Parent Report and Audio Recording Patterns in Bilingual Children Who Speak a Minority Language at Home

Angela BRICKMAN

UNIVERSITY OF COLORADO BOULDER

SPEECH LANGUAGE AND HEARING SCIENCES

COLLEGE OF ARTS AND SCIENCES

Abstract

Bilingualism is when an individual is able to utilize at least two languages. If a person speaks two languages, it is beneficial for them both socially and for their brain development. This study focuses on the language input of sequential bilingual children in a bilingual Head Start program that promotes early childhood development and school readiness for low income families. There were six participants and their families in this study. To examine these children's use of language, parental questionnaires and auditory recordings from the children's day were utilized. Each family that participated in this study completed a questionnaire about their child which included family history and information on languages that are spoken at home. These parent reports gave insight to clinicians about the child directly from someone who was close to them. The audio recordings were retrieved from a device that the children were required to wear that depicted a typical day of language use. These recordings included linguistic information from both their school and home environments. Both the audio recordings and the parent reports were examined to compare data from a child's day as recorded in the audio recordings to what the parent report stated. The results provides some information about children's language input in L1 (Cantonese) and in L2 (English) at home and in school. The findings suggested that there are advantages and disadvantages of auditory recording and questionnaires. The results could be useful for clinicians and teachers who work with young bilingual children.

Contents

1 Introduction

The purpose of this study is to examine language input of preschool children who are exposed to a minority language at home and who learn English as a second language in school settings. A minority language is a language that has fewer opportunities to develop and has less social value compared to the majority language (Kohnert, 2010). The minority language for the children in this study is Cantonese. Parent reports can be a useful tool in determining information about a child's language input (Duursma et al., 2007). In addition to the children's personal information, the parent reports require the caregiver (parent) to answer questions regarding the language produced by their child and the language which the child hears. However, these reports might involve a parental bias (Paradis, 2017). In order to understand parent reporting, this study compares auditory data collected directly from a child's day to the parent reports. Data for this study was collected from parent reports and auditory recordings. The audio recordings include the children's verbal input and output from their natural school and home environments. This study is part of a larger effort to examine language acquisition of children from families who speak a minority language at home. My study also looked at how children's home and school environments affect their overall language development. We will first examine language input and how this input varies in different environments. Then we will analyze characteristics of sequential bilingual children through the language input they receive.

1.1 Language Input

Language input refers to the exposure that individuals have with a language (De Houwer, 2017). This can be orally, in print, or involve other exposure to the language. Language input is like a target grammar database that children are able to pull information from when learning to form the structure of that language (De Houwer, 2017). It is important for children to learn, understand, and acquire the grammar rules of the languages they are

learning. The more advanced the language input is, the more opportunities there are for the child's personal growth in the language.

Previous studies on language input provided information on monolingual and bilingual speakers. One type of study is the parent report. Parent reports are questionnaires that are completed by the parents or caregivers of the individual in question. Parent reports have been used to identify the amount of language input that bilingual children receive in research and clinical settings (De Houwer, 2017). These reports usually consist of what the caregiver sees the child doing including language use, social activities, and living situations. Information regarding the child and their family situation is also retrieved. Parent reports are used to predict children's language abilities and do not reveal any information regarding temporal units about language input (Paradis, 2017). To understand the parent reports, it is important for clinicians and researchers to dissect and question the data obtained. Due to their own individual bias', parents and caregivers may exaggerate or underestimate the child's ability. Parent's want the best for their children which could lead them to exaggerate the truth about their child's ability. This would lead to false information being provided to clinicians and teachers. A previous study by Mancilla-Martinez (2016) which looked at bilingual children's productive vocabulary found that parent reports present a cost effective and valid measure for monitoring vocabulary in early childhood education settings (Mancilla-Martinez, Gámez, Vagh, & Lesaux, 2016).

The role of language input in language development is a highly debated and heavily studied topic. Bialystok (2010) found that there are multiple elements affecting early language development in preschool children including neurophysiologic factors, early language sequences, and emotional and environmental factors (Bialystok, 2010). Research also suggests that the rate at which children learn and develop language(s) depends on the "quantity and quality of their language exposure" (Place & Hoff, 2016). Past evidence states that the use of the home language (L1) in an educational setting promotes bilingualism, biliteracy, and academic development (Cha & Goldenberg, 2015). It has been found that the age of

the child at first exposure compared to the age the second language is acquired seems to affect L1 skills (Duursma et al., 2007). Another study illustrates that for bilingual learners, oral language input has a strong influence on the growth of pre-literacy skills and language competence in both languages (Zhang, 2016).

The environments in which a child learns language is also important. Duursma et. al (2007) have examined the role of language environments and home literacy activities on bilinguals vocabulary development. They found that examining environmental supports for literacy activities, including determining the number of books found in the household, were factors related to use of the languages in the home (Duursma et al., 2007).

There is a lack of research done on bilingual children's language input from both languages. Consequently, this means that there is a lack of data on the weaker language of simultaneous bilinguals (Bernardini, 2017). Data is also lacking regarding when a child begins to receive language input in their multiple languages. Due to the large range of when a child may begin to hear multiple languages, data regarding when a language is introduced is difficult to represent. However, learning more in this area may provide important results regarding when the best period for dual language learning is. For my study, language input is referred to as significant language the child receives from the individuals they interact with. This includes input from the child's peers, teachers, siblings, family members, and caretakers. Language exposure from multiple sources is also noted in this study including television, radio, and other technologies. The research assistants and myself defined significant language input as when someone was speaking directly to the child or the child reacted to something spoken around them. Due to our data being auditory and not visual, it was difficult to differentiate when the child was actually being spoken to and when someone was speaking in the proximity around the child. Therefore, individual interpretations regarding this issue were left to the individual coders. However, the coders and myself were able to listen to cues from the children's communication partners to determine if the child was listening.

I will also investigate the children's linguistic relationships in respect to the natural home and school environments the child is in. Research on bilingual environments is currently only found in a small subset of literature. Place and Hoff (2016) conducted a study to examine properties of input for bilingual children that influence the rate of language development. Previous studies suggest that the proportion of input from native speakers and the number of different speakers who provide input are positive predictors of a child's language skills, while the findings are unclear about the benefit of language mixing compared to language separation. This study concludes that the amount of language input a child receives in each of their languages is correlated to the skill they have in that language (Place & Hoff, 2016).

No studies have been found that directly examine the environment the child was in or of the parent reports regarding these environments. My research will delve deeper into the effects of different environments on bilingual children and determine if these reports on children's environment are as the parents reported. I plan to investigate the relationship between bilingual children's language input and their language output. Then, I will compare the children's language input to the answers provided on the parent reports. A main focus of this study is to examine the bilingual input relationship between a child's L1 and L2. In summary, the study of children's language input is a growing and diverse field. Understanding the factors that contribute to the development of vocabulary in bilingual children will lead to exercises and curriculum that better educate a dual language learner. Studies have emphasized the importance of exposure for further learning of the grammar structure of a language, influence of the growth of pre-literacy skills, and highlighting the need for further research.

1.2   Characteristics of Sequential Bilingual Children

Bilingualism is when an individual is able to use two languages (ASHA, 2004). It includes oral or written use. The phenomenon of bilingualism can occur in childhood and adulthood. However, this study focuses on children. Bilingual children are children who during their

"most dynamic period of communication development" receive input in at least two languages (Kohnert, 2010, p. 457).

Bilingualism is perceived differently around the world. Communities have a range of emotions regarding bilingualism. In the United States, it is mostly prevalent in immigrant communities (Finegan, 2012). A study about adolescent immigrant children in the United States who speak both their community and home language indicates that these adolescents are more likely to graduate high school than their counterparts who only speak English. These dual language speakers also develop close connections with their family and culture which is linked to better emotional and social health (Kohnert, 2010). There are other positive attributes of bilingualism including linguistic, cognitive, and economic advantages (Cha & Goldenberg, 2015). Bialystok (2010) in her research found that over a wide range of testing, bilingual children performed better. Research included tests for nonverbal intelligence and mental reorganization (Bialystok, 2010). Bilinguals also have enhanced functioning of their executive control system. It has also been found that the benefits of bilingualism extend past inhibitory control and into working memory (Verhagen, Mulder, & Lesseman, 2017). However when compared to monolinguals, bilinguals have poorer performance with tasks "based on rapid lexical retrieval and processing" (Bialystok, 2010).

My research will focus on children who learn two languages sequentially in multiple settings. The settings studied include home and school. Sequential bilingualism refers to when a child acquires a first language, and then acquires a second language (Auer, 1984). This can occur when a child is exposed to multiple languages. A common example occurs when a child learns a language at home and then acquires another language when they begin to attend school. The children in this study are assumed to be sequential bilinguals since they are sequentially using Cantonese (L1) and English (L2) in their school and home environments.

As explained by Kijoo Cha and Claude Goldenberg (2015), there are two types of bilingualism, additive and subtractive. Additive bilingualism is when learning a second language

is not likely to replace the individual's L1. Subtractive bilingualism is when the individual's original language is gradually replaced by the new language (Cha & Goldenberg, 2015). The children in this study are assumed to be additive bilinguals since they are learning another language without the intent to replace their L1 or Cantonese.

Three general characteristics of typical developing bilingual learners are stated by Kohner (2010). These include "uneven ability or distributed skills within and across linguistic domains, the variable presence and nature of cross-language associations, and individual variation in language performance even in response to relatively similar circumstances (p. 458)." Uneven ability references the difference of ability between the two languages a child is learning. For example a child may perform better at one task in their L1 and worse on a different task in their L2. This distributed skill set becomes apparent if the developing bilingual uses their two languages in different settings. It has been found that both of a bilingual's languages are functionally independent (Kohnert, 2010). For this study, all children were marked as being more comfortable speaking their L1 instead of English. This may provide domains that create distributed and uneven skills across their languages.

An individual's fluid language environments contribute to their proficiency in both their languages. Language input environments include a person's nonverbal communications as well as their conversational interactions. Increased exposure to a language, increases an individual's language intake which will help them to acquire that language (De Houwer, 2017). Language environments vary daily. They depends on each speaking partner and depend on the child and the activities they are involved in. For the children in this study, both their home and school language environments were included. Factors that affect language environments include exposure to a language, ability to practice, and use of linguistic skills (De Houwer, 2017). The amount of exposure is said to be linked to the development of language dominance meaning the child is more comfortable in the language they have more exposure too. Language dominance is then linked cognitively with having a strong and a weak language (Carroll, 2017). New research suggests that not all types of input are "equally

supportive of language acquisition" (Place and Hoff 2016). Evidence has also been found that input from native speakers is more beneficial for language development than from nonnative speakers (Place & Hoff, 2016).

As learned above, there are multiple variations of bilingualism, positive attributes, and how language environment affects both languages. Learning about the different types of bilingualism and issues that affect it highlights how important the results of this study could be. This study will be focusing on the sequential bilingual children at the Kai Ming Head Start program.

1.3 Language Impairment and Bilinguals

Primary language impairment (PLI) is characterized as someone having slow language acquisition and weaknesses in nonlinguistic processing. It is a developmental disorder that has many labels such as language-based learning disabilities and specific language impairment (SLI). PLI does occur with bilingual children and affects both languages (Kohnert, 2010). This means that when compared to typically developing bilingual children, someone with PLI will acquire both language at a decreased pace. Due to a lack of studies, the percent of bilingual children who have PLI is unknown. However, it is believed to mimic the incidence rate of PLI in monolinguals which is predicted to be nearly seven percent (Kohnert, 2010). It is also believed that bilingual environments do not aid or hinder a bilingual ability to learn.

This study has three children with language impairment. The children will not be referred to as having specific language impairment (SLI) because they do not have formal diagnosis'. To see if there are any differences between the two populations, the linguistic input and output from the bilingual children with impairment will be compared to the children without impairment through the use of the children's parent reports and audio recordings.

2 The Current Study

The purpose of this study is to examine the relationship between parent reports and bilingual children's language exposure in different environments. Previous studies examine monolingual language exposure and parent reports. Parent reports have been shown to accurately predict bilingual development (Paradis, 2017). This is not an exhaustive study on bilingual input in various environments. This is due to the low sample size and large standard deviations present in the time spent in activities per child.

Little is known about the relationships between parent reports and a child's actual language input. This means clinicians and those who use parent reports do not know the reliability of the data obtained in the reports. This lack of information could compromise therapy and the work done through the use of parent reports (Kohnert, 2010). This study aims to determine how those using parent reports will treat the answers provided by the parents.

Even fewer studies have examined language input of children who speak more than one language. Bilingual children are a population that do not have much statistical information on them. This lack of information leads to restrictive and uninformed decisions from caretakers and teachers. Therefore, this research will highlight new avenues for research and will serve as a pioneer study for the study of bilingual populations. Collecting information on the bilingual's environments and comparing it to the parent reports makes this study the first of its kind.

Another focus of this study will be to determine if language input in parent reports accurately reflects what actually occurs in bilingual homes. By listening to the child's language input at home, we will be able to directly compare these findings with the parent reports. The data from the parent reports and the recordings will be validated and examined in order to highlight any observed patterns in the bilingual homes. This study will serve as a spearhead study of language input and children. It will analyze the relationship between parent reports and language input by observing children in their natural home and school

settings. In particular, there are five primary research questions within this study.

2.1   Environmental Input Patterns

What are the different significant language input patterns that are present in a bilingual child's environments?

Environments are key factors in how a child learns language as seen in Duursma's 2007 paper. Therefore I expect to see distinct differences in the linguistic input in children's home and school environments across all activities (Duursma et al., 2007). I predict that we will see a higher, if not exclusive use of the child's L1 in the family's home environment across all family members (Cheung, 2015). I will compare the answers of which language is used at home on the questionnaire to the data found in the audio recordings.

Due to the lack of information regarding both languages of sequential bilinguals, this study will be able to highlight the patterns seen across a bilingual child's environments (Bernardini, 2017). Cheung's (2015) study states that L1 is maintained in the home environment while L2 is developed in the educational setting. I expect to see similar results in my data (Cheung, 2015).

2.2   Child L1 and L2 use Patterns Based on Home Language

What are the patterns of L1 and L2 use for children who use both languages in their home environments?

In this study, language output is defined as the amount of language a child produces. This output may be in response to a communication partner's comment or it may occur randomly. Predictions for the results of a child's use of L1 in their home environments are to see an increase of language output the more that L1 is used in the home.

The amount of language output a child produces is also predicted to increase as they hear more language input from adults (Cheung, 2015). This may include extended family (e.g., grandparents, siblings), parents, and teachers. Therefore, patterns of home language

use should mimic the adults.

If the children have siblings as identified in the parent reports, it is also expected that the child will have a higher percentage of language output compared to their peers without siblings. This is hypothesized due to the expected increase in language input from the sibling (Brody, 2016). It is also a prediction that having more family members who reside in the child's house, will increase the child's output since they will hear more input.

2.3   Child Language Input and Output in Various Environments

What are the similarities and differences between the children's significant language output in home and school environments?

In different environments there are variables that affect social interactions. These variables shape linguistic input and output for a child. Predictions of the children's language output across different environments include; to see an increase in the child's output of L1 in their home environments (Cheung, 2015). Due to the evidence that the use of home language use in school settings promotes bilingualism, I expect to see a consistent pattern of L1 use in school but less than that seen in home environments (Cha & Goldenberg, 2015). It will be interesting to see the relationship between the peers at school and the peers or the siblings at home.

2.4   Parent Reports and Recorded Data

Are the parent reports questions about language use in a home setting consistent with the recorder data collected from the child's day?

Due to Mancilla-Martinez (2016) study, I expect the parent reports to accurately depict the child's vocabulary (Mancilla-Martinez et al., 2016). Predictions include to see similarities highlighted between the recorder data and the parent reports. I will specifically look at the adult input in settings like during play time, eating, and technology use. Due to parents extended exposure to their children, it is thought the parents are a reliable source of infor-

mation. However, because the parents are not in the classroom with the children there will likely be differences between the parent reports and the data collected from the classroom (Paradis, 2017). Due to the audio recordings not being visual, some activities asked about in the parent reports will not be able to be analyzed. An example of information that is not accessible is information regarding a child's reading activities.

2.5 Differences in Language Input and Output For Children with Language Impairment

What are the similarities and differences in language input and output for bilingual children with and without language impairments?

For children with language impairments it has been assumed that they have difficulties with their abilities to process or produce language (Armon-Lotem, 2017). On the parent reports, if a child has a language impairment then the parent identifies that the child has an Individualized Education Plan (IEP). Therefore, it is predicted that children with a language impairment will have a significantly smaller language output compared to their peers without a language impairment. It is also predicted that the input a child with a language impairment receives will be the same as a child without a language impairment.

3 Methods

3.1 Participants

Participants in this study were six children who learned Cantonese (L1) as a home language from birth and then in the classroom setting they began to learn English (L2). The six different participants are described in Table 1. All participants for this study were selected from the Kai Ming Head Start bilingual preschool in San Francisco California. The Kai Ming Head Start program does not follow any specific bilingual education curriculum.

However, it does provide a bilingual educational environment for their children. Some of the staff only speak Cantonese or only English while some of the staff are able to speak both languages. This bilingual and multicultural environment is beneficial for the children because it brings cultural diversity into the program. This school was selected for this study because the Child, Language, and Learning Lab has established a trusting relationship with the Kai Ming Head Start program and it provides a convenience sample. The children are systematically exposed to English in Kai Ming's Head Start educational setting.

Students that attend the Kai Ming Head Start program are likely to have similar socio-economic status due to one of the requirements to be in the preschool program being yearly income. Like other Head Start programs, Kai Ming services children who come from a low income household.

Six children and their families participated in this study. They are composed of two females and four males. The mean age of when the children began school is 53.66 months old with a standard deviation from the mean of 5.05 months (reference Figure 1). For the safety and privacy of the participants and their family, each participant was given an identification number. There names were never distributed, only their identification numbers among lab members. For this study, to further protect their identity and keep easily organized every child is given a letter, A-F. To ensure the privacy of these families information, every volunteer and lab employee underwent an online Collaborative Institutional Training Initiative (CITI) program. This course's curriculum was based on Human Research and social behavioral research.

Table 1: Age, gender, and highest education of parents for each child.

Child Identiifer	Age (Months)	Gender	Highest Education Mother	Highest Education Father
A	48	Female	Middle School	Middle School
B	59	Male	College	College
C	47	Female	High School	High School
D	57	Male	High School	Junior High
E	57	Male	High School	Junior High
F	54	Male	High School	High School

3.2 Questionnaire

To collect the data on the children, a questionnaire was used that was designed and used by Shirley Cheung (Cheung, 2015). The questionnaire was available in both Cantonese and English for the parents or caregivers. Prior to their participation in the study, this questionnaire was given to the participating families. It was answered by six of the children's mothers. The questionnaire titled *Language Input: Parent/Caregiver Questionnaire* was in three parts. Part one was about the parents and included; a description of the study for the participants and their family, contact information, and questions about the child who is participating. Questions about the relationship to the child, the child's date of birth, gender, and classroom information were also asked. Questions regarding the child's language input were asked to be completed to the best of the parents ability and knowledge. This portion of the questionnaire also asked questions regarding the education of the parents, where the child was born, and when the child started school. Additional information asked included what language was spoken at home, the language the child feels more comfortable learning at school, and the percent for each language the parent believes the child learns at Kai Ming.

Part two of the questionnaire asked about the child's use of their L1 and L2 language with different family members. Family members included the parents, grandparents, and siblings. The majority of the questionnaire asked what percentage of each language each family member speaks to the child at home per week. The questionnaires were coded and represented on a seven point scale as shown in Table 2.

Table 2: 7 Point scaling system for defining language input.

Scale Rating	Language Percent
1	100% English
2	20% Cantonese/Toisan, 80% English
3	40% Cantonese/Toisan, 60% English
4	50% Cantonese/Toisan, 50% English
5	60% Cantonese/Toisan, 40% English
6	80% Cantonese/Toisan, 20% English
7	100% Cantonese/Toisan

For coding purposes, 100 percent English was coded as 1 while 100 percent Cantonese/Toisan was input at 7. These are random numerical values assigned for the purpose of coding and defining language input.

In this portion of the questionnaire for each member in the home it asked if they spoke any other language at home. The estimated percentage of Toisan and Cantonese spoken in the home was also asked. It was a possibility to mark if the family member presented did not live at home. The third and final part of the questionnaire asked about the communication across activities with the child. Activities in question include eating, watching television, and playing word games. Eating was broken into breakfast, lunch, and dinner sessions.

3.3   Language ENvironment Analysis

The Language ENvironment Analysis (LENA) is a research foundation focused on children's language development based out of Boulder Colorado. For this study, a product manufactured by the LENA Research Foundation called LENA Pro was utilized. The LENA Pro system was used to collect the auditory data from a sample of these children's most natural environments. To record the data, the preschoolers wear shirts with front pockets in the middle of the shirt. The LENA Pro systems were held in the front pocket. The children wore the shirt with the recorder in it until their bedtime. The software in the recording devices is a system that is able to measure children's expressive language at its development. These devices were able to record twelve to sixteen hours of audio recordings of the children in their natural home and school environments. The LENA Pro recorder utilizes automatic vocalization assessment (AVA™) software that is able to recognize speech and the wearers vocalizations. The LENA software is able to illustrate and compare audio data to standard expressive language assessments (Foundation, 2012).

This study was unable to use the AVA statistics provided by the LENA Pro software due to the inability of this program to distinguish between the use of multiple languages. This is why the six sound files were coded by hand with the help of my fellow research assistants

and undergraduate volunteers. To begin the coding process, the high-quality audio files were downloaded from the LENA Pro recorder. To code the sound files into the multiple languages, speakers, and durations the software program Praat was utilized. Praat is a software package available for download written by Paul Boersma and David Weenink. The program helps in the analysis of phonetics in speech. With Praat the data was able to be organized and sorted into clear distinct categories.

3.4 Introduction to the LENA Coding System

We received the data in auditory form from the LENA recorders. To process this data it was then plugged into a University computer with the specific LENA program. It provided the number of adult words heard, the estimated number of conversational turns with an adult, information on the audio environment, and information based on the child's vocalizations. This data was unable to be used from the LENA system because this system is unable to distinguish between multiple language use. To achieve accurate data from the audio recordings, myself and other CITI certified undergraduate students coded all the sound files by hand. We created two coding systems on a downloadable computer software package called Praat. With this program we designed a script that we are able to collect information on the duration of time spoken, what language, and by who. The other coders and myself only coded significant speech which we identified as when the child was being spoken to, when we believed the child was listening to a conversation, and when the child was interacting with others including during group work. However, because we are not physically with the participants while they are being recorded, the coders were not able to mark down exactly what the child was listening or attending to compared to what was in close proximity to them.

Due to the data being recorded in audio form, it is impossible to distinguish between which exact individual is speaking. For example the main teacher and the teacher's assistant are unable to be distinguished. Therefore, to code the different individuals the child was

listening to, they are grouped in three groups, adults, peers, and the child themself. When the child is in the home environment, the group peers includes the siblings and whomever the child interacts with after school. These groups are then divided into which language is spoken. The two options for language spoken are Cantonese and English. The change between the home and school environment is noticeable for the coders. There is less background noise and less input from their peers in the audio recordings. Overall the duration, language, environment, and who is speaking is recorded.

3.5   Coding System

Due to the LENA Pro system being unable to distinguish input and output from multiple languages the audio data was coded by hand. For this study, on Praat we utilized six differently labeled tiers (Figure 1). Each participant (peers, adults, and the child) had two tiers for each of their languages (Cantonese and English). When coding, the tiers were labeled as follows (Table 3).

Table 3: Participant and language as coded on Praat on both coding systems.

As Coded	Participant and Language
AC	Adult Cantonese
AE	Adult English
CC	Child Cantonese
CE	Child English
PC	Peer Cantonese
PE	Peer English

There were three sound files completed with this coding system (A, B, D). The coders and myself did not code the syllables or number of words in this system because we were interested in the duration of speech and not the word count. Instead we looked at the amount of time that each participant spoke significant speech and in what language it was spoken.
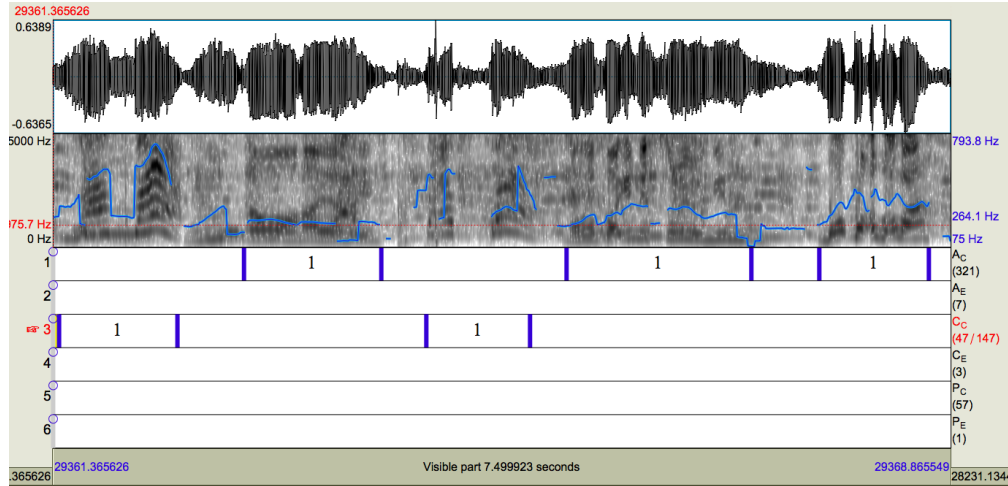
16

Figure 1: New Coding System on Praat

For three of the sound files (participants C, E, F) we used a differently designed coding system on Praat that included nine tiers (Figure 2). Each participant (peers, adults, and the child) had three tiers. One tier is for when a participant speaks Cantonese, the other two are used when English is spoken. Cantonese only needs one tier because each word in Cantonese is one syllable making the syllable and the word count the same. This is not the same in English. Therefore, one of the tiers for English was used to count syllables while the other was used to count words. This then provided the same information for both languages.
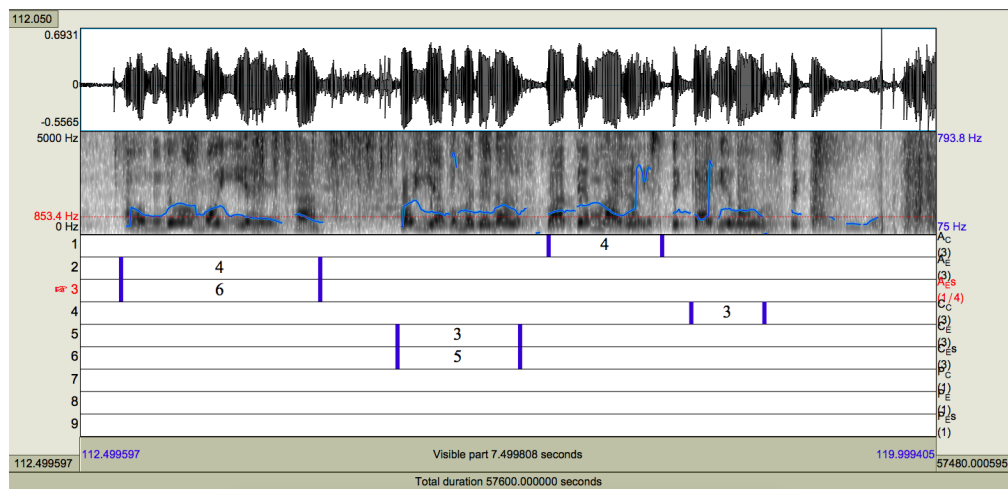

Figure 2: Old Coding System on Praat

4 Results

4.1 Questionnaire

For the six children in this study, the results of their questionnaires were compiled. Out of the questions asked to parents, this study focuses on whether the child has a language impairment, whether the parent is concerned about their speech and language, and what is the predominant language spoken at home. The results of this questionnaire for each child is tabulated below in Table 4.

Table 4: Questionnaire results for the 6 children in this study.

Child Identifier	Concerns	IEP	Language Spoken at Home
A	Speech and Language	No	100% Cantonese/Toisan
B	Speech and Language	Yes	80% Cantonese/Toisan, 20% English
C	None	No	100% Cantonese/Toisan
D	None	No	100% Cantonese/Toisan
E	Speech and Language	Yes	100% Cantonese/Toisan
F	Speech	Yes	80% Cantonese/Toisan, 20% English

The questionnaires also include information regarding the children's amount of time spent reading, playing games, and telling stories out loud. However, due to the data being in audio form these activities were not able to be recognized. Other data from the questionnaires include percentage of what language was spoken around the child from the mother, father, grandparents, and siblings.

4.2 Recorded Audio Data

Due to the small sample size, I analyzed the patterns of the amount of L1 and L2 use by the child (output), adults (input by all adults, including parents, teachers, grandparents), and peers (input, including siblings and peers). In what follows, I first present the data from the children as a group, and then qualitatively analyze the data for each child.

4.3 Input and Output of Participants as a Group

I also attempted to understand the patterns of the children as a group. Thus, from audio sound files the language spoken, the duration of significant language, the speaker, and the environment were recorded. All data was normalized by dividing the amount of time from the total sound file length. For example, to find the total percent time of adult Cantonese (AC):

$$AC_\% = \frac{Time_{AC}}{Length_{File}} \tag{1}$$

The average percent time spoken was found for each language and environment based on the speaker. To analyze general trends of the language input and output, the average percent time across multiple children was found. For example, the average percent adult Cantonese time ($AC_{\%,Ave}$) can be found by:

$$AC_{\%,Ave} = \frac{AC_{\%,1} + AC_{\%,2} + \cdots + AC_{\%,n}}{n_{children}} \tag{2}$$

Using this information, the standard deviation of the sample (σ) was calculated. For example, to find the standard deviation of the percent time in adult Cantonese:

$$\sigma = \sqrt{\frac{\sum_{n=1}^{n} \left(AC_\% - AC_{\%,Ave} \right)^2}{n_{children} - 1}} \tag{3}$$

This allows the data recorded to be generalized to a population and sets the upper and lower bounds on the error bars.

4.3.1 Overall Data

In an effort to increase readability, the recorded data has been averaged across all children (Figure 3) as described above. The percent times in Figure 3 represent the average total percent time all six children received input or produced output across their individual day.

This illustrates the average percent time the children spent participating with language.
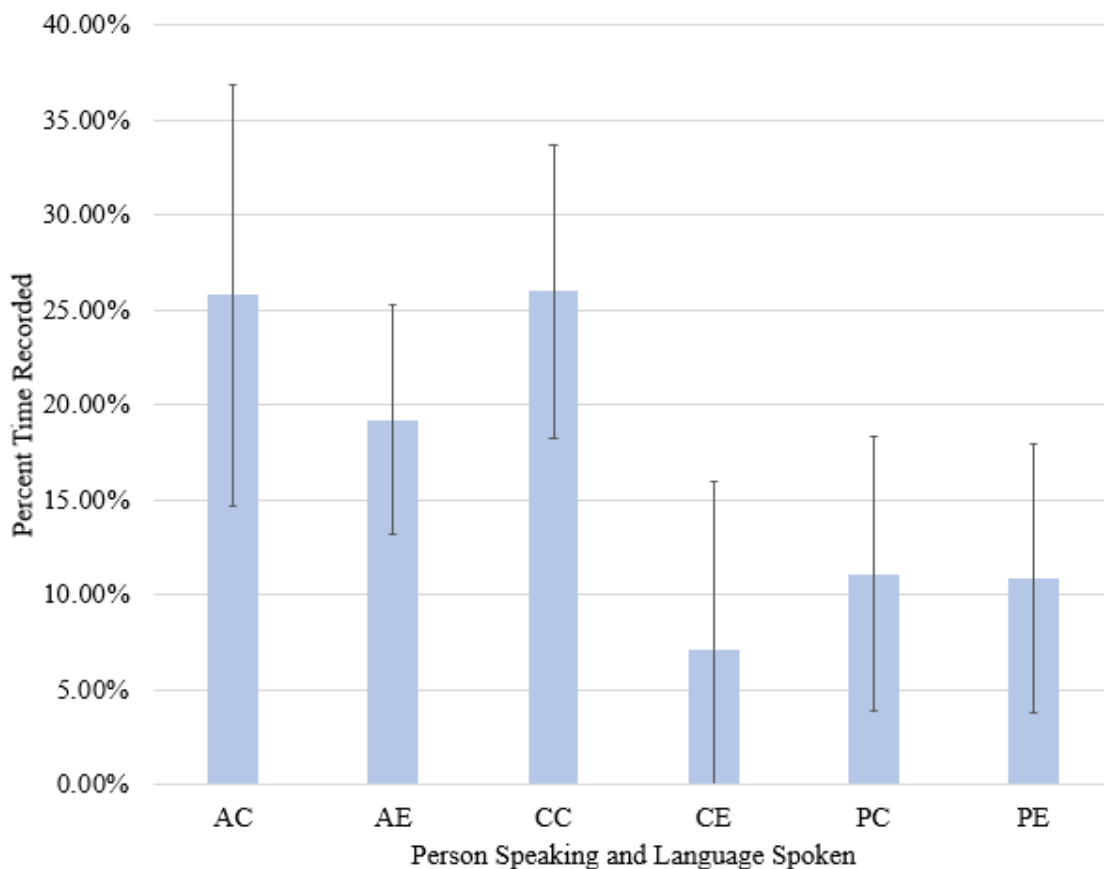

Figure 3: Average total percent time recorded for all 6 Children. Error bars represent the standard deviation from the mean as described in Section 4.3.

In addition, the data collected was on all six children, but only three children (A, B, D) had information collected regarding their home and school environments. Only three children are used to illustrate input and output in different environments because a new coding system was used where environment coding was allowed. This coding system was implemented part way through the audio recording coding. This data is shown below in Figure 4.
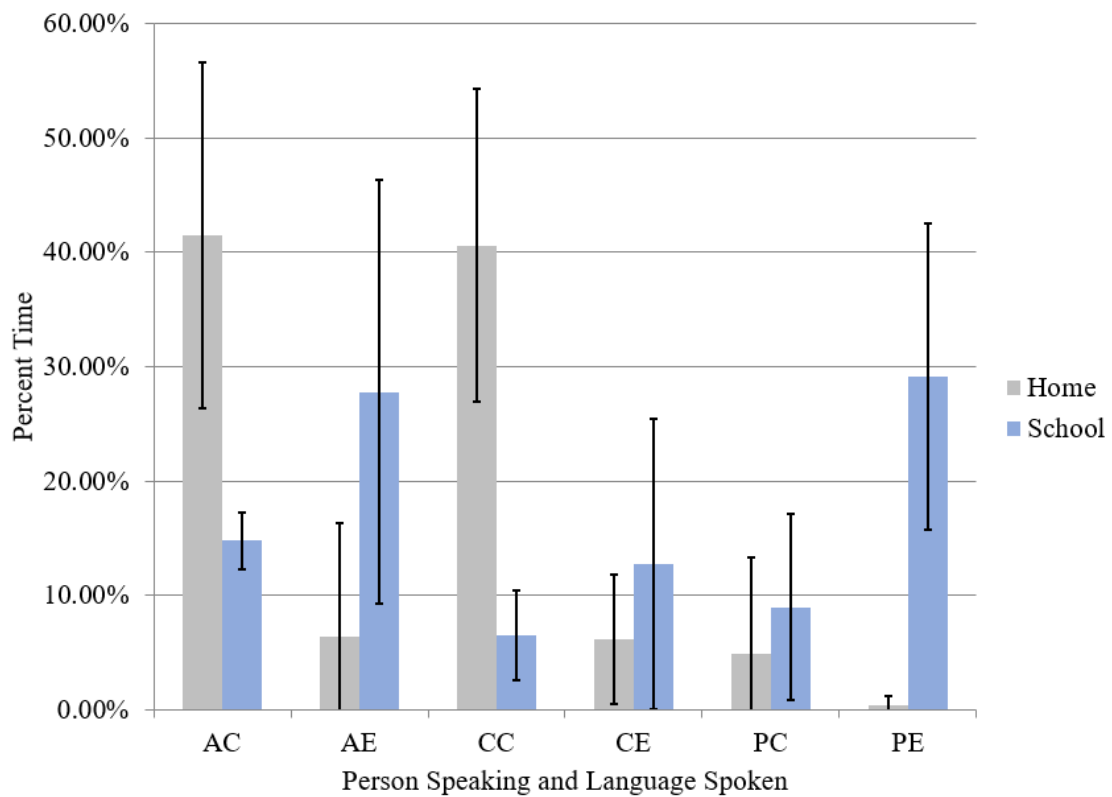
Figure 4: Average time in home and school environments a person speaks certain language. Error bars represent the standard deviation from the mean as described in Section 4.3.

The relationship between the different groups of individuals and their language input and output is represented. There are considerably large standard deviations because each child's day is different. Standard deviations do not go past zero because you can not have a negative amount of language input. As a group the adults produced more input in their L1 language (Cantonese). Peers produced more input in their L2 language (English) as a group. Finally, as a group, children used their L1 (Cantonese) to produce more output.

4.4   Input and Output of Each Child

Each child's individual input and output from each participant (adult and peers) is seen in Figure 3 in the appendix. There is a large variability between the input and output of each child. Every child did have input and output present in every language from every participant. On average there is more output by the child in Cantonese (L1). For peers input there is not a distinguishable pattern present among all participants. The adults on average produced more input in Cantonese (L1) but are close to equal to their input in English (L2).

4.4.1   Children Without Language Impairment

Three out of the six children in this study are notified as not having a significant language impairment by not having IEPs on their individual parent reports (children D, E, and F). The results of this are shown below for child A, B, and C in Figures 5, 6, and 7 respectively.
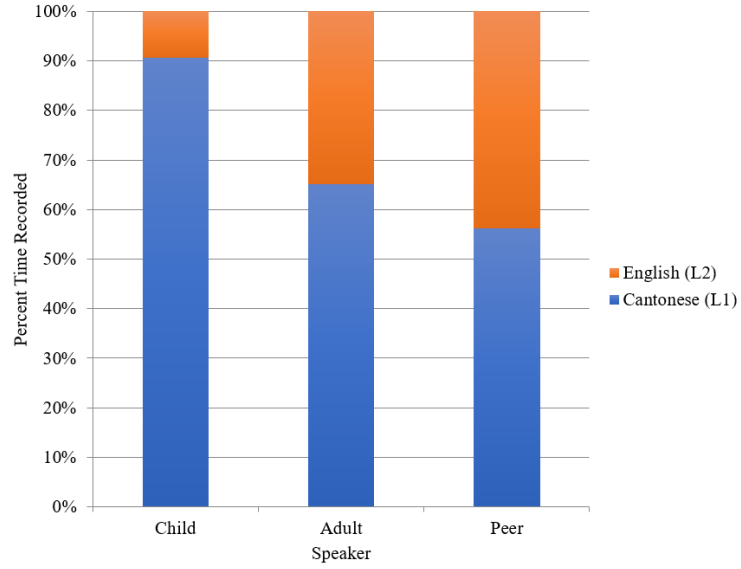
22

Figure 5: Percent of time recorded and the person speaking per language for Child A.
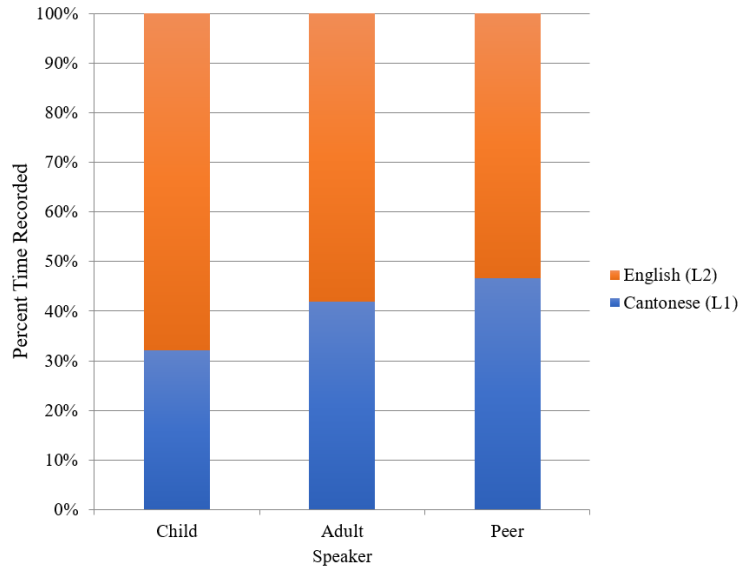

Figure 6: Percent of time recorded and the person speaking per language for Child B.
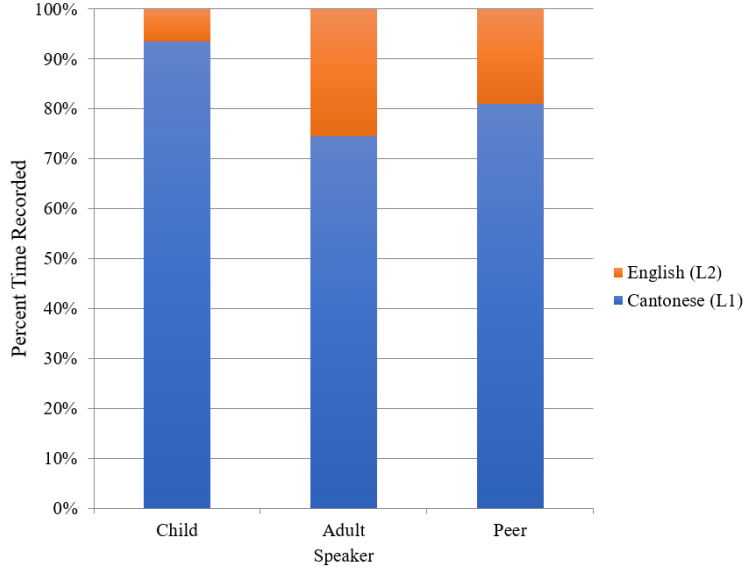
Figure 7: Percent of time recorded and the person speaking per language for Child C.

Unlike the easily spotted patterns from children with language impairments, the children without language impairment's do not have easily spotted patterns. Adults still provide a large amount of input in both languages. Two out of the three children without language impairment produce more output in their L1 language than their L2 language.

4.4.2 Children With Language Impairment

Three out of the six children (A, B, C) in this study are children with a language impairment. For all three children, the time when significant language was spoken was summed. This was then divided by the amount of time each speaker spoke in Cantonese (L1) and English (L2). This then equates to 100%.

$$Time_{Percent} = \frac{AC + AE + CC + PC + PE}{Time_{Significant}} \tag{4}$$

The results of this are shown below for children D, E, and F in Figures 8, 9, and 10 respectively.
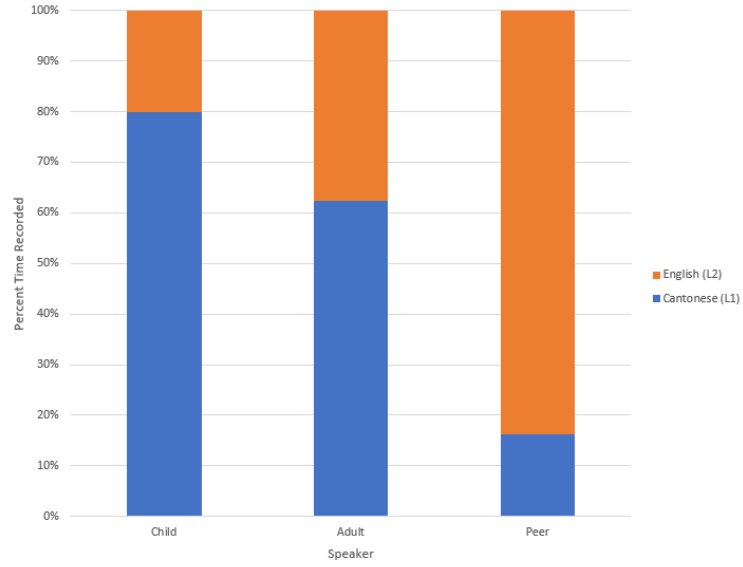
24

Figure 8: Percent of time recorded and the person speaking per language for Child D.
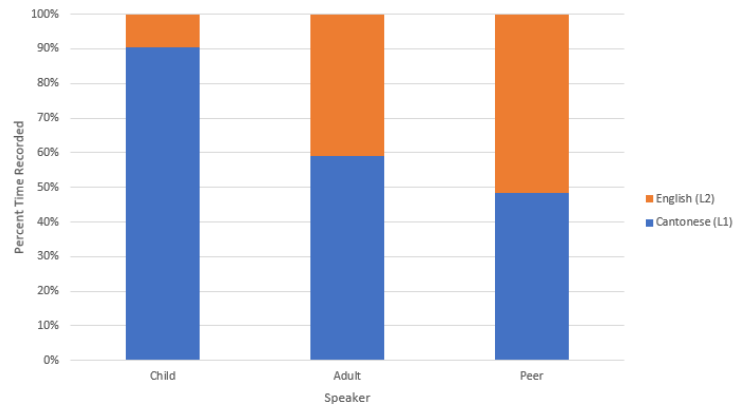

Figure 9: Percent of time recorded and the person speaking per language for Child E.
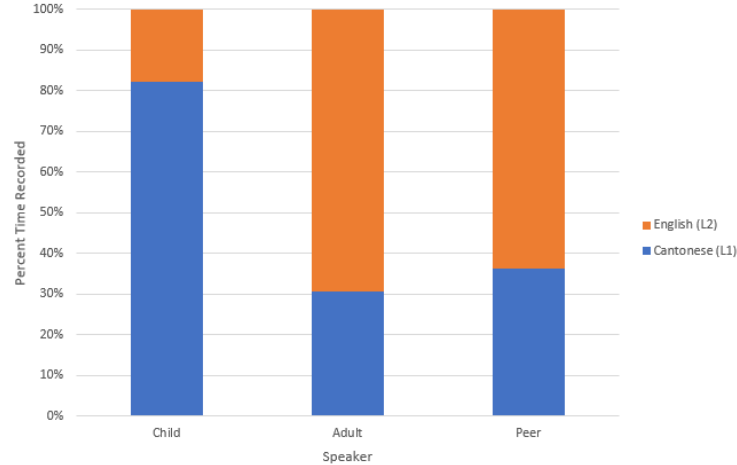
Figure 10: Percent of time recorded and the person speaking per language for Child F.

From the three previous graphs based on the results from the children with language impairment, patterns regarding language use are seen. All three children produced a greater output in Cantonese (L1) than English (L2). However, the children's peers produced a greater amount of English (L2) than Cantonese (L1) for all three children. Adults produced a large amount of input in both languages.

4.5 Comparison Between Questionnaire and Audio Recording Results

A relationship between the parent questionnaires and the audio recording for each child is the assumption that the child is more comfortable speaking their L1. On all the children's questionnaires it marked they were more comfortable speaking in their L1 (Table 4). This was highlighted through the audio recordings by five out of the six children speaking more Cantonese (L1) than English (L2).

On the questionnaires, it asked about each family members and their speaking patterns seen in Table 5. Ten out of the twelve parents (mothers and fathers) responded that they speak 100% Cantonese. With audio recording we were unable to distinguish between the different speakers (mother, grandma, father, teacher...). However looking at the percent

time language input was heard, three out of the six children had more input from adults in Cantonese (L1). In Figure 4 where all three of the children with environment coded data were averaged to compare percent time across the three children, adults produced more input in Cantonese (L1) in the home. This compares directly with the information provided by the parent reports.

5   Discussion

With the intent to compare the parent reports and recordings, reports were analyzed and compared with the audio files of the children's daily and natural environments. The purpose of this discussion is to take these data findings and examine the language input and output of bilingual preschool children who learn a second language in their school settings.

5.1   Environmental Input and Output Patterns

In the parent reports, there were not specific questions regarding the input from the children's environment. There were questions regarding input from various family members. Analysis of the reports concluded that on average the child received more L1 input at home from their older family members. Four out of the six children had older siblings that were said to give input 80% of the time in Cantonese (2 respondents) and 50% of the time in Cantonese and 50% in English (2 respondents).

To compare different environments, information from the audio recordings was obtained. Three children were used to analyze the different environments they were in. Children A, B, D's audio recordings were coded using the new coding system. This allows us to determine information regarding their environments. Differences between the children's home and school environments are illustrated in Figure 4. In this figure, standard deviations are not shown past zero because it is not possible to have a negative amount of time.

There is an increase in the variation of the average time spoken in school compared to

27

at home environments. All speakers who spoke English (AE, CE, PE) had more significant language spoken at school than in their home environments. Cantonese significant language use in the home was larger than the use of English for both adults and children. This finding is consistent with previous findings regarding an increase in L1 language use in the home (Cheung, 2015). On average peer Cantonese input was larger in school environments. This result is consistent to findings in similar studies (Cha & Goldenberg, 2015). Adult use of Cantonese and the average peers use of English had the smallest standard deviation from the mean.

5.2   Child Language Patterns Based on Home Language

Results showed that every child's home language use is different. They participate in different activities, live with different family members, and use technology differently. To account for the fact that every child is different, participates in different family activities, and has a different schedule, the results are normalized so they are comparable. The children who were coded using the new coding system (children A, B and D) were used because they have information regarding different activities in different environments. The different activities recorded are bed/story time, commute, eating, free time, and technology time. The percent time spoken by each speaker in their L1 and L2 languages is seen in Figure 4.

Across all the participants data, no significant amount of time was spent doing one task while they were in their home environments. This is due to the children having diverse home schedules. This result helps to highlight that every child has different experiences. For example, Child B spent 54.8% commuting (Figure 6) while Child A only spent 3.9% of their total time commuting (Figure 5).

5.3   Child Language Input and Output in Various Environments

Input from these children's home and school settings are very diverse. The children were the same ones who were coded under the new coding system; children A, B, and D. To

compare the three children's environments, their results were normalized (Figure 4).

On average the input at home from adults in Cantonese (41.5%) is very similar to the child output in Cantonese (40.6%) in the same environment. There is a similar result with adults who speak English (6.5%) and the children's output in English (6.1%). The largest contributors to a child's source of input on average are adults speaking Cantonese (41.5%) and adults speaking English (6.5%). These results are consistent with the results found in the questionnaires for these participants. The questionnaire results include that the parents believe the children hear more Cantonese (L1) in the home by all adults (grandparents and parents) and peers (siblings) present at home. These results are also similar to Cheung's (2015) findings that as input from adults increase, output from the children increases (Cheung, 2015).

In the school environment there is a significant increase in the children's peers English input (29.1%) with a small standard deviation. This is compared to peers English input at home (0.5%). On average, there is more English output from the child at school (12.8%) compared to at home (6.5%). The largest on average source of input in school environments is peers speaking English (29.1%) and adults speaking English (27.8%).

5.4 Parent Reports and Recorded Data

All children were born in the United States and all participants (6) spoke languages other than English in their home environments. Their responses included Toisan, Cantonese Zhong Shan, Cantonese Toisan, and Cantonese. The questionnaire also shows that the children are said to be more comfortable speaking languages other than English. The languages they are more comfortable speaking include Toisan, Cantonese Shan, Cantonese, and Cantonese English.

The percent of language spoken at home by family members was hypothesized in the parent reports. All of the mothers (6) in this study are reported as speaking 100% Cantonese/Toisan. Four out of the six (66.6%) are recorded as speaking 100% Cantonese/Toisan.

One father is reported as speaking 80% Cantonese/Toisan, 20% English and one father is reported as speaking 50% Cantonese and 50% English. Only two out of the six participants had older siblings. One older sibling was recorded as speaking 80% Cantonese/Toisan, 20% English and the other older sibling was recorded as speaking 50% Cantonese and 50% English. No younger siblings were reported. All grandmothers (6) and all grandfathers (6) reported spoke 100% Cantonese/Toisan. Below is Table 5 summarizing this data according to how it is coded. How it is coded is referenced in Table 2.

Table 5: The amount of each language the members of the children's family are reported to produce.

Child Identifier	Gender	Older Sibling	Mother	Father	Grandmother	Grandfather
A	Female	7	7	7	7	7
B	Male	6	7	7	7	7
C	Female	N/A	7	6	7	7
D	Male	N/A	7	7	7	7
E	Male	N/A	7	7	7	7
F	Male	N/A	7	4	7	7

The amount of Cantonese compared to the amount of English that is spoken decreases across the three generations represented. No child speaks 100% Cantonese, while all grandparents spoke 100% Cantonese. There was not a difference found regarding language use across the different genders and generations. Information found in the parent reports are consistent with that found in the audio recordings. However, audio recordings are more variable. This finding is similar to previous results that determine that parent reports are a reliable source of information to predict language exposure and bilingual development (Paradis, 2017). However, bilingual development was not a focus in this study but it could be a future focus of study.

The total amount of hours from the six sound files is 65.5 hours. The mean length of the average sound file is 10.9 hours. The longest recording is 16.0 hours long and the shortest sound file was 6.5 hours in length. The sound files are different lengths are due to different start times for school, leaving the microphone on when the child is sleeping, and other human

errors. In order to compare the results, the differences in lengths of the clips were accounted for when processing data by taking the average of the results.

5.5   Differences for Children with Language Impairment

There was not a significant difference in the amount of language output from the children with a language impairment compared to children without an language impairment. The same result occurs with what they heard or their language input. For children with language impairments, more English input from the adults was prevalent. Five out of six children had more Cantonese (L1) output than their English output, except for Child B. Overall, there was not a significant difference between the children with language impairments input and their produced language output. A common symptom of having a language impairment is having decreased language skills (Armon-Lotem, 2017). The impairment creates a chain reaction where the impairment affects the amount of input the child receives. Due to the child having low language skills and therefore not responding as much as a child without a language impairment, parent responses might be affected which could lead to a decrease in input heard and therefore creating less opportunities to produce language output.

Due to the lack of studies of bilingual children who have language impairments and the small sample size presented in this study, the incidence of language impairment in this population is unknown (Kohnert, 2010). Further research is needed. Similarities in this study compared to previous findings suggest that bilingual environments do not seem to have any adverse or positive effect on these children (Kohnert, 2010).

6   Conclusion

All children were assumed to be more comfortable with Cantonese (L1) then English (L2) as answered in the parent reports. This can be illustrated by looking at the duration of output per child. Five out of six children spoke more Cantonese than English. The

only child who did not meet this norm just spoke English more by 7.83%. It can now be assumed the parent or caretaker who fills out the parent reports is reliable when they state the child's preferred language. By comparing the parent reports to the audio recordings, audio recordings highlight more variability that from the parent reports regarding the child's familiar relations and language use. On the other hand, parent reports did provide baseline information about children's general language patterns. These findings are similar to the results of Mancilla-Martinez et al. (2016) who found that parent reports are a cost effective and valid measure to analyze early childhood education settings (Mancilla-Martinez et al., 2016).

After comparing children without language impairments to an equal number of children with language impairments, there is not any significant difference with the children's output or input. Due to the children having different daily environments, more participants are needed to be able to accurately assess. Then this would determine if there is a significiant difference between children's language input and output with language impairments and those without.

A significant finding in this study is the strong relationship between adult input and child output in both Cantonese and English. This relationship is illustrated in both Figure 3 and Figure 4. This finding also relates to which language the child is stated to be more comfortable with and the language the mother speaks. This finding is consistent with previous studies stating speakers who provide input are positive predictors of a child's language skills (Place & Hoff, 2016).

6.1   Implications

An inference from the findings in this study is that the language a child hears for a longer period of time in the home may become the language the child is more comfortable interacting with. This result could lead to families learning a second language so that their children can learn to be comfortable using multiple languages. The importance of home language use by

all family members to improve a child's use of that language is also important. The more a child hears a language and interacts with others through different activities with a language, the more they will be comfortable using it.

Studying the PLI in bilingual children will lead to accurate and timely identification (Kohnert, 2010). Correct identification of bilingual children with language impairments will help others to provide proper support to children who need it instead of just identifying the severe cases. An increase in research of this topic will also decrease the delay of identification of language concerns.

If a parent or caregiver of a child wanted to increase their child's English (L2) language output, they need to increase their auditory input in that language in various environments with an emphasis in their home environment. This auditory input can come from many different sources such as through conversation, technology use, or from new activities.

Illustrating the relationship between adult input and children's output is important. The more adults speak, the more output children will produce which can lead to these children having a larger vocabulary than their counterparts who received less adult input. This input difference has important implications at home and in classroom settings. At home, parents and caregivers would be able to interact and speak more with the children using a mature vocabulary. This result also implies that having varied language input throughout the home environment is important. Exposing a child to different settings and situations could be beneficial because it would increase the amount of language that is heard while exposing the child to an expanded vocabulary. In a classroom setting these findings could lead to new lesson planning that promotes speech activities and leads to more children reading out loud. This also implies having a reduced teacher to student ratio is beneficial. The more adults that are present within a classroom giving input, the more input children will receive which will increase a child's output.

The implications of this study for the bilingual community and future research are extensive. School's could help provide a more social environment that is friendly to bilinguals

and if possible they could reduce teacher to student ratios. Home environments could make adjustments to better suit a dual language learner where the environment provides an abundance of input in both languages. Finally, researchers could use this data to further examine language input and specific activities that children participate in. Another future research topic could be how reading out loud by adults or peers (siblings) effects a child's linguistic output.

6.2 Limitations

A limitation of this research was that the data was collected through auditory recordings and not through visual observations. When coding the sound files it was difficult for the research assistants and lab volunteers to distinguish between when voices were providing significant input for the child. For example, if an adult has a loud voice it may seem to the coders that the adult is in close enough proximity to the child and that they are providing significant linguistic input to the child. Due to not directly observing the child, the coders of this data do not know if the child is fully attending to what an individual is saying. If the data was in video format or there was direct observation, the coder could visualize if the child was reacting (turning a head or making eye contact) silently to stimuli which is missed through audio recordings. The research assistants and I made educated guesses to solve this problem by listening for information cues to determine if the child responds which included determining if the child makes noises immediately after language input as if they are participating in a conversational turn. For future research I recommend having a video recording of the child. This would aid researchers when coding the data to visualize if the child is indeed communicating with someone. Having the data visually presented would also provide information on the specifics of who is speaking to the child.

Another limitation of this study is the lack of questions on the questionnaire regarding parents perspectives on how often their children participates in certain activities. The questionnaires were paired with the sound files. If possible it would have been intriguing to write

a questionnaire that would specifically ask questions relating to the children's different environments, habits, and obtaining more of the parent's perspective on the children's language use and input.

It would be interesting to have the children partake in formal language testing previous to participation in this study. This would give an insight into where the child stands compared to an average child and their peers. Formal language testing would also provide a baseline for the children with language impairments in this study which would provide more information on the bilingual language impairment population.

6.3   Further Research

Due to this study being the first study of its kind, further research is needed and encouraged. I would recommend a more detailed questionnaire highlighting environment and input patterns. Visual data instead of audio recordings are recommended for future studies. Further research could also include having more bilingual children from different schools participate. Testing the participating children for language impairment and expressive vocabulary in both languages would also be a beneficial tool for study.

This study only took audio recording from one day for an average of 10.9 hours. This may not represent a typical day for a child. For example a child may have sports practice on certain days or live at a different parent's house on different days. For future research, researchers should have the child wear the LENA Pro device over an extended period of time. Then the daily input and output can be averaged to illustrate a typical day. This would demonstrate the changes that bilingual children undergo over a longer period of time.

I expect to see future undergraduate students in the Child Language and Learning Laboratory use this study as a springboard to further investigate bilingual environments and their input patterns. The increase of compiled data would lead to a study that solidifies or disproves the finding in this paper and lead to more conclusive findings that can be used for bilingual children and their families.

I hope the findings in this paper help researchers to develop more research questions and new ideas that could lead to more studies of the bilingual children population. The more studies that are completed, the more bilingual children will benefit as they will have data just for their population that highlights what they need to focus on regarding language input, their environments, and learning.

References

Armon-Lotem, S. (2017). Disentangling bilingualism from sli:dissociating exposure and input. *Bilingualism: Language and Cognition*, *20*, 33-34.

ASHA. (2004). *Knowledge and skills needed by speech-language pathologists and audiologists to provide culturally and linguistically appropriate services.*

Auer, P. (1984). *Bilingual conversation.* John Benjamins Publishing Company.

Bernardini, P. (2017). Weak interest in the weaker language. , *20*(1), 29–30. Retrieved from `http://dx.doi.org/10.1017/S1366728916000341`

Bialystok, E. (2010). Bilingualism. *WIREs Cognitive Science*, *1*, 559-572.

Brody, G. H. (2016, June). Siblings' direct and indirect contributions to child development. *Current Directions in Psychological Science*, *13*, 124-126.

Carroll, S. E. (2017, January). Exposure and input in bilingual development. *Bilngualism: Language and Cognition*, *20*, 3-16.

Cha, K., & Goldenberg, C. (2015, November). The complex relationship between bilingual home language input and kindergarten children's spanish and english oral proficiencies. *Journal of Educational Psychology*, *104*, 935-953.

Cheung, S. M. (2015). Comparisons on the cantonese-english language use of family members in bilingual home environments and of teachers in bilingual classroom environments. *Undergraduate Honors Theses*, *852*.

De Houwer, A. (2017, January). Bilingual language input environments, intake, maturity and practice. *Bilingualism: Language and Cognition*, *20*, 19-20.

Duursma, E., Romero-Contreras, S., Szuber, P., Annaand Proctor, Snow, C., August, D., & Calderon, M. (2007, January). The role of home literacy and language environment on bilinguals' english and spanish vocabulary development. *Applied Psycholinguistics*, *28*, 171-190.

Finegan, E. (2012). *Language: Its structure and use.* Micheal Rosenberg.

Foundation, L. R. (2012). *LENA Pro Advanced technology to accelerate language devel-*

opment of children 0-5 and for research and treatment of language delays and disorders. Brochure. Retrieved from `http://lena.org/wp-content/uploads/2016/07/`
`LTR-11-1_LENA-Pro-Brochure.pdf`

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

Mancilla-Martinez, J., Gámez, P. B., Vagh, S. B., & Lesaux, N. K. (2016, January). Parent reports of young spanish–english bilingual children's productive vocabulary: A development and validation study. *Journal of communication disorders*, *47*, 1-15.

Paradis, J. (2017, January). Parent report data on input and experience reliably predict bilingual development and this is not trivial. *Bilingualism: Language and Cognition*, *20*, 27-28.

Place, S., & Hoff, E. (2016). Effects and noneffects of input in bilingual environments on dual language skills in 2 1/2-year-olds. *Bilingualism: Language and Cognition*, *19*, 1023.

Verhagen, J., Mulder, H., & Lesseman, P. P. (2017, January). Effects of home language environment on inhibitory control in bilingual three-year-old children. *Bilingualism: Language and Cognition*, *20*, 114-127.

Zhang, H. S. (2016, November). Early language input and later reading development in chinese as heritage language (chl) learners. *International Journal of Applied Linguistics*, *26*, 437-448.

A About the Author

I am personally drawn toward the topic of the validity of parent reports because I am going to school at the University of Colorado Boulder to become a childhood speech language pathologist. Being a speech pathologist requires a relationship with parents, caregivers, and teachers. Therefore, having a deeper understanding of the validity of the results of parent reports will assist me, and other speech pathologists, in comprehending what is best for the child regarding testing and treatment. I find bilingualism to be an interesting topic since many of my friends and family grew up in a bilingual household. I was surrounded by many languages including Hebrew, Arabic, and Spanish. I saw my best friends use common patterns of use like using L1 at home and code switching. Personally, to receive the benefits of being a multilingual, I hope to expand my knowledge of languages in both French and American Sign Language.

B Recorded Data

Table 6: Percent time spoken by speaker in certain language for each participant.

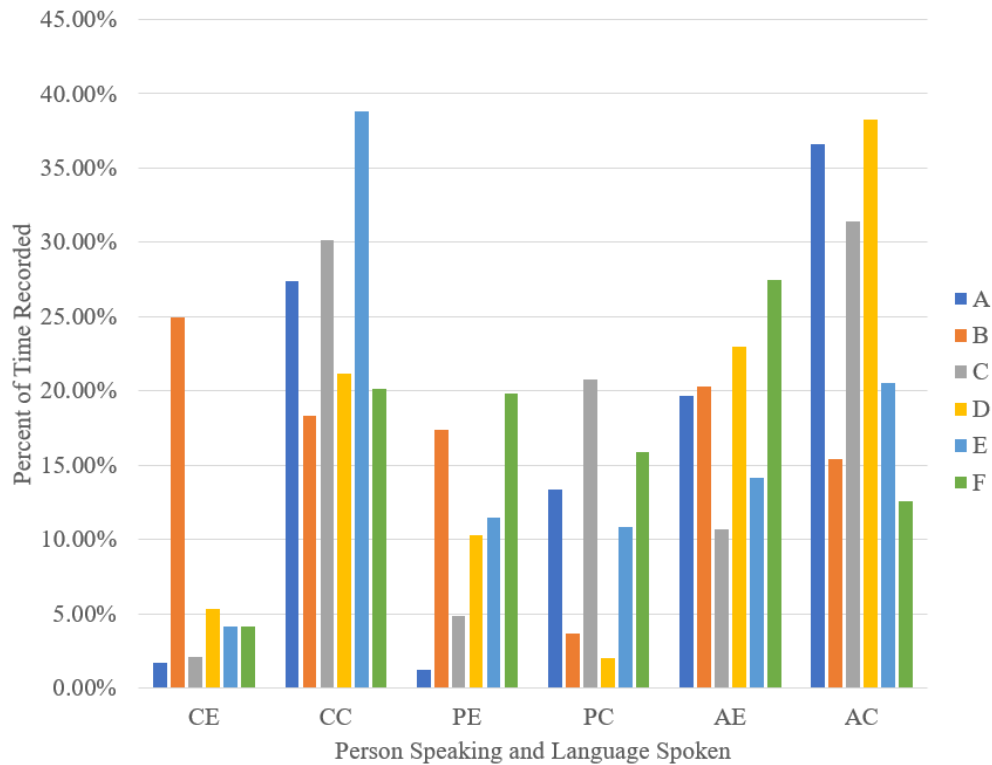| Coded ID	CE	CC	PE	PC	AE	AC
A	1.71%	27.42%	1.26%	13.36%	19.67%	36.59%
B	24.94%	18.30%	17.35%	3.64%	20.33%	15.44%
C	2.08%	30.15%	4.89%	20.79%	10.69%	31.40%
D	5.34%	21.17%	10.27%	1.99%	22.98%	38.25%
E	4.15%	38.81%	11.50%	10.84%	14.18%	20.53%
F	4.16%	20.14%	19.86%	15.86%	27.45%	12.54%

Figure 11: Percent time spoken by speaker in certain language for each participant.