

2021

## IS PARENT REPORT OF BILINGUAL CHILDREN'S LANGUAGE ENVIRONMENT AS VALID AS WE THINK?

Gabrielle Zeyl  
University of Rhode Island, ellazeyl@gmail.com

Follow this and additional works at: <https://digitalcommons.uri.edu/theses>

---

### Recommended Citation

Zeyl, Gabrielle, "IS PARENT REPORT OF BILINGUAL CHILDREN'S LANGUAGE ENVIRONMENT AS VALID AS WE THINK?" (2021). *Open Access Master's Theses*. Paper 1951.  
<https://digitalcommons.uri.edu/theses/1951>

This Thesis is brought to you for free and open access by DigitalCommons@URI. It has been accepted for inclusion in Open Access Master's Theses by an authorized administrator of DigitalCommons@URI. For more information, please contact [digitalcommons@etal.uri.edu](mailto:digitalcommons@etal.uri.edu).

IS PARENT REPORT OF BILINGUAL CHILDREN'S LANGUAGE ENVIRONMENT  
AS VALID AS WE THINK?

BY  
GABRIELLE ZEYL

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE  
REQUIREMENTS FOR THE DEGREE OF  
MASTER OF SCIENCE  
IN  
SPEECH LANGUAGE PATHOLOGY

UNIVERSITY OF RHODE ISLAND

2021

MASTER OF SCIENCE IN SPEECH LANGUAGE PATHOLOGY

OF

GABRIELLE ZEYL

APPROVED:

Thesis Committee:

Major Professor

Alisa Baron

Vanessa Harwood

LeAnne Spino-Seijas

Brenton DeBoef  
DEAN OF THE GRADUATE SCHOOL

UNIVERSITY OF RHODE ISLAND  
2021

## **ABSTRACT**

Researchers have defined bilingual language proficiency and dominance in different ways for decades; however, to date there is no standard, systematic description used by the scientific community to quantify bilingual children's language skills. Parent report questionnaires have been regarded as a valid measure of a child's language skill. The purpose of this study is to explore the validity of parent report of their child's language use when compared to hand coded audio segments retrieved using the Language ENvironment Analysis (LENA) System. Additionally, this study explores potential differences in the number of conversational turns and amount of child-directed speech between monolingual and bilingual early elementary-aged children. Two third-grade students (1 monolingual and 1 Spanish-English bilingual) wore the LENA System for two days. Their parents completed the Bilingual Input-Output Survey (BIOS) to report on their child's language environment, including the amount of language their child hears and speaks. The amount of exposure and use in each language was calculated based on the hand coded segments recorded using the LENA device. These calculations were compared to parents' reported estimates of their child's language environment using the BIOS. For the bilingual participant analyzed, parent report did not appear to accurately estimate the child's language environment. Parents overestimated Spanish language exposure and use when compared to actual data retrieved from the LENA recordings. The monolingual participant heard more child-directed speech than the bilingual participant, yet the bilingual participant had nearly double the number of conversational turns. In order to

generalize these findings beyond these two participants, this question needs to be explored on a larger scale.

## **ACKNOWLEDGMENTS**

I want to thank my parents for their love and encouragement, who taught me the value of education and investing in myself. I would like to thank my Opa, who infected me with a love for critical thinking and philosophy. I would also like to thank my partner Jacob, who has supported me throughout this entire process. Additionally, I would like to thank Alisa Baron for her patience, guidance, encouragement and advice throughout my time as a graduate student. Finally, I would like to thank Victoria Garcia, who dedicated her time and bilingual brain to hand coding segments for this study.

## TABLE OF CONTENTS

<b>ABSTRACT</b> .....	ii
<b>ACKNOWLEDGMENTS</b> .....	iv
<b>TABLE OF CONTENTS</b> .....	v
<b>LIST OF TABLES</b> .....	vi
<b>CHAPTER 1</b> .....	1
INTRODUCTION .....	1
<b>CHAPTER 2</b> .....	2
REVIEW OF LITERATURE .....	2
<b>CHAPTER 3</b> .....	9
METHODOLOGY .....	9
<b>CHAPTER 4</b> .....	15
FINDINGS .....	15
<b>CHAPTER 5</b> .....	19
CONCLUSION .....	19
<b>BIBLIOGRAPHY</b> .....	23

## LIST OF TABLES

TABLE	PAGE
Table 1. Labels for Coding Segments .....	14
Table 2. Predicted (BIOS) vs Observed (LENA) Language Environment .....	16
Table 3. Conversational Turn Counts (CTC) .....	22
Table 4. Percent of Child-Directed vs Overheard Speech Given Total Language Input .....	22

# CHAPTER 1

## INTRODUCTION

Researchers have defined bilingual language proficiency and dominance in different ways for decades; however, to date there is no standard, systematic description used by the scientific community to quantify bilingual children's language skills (cf. Bedore et al., 2012; Fishman & Cooper, 1969; Grosjean, 1998).

Operationalizing a definition of bilingualism is challenging due to the fact that there is significant variability in the environment bilingual children are exposed to and those who acquire two languages simultaneously (learning two languages at the same time) or sequentially (learning one language after the other), creating a wide spectrum of possible bilingual profiles. Bilingual children receive spoken input that is distributed across two languages and they must discriminate and differentiate this dual language input. A child's language profile is dependent on exposure to and use of each language, including the amount of child-directed speech from caregivers within each language (Gollan et al., 2015). Child directed speech is when parents explicitly attempt to guide their child's attention or behavior or engage their child in verbal responses and turn-taking (Shanks, 2016). This is different than overheard speech, which is adult speech that is not directed to the child but instead directed to another child or an adult; this is overheard because the adult is speaking within the child's earshot. Although bilingual children hear both languages, they do not necessarily receive equal exposure or opportunity to use the two languages, creating a complex

language profile that is unique to each individual bilingual child (Grosjean, 2011). Therefore, bilingualism is on a continuum and children have a variety of language profiles that researchers describe and measure in different ways. Traditional measures to explain bilingual language skills in children rely on age of first exposure, cumulative language exposure, and/or current language use patterns through parent report (Bedore et al., 2012).

## CHAPTER 2

### REVIEW OF LITERATURE

#### **The Use of Parent Report to Study Language Environment**

To date, research has shown that parent report questionnaires of a child's language exposure and use are a valid mechanism for vocabulary monitoring and measuring language skills (Mancilla-Martinez Jeannette et al., 2016). Through a parent report, researchers may ask a caregiver to describe their child's language profile to the best of their ability; for bilingual children, parents report on both languages. One type of parent report asks the caregiver to identify the languages their child typically hears and speaks on an hour-by-hour basis, then extrapolates this data to a full week to quantify the language environment. This report provides rough estimates of the child's exposure and use of each language (Peña et al., 2018). Some parent reports also include more intensive, diary-based measures of language in which caregivers keep a week-long log of the children's language exposure and specific caregiver interactions throughout the day, providing a more detailed, authentic depiction of children's bilingual experience than a typical caregiver report that asks parents to estimate exposure based on past experiences (Marchman et al., 2017). Parent report surveys are cost-effective and time-efficient, but it is necessary to analyze the validity of parent language questionnaires. This study aims to determine the accuracy of a parent's ability to report on their child's language exposure and use.

## **The Use of Parent Report with Bilingual Populations**

Due to variability in daily caregiver interactions and language contexts for listening and speaking two languages, it is possible that parents differ in the accuracy of explaining their bilingual child's listening environment within parent report measures. While parent report is strongly related to diary-based measures of language use and bilingual children's language skill (Hoff et al., 2012; Place & Hoff, 2011), it may underestimate the amount of language the child actually hears, making it difficult to gauge the child's true language environment (Houwer et al., 2005; Ribot et al., 2018). A persistent and somewhat puzzling finding in the literature regarding Spanish-English bilingual development in the United States is that it appears to require more Spanish input to learn Spanish than English input to learn English (Hoff et al., 2012; Pearson et al., 1997; Ribot & Hoff, 2014). Studies that have assessed Spanish-English bilingual children's expressive language skills in both languages found that those who are reported to hear a balanced amount of each language appear to be more English-dominant in their conversational skills while those who come from Spanish-dominant environments tend to be balanced in their language abilities across the two languages (Hoff & Ribot, 2016). This could be due to measures of input underestimating English exposure in other language environments, like school or television, or the effects of living in the United States where the majority language is English; most of what children read or hear outside the home will be in English. Essentially, this means that parent report may underestimate the amount of English a child actually hears, not that English requires less input or is easier to learn than Spanish.

This finding may also be due to a cultural practice of adults expecting children to speak less in Spanish language interactions, reducing children's levels of Spanish output relative to their Spanish input. There have been noted differences between bilinguals and monolinguals in the number of conversational turns taken between mothers and their children. A study of mother-child conversations in bilingual Latino mother-child pairs speaking Spanish, and monolingual Anglo mother-child pairs speaking English, found that Anglo children took more conversational turns than Latino children when speaking with their mothers (Shanks, 2016). This would explain the finding that children may have reduced levels of Spanish output relative to their input. Language output directly contributes to expressive language skill and development (Ribot et al., 2018); this could explain the gap between the amount of Spanish input children hear and the level of Spanish expressive skill they display. Essentially, children from Spanish-dominant environments may be balanced in their expressive language abilities across the two languages because there is a cultural expectation that the child speaks less in Spanish. However, it is important to note that this could be due to a difference between monolingual and bilingual language learners rather than being attributed to a cultural difference. Since the study included bilingual native Spanish speakers, the observed difference could be attributed to bilingualism itself (bilingual vs. monolingual) and not culture (Latino vs. Anglo) or language (Spanish vs. English).

In support of there being cultural differences, Peredo et al., (2020) found that Spanish-speaking Latino caregivers from low-socioeconomic backgrounds use a more directive interaction style than English-speaking caregivers from high-socioeconomic

backgrounds. A directive interaction style is high in intrusive behaviors that attempt to control the child (such as taking an item from the child's hands) and gives the child few opportunities to take communicative turns. Non-Latino, English-speaking caregivers from high-socioeconomic backgrounds spent more time observing and narrating play and providing opportunities for cognitive stimulation (Peredo et al. (2020). It is possible that this cultural difference in interaction styles may result in less opportunities for conversational turn-taking for Spanish-speaking Latino children. This cultural parenting style of decreased expectations and opportunities for turn-taking may further explain why children's levels of Spanish output are reduced relative to their input, per parent report measures.

However, as previously stated, a bilingual child's language profile depends on exposure, use and amount of child-directed speech within each language (Gollan et al., 2015). All of these factors (exposure, use, and child-directed speech) influencing a child's language profile may make language reporting from parents flawed, biased or difficult to precisely measure. The difference in Spanish output relative to Spanish input may, in fact, be attributed to parent report measures inaccurately estimating the child's language profile.

### **The Limitations of Parent Report**

Parent report has been validated for measuring a child's language skill. It has become common practice to use parent report questionnaires as a measure of a child's language use (Mancilla-Martinez et al., 2016; Marchman et al., 2017; Ribot et al., 2018). Parent report began with diary-based measures, which asked the caregivers to

keep a written record of their children's dual language exposure for a predetermined amount of time. While diary-based measures of relative amount of dual language exposure seem to predict language outcomes, this is a time-consuming, unrealistic process for caregivers (Place & Hoff, 2011). Later forms of parent report asked caregivers broad questions regarding patterns of their child's language use in conversation; this did not provide researchers with specific enough information regarding quantity of language exposure (Ribot et al., 2018). Current forms of parent report ask the caregivers pointed questions about the amount of language their child hears and speaks; researchers inquire about both a typical weekday and weekend day with specific questions like, "What does your child do at 4 pm on a typical weekday?", "What language is he/she speaking?", "Who is he/she with?" and "What language is the other person speaking in?" (Peña et al., 2018).

Measures of parent report that use specific, pointed questions have been found to be reliable, cost-effective, and time-efficient, however it still relies on caregivers to reflect on their child's past language experiences to estimate current abilities. Parent report questionnaires are considered valid measures of a child's language skills, yet researchers have suggested future studies are necessary to determine alternative ways to more accurately measure a child's current language environment, including measures of language specific input.

Marchman et al. (2017) revealed limitations and made applicable suggestions to improve future studies involving parent report. This study with 3-year-old Spanish-English bilinguals revealed a moderate correlation between parent report estimates of language exposure and the actual amount of child-directed speech calculated from

naturalistic recordings. However, there were significant limitations to generalizability of this study due to a small, mostly Spanish-dominant sample size. The findings are difficult to generalize as English-dominant bilinguals and balanced bilinguals were not included, participants were only tested once at a young age, and the average recording length was just 8 hours. Future studies should utilize recording devices and gather extensive data over a longer period of time with a more varied range of bilinguals.

Ribot et al. (2018) also used parent report and naturalistic recordings to explore the effects of language use on language growth. The researchers used a recording device, the Language ENvironment Analysis System (LENA) to analyze parent-child interactions within the home. Parent report was used to measure the child's language output. This study, however, used a binary measure of parent report of their child's language profile. Caregivers were asked what language their child uses to reply in when addressed in English and in Spanish as either "Always responds in the same language" or "Does not always respond in the same language". The researchers then created two groups of children that differed in language output. Instead of using binary measures of code-switching to categorize children based on preferred language, Ribot et al. (2018) suggested using more extensive measures of children's expressive language skills in future studies.

Another study with applicable suggestions for future research validated the use of parent report to estimate children's vocabularies when compared to scores from standardized expressive and receptive vocabulary assessments (Mancilla-Martinez et al., 2016). The findings of this study suggest that parent report of their child's vocabulary is valid for monitoring productive vocabulary knowledge in both of the

child's languages. There were limitations to the study's participant demographic: the researchers focused solely on toddler and preschool-age children from Spanish-speaking, low-income households who attended Head Start programs. Future studies should investigate the use of parent report with other Spanish-speaking populations so that the findings are more generalizable to a larger demographic. Additionally, the authors suggested further studies should measure expressive and receptive language in the home and in classrooms and not solely through vocabulary assessments. Overall, to accurately report language profiles, researchers have suggested that a more nuanced measure to calculate language input and output may be necessary, providing additional information regarding the quality of the child's language environment including conversational turn-taking (Mancilla-Martinez et al., 2016; Marchman et al., 2017; Ribot et al., 2018).

### **Using the Language ENvironment Analysis Device**

The LENA System records all audio within the child's earshot, including adult and other child speech and environmental sounds. The LENA device was originally designed to analyze English but has recently been validated for all language contexts, including its use with the Spanish-speaking population (Orena et al., 2019). The LENA Pro Software analyzes the recordings using a specialized speech-recognition algorithm to estimate adult word count (AWC), child vocalization count (CVC), conversational turn count (CTC). It can automatically produce CTC, the number of times a child engages in conversation with an adult, however, it cannot differentiate between languages or distinguish child-directed from overheard speech (Ganek &

Eriks-Brophy, 2018). Researchers can further analyze the LENA recordings by hand coding the language used and distinguishing child-directed from overheard speech to get a better understanding of the quality of the child's language environment (Gollan et al., 2015). Once coded, the LENA recordings can provide a clearer understanding of the quality of the child's language environment by calculating the actual amount of each language a child uses and is exposed to; this can be compared to the parent report estimates to assess the validity of the BIOS. Since the BIOS is a cost-effective and time-efficient form of parent report, it is critical to analyze its validity by comparing it to hand coded, naturalistic audio segments retrieved using the LENA System.

### **Motivation of the Current Study**

Previous studies using the LENA device have focused on younger populations and a narrow range of bilinguals. This research study intended to analyze an older population of early school-aged children and include a wider range of bilinguals from a larger demographic, however it was limited to 1 bilingual and 1 monolingual participant due to COVID-19. Unlike previous studies that analyzed shorter naturalistic recordings, families were encouraged to record for a longer period of time to provide a more representative sample of the language environment. Instead of using parent report with a binary measure of language use like code-switching, the researchers used a comprehensive parent report questionnaire, the Bilingual Input-Output Survey (BIOS), to inquire about the language environment (Peña et al., 2018). The BIOS questionnaire uses pointed questions through which researchers can assess the amount of language a child hears and speaks (Peña et al., 2018). This study

evaluates the relationship between parent report of a child's language environment using the BIOS, and an analysis of naturalistic recordings produced by the LENA System, a digital recording technology that will be used to quantify the child's language environment. To our knowledge, this is the first study to assess the validity of the BIOS questionnaire; we intend to compare the parent's report to naturalistic recordings from the LENA device.

While using the LENA System, not only can researchers gather information that may validate or invalidate the BIOS, but they can also learn about the richness and quality of the child's language environment and explore potential cultural differences that may not be captured on parent report. Researchers intend to explore the quality of a child's language environment through analyzing child directed speech and CTCs. By examining turn-taking opportunities and child-directed speech within the naturalistic recordings, researchers can explore differences between bilingual and monolingual speakers. CTCs will be explored in addition to child directed speech as turn taking impacts children's language processing more than the sheer quantity of words that they are exposed to (Gilkerson et al., 2018; Romeo et al., 2018). Researchers intend for a qualitative group analysis to be run to examine how potential differences between monolingual and bilingual children may affect conversational turn-taking once a large enough sample size of monolingual and bilingual children participate in the study (Shanks, 2016). Therefore, there are two prominent research questions in this study:

### **Research Questions**

1. Is parent report of their child's language use as valid as the hand-coded audio segments retrieved using the LENA System in early elementary school-aged children?
2. Are there differences in the number of conversational turns in monolingual and bilingual early elementary-aged children? If so, to what extent?

## CHAPTER 3

### METHODOLOGY

#### **Participants**

In the winter of 2021, researchers invited families to participate in an at-home study using the LENA device. The participants were recruited from a larger, 2-year longitudinal study at their elementary school. Due to COVID-19, on-site research in the elementary school was halted and the researchers modified the research plan for the participants. Due to time constraints of the master's program and difficulties in recruiting participants during this unprecedented time, we analyzed 1 8 year 9-month-old bilingual male (will be referenced as Julian) and 1 9 year 0-month-old monolingual male (will be referenced as Nick). Researchers were careful to select participants who were closely matched on critical factors such as gender and socioeconomic status. Both participants are in the 3<sup>rd</sup> grade at the same school in Rhode Island.

#### **Procedure**

##### *BIOS Questionnaire*

Due to COVID-19, participants completed the consent forms, conducted interviews, and behavioral assessments via Google Meet conferencing and the Qualtrics Survey Software. Both families participating in the study completed the Bilingual Input-Output Survey (BIOS) to examine the child's language environment.

The researchers completed the BIOS virtually with one parent per participant. Parents were first asked to report which language(s) the child was exposed to on a yearly basis from birth to their current age (as there is variation in when children are first exposed to English). Parents were asked to report information on an hourly basis, reporting the activities the child typically participates in throughout a given day, the typical caregiver/adult or peer who interacts with the child, and the language(s) the child hears and uses during that time. The parents were asked to do this for a typical weekday and a typical weekend day, and the researchers then extrapolated the data to a 7-day week to estimate current language input (exposure) and output (use) (Baron et al., 2018). Researchers followed guidelines set forth by Bedore et al. (2012) to classify the bilingual participant as Bilingual English Dominant because his use was between 60% and 80% in his dominant language (English) and between 20% and 40% in his other language (Spanish).

Parents of monolingual children completed the BIOS as well, however, they reported solely on English use and exposure (100% English and 0% Spanish). Once completed, the BIOS was used to calculate cumulative language exposure to reveal language experience and language contexts, in addition to current language exposure and use, revealing a relative percentage of input and output in each language on a regular basis.

#### *LENA Device Recordings*

Following the administration of the BIOS, participants were given LENA devices to wear with a specially designed t-shirt with a pocket. Two devices and a t-shirt were

sanitized and dropped off to both participants at their home address. Researchers conducted video calls to train parents on how to use the device and answer any questions. The LENA device can record up to 16 hours of audio, measures about 3” x 2” and weighs two ounces (Ford et al., 2008). Caregivers were instructed to record for a minimum of 10 hours a day for each day (up to 16 hours); one device was for a week day and the second was for a weekend day. The researchers additionally provided a LENA document explaining how to read the display screen and use the two buttons on the device (power on and off) and how to place the device in their child’s shirt. Specific guidelines for recording were explained with a checklist and the caregivers were asked to keep a daily activity diary to outline the events that took place during the recording. Caregivers were encouraged to complete their typical schedule for the day, including any community or social events they participated in, as these settings have not been explored as much as the home or school settings in past studies using the LENA System (Greenwood et al., 2018). For the weekday, if the child was in school (either virtual or in person) the parents were informed to turn off the device for the duration of the school day as teacher permission to record during the school day was not granted. Following the completion of the two recording days, the devices were picked up and sanitized.

### *Analysis*

The recordings were uploaded to the LENA Pro Software and automatically segmented and analyzed using specialized speech-recognition algorithms to estimate Adult Word Count (the number of words spoken by an adult within the child’s

earshot), Child Vocalization Count (the number of words spoken by the child who is wearing the LENA device,) and Conversational Turn Count (the number of turns taken between the child wearing the device and the adult speaker) (Ford et al., 2008). Next, data for each participant was exported from LENA Online in an Excel spreadsheet in rows of 5-minute segments. From here, the majority of researchers using the LENA device do not state how they selected segments to analyze or decided the time interval to use for analysis, so it is generally assumed researchers analyzed the entire day's recording. However, some researchers selected segments based on areas of interest given the LENA algorithm estimates (such as high Child Vocalization Count (CVC)) and chose to limit the amount of recording that was used in analysis to the first 30 seconds of audio from the 5-minute segments (Ganek & Eriks-Brophy, 2018). For this study, researchers divided the rows of 5-minute segments into 2-hour time blocks and then split the segments into quartiles based on conversational turn counts and randomly selected one segment from each quartile within the block. This yielded 4 segments to analyze per 2-hour time block which is thought to be a representative sample of the language environment during that time of the day. CTC was selected as the area of interest because researchers wanted to ensure there would be adult speech to code within the audio clip. Segments that were less than 300 seconds (5 minutes) were not included (excluding 39 segments from Nick's total count and 6 segments from Julian's); if the device was paused, the 5-minute segment would have been interrupted, resulting in less than 300 seconds. Segments that had 0 conversational turn counts were excluded (excluding 29 from Nick's total count and 26 from Julian's) (Ganek & Eriks-Brophy, 2018). Time blocks that did not have 2 complete hours of

segments (24 segments) were not included (excluding 23 segments from Nick's total count and 8 from Julian's). Rather than set a predetermined cutoff for the number of segments selected to analyze, researchers wanted to capture a representative sample based on the length of recordings for each child. The recordings were spliced into 5-minute segments using Audacity® software.

Researchers were then able to listen to the 5-minute audio segments in Audacity® and hand code using labels to yield quantitative and qualitative information that the LENA could not provide. A coding manual was created to ensure consistency across coders. Prior to formally coding the segments, the monolingual and bilingual researchers coded the same sample segment to ensure similar coding methods; they each coded the segment and compared label tracks for similar onsets and duration times. The labels shown in Table 1 were used to identify speaker identity (the child wearing the LENA device, an adult or another child), the language being used (Spanish or English), and child-directed vs overheard speech; a label was added at the onset of speech and continued for the duration of the utterance. A bilingual Spanish-English undergraduate student coded the bilingual participant's recordings to ensure any Spanish spoken was coded correctly (Weisleder & Fernald, 2013). After coding the segments, all labels were exported to an Excel workbook and duration of utterances was calculated, including separate sheets to total output for "Child English" and "Child Spanish" and input for "Adult English and Other Child English" and "Adult Spanish and Other Child Spanish". Next, the amount of exposure and use in each language was totaled in seconds based on the exported durations from the hand coded segments. Then, the percent of exposure and use of both Spanish and English

was calculated. Then, researchers used descriptive statistics to analyze parents' predicted language exposure and use from the BIOS questionnaire and compared it to observed naturalistic recordings from the LENA System.

In order to explore the second research question regarding differences in CTC between monolinguals and bilinguals, the CTC estimates were totaled from the specialized speech-recognition software for the segments that were selected and analyzed for each participant. Researchers also looked at the breakdown of child-directed and overheard speech after hand coding the segments for speaker identity and language. Recall that child-directed speech includes explicit parental attempts to guide a child's attention or behavior or engage them in verbal responses and turn-taking (Shanks, 2016). While all CTC must involve child-directed speech, not all child-directed speech involves conversational turns between the child and adult.

Table 1. Labels for Coding Segments

Label	Meaning
AE	Adult English
AS	Adult Spanish
CE	Child English
CS	Child Spanish
XE	Other child/peer English
XS	Other child/peer Spanish
CDSAE	Child-Directed speech Adult English
CDSAS	Child-Directed speech Adult Spanish
OHAE	Overheard speech Adult English
OHAS	Overheard speech Adult Spanish

CDXE	Child-Directed other child/peer English
CDXS	Child-Directed other child/peer Spanish
OHXE	Overheard other child/peer English
OHXS	Overheard other child/peer Spanish

## CHAPTER 4

### FINDINGS

#### **Research Question 1**

##### *Summary of The BIOS Questionnaire*

Researchers were aware that the parents of the monolingual participant, Nick, would report hearing and using 100% English and 0% Spanish. After analyzing the hand coded audio segments from the LENA, it was confirmed that Nick heard and used 100% English. When the parents of Julian completed the BIOS questionnaire, they reported that his language environment was comprised of 60.42% English input and output and 39.58% Spanish input and output (as shown in Table 2). The mother reported that she pretends not to understand Julian's replies if she addresses him in Spanish and he replies in English. She also reported Julian uses Spanish with his Spanish-speaking family on the phone. This could explain the similarities between reported input and output.

##### *Summary of the LENA Device*

Nick was recorded for 16.3 hours of audio between the LENA devices; 24 segments were selected and coded, totaling 2.0 hours (7,200 seconds) of audio. Of those total seconds, 1,341 seconds was total input and 1,798 seconds was total output (all communication was in English). Julian was recorded for a total of 18.9 total hours of audio (9.5 hours on the week day and 9.4 hours on the weekend day). Researchers

selected 30 representative, 5-minute segments from the overall audio (15 from each day), totaling 2.5 hours (9,000 seconds) of segments to be hand coded. More segments were selected from Julian’s audio since he recorded for a longer period (and perhaps had more waking hours); rather than setting a predetermined cutoff, researchers wanted to try and capture a representative sample based on the recording length for each child. After hand coding, researchers found that Julian’s language environment was comprised of 1,504 seconds of language input and 1,515 seconds of language output. When analyzed by language, there was 1,347 seconds of Child English use (88.89% of total language use) and 168 seconds of Child Spanish use (11.10% of total language use). There was 1,352 seconds of Adult English use (89.92% of total language exposure) and 151 seconds of Adult Spanish use (10.09% of total language exposure). This can be seen in Table 2.

Table 2: Predicted (BIOS) vs. Observed (LENA) Language Environment

Bilingual Participant: Julian		BIOS	LENA
Input	Spanish	39.58%	10.09%
	English	60.42%	89.92%
Output	Spanish	39.58%	11.10%
	English	60.42%	88.89%

*Comparing the BIOS and the LENA Findings*

To answer Research Question 1, researchers calculated the percentages of input and output in each language measured by the LENA system after hand coding the segments and compared it to parents' reported values from the BIOS. As shown in Table 2, the parents of Julian reported that their child was exposed to more Spanish than was observed based on the hand coded segments from the LENA. According to the BIOS, parents reported 60.42% English input and 39.58% Spanish input, whereas the observed input breakdown was actually 89.92% English and 10.09% Spanish. Calculations from the BIOS reported child output as 60.42% English and 39.58% Spanish, yet hand coding the segments revealed the actual percentages as 88.89% English and 11.10% Spanish. These estimates show that the parents reported more Spanish language use and exposure than was actually observed. For this participant, parent report does not appear to accurately estimate the child's language environment; Spanish language exposure and use was overestimated when compared to actual data retrieved from analyzing recordings.

## **Research Question 2**

### *Summary of Conversational Turn Counts*

To explore the second research question, researchers totaled the automated CTC estimates from the LENA Pro Software for the segments that were selected and analyzed for each participant, as shown in Table 3. For Nick, there were 528 CTCs, with an average of 22 (range: 5-59), given the 24 segments analyzed. For Julian, there were 1,310 CTCs, with an average of 41 (range: 6-91), given the 30 segments

analyzed. There is a difference of 19 between the average CTC for each participant. It is important to note that 6 more segments were selected and analyzed for Julian than for Nick. When CTC per segment was calculated, Nick had 66 and Julian had 131.

Table 3. Conversational Turn Counts (CTCs)

	Nick	Julian
CTC	528	1,310
CTC Per Segment	66	131
Average CTC Per Segment	22	41
Range	5-59	6-91
Segments analyzed	24	30

*Summary of Child-Directed and Overheard Speech*

The breakdown of child-directed speech and overheard speech in both Spanish and English was also analyzed. For Nick, 62.86% of his total language input was child-directed English, while 37.13% was overheard English. For Julian, 38.36% of his total language input was child-directed English, and 2.82% was child-directed Spanish; 51.55% was overheard English, and 7.26% was overheard Spanish.

Table 4. Percent of Child-Directed Vs Overheard Speech Given Total Language Input

	Nick	Julian
Child-Directed Speech Adult English (CDSAE)	62.86%	38.36%
Overheard Adult English (OHAE)	37.13%	51.55%
Child-Directed Speech Adult Spanish (CDSAS)	N/A	2.82%
Overheard Adult Spanish (OHAS)	N/A	7.26%

### *Comparing Conversational Turn Count and Child-Directed Speech*

When comparing the automated counts that the LENA produces using the specialized speech-recognition algorithm, there was a noticeable difference between the average CTC of both participants, as seen in Table 3. It is evident that Julian, the bilingual participant, has a greater CTC than Nick, the monolingual participant. However, when looking at the child-directed speech data, Nick heard 21.68% more child-directed speech than Julian did. While it is interesting to note this difference between CTC and child directed speech, research shows that the number of conversational turns is a better measure of the quality of a child’s language environment than the quantity of words that the child is exposed to as this impacts their language processing more (Gilkerson et al., 2018; Romeo et al., 2018). While child-directed speech is a measure of the quantity of language that was specifically directed towards the child wearing the LENA device, CTC provides more information on the quality of the language interaction because it encompasses a verbal response from the child.

## CHAPTER 5

### CONCLUSION

#### **Discussion**

##### *Research Question 1*

The purpose of this study was to explore the validity of parent report of their child's language use when compared to hand coded audio segments retrieved using the LENA System. Additionally, this study explored potential differences in the number of conversational turns between monolingual and bilingual early elementary-aged children. While for Julian, parent report did not accurately estimate his language profile, it is striking to note that the mother reported equal percentages for language specific input and output and when directly observed, the child's Spanish input differed only by .96% to Spanish output, and English input differed by 3.03% from English output. As stated before, this is most likely seen because Julian is expected to respond to others using the same language that he was addressed in. When completing the BIOS, Julian's parents underestimated the amount of English and overestimated the amount Spanish that their child hears. It is not uncommon for parent report to underestimate English exposure, likely due to the effects of the majority language influence (Houwer et al., 2005; Ribot et al., 2018).

It became evident through analyzing hand coded segments from the LENA that the BIOS was not a valid measure of this bilingual child's language profile. It should be noted that Julian's father completed the BIOS questionnaire, and his father was

with him for the majority of the recordings, however he is not a heritage Spanish speaker. It is possible, however, that the 2.5 hours of segments selected to be analyzed for Julian were not a representative sample of his language environment. There were 18.9 hours of audio captured on Julian's LENA devices; it is possible that the validity of the BIOS would have increased if a more representative way to select segments was used, or more segments were hand coded. Additionally, researchers were unable to record during the school day, yet the BIOS asks parents to report on expected language input and output during both a weekend and a full weekday, during the child's school day. The children's school is a monolingual environment, and the parent report included this expected English input and output for the school day. Both Nick and Julian were attending school in person; many districts were remote learning due to COVID-19. As researchers were unable to record during the school day, this additional time in English was not reflected in the LENA segments. For Julian, parental estimates of language exposure and use appear to not be valid, however, due to a very limited sample size, we cannot extrapolate the results of this study beyond the participants that were tested.

### *Research Question 2*

It is interesting to note the significant CTC discrepancy between the two participants; especially since Nick heard more child-directed speech than Julian. Therefore, it can be concluded that while Nick hears more child-directed speech than Julian, Julian is provided with more conversational turn-taking opportunities (keeping in mind that there were 6 more segments selected and analyzed for Julian than for

Nick). Not only is it important to talk to your child, but it is critical to offer him or her turn-taking opportunities. Providing your child with a chance to engage in conversation directly contributes to their expressive skills in that language. These findings are inconsistent with previous research showing that monolingual mother-child pairs take more conversational turns than bilingual children (Shanks, 2016). It is important to note that the participants in the Shanks (2016) study were significantly younger (2-year-olds) than participants in this study which may also contribute to the difference in findings. Additionally, this is inconsistent with findings from Peredo et al. (2020) who found that Spanish-speaking Latino caregivers from low-socioeconomic backgrounds used a directive interaction style that resulted in few communicative opportunities. It is important to note, however, that Julian was matched on SES with Nick, and that Julian is an English dominant bilingual, not a Spanish-speaking monolingual. When Dyer et al. (2014) examined the interaction styles of English and Spanish-speaking Latina mothers with their 2 and 3-year olds, they found that 50% of mothers identified with a “child-oriented” interaction style and 45% of mothers identified with “directive”. While interaction style and language status are not reported specifically, 60.5% of Latina mothers were Latina-oriented, 29.5% were bicultural and 5.7% were American-oriented; 80% of all Latina mothers completed the home interview in Spanish, but the percent of mothers who are bilingual is not reported. While the majority of mothers used Spanish and were Latina-oriented, it is not possible to attribute interaction style to language or culture. This research question exploring opportunities for turn-taking needs to be analyzed with a

larger group of monolinguals and bilinguals before making conclusions about potential cultural differences.

### **Limitations & Future Directions**

Due to COVID-19, the sample size for this study was limited to a case study. Researchers had intended to select 10 monolingual and 10 bilingual participants who were closely matched on critical factors such as number of siblings in the house and birth order (Bridges & Hoff, 2014), socioeconomic status and gender. However, given necessary modifications to the study methods due to COVID-19, Julian and Nick were not matched on all critical factors mentioned above. Nick is an only child and Julian has a younger sibling. While this did not affect the other child language input since Julian's sibling is an infant, it is possible that it may have affected the amount of overheard speech that Julian was exposed to. Julian's parents directed some speech to the infant, creating more opportunities for overheard speech for Julian. It is also possible that having a sibling affected the amount of child-directed speech that Julian heard; the parents potentially split their attention and speech between Julian and his sibling.

An additional qualitative group analysis was intended to be conducted to examine potential differences in cultural expectations for conversational turn-taking, however this group analysis did not occur during the duration of this thesis due to the small sample size. One monolingual participant and one bilingual participant have been coded and analyzed so far, distinguishing overheard speech from child-directed speech in each language. It is critical to identify amount of child-directed speech and explore

number of conversational turns between the child and the language speaker in order to obtain a better understanding of the quality of the child's language environment (Gollan et al., 2015; Shanks, 2016). This qualitative group analysis exploring CTC and child-directed speech will be conducted once a larger group of monolingual and bilingual children have participated in the study.

A second limitation is that we only had one bilingual researcher hand coding the LENA audio recordings in the lab and had a limited number of personnel that had been properly trained to hand code segments as this is a very time-consuming process. While Julian recorded 18.9 total hours of audio, 2.5 hours of representative segments were selected, coded and analyzed due to limited bilingual personnel and time constraints. It is possible that if researchers were able to code and analyze more segments, the findings may reveal a stronger correlation between parent report and observed language environment.

Another limitation was that our participants were older than the age that the LENA device was initially intended to be used for. Nick and Julian are 9, whereas the LENA device is validated for use with children up to age 6 (Romeo et al., 2018). While hand coding was ultimately used in our comparison between the two measures (not automated word counts) and our hand coders were presumably accurate in identifying speaker identity as adult or child and, researchers tried to align automated CTC estimates to totals from the corresponding hand coded audio segment and the counts did not align (the LENA reported one of Nick's segments had 16 conversational turn counts, but when hand coded, researchers only identified 7 conversational turns between the child and adult). It is possible that the LENA device

coded the participant's vocalizations as adult female in the automatic speech-recognition word count estimates, altering the adult word count and impacting the CTC totals. It is important to note that the samples analyzed for each participant were not equivalent; Nick 24 segments and Julian had 30.

Given the limitations mentioned above, future research should explore these questions with a larger group of monolingual and bilingual children. Additionally, researchers should strive to record the child's school day environment since the BIOS asks parents to report on a typical weekday. Following recording with the LENA device, parents should be provided with an opportunity to report on how representative they feel that specific day was when compared to their typical schedule. Researchers should train more bilingual personnel to code so that a larger number of segments can be analyzed per participant. Lastly, CTCs may need to be calculated for older children by hand coding rather than estimated by the LENA Pro Software. Future studies might also explore this question with bilingual participants who attend a school with a dual language program rather than an English-only school.

## BIBLIOGRAPHY

- Baron, A., Bedore, L. M., Peña, E. D., Lovgren-Urbe, S. D., López, A. A., & Villagran, E. (2018). Production of Spanish grammatical forms in U.S. bilingual children. *American Journal of Speech-Language Pathology*, *27*(3), 975–987. [https://doi.org/10.1044/2018\\_AJSLP-17-0074](https://doi.org/10.1044/2018_AJSLP-17-0074)
- Bedore, L. M., Peña, E. D., Summers, C. L., Boerger, K. M., Resendiz, M. D., Greene, K., Bohman, T. M., & Gillam, R. B. (2012). The measure matters: Language dominance profiles across measures in Spanish–English bilingual children. *Bilingualism (Cambridge, England)*, *15*(3), 616–629. <https://doi.org/10.1017/S1366728912000090>
- Bridges, K., & Hoff, E. (2014). Older sibling influences on the language environment and language development of toddlers in bilingual homes. *Applied Psycholinguistics*, *35*(2), 225–241. <https://doi.org/10.1017/S0142716412000379>
- Dyer, N., Owen, M. T., & Caughy, M. O. (2014). Ethnic differences in profiles of mother–child interactions and relations to emerging school readiness in African American and Latin American children. *Parenting, Science and Practice*, *14*(3–4), 175–194. <https://doi.org/10.1080/15295192.2014.972756>
- Fishman, J. A., & Cooper, R. L. (1969). Alternative measures of bilingualism. *Journal of Verbal Learning and Verbal Behavior*, *8*(2), 276–282.
- Ford, M., Baer, C., Xu., Yapnel, U., & Gray, S. (2008). The LENA language environment analysis system: audio specifications of the DLP-012 (Technical Report LTR-03-2). Boulder, CO: LENA Foundation.
- Ganek, H., & Eriks-Brophy, A. (2018). Language ENvironment analysis (LENA) system investigation of day long recordings in children: A literature review. *Journal of Communication Disorders*, *72*, 77–85. <https://doi.org/10.1016/j.jcomdis.2017.12.005>

- Gilkerson, J., Richards, J. A., Warren, S. F., Oller, D. K., Russo, R., & Vohr, B. (2018). Language experience in the second year of life and language outcomes in late childhood. *Pediatrics*, *142*(4). <https://doi.org/10.1542/peds.2017-4276>
- Gollan, T. H., Starr, J., & Ferreira, V. S. (2015). More than use it or lose it: the number of speakers effect on heritage language proficiency. *Psychonomic Bulletin & Review*, *22*(1), 147–155. <https://doi.org/10.3758/s13423-014-0649-7>
- Greenwood, C., Schnitz, A., Dwight, I., Tsai, S., & Carta, J. (2018). Automated language environment analysis: A research synthesis. *American Journal of Speech-Language Pathology*, *27*(2), 853–867. [https://doi.org/10.1044/2017\\_AJSLP-17-0033](https://doi.org/10.1044/2017_AJSLP-17-0033)
- Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition*, *1*(2), 131-149.
- Grosjean, F. (2011, December 7). Visualizing One’s Languages. *Life as a Bilingual*. <http://www.psychologytoday.com/blog/life-bilingual/201112/visualizing-ones-languages>
- Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language*, *39*(1), 1–27. <https://doi.org/10.1017/S0305000910000759>
- Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language*, *39*(1), 1–27. <https://doi.org/10.1017/S0305000910000759>
- Houwer, A. D., Bornstein, M. H., & Leach, D. B. (2005). Assessing early communicative ability: A cross-reporter cumulative score for the MacArthur CDI. *Journal of Child Language*, *32*(4), 735–758. <https://doi.org/10.1017/S0305000905007026>
- Mancilla-Martinez, J., Gámez, P. B., Vagh, S. B., Lesaux, N. K., Nippold, M., & Alt, M. (2016). Parent reports of young Spanish-English bilingual children’s productive vocabulary: a development and validation study. *Language, Speech & Hearing Services in Schools*, *47*(1), 1–15. [https://doi.org/10.1044/2015\\_LSHSS-15-0013](https://doi.org/10.1044/2015_LSHSS-15-0013)

- Marchman, V. A., Martínez, L. Z., Hurtado, N., Grüter, T., & Fernald, A. (2017). Caregiver talk to young Spanish-English bilinguals: Comparing direct observation and parent-report measures of dual-language exposure. *Developmental Science*, *20*(1), e12425. <https://doi.org/10.1111/desc.12425>
- Orena, A. J., Byers-Heinlein, K., & Polka, L. (2019). Reliability of the language environment analysis recording system in analyzing French-English bilingual speech. *Journal of Speech, Language, and Hearing Research*, *62*(7), 2491–2500. [https://doi.org/10.1044/2019\\_JSLHR-L-18-0342](https://doi.org/10.1044/2019_JSLHR-L-18-0342)
- Pearson, B. Z., Fernández, S. C., Lewedeg, V., & Oller, K. (1997). The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics*, *18*(1), 41–58. <https://doi.org/10.1017/S0142716400009863>
- Peña, E. D., Gutierrez-Clellen, V. F., Iglesias, A., Goldstein, B. A., & Bedore, L. M. (2018). *Bilingual Input/Output Survey (BIOS)*. Baltimore, MD: Brookes Publishing.
- Peredo, T. N., Dillehay, K. M., & Kaiser, A. P. (2020). Latino caregivers' interactions with their children with language delays: A comparison study. *Topics in Early Childhood Special Education*, *40*(1), 52–63. <https://doi.org/10.1177/0271121419900269>
- Place, S., & Hoff, E. (2011). Properties of dual language exposure that influence two-year-olds' bilingual proficiency. *Child Development*, *82*(6), 1834–1849. <https://doi.org/10.1111/j.1467-8624.2011.01660>
- Ribot, K. M., Hoff, E., & Burrige, A. (2018). Language use contributes to expressive language growth: evidence from bilingual children. *Child Development*, *89*(3), 929–940. <https://doi.org/10.1111/cdev.12770>
- Romeo, R. R., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Rowe, M. L., & Gabrieli, J. D. E. (2018). Beyond the 30-million-word gap: Children's conversational exposure is associated with language-related brain function. *Psychological Science*, *29*(5), 700–710. <https://doi.org/10.1177/0956797617742725>
- Shanks, K. F. (2016). Properties of child-directed speech in mothers' native and nonnative languages: A comparison of English monolingual and Spanish-

English bilingual mothers' talk to 2-year-olds (Unpublished Master's thesis).  
Florida Atlantic University, Davie.

Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science*, *24(11)*, 2143–2152. <https://doi.org/10.1177/0956797613488145>