throughout this study better?

_____
_____
_____
_____

5

Follow up assessment questionnaire

Subject ID#: _____   **Date:** _____   **VISIT: 1 2 3 4 5**

Person completing form (please circle one):

Mother   **Legal Guardian**   **Other:** _____

The following questionnaire requests case history information which may be relevant to the research questions being examined in the study. This information will remain completely confidential and will only be available to the researchers conducting the study. If any of this information is used in the final research report, all identifying information will be removed.

Please fill out the following information as completely as possible. <u>We have asked most of these questions before, but we realize that some things may have changed since we last asked these questions.</u>

1. Mother: What is your highest completed educational level? Has it changed since your last visit? (Circle one)
   A. No formal schooling
   B. Grade school: (Some / Completed)
   C. High school: (Some / Completed)
   D. College or associate degree: (Some / Completed)
   E. Graduate degree(s):   (Masters / Doctorate)   Other:_____

   1.1 Number of caregivers (e.g., mother, father, grandparent, nanny, others) <u>currently</u> in household: ____

   1.2 Number of siblings: ___
       1.2.1 Ages: _____

2. Is your child now exposed to any other languages during the day?   Y / N
   2.1.1   If so, which one(s)? _____
   2.1.2   For approximately how many hours a day? _____

   2.2 Has your child spent one month or longer outside of the U.S. during the past year? Y / N
       2.2.1 Where? _____
       2.2.2 For how long? _____

3. Where does your child spend most of the day Monday through Friday?

   3.1 __ at home with me and/or another primary caregiver (please name: _____)
       If there is an additional primary caregiver, is this person:
       3.1.1   __ a relative   __ a "nanny" or unrelated child care provider
       3.1.2   Is this person a native speaker of English? Y / N
       3.1.3   If not, what is this person's native language? _____

1

3.2 ___ in a day care with ____ other children ages ___ to ___ (if your child is in a classroom inside a larger facility, estimate the number of other children in your child's "classroom")

      3.2.1 Are languages other than English spoken in the day care Y / N

4   How many TVs are in the household today? _____

    4.1   Please estimate how many hours per day the TV is on _____

    4.2   Please estimate how many hours of screen time (TV/movies/tablet animations) your child watches per day_____

    4.3   What is your child's current favorite TV show(s)? _____

    4.4   What is your child's current favorite DVD or video? _____

5   Does your child play computer/tablet games? Y / N

    5.1   If so, how many hours per day? _____

    5.2   What is your child's favorite computer game? _____

6   Does your family own any pets? Y / N

    6.1   If so, what kind(s)? _____

    6.2   How many? _____

7   Do you and your child read books together? _____ How many times per week? ____

    7.1 On average, *how many different books per week* does your child have read to him/her? _____

    Please estimate:

    7.2 how many books you own _____

    7.3 how many children's books your child owns _____

        (do not count books that are primarily used by your child's siblings)

    7.4 Does your child ever pretend to read books? Y / N

    7.5 Does your child have a favorite book? Y / N

      7.5.1 If so, what is it? _____

8   On average, how much time per day do you (and other adults in your household) spend in one-to-one conversation with your child? _____

9   On average, how many hours per day do you (and other adults in your household) spend in one-to-one play with your child? _____

10 On average, how many hours per day does your child spend playing alone? _____

11 On average, how many hours per day does your child spend playing with other children? _____

12 Does your child have a history of ear infections? Y / N

    12.1 How many? _____

    12.2 How many in the past year? _____

2

13 Has your child been identified with any major conditions (e.g., medical, developmental, learning) since the last visit? Y / N

    13.1 If so, please describe: _____

    _____

    _____

    13.2 Is your child receiving any treatment or intervention for this condition/these conditions? Y / N

    13.2.1 If so, please describe: _____

    _____

    _____

14. Since the last visit, have you noticed any change in your child's development? Y / N

    14.1 If so, please describe: _____

    _____

    _____

15. Please rate the improvement of your child's development from last year:

1	2	3	4	5	6	7
very low			normal for age			very high

16. Since the last visit, has your child had any major medical events? Y / N

    16.1 If so, please describe: _____

    _____

    _____

    16.2 At what age(s)? _____

    16.3 Number of hospitalizations: _____

    16.4 Length of hospital stay(s): _____

17. Do you have any current concerns about your child's speech and language development?

    17.1 If so, please describe: _____

    _____

    _____

18. Have any of your child's biological siblings had delays in meeting developmental milestones?

    18.1 If so, please describe: _____

    _____

    _____

3

19. If your child is learning more than one language, please rate your child's proficiency, from beginning learning stage (1) to sounding "native-like" (10)

    19.1.1 My child's proficiency in learning English is

1	2	3	4	5	6	7	8	9	10
Learning									native
stage									like

    19.1.2 My child's proficiency in learning Spanish is

1	2	3	4	5	6	7	8	9	10
Learning									native
stage									like

    19.2 Has the relative balance between my child's use of English/Spanish changed since our last visit?
        19.2.1 If so, please describe: _____
        _____
        _____

20. If were concerned about your child's language or fluency development when you last saw us, has this changed since your last visit? (please circle your response)
    a. My child's expressive language or fluency (e.g., stuttering) has improved since the last visit
    b. My child's expressive language or fluency remains pretty much the same
    c. I am more concerned by my child's language or fluency than I was at the last visit
        20.1 If so, please describe your concern or question:
        _____
        _____
        _____

4

SLAS1.   Please rate your child's language skills in English relative to those of other native-English-speaking children his or her own age:

1	2	3	4	5	6	7
very low			normal for age			very high

SLAS2.   If your child is learning Spanish
Please rate your child's language skills in Spanish relative to those of other Spanish-speaking children his or her own age:

1	2	3	4	5	6	7
very low			normal for age			very high

SLAS3.   Please rate your child's social skills relative to those of other children his or her own age:

1	2	3	4	5	6	7
very low			normal for age			very high

SLAS4.   Please rate your child's ability to say sentences clearly enough to be understood by strangers:

1	2	3	4	5	6	7
very low			normal for age			very high

SLAS5.   Please rate the number of words your child knows, compared to other children his or her age:

1	2	3	4	5	6	7
very low			normal for age			very high

SLAS6.   Compared to other children, please rate how well your child responds when you call his or her name:

1	2	3	4	5	6	7
very low			normal for age			very high

SLASF.   Compared to other children his/her age, how fluent is your child's speech (free of repetitions, hesitations, and use of fillers such as "um" or "uh")

1	2	3	4	5	6	7
very low			normal for age			very high

Is there anything else you'd like us to know about your child, so that we can appreciate his/her development throughout this study better? If you have any questions about your child's communication or development, or about our study, please feel free to ask us here:

_____

_____

5

52

Appendix C: English & Spanish RWR and NWR words

	bɪm
	mʌp
dɔg	kɔg
dʒus	bus
kæt	dʒæt
bʊk	dʊk
bəlun	tʃʌlun
kʊki	pʊki
pʌpi	kʌpi
tʃɪkən	bɪkən
bənænə	bədʒæpɔp
tɛləfoʊn	tɛlɪnə
lalipap	laləməz
pədʒaməz	pənæfoʊn

	col
pan	mon
luz	
tren	lan
sol	truz
vaca	sen
gato	pol
mesa	vato
leche	meca
muñeca	lesa
gallina	gache
caballo	gañeca
pelota	mullina
	peballo
	calota

Bibliography

Armon-Lotem, S., & Meir, N. (2016). Diagnostic accuracy of repetition tasks for the identification of specific language impairment (SLI) in bilingual children: evidence from Russian and Hebrew. *International Journal of Language & Communication Disorders, 51*(6), 715-731.

Bedore, L., Peña, E., Anaya, J., Nieto, R. Lugo-Neris, M., & Baron, A. (2018). Understanding disorder within variation: production of English grammatical forms by English language learners. *Language, Speech, and Hearing Services in Schools. 49*(1), 277-291.

Blom, E., de Jong, J., Orgassa, A., Baker, A., & Weerman, F.(2013). Verb inflection in monolingual Dutch and sequential bilingual Turkish–Dutch children with and without SLI. *International Journal of Language & Communication Disorders, 48*(1), 1–12.

Boerma, T., Chiat, S., Leseman, P., Timmermeister, M., Wijnen, F., & Blom, E. (2015). A quasi-universal nonword repetition task as a diagnostic tool for bilingual children learning Dutch as a second language. *Journal of Speech, Language, and Hearing Research*, *58*(6), 1747–1760. https://doi.org/10.1044/2015_JSLHR-L-15-0058

Brown, R. (1973). Development of the first language in the human species. *American Psychologist*, *28*(2), 97–106. doi: 10.1037/h0034209

Campbell, T., & Dollaghan, C. (1997). Reducing Bias in Language Assessment: Processing-Dependent measures. *Journal of Speech Language Hearing Research, 40*(1), 519-525.

Chiat, S., Roy, P. (2007). The Preschool Repetition Test: An evaluation of performance in typically developing and clinically referred children. *Journal of Speech Language Hearing Research, 50*(2). 429-443.

Cummins, J. (1981). Empirical and theoretical underpinnings of bilingual education. *The Journal of Education, 163*(1), 16–29.

de Almeida, L., Ferré, S., Morin, E., Prévost, P., dos Santos, C., Tuller, L., Zebib, R., & Barthez, M. (2017). Identification of bilingual children with specific language impairment in France. *Linguistic Approaches to Bilingualism, 7*(3). 441-358.

DeAnda, S., Bosch, L., Poulin-Dubois, D., Zesiger, P., & Friend, M. (2016) The Language Exposure Assessment Tool: Quantifying language exposure in infants and children. *Journal of Speech, Language, and Hearing Research, 59*(6). 1346-1356.

Dispaldro, M., Deevy, P., Altoé, G., Benelli, B., & Leonard, L. (2011). A cross-linguistic study of real-word and non-word repetition as predictors of grammatical competence in children with typical language development. *International Journal of Language & Communication Disorders, 46*(5). 564-578.

Dollaghan, C., & Campbell, T. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research, 41*(5), 1136-1146.

Dunn, M., Flax, J., Sliwinski, M., & Aram, D. (1996). The use of spontaneous language measures as criteria for identifying children with specific language impairment: an attempt to reconcile clinical and research incongruence. *Journal of Speech and Hearing Research, 39*(3), 643–54.

Eaton, C., Newman, R., Ratner, N., & Rowe, M. (2015). Non-word repetition in 2-year olds: Replication of an adapted paradigm and useful methodological extension. *Clinical Linguistics & Phonetics, 2015*(1). 1-13.

Gibson, T., Summers, C., Peña, E., Bedore, L., Gillam, R., & Bohman, T. (2015). The role of phonological structure and experience in bilingual children's nonword repetition performance. *Bilingualism: Language and Cognition, 18*. 551–560.

Guevara, S. (2020). *An analysis of code switching events in typically developing Spanish-English bilingual children.* [Unpublished master's thesis]. University of Maryland.

Guiberson, M., & Rodríguez, B. L. (2013). Classification accuracy of nonword repetition when used with preschool-age Spanish-speaking children. *Language, Speech & Hearing Services in Schools*, *44*(2), 121–132.

Gutierréz-Clellen, V., Restrepo, M., Bedore, L., & Peña, E. (2000). Language sample analysis in Spanish-speaking children. *Language, Speech, and Hearing Services in Schools, 31*(1), 88-98.

Gutierréz-Clellen, V. (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*(1), 267–288.

Gutiérrez-Clellen, V., Simon-Cereijido, G., & Wagner, C. (2008). Bilingual children with language impairment: A comparison with monolinguals and second language learners. *Applied Psycholinguistics, 29*(1), 3-19.

Gutiérrez-Clellen, V., & Simon-Cereijido, G. (2012). The use of spontaneous language measures as criteria for identifying children with specific language impairment: an attempt to reconcile clinical and research incongruence. *Seminars in Speech and Language, 30*(4), 234–234.

Hamann, C., & Abed Ibrahim, L. (2017). Methods for identifying specific language impairment in bilingual populations in Germany. *Frontiers in Communication, 2*(16). 1-19.

Hoff, E., Core, C., & Bridges, K. (2008). Nonword repetition assesses phonological memory and is related to vocabulary development in 20- to 24-month-olds. *Journal of Child Language, 35*(1). 903-916

Hoff, E., & Ribot, K. (2017). Language growth in English monolingual and Spanish-English bilingual children from 2.5-5 year. *Journal of Pediatrics, 190*(1). 241-245.

Individuals with Disabilities Education Act, 20 U.S.C. § 1412(a)(5) (2004)

Jacobson, P., & Walden, P. (2013). Lexical Diversity and Omission Errors as Predictors of Language Ability in the Narratives of Sequential Spanish–English Bilinguals: A

Cross-Language Comparison. *American Journal of Speech-Language Pathology, 22*(1). 554-565

Jalilevand, N., & Ebrahimipour, M. (2014). Three measures often used in language samples analysis. *Journal of Child Language Acquisition and Development, 2*(1). 1-12.

Kohnert, K., Windsor, J., & Yim, D. (2006). Do language-based processing tasks separate children with language impairment from typical bilinguals? *Learning Disabilities Research & Practice, 21*(1). 19-29.

Kohnert, K. (2010). Bilingual children with primary language impairment: Issues, evidence and implications for clinical actions. *Journal of Communication Disorders, 43*(2010), 456-473.

Laing, S., & Kamhi, A. (2003). Alternative assessment of language and literacy in culturally and linguistically diverse populations. *Language, Speech, and Hearing Services in Schools, 34(1), 44–55.*

Langdon, H. (1992). Speech and language assessment of LEP/bilingual Hispanic students. In H. W. Langdon (Ed.), *Hispanic children and adults with communication disorders: Assessment and intervention* (pp. 201–271). Gaithersburg, MD: Aspen

Libertus, M., Odic, D., Feigenson, L., & Halberda, H. (2015). A developmental vocabulary assessment for parents (DVAP): Validating parental report of vocabulary size in 2- to 7-year-old children. *Journal of Cognition and Development, 0*(0), 1-13.

MacWhinney, B. (2000). *The CHILDES project: tools for analyzing talk* (3rd ed.). Lawrence Erlbaum.

Malvern, D., & Richards, B. (2002). Investigating accommodation in language proficiency interviews using a new measure of lexical diversity. *Language Testing, 19*(1), 85-104.

Marchman, V., & Martínez-Sussmann, D. (2002). Concurrent validity of caregiver/parent report measure of language for children who are learning both English and Spanish. *Journal of Speech, Language, and Hearing Research,* 45. 983–97

Meir, N., Walters, J., & Armon-Lotem, S. (2017). Bi-directional cross-linguistic influence in bilingual Russian-Hebrew children. *Linguistic Approaches to Bilingualism, 7*(5), 514–553. https://doi.org/10.1075/lab.15007.mei

National Center for Education Statistics. (2020a). *English Language Learners in public schools.* Retrieved from https://nces.ed.gov/programs/coe/indicator_cgf.asp.

National Center for Education Statistics. (2020b). *Students with disabilities.* Retrieved from https://nces.ed.gov/programs/coe/indicator_cgg.asp.

Ooi, C., Wong, A. (2012). Assessing bilingual Chinese-English young children in Malaysia using language sample measures. *International Journal of Speech-Language Pathology, 14*(6), 499-508.

Paradis, J. (2007). Early bilingual and multilingual acquisition. In L. Wei & Auer, Peter, *Handbook of Multilingualism and Multilingual Communication* (p. 30).

Paradis, J., Emmerzael, K., & Duncan, T. S. (2010). Assessment of English language learners: Using parent report on first language development. *Journal of Communication Disorders, 43*(6), 474–497.

Paradis, J., Schneider, P., & Duncan, T. S. (2013). Discriminating children with language impairment among English-language learners from diverse first-language backgrounds. *Journal of Speech, Language, and Hearing Research, 56*(3), 971–981.

Parra, M., Hoff, E., & Core, C. (2011). Relations among language exposure, phonological memory, and language development in Spanish–English bilingually developing 2-year-olds. *Journal of Experimental Child Psychology, 108*(1), 113-125.

Place, S., & Hoff, E. (2011). Properties of dual language exposure that influence 2-year-olds' bilingual proficiency. *Child Development, 82*,1834–1849. doi:10.1111/j.1467-8624.2011.01660.x

Restrepo, M., Silverman, S. (2001). Validity of the Spanish Preschool Language Scale-3 for use with bilingual children. *American Journal of Speech-Language Pathology, 10*(1). 382-393.

Rojas, R., & Iglesias, A. (2009). Making a case for language sampling: Assessment and intervention with (Spanish-English) second language learners. *The ASHA Leader, 14*(3), 10–13. https://doi.org/10.1044/leader.FTR1.14032009.10

Roy, P., & Chiat, S. (2004). A prosodically controlled word and nonword repetition task for 2- to 4-year-olds. *Journal of Speech, Language, and Hearing Research, 47*(1). 223-234.

Scarborough, H. (1990). Index of Productive Syntax. *Applied Psycholinguistics, 11*(1), 1-22.

Scheidnes, M. (2018). A longitudinal comparison of object clitic production in the spontaneous language of L2 children and children with SLI. *Proceedings of the 42nd annual Boston University Conference on Language Development,* 694-797.

Sheng, L., Before, L., Peña, E., & Fiestas, C. (2012). Semantic development in Spanish-English bilingual children: Effects of age and language experience. *Child Development, 84*(3), 1034-1045.

Suni, E. (2020, September 24). *How much sleep do babies and kids need?* Sleepfoundation.org. Retrieved from https://www.sleepfoundation.org/children-and-sleep/how-much-sleep-do-kids-need.

Thordardottir, E., & Brandeker, M. (2013). The effect of bilingual exposure versus language impairment on nonword repetition and sentence imitation scores. *Journal of Communication Disorders, 46*(20131), 1-16.

Valenzuela, J., Copeland, S., Huaqing, C., & Park, M. (2006). Examining educational equity: Revisiting the disproportionate representation of minority students in special education. *Council for Exceptional Children, 72*(4), 425-441.

62

Windsor, J., Kohnert, K., Lobitz., K., & Pham, G. (2010). Cross-language nonword repetition by bilingual and monolingual children. *American Journal of Speech-Language Pathology, 19*(1), 298-310.

Winsler, A., Diaz, R., Espinosa, L., & Rodriguez, J. (1999). When learning a second language does not mean losing the first: Bilingual language development in low-income, Spanish-speaking children attending bilingual preschool. *Child Development, 70*(1), 349-362.

Yamasaki, B., & Luk, G. (2018). Eligibility for special education in elementary school: The role of diverse language experiences. *Language, Speech, and Hearing Services in Schools, 49*, 889-901.

Yang, J., MacWhinney, B., & Ratner, N. (2021). *The Index of Productive Syntax (IPSyn): Psychometric properties and suggested modifications* [Manuscript submitted for publication]. *Department of Hearing and Speech Sciences, University of Maryland*.