DIFFERENCES IN ADULT SPEECH: DO BOY-DIRECTED AND GIRL-DIRECTED SPEECH EXIST?

Tori Keeping
Bachelor of Science, University of Lethbridge, 2020

A thesis submitted
in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

in

PSYCHOLOGY

Faculty of Arts and Science
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

© Tori Keeping, 2023

DIFFERENCES IN ADULT SPEECH: DO BOY-DIRECTED AND GIRL-DIRECTED SPEECH EXIST?

TORI KEEPING

Date of Defense: April 6, 2023

Dr. F. Li Thesis Supervisor	Professor	Ph.D.
Dr. D. Logue Thesis Examination Committee	Associate Professor	Ph.D.
Dr. J. Mather Thesis Examination Committee	Professor	Ph.D.
Dr. J.B. Leca Thesis Examination Chair	Associate Professor	Ph.D.

ABSTRACT

It has been demonstrated that boys and girls speak differently; however, the question of whether this is directly a result of the input they receive from adults is yet to be determined. Forty university students participated in a speech production study in which they were asked to read a short passage to different gendered images of babies. Acoustic analyses were conducted on their recorded speech, with measures including the centroid frequency of the sound /s/, as well as f0 and intensity. Although there were no significant findings in m1 values, f0 values were significantly different, indicating that adults greeted photos of baby boys with a higher f0. The results also suggest a trend indicating adults may greet photos of baby boys with greater intensity than girls. Although these findings were not robust, they may suggest differences in speech may exist in the ways adults converse with boys versus girls.

Keywords: gendered speech production, acoustics, phonetics, child-directed speech

ACKNOWLEDGEMENTS

There are several amazing people who have contributed to my thesis and helped me along this journey who I owe many thanks. First and foremost, I want to thank my supervisor, Dr. Fangfang Li, who has given me an immense amount of support and guidance over the past two and a half years. Her mentorship has helped me become a better researcher and I would not have been able to accomplish this without her.

Secondly, I would like to show gratitude to my committee members, Dr. David Logue and Dr. Jennifer Mather. They both have encouraged me to think about things with a new perspective and have been so encouraging along the way. I want to thank them both for their time, guidance and support. I would also like to acknowledge Dr. Tyler Bonnell who was always there to help me untangle the complexities of my statistical models. His help never went unnoticed.

To all my lab members, thank you for always being supportive. I would like to especially thank Claire, Sydney and Ginger who have been there to keep me motivated through all the ups and downs. I would also like to extend an extra special thank you to my research assistant Cassie. She has been a beacon of support for me, and I would not have been able do this without her.

Finally, I want to thank my loving family. I could not have done this without their constant support. I am forever grateful for their never-ending confidence in me as it gave me the strength to persevere.

TABLE OF CONTENTS

Abstract……………………………………………………………………..………...iii

Acknowledgements……………………………………………………..……….iv

List of Tables…………………………………………………………………........vii

List of Figures…………………………………………………………..………...…viii

List of Abbreviations……………………………………..……………..……....ix

Chapter 1: Introduction……………………………………………………………1

1.1 Gender Differences in Adult Speech………………………………………….2

1.2 Gender Differences in Child Speech…………………………………………..5

1.3 A Possible Mechanism for Gendered Speech in Children……………………7

1.3.1 General Accommodation Differences………………………………7

1.3.2 Input Differences for Boys and Girls………………………………..9

1.4 Purpose and Hypothesis……………………………………………………...11

1.5 Implications…………………………………………………………………..12

Chapter 2: Methods……………………………………………………….………....13

2.1 Participants…………………………………………………………………...13

2.2 Study Design…………………………………………………………………13

2.3 Procedure…………………………………………………………………….…15

2.4 Analysis………………………………………………………………………17

Chapter 3: Results…………………………………………………..…………..22

3.1 Spectral Mean Frequency/ Centroid of /s/ Production……………………….22

3.2 Average Fundamental Frequency (f0) Results……………………...…….26

3.3 Average Intensity Results……………………………………………………32

Chapter 4: Discussion…………………………………………………..…………38

4.1 Limitations and Future Considerations……………………………………………40

4.2 Conclusion…………………………...……………………………………….…41

References……………………………………...……………………………….44

Appendix A: Passages Used in the Current Study……………………………..…..……..52

Appendix B: Baby Photos, Names and Facts Used in the Current Study………..…..54

Appendix C: Face ID's of Photos………………………………..…………………...55

LIST OF TABLES

Table 1: Descriptive Statistics for m1 Values…………………………………………....22
Table 2: Population Level Effects for m1 Results…………………………...…………..25
Table 3: Descriptive Statistics for f0 Values…………………………………………….26
Table 4: Population Level Effects for f0 Results………………………………….…..30
Table 5: Descriptive Statistics for Intensity Values…………………………..…………32
Table 6: Population Level Effects for Intensity Results………………………………....36

LIST OF FIGURES

Figure 1: Spectrogram in Praat……………………………………………..…………..18
Figure 2: Posterior Predictive Check for m1………………………………………..….23
Figure 3: Conditional Effects Plot for m1……………………………………………..24
Figure 4: Posterior Predictive Check for f0…………………………………….....….28
Figure 5: Conditional Effects Plot for f0……………………………………….…...29
Figure 6: Box Plot of f0 Differences for Speaker Gender, Baby Gender and Task…..….31
Figure 7: Posterior Predictive Check for Intensity…………………………………….34
Figure 8: Conditional Effects Plot for Intensity…………………………………...…..35
Figure 9: Box Plot of Intensity Differences for Speaker Gender, Baby Gender and Task…………………………………………………………………………………37

LIST OF ABBREVIATIONS

ADS	Adult-directed speech
CDS	Child-directed speech
CI	Credible interval
FDS	Foreigner-directed speech
JND	Just noticeable difference
IDS	Infant-directed speech
RMS	Root mean square
SD	Standard deviation
VOT	Voice onset time

CHAPTER 1: INTRODUCTION

It has been evidenced in the literature that males and females tend to have gender specific differences in their speech production (Labov, 1990; Leaper & Ayres, 2007; Pépiot, 2014; Simpson, 2009; Weirich & Simpson, 2018). Some of these gendered differences have also been found to emerge in early childhood. The ongoing debate surrounding nature versus nurture is one aspect to consider when looking into the language acquisition of children. It may be that biological factors, environmental factors or a combination of both, play a role in why children develop gender specific patterns in their speech production. The role of both nature and nurture in regard to child development is one largely unanswered question. Over the years, many people have added to the nature versus nurture debate and seemingly the conclusion is that there is a combination of both aspects (the environment and biological underpinnings) that may influence child development in different ways (Eagly & Wood, 2013). However, this thesis will take the perspective that language acquisition is influenced by the environment (nurture) in a more substantial way than nature. The purpose of this thesis is to investigate one of the possibilities that could explain the early emergence of gendered speech patterns in childhood. More specifically, the aim is to investigate whether the input children receive from adults differs depending on the child's gender, that is, do adults speak differently to boys than they do girls. The following sections will provide an overview of the gender differences that have been discovered in both adult and children's speech and highlight some similarities that have emerged. Then, a detailed outline of the present study will be described, and the results will be discussed.

1.1 GENDER DIFFERENCES IN ADULT SPEECH

Noticeable gendered speech differences are prevalent in the literature when comparing the speech production of male and female adults (Leaper & Ayres, 2007). Although not all differences are relevant to speech acquisition, male and female adults exhibit many gendered differences in speech. For instance, the two genders differ in speech clarity (Chen et al., 2016; Kwon, 2010), speed (Jacewicz et al., 2009) or style (Labov, 1990), depending on the language they speak. Women produce clearer speech than men in languages such as Korean (Kwon, 2010), English (Yoho et al., 2019), French (Pépiot, 2014) and Portuguese (Albuquerque et al., 2021). Men speak faster than women in languages such as English (Jacewicz et al., 2009; Lee & Doherty, 2017; Pépiot, 2014), Turkish (Cangi et al., 2020) and French (Pépiot, 2014).

In addition to speech clarity and speed, males and females tend to employ distinct stylistic choices in speech (Fitzpatrick et al., 1995; Janssen & Murachver, 2004). For example, men tend to use more assertive language, defined as an advancement for one's own gain or power (Leaper & Ayres, 2007) whereas women use more affiliative language, which describes the tendency to positively affirm or engage with others (Leaper & Ayres, 2007) in conversations (Labov, 1990). Further differences exist in which women tend to utilize standard forms of words, whereas males use more nonstandard versions (Labov, 1990) and have more polite and cooperative conversations than men (Hannah & Murachver, 1999). Moreover, the types of conversations adults have can impact their language use. For instance, women tended to gossip more frequently and in a more positive manner than men who often take a negative perspective when gossiping (Eckhaus & Ben-Hador, 2019).

The most prominent gender specific speech patterns have been exemplified through distinctive phonetic characteristics. One notable difference between the speech of male and female adults is their pitch range. In particular, females have a higher pitch range than males (Kwon, 2010; Pépiot, 2014; Zimman, 2018). Further gendered differences in speech exist in formants of vowels such as /æ/ (as in *bat*), /a/ (as in *father*) and /i/ (as in *seed*) in American English. For instance, females produce clearer separation between prenasal /æ/ (the vowel comes before a nasal sound) and pre-oral /æ/ (the vowel comes before a sound that comes from the mouth) than males (De Decker & Nycz, 2012). In addition, females exhibit higher second formant (F2) values for /i/ (Kwon, 2010; Simpson, 2009) and higher first formant (F1) values for /a/ (Kwon, 2010; Simpson, 2009; Weirich & Simpson, 2018) than males. These differences are non-uniform, suggesting that vocal tract length may not be the only factor that produces differences in vowel production between men and women (Simpson, 2009).

Gender differences in consonant production have also been established in the literature. For example, females have shorter fricative duration (a rapid flow of air through a narrow constriction in the oral cavity; Jongman et al., 2000) for /θ/ (as in *thought*) than men and a more anterior articulation for /s/ (as in *sun*), and /ʃ/ (as in *shoe*) than men (Fox & Nissen, 2005). Li et al. (2016) suggest that social processes, particularly gender identity development, give rise to gender differences in the production of the fricative sound /s/. In addition, Mack & Munson (2012) found a link with /s/ production and sexual orientation that showcased gay men as producing /s/ less accurately than heterosexual men. Steensma et al. (2013) suggest that gender identity is more malleable in early childhood. This suggests that differences in sexual orientation may also be a

considerable contributing factor to gender differences in speech production echoed in young children.

Further differences in speech production between men and women can be seen in voice-onset time (VOT). VOT can be described as the time lag between the release of a plosive (an articulatory closure of the mouth followed by a burst of sound due to pressure build up in the mouth; Ladefoged & Johnson, 2015) and the onset of vocal fold vibration (Simpson, 2009). Females have longer VOT for plosives /p, t, b, d, k, g/ than men but the greatest differences in production were exhibited in /t/ production (Morris et al., 2008).

Anatomical differences account for a sizable portion of the differences between male and female adult speech. For instance, females tend to have shorter and smaller vocal tracts (Simpson, 2009) and a smaller larynx (Zimman, 2018) in comparison to men. Further, women do not have a full vocal fold closure; therefore, a constant airflow is achieved in typical woman voices (Simpson, 2009) causing a distinction between their vocal quality and that of men. In particular, females tend to have a breathier voice quality than men (Pépiot, 2014; Simpson, 2009; Suire et al., 2020).

Gendered speech can be attributed to a combination of factors such as social processes (Tripp & Munson, 2021) and anatomical differences (Simpson, 2009). The social situations in which men and women find themselves often dictate their language usage (Leaper & Ayres, 2007). This is in part due to gender norms which dictate acceptable behaviour for men and women (Seguino, 2007). Fitzpatrick et al. (1995) suggest that traditional sex-typed men have a harder time switching to a more feminine speech style. This may be due, in part, to societal standards which are entrenched in the beliefs of how men and women should behave. These social processes in tandem with anatomical differences influence language such that gender specific speech differences

are produced in adult speech. In other words, both nature and nurture influence the differences in phonetic characteristics outlined above, but it may be social processes in particular that influence the development of children's language acquisition.

1.2 GENDER DIFFERENCES IN CHILD SPEECH

Gender differences in children's speech, although they emerge slowly, are also evident, and may be detectable at a young age. During a listening task of children aged 2.5-5.5 years old, adults were able to accurately identify whether a recording was of a boy or girl speaker (Fung et al., 2021). The authors suggest that due to this finding, both boys and girls may be able to produce gender-specific speech patterns as young as 2.5 years of age (Fung et al., 2021).

Further evidence of gendered speaking patterns in toddlers have been identified in studies in which infants were found to begin correctly using gender labels (girl, boy, man, lady) by 18-21 months old (Zosuls et al., 2009) and by two years of age, both boys and girls use more same-gendered words (such as dress, sister, necklace for girls and cowboy, brother, hammer for boys) than other-gendered words (Stennes et al., 2005). Despite these findings, there is no robust evidence of gendered patterns of speech at such a young age as a larger portion of research focuses on gender differences in preschool aged children.

Similarly to adult speech, there are aspects of children's speech such as clarity, speed and style of speech that vary depending on the gender of the child. In regard to clarity of speech, girls were determined to be easier to understand at 4 and 5.5 years of age in Australian English (McLeod et al., 2015) and at 11-13 years of age in Italian (Prodi et al., 2019) compared to boys of the same age. In contrast, American English speaking children aged 30-47 months old did not produce similar findings in regard to clarity of

speech (Hustad et al., 2020). This suggests that patterns similar to the speech clarity of adults may not arise in children until they reach preschool age.

Divergent results in girl versus boy speech rate have also been discussed in the literature. Akin to the findings of speed in female versus male adult speech, toddler girls have also been found to have a slower speech rate than toddler boys in American English (Tenders et al., 2019). Further, a trend suggesting girls spoke slower than boys in English at ages 4-6 has also been evidenced in the literature (Walker & Archibald, 2006).

In addition to clarity and speech rate, speech style is another factor in which children display gender differences that parallel the patterns seen in adult speech. Girls tend to use more standard forms of past tense verbs (Ladegaard & Bleses, 2003), more cooperative and friendly language styles (Bablekou, 2009) and more affiliative language overall (Leaper & Smith, 2004) compared to boy who used more nonstandard past tense verbs (Ladegaard & Bleses, 2003) and more dominant and assertive language styles (Bablekou, 2009; Leaper & Smith, 2004).

There are several gender specific speech acoustic differences that emerge in early childhood that parallel the differences found in adult speech. These differences, which are of paramount interest in the current study, include pitch, consonant sounds and VOT. Regarding pitch, boys aged 6-12 had significantly lower pitch than females of the same age (Nicollas et al., 2008). Females were found to have a significantly higher centroid frequency (more anterior constriction) for /s/ and /ʃ/ than males (Fox & Nissen, 2005).

Further parallels exist in the VOT of children. In a study of children aged 7-11 years old, Whiteside & Marshall (2001) found that girls had longer VOT values than boys for plosives; however, the phonological contrast between /p/ and /b/, and /t/ and /d/ was more marked for the 11-year-old girls compared to males of the same age. A follow-up

study showed similar results in which girls exhibited longer VOT than boys and that the most pronounced difference in VOT was in 13-year-old females (Whiteside et al., 2004). One explanation of these differences could be due to sexual dimorphism; however, since these results parallel the VOT differences in male and female adults, it may also suggest that children are attuned to specific differences in speech of the same sexed adults in their environment.

Even though anatomically speaking, boys and girls have similar vocal tract lengths in early childhood (Barbier et al., 2015), gender specific differences that mimic those of adults, are still evident in children's speech. This may suggest that gender differences in speech may have less to do with anatomical structures and more to do with social processes. For example, adults may encourage specific behaviours because they hold boys and girls to different expectations. This may account for why girls and boys have specific gendered behaviours that are then reflected within speech production (Zosuls et al., 2009).

1.3 A POSSIBLE MECHANISM FOR GENDERED SPEECH IN CHILDREN

One possible mechanism that may explain why gendered patterns of speech production emerge in early childhood and why these patterns mirror the gendered patterns of adult speech may be that children receive different input based on their gender. In other words, adults could speak differently to boys than they do to girls. Based on previous literature, adults have been shown to treat boys and girls differently (Leavell et al., 2012) and use different speech styles.

1.3.1 GENERAL ACCOMODATION DIFFERENCES

Input differences affect the level of conversation in typical social interactions in many instances. Adults in particular have been found to accommodate their

conversational partners based on gender (Fitzpatrick et al., 1995), attractiveness (Hughes et al., 2010), language background (Rodriguez-Cuadrado & Costa, 2018; Rothermich et al., 2019) and age (Cockrell, 2020; Masataka, 2002; Shaw & Gordon, 2021).

Gender is one aspect that may influence the speech styles of adults. In basic conversations with strangers, both men and women used a speech style in accordance with their respective gender with same sex adults and decreased their preferred style when talking to opposite sex adults (Fitzpatrick et al., 1995). When conversing with their spouses, men adjusted their speech to their wives by using a more feminine style (Fitzpatrick et al., 1995). Evidence also suggests that relative attractiveness of a conversational partner may influence the ways in which adults address other adults. In a study conducted by Hughes et al. (2010), it was discovered that both males and females tended to lower the pitch of their voices when communicating with more attractive individuals.

When addressing older individuals, adults tend to use elderspeak. Elderspeak can be defined as a form of overaccommodation used when speaking to elders that utilizes juvenile lexical choices and exaggerated prosody (Shaw & Gordon, 2021). One study found that collective pronoun usage (saying "should we eat supper now?" instead of "can you eat your supper now") was the most common type of elderspeak (Cockrell, 2020). In addition, Masataka (2002) found that women significantly increased their pitch when talking to elders.

Another form of modification used by adults is foreigner-directed speech (FDS). FDS is used towards non-native speakers of a language and can be characterized as having vowel hyperarticulation (Knoll & Costall, 2015; Uther et al., 2007), louder and

slower speech (Rodriguez-Cuadrado et al., 2018) and simpler syntactical construction compared to speech between two native speakers (Rothermich et al., 2019).

The most relevant speech accommodation to the current research is when adults are conversing with children. When addressing infants or children, adults use shorter and simpler words and speak in a higher pitch than when they use adult-directed speech (ADS) (Jacobson et al.,1983; Masataka, 2002). In addition, experience with children can also affect the extent of speech modification such that mothers were found to modify their speech more than non-mothers while using child-directed speech (CDS) (Masataka, 2002) and individuals who grew up with siblings were more likely to use CDS than only children (Ikeda & Masataka, 2002).

1.3.2 INPUT DIFFERENCES FOR BOYS AND GIRLS

In relation to input differences, adults accommodate their speech by using infant-directed speech (IDS) or CDS when conversing with children which may result in speech pattern differences. These differences may lead to different input based on children's gender. In this regard, there is substantially more evidence demonstrated in the speech of mothers. For instance, mothers engage more with their daughters than their sons. Further, mothers ask more questions and talk more overall with their daughters than their sons in the first year of their life (Clearfield & Nelson, 2006).

Mothers and fathers have also been shown to have different types of conversations with their children. In regard to emotions emotion-related words, both mothers (Fivush et al., 200) and fathers (Denham et al., 2010) discussed these topics more frequently with their daughters than their sons. Not only do the topics of conversation adults have with their daughters compared to their sons differ, but they also use different types of words when speaking to children of different genders. One example of this is the use of spatial

words, that is, words describing the size, shape and spatial properties of spaces and objects. Pruden & Levine (2017) found that on average, parents used more spatial words with their sons than their daughters which was shown to produce better spatial skills in boys later in their development. Kuchirko et al. (2020) found that mothers of boys used more regulatory language than mothers. That is, mothers gave more attention directives (utterances to get the infants attention), action directives (utterances that regulate infants behaviour) and prohibitions (utterances used to stop infants' behaviour) to their sons compared to their daughters (Kuchirko et al., 2020).

Another factor that could explain why male and female children produce different speech may be in part due, not only to the input they receive from their parents but also the input they receive from media. For example, commercials targeted towards boys had more action words (run, race, jump), destruction words (crush, slam, stomp) and power words (control, rule, defeat) whereas commercials targeted towards girls had more nurturing words (cuddles, love, taking care of) and limited action words (talk, watch, look) (Johnson & Young, 2002). This shows how different words can become gendered when they are targeted towards girls or boys which, in turn, may influence gendered speech production.

Evidence regarding specific phonetic details of individual speech sounds that adults produce when conversing with children is sparse. One notable study is Foulkes and Docherty's (2006) study in which it was discovered that mothers produce significantly different variants of /t/ when speaking to boys versus girls. More specifically, they identified that mothers of boys used standard (plain) and nonstandard (laryngealized) variants of /t/ in word-medial position equally whereas mothers of girls used plain [t] twice as frequently as the laryngealized form in word-medial positions in their speech

(Foulkes & Docherty, 2006). That is, mothers use the standard production of /t/ more frequently with their daughters than with their sons. Although Foulkes and Docherty (2006) were able to find this gender specific speaking pattern in mothers, other evidence surrounding these patterns in relation to phonetic level differences is lacking. Thus, it still remains a question as to whether differences in input between boys and girls at a phonetic level are a true mechanism of the emergence of gendered speech patterns that emerge in childhood.

Overall, there are many instances in which mothers and fathers use language differently when they are talking to their sons compared to their daughters. The gender differences in input that children receive offers an example of how nurture can affect language acquisition. Children may learn to produce gender specific differences in speech as a result of the input they receive from the adults in their environment. This process helps to provide some understanding of a possible mechanism that may account for the emergence of gendered differences in early childhood. It is also this mechanism that was the motivation of the current research.

1.4 PURPOSE AND HYPOTHESIS

The purpose of the current research is to analyze the extent to which an environmental influence can affect speech development in relation to gendered patterns of behaviour; specifically, how does the input that children receive in their environment differ between boys and girls. It still remains a question in the literature as to whether adults further modify their speech (beyond IDS and CDS) to children depending on whether they are girls or boys. That is, do adults modify their speech to an extent that girl-directed and boy-directed speech exists? A speech pattern may be defined as girl-directed or boy-directed based on the context surrounding the speech. Males and females

may prefer certain linguistic features that allow them to express certain feelings or motivations (Ladegaard & Bleses, 2003).

Prior research has been able to contribute to the vast knowledge compiled on the differences between male and female speech ranging from young children to adulthood. In regard to the gendered speech patterns of children specifically, there has not been, to my knowledge, any research conducted to determine whether girl-directed and boy-directed speech exists. It is for this reason that the following paper will attempt to fill in the gap in this specific type of conversation style that may be utilized in adult speech. This research was motivated by Foulkes & Docherty's (2006) study in which they found differences in consonant production when mothers addressed boys versus girls and by the study from Li et al. (2006) in which they found that boys and girls have different productions of /s/. The current study utilizes these two studies as a foundation for testing whether adults have differing productions of /s when talking to boys than when they are talking to girls. The hypothesis of this research is that adults will talk to boys in a different manner than they talk to girls which could suggest evidence of the existence of girl- and boy-directed speech.

1.5 IMPLICATIONS

The implications of this research can help to understand subtle differences between the speech of adults when talking to boy and girl children. This in turn can help to understand the learning process of boys and girls and how that process may differ during child development. In addition, this research can also help contribute to the never-ending debate of nature versus nurture; are people born with gender differences ingrained in their behaviour or is it the environment that has a bigger sway on gendered patterns of speech that we find to exist in children's early speech production? The current research

aims to contribute to this debate in addition to helping explain the emergence of gendered speech in early childhood development.

CHAPTER 2: METHODS

2.1 PARTICIPANTS

Forty university students aged 17-36 years old (M = 21.8, SD = 3.9; females = 26; Caucasian = 58%) were recruited to participate in this study. Students were recruited using the psychology participant pool through the University of Lethbridge. Students who successfully completed the study were compensated with course credit.

2.2 STUDY DESIGN

This study required participants to read a passage to photographs of Caucasian babies (N = 6; females = 3) in a randomized order on the basis of apparent gender of the stimuli. The images of babies were grouped into three male-female pairs. In these pairs, one baby photo was left in its original condition and the second baby photo in the pair was manipulated in Photoshop so that it shared the same face as the original baby photo (see Appendix C). In other words, there were a total of six stimuli sharing three faces which were each attached to a pair of bodies that were gendered by clothing. This step was completed to eliminate confounds that could arise from gendered facial features. The photos left in their original condition were chosen at random by an online random number generator. Caucasian images were chosen as I expected that the majority of subjects recruited for this study from the university would be Caucasian. All the photos chosen for this study were retrieved through Pexels (Pexels, n.d.) through licensing from the University of Lethbridge.

All the stimuli were showcased wearing stereotypically male or female outfits. In addition, the stimuli were given stereotypical names as gendered names help elicit a more masculine or feminine representation (Bauer & Coyne, 1997). The photos of the babies were given names from a list of the most popular boy and girl names in some of the provinces in Canada in 2019 (Government of British Columbia, n.d.) or 2020 (Government of Alberta, n.d.; Government of Saskatchewan, 2021). For instance, Noah was used for a boy and Olivia for a girl. The names were chosen based on popularity and clarity to eliminate any ambiguity as to whether the given names were male or female. Any names that I deemed gender neutral were not used in this study. In addition to a name, each stimuli was supplied with a stereotypically male or female fact to give the photo an essence of personality that participants could relate to or understand (eg. for a boy, he would love frogs and for a girl, she would love to play with dolls). The stereotypical facts were also chosen on the basis of clarity and popularity to eliminate any confusion about gender typicality of the nature of the baby photos. For the full list of names and facts used for the stimuli in this study, refer to Appendix B.

The shortened version of "The North Wind and the Sun" (Kotami & Yoshimi, 2018) was chosen for this study because of its past use in phonological assessments (Deterding, 2006). The words in the story were chosen to elicit /s/ which is the sound of interest in the current study. The short story was printed and made into a physical book format with hand drawn images that corresponded to the words on each page (see Appendix A).

The study was conducted in one of the rooms in the Speech Production and Development lab at the University of Lethbridge which was designed to replicate the environment of a dayhome. The lab space was decorated with several coloured drawings

done by children taped to the walls to elicit a dayhome environment. The dayhome was named “Sunshine Dayhome” and signs with this name were displayed inside the lab space and on the outside of the door where participants would enter to complete the study. Some children’s toys such as a play kitchen, little people and cars were also placed in the lab. The reasoning for using a dayhome setting was as follows; a) dayhomes tend to have the ages of children used in this study (1-2 years old), b) dayhomes have multiple children who may be the same age, and c) having participants read to children mimics common activity in a dayhome setting.

During the time of data collection, there were strict COVID-19 restrictions in the university and the province of Alberta, so participants were required to show their vaccine passport for proof that they were fully vaccinated (two doses of an approved vaccine in Canada) prior to participating in the study. Participants were then required to read and sign a consent form about the study. To hide the true intent of the study, which was to examine if adults speak differently to boy versus girl children, participants were instead told that the researcher was testing for individual speech differences. This would deter participants from knowingly adjusting their speech by understanding the true intent of the research. Participants also had to acknowledge that they were required to take their mask off for the majority of the study for appropriate audio quality during recording. When participants had their mask off, they were seated far away from the researcher or research assistant conducting the study, and the researcher or research assistant kept their mask on for the entire study.

2.3 PROCEDURE

The first portion of the study required getting a baseline measure of every individual’s speech production. To achieve this goal, participants were informed that they

should read a short passage in their normal speaking voice so that the researcher could make sure the mic was recording properly. The passage was titled “Limpy the Duck” (Berggren et al., 2018). Participants were then given the scenario that they were a volunteer at Sunshine Dayhome and they would be required to read some passages to the babies at the dayhome.

Participants were told that they would be reading a passage to 6 different babies. In this case, it was the shortened version of “The North Wind and the Sun.” Each baby (stimulus) was introduced by name and with a sentence regarding something they liked. For example, “This is Sophia, she really likes to play with dolls.” Participants were asked to greet each baby photo by name after they were introduced. All six stimuli were presented in a randomized order (determined by an online random number generator) and participants' speech was recorded throughout for each of the six stimuli.

An Audio-Technica ATR 1002 microphone placed approximately 15-20 cm away from the participants mouth was used in this study. A Shure Beta Zoom H1n recording device was used to record the audio (44.1 Khz and 16 bits) of the participants during the task.

Two general questionnaires were given to participants to fill out at the end of the study. The first survey was a general demographics questionnaire and assessed participants' experience with children. The second survey, used by Kennison & Byrd-Craven (2015) was adopted and used to assess the participants' communication and interaction styles with children. After the two surveys were completed, participants were debriefed of the true intent of the study and, if they had no further questions, were allowed to leave.

2.4 ANALYSIS

The audio of all 40 participants was analyzed using Praat (Boersma & Weenink, 2001). Three acoustic parameters were chosen, including the centroid frequency of the sound /s/, the average f0 and the intensity. A script was designed to run in Praat to annotate the instances of /s/ in 9 tokens that began with /s/ in word initial position and to identify boundaries for the fricative [s] throughout the passage participants read (Fig. 1). A subset of this data were analyzed which only included instances where the [s] sound was produced correctly. It should be noted that some adults produce /s/ with a [ʃ] variant in words such as “stronger.” All instances of [ʃ] were eliminated from further analysis and centroid frequency values were calculated for correct pronunciations of /s/. The centroid frequency, otherwise known as m1, is one spectral characteristic of the frication part of the /s/ sound and is the first spectral moment in moments analysis (Lathrop-Marshall et al., 2022). M1 values reflect the mean frequency of the sound energy (Jiang et al., 2016) and higher m1 values suggest a more anterior articulation for /s/ (Fox & Nissen, 2005). Centroid frequency values of the /s/ snippet were calculated from the middle 40-ms window extracted from the frication portion of the /s/ sound. In addition, average f0 and RMS intensity of the reading passages as well as the greetings were also measured.

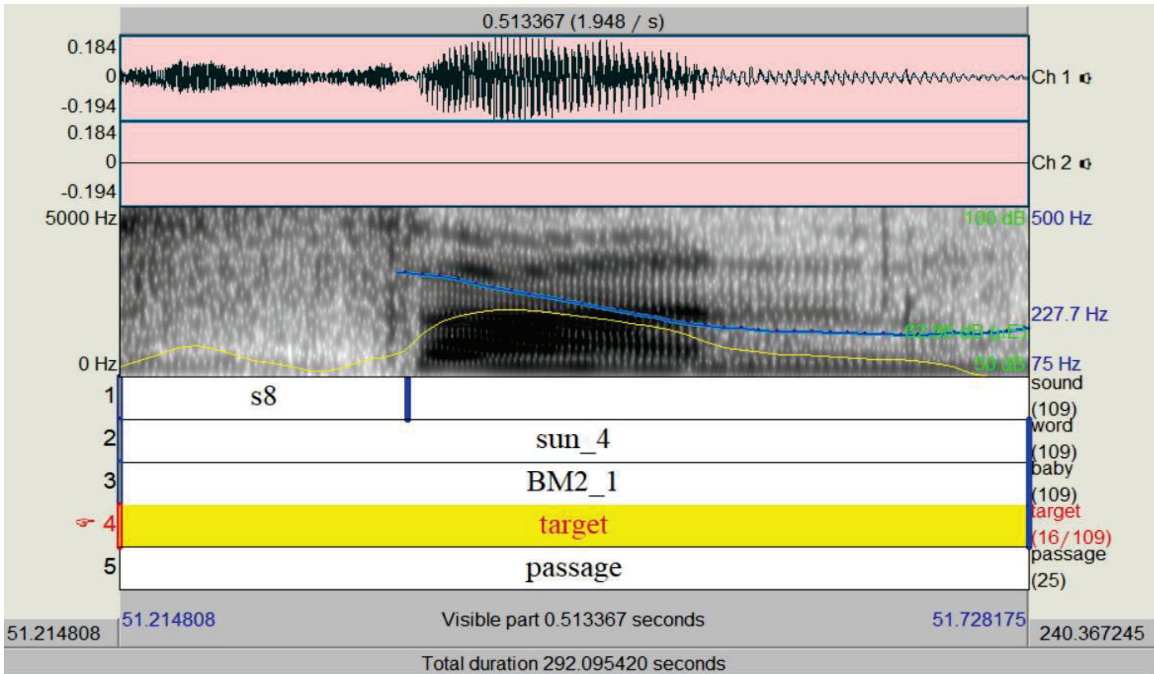


Figure 1: A spectrogram of the word “sun”, one of the tokens in the current study, in which boundaries identify the [s] sound. Pitch (blue line) and intensity (yellow line) contours are also shown.

All the data were analyzed with a mixed effects Bayesian model using the “brms” package in R (Bürkner, 2018). A Bayesian approach was used over the traditional Frequentist approach because Bayesian statistics use a more direct approach to answering research questions as it estimates the variable of interest directly from the population distribution (Hespanhol et al., 2019) and considers all the parameters in a model as random variables (Nalborczyk et al., 2019). Using random effects is important as it decreases the chances of models with false positives (Oberauer et al., 2022) and gives more overall strength to models (Rubin, 1984). In total, three separate models were used for the analysis, one for m1 values, one for f0 and one for intensity.

Model for m1 values

A Bayesian model (4 chains, 2000 iterations) was constructed to analyze m1 differences for the effect of baby gender in male and female speakers. Specifically, the model estimates the probability of m1 differences when individual adults are addressing male or female baby photos.

The dependent variable in this model is the m1 values, and the fixed effects are speaker gender (male vs. female), baby gender (male vs. female), the interaction between baby gender and speaker gender, and presentation order of the photos (1st to 6th presentation). The model in the analysis has random slopes for baby gender and presentation order that vary with individual speakers, as well as random slopes for speaker gender and baby gender that vary with individual words. The model also has a random intercept of face ID (baby faces 1-3), participant ID and individual words.

The fixed effects of baby gender, speaker gender, the interaction between baby gender and speaker gender were chosen because these variables are factors that are being controlled for in the model and are variables that I am interested in analyzing further. Likewise, presentation order was chosen as a fixed effect because it is a measure that is being controlled for in the model. Random effects were implemented into the model as they allow for individual variation in each selected measure. Speaker gender and baby gender were chosen as random slopes because m1 values are likely to have individual differences. In addition, they were chosen to vary with words as it is also likely that individual words can cause variations in m1 values. By adding these as random slopes, it allows for individual differences to be analyzed for potential relationships. Similarly, the presentation order and baby gender were chosen as random slopes that vary with individuals as participants are also very likely to show individual differences and

potential relationships can be analyzed. Face ID and words were included as random intercepts because these measures were repeated across participants. Participant ID was also included as an intercept because each individual is likely to represent a different intercept for m1 values due to individual differences. The equation used for this model in R can be seen below.

```
brm(m1 ~ babyGender * speakerGender + babyPresOrder +                    (1)
(1+babyPresOrder+babyGender|ParticipantID) + (1|faceID) +
(1+babyGender+speakerGender|word))
```

Model for f0

A similar Bayesian model (4 chains, 2000 iterations) was also constructed to analyze the f0 values. This model estimates f0 when adults are speaking to photos of different genders.

The dependent variable in this model is the f0 and the fixed effects are baby gender (male or female), speaker gender (male or female), the order the baby photos were presented in (1st through 6th presentation), task (either reading to or greeting photos of babies) and an interaction between baby gender and the task. There is a random slope for presentation order which varies with participant ID and random intercepts for participant ID and face ID (faces 1-3). The fixed effects and the random effects in this model were chosen for the same reasons as described previously, only the random slopes allow for individual differences in f0, rather than m1 values. In addition, task (reading to or greeting baby photos) was added as a fixed effect in this model. This was chosen as this is

also a measure of interest and is also a measure that is being controlled for in the model. The equation used for this model in R can be seen below.

```
brm(f0 ~  speakerGender + Order +  babyGender * Task +
(1+Order|ParticipantID) + (1|faceID))
```
(2)

Model for intensity

Another Bayesian model (4 chains, 2000 iterations) was used to examine intensity. In particular, this model estimates the probability of intensity differences when adults are speaking to boy versus girl baby photos.

This model estimates the probability of intensity differences when adults are speaking to boy versus girl baby photos. The dependent variable for this model is the intensity as measured by RMS averaged over each reading passage and the fixed effects are baby gender (male or female), speaker gender (male or female), the order the baby photos were presented in (1st through 6th presentation), task (either reading to or greeting images of babies) and an interaction between baby gender and the task. There is a random slope for presentation order which varies with participant ID and random intercepts for participant ID and face ID (faces 1-3). The fixed effects (baby gender, speaker gender, presentation order and the task) and random effects were chosen for the same reasons described previously, only the random slopes in this model allow for individual differences of intensity rather than m1 or f0. The equation used for this model in R can be seen below.

brm(intensity ~ speakerGender + Order + babyGender * Task + (3)

(1+Order|Participant.ID) + (1|faceID))

Chapter 3: Results

3.1 SPECTRAL MEAN FREQUENCY/ CENTROID OF /S/ PRODUCTION

Descriptive statistics for the mean and standard deviations of m1 values of the sound /s/ produced by adults addressing the different gender stimuli and face ID (see appendix C) are reported in Table 1. Model comparisons were used in the analysis using the leave one out cross validation (LOO) function which determined which of the models in this analysis were a better fit in analyzing the current data. This model ran using 4 chains and 2000 iterations in which all R^s = 1.00 which show that the model converged. A posterior predictive check (pp_check) was assessed as well, which showed that the model fits well with the data (Fig. 2). The output for this model is visually represented in Fig. 3.

Table 1

Mean and standard deviation of m1 values for male and female speakers when addressing different gender stimuli and face ID

Participant Gender	**Baby Photo Gender**	**Baby Face ID**	**M1 (centroid frequencies of /s/)**	
			Mean	SD
Male	Male	1	6227.52	988.58
		2	6272.69	939.75
		3	6269.49	1009.07
	Female	1	6197.79	982.64
		2	6136.62	1017.22

		3	6244.60	937.15
Female	Male	1	7995.27	1154.98
		2	8140.57	1133.08
		3	7957.68	1198.20
	Female	1	8101.49	1102.62
		2	7960.44	1241.54
		3	8181.19	1158.36

Note. M1 values are the first spectral moment in moments analysis and represent average frequency of the sound energy of /s/.

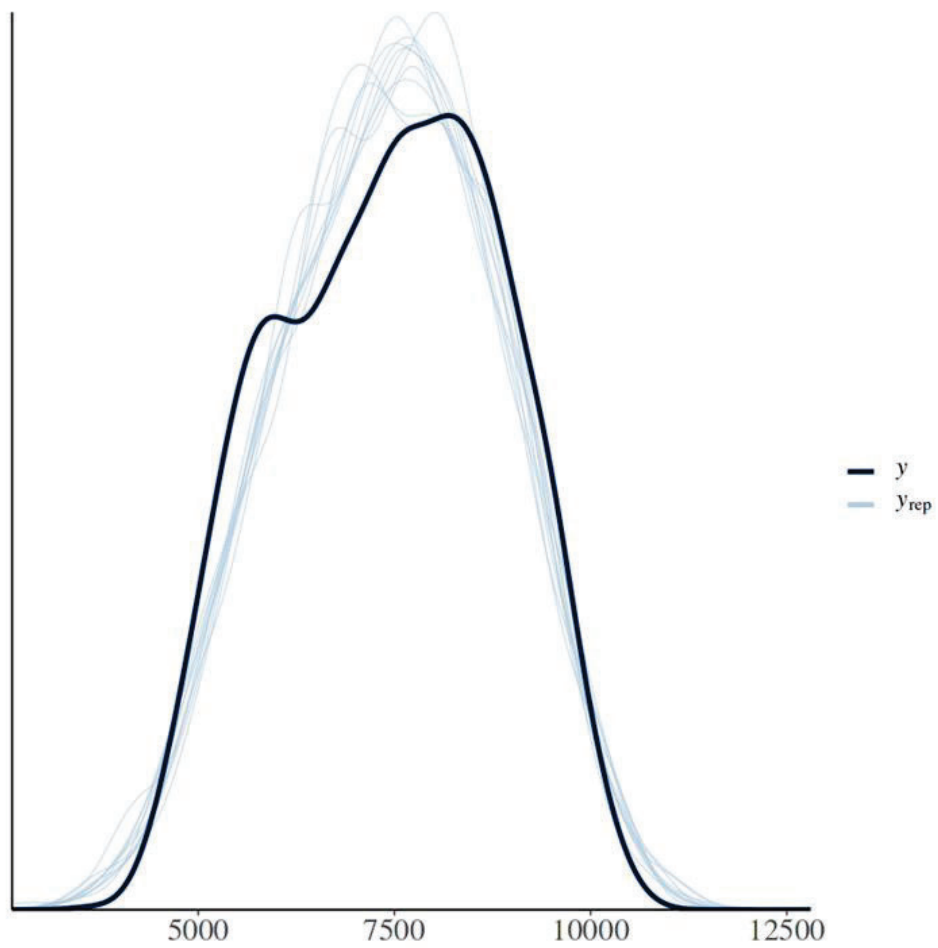


Figure 2: A posterior predictive check of the model where the light blue line represents predicted data and the dark blue line indicates actual data. M1 values are represented on the x axis.

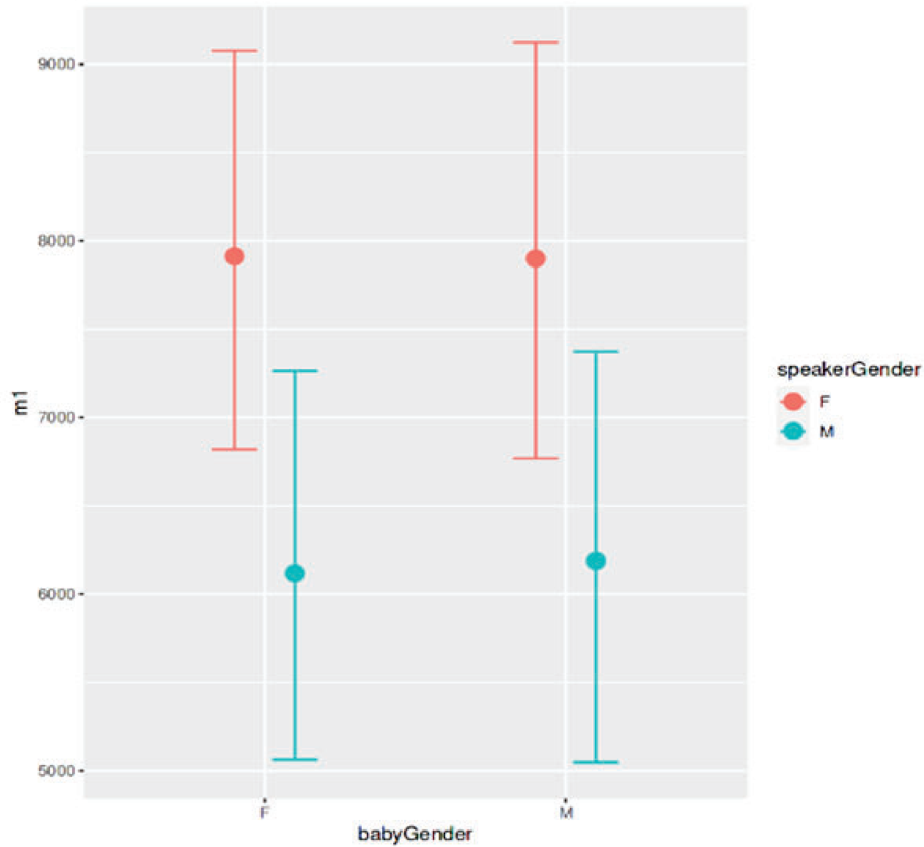


Figure 3: Conditional effects plot from the mixed effects Bayesian model output which reflects m1 values for male and female speakers when they are addressing different stimuli genders.

Population level effects of this model are reported in Table 2. The summary shows that baby gender (-10.38, 95% CI: -230.14, 273.77) and the presentation order of baby photos (-4.47, 95% CI: -25.21, 15.80) are not significant. In addition, the interaction

between baby gender and speaker gender is also not significant (81.62, 95% CI: -53.24, 219.13). This is because the credible intervals cross zero, indicating a non-significant result as the model is not able to accurately estimate the probability of the effect given the observed data. In contrast, speaker gender was significant as the credible intervals do not cross zero (-1657.03, 95% CI: -2350.90, -954.10). The estimate for male speakers (-1657.03) indicates that on average, the frication of /s/ in males is -1657.03 Hz less than frication in female speakers, indicating a significant difference between male and female speaker frication.

Table 2

Population level effects of the current model that assessed stimuli gender, speaker gender, the interaction between speaker gender and stimuli gender and presentation order

Parameter	Estimate	Est. error	Lower 95% CI	Upper 95% CI
Intercept	7801.76	534.72	6737.98	8855.97
Baby gender (M)	-10.38	121.63	-230.14	273.77
Speaker gender (M)	-1657.03	351.49	-2350.90	-954.10
Presentation order	-4.47	10.57	-25.21	15.80
Interaction (baby gender (M) and speaker gender (M)	81.62	70.41	-53.24	219.13

3.2 AVERAGE FUNDAMENTAL FREQUENCY (F0) RESULTS

Since this study required participants to greet and then read to images of babies (stimuli), f0 for this audio was analyzed. F0 was calculated for the greeting and reading to the stimuli as this elicited two different types of speech: spontaneous and scripted. Participants were not directed on how to greet the stimuli, which allowed for more spontaneous speech, but were given specific words to a story (as opposed to a wordless picture book) enforcing a more scripted dialogue when reading to the baby photos. Descriptive statistics for those f0 values when adults read to or greeted different gendered stimuli and face ID's were calculated and reported in Table 3. This model which ran using 4 chains and 2000 iterations, had all R^s less than 1.01 which showed that the model converged. A posterior predictive check (pp_check) was conducted to analyze how well the model fit with the data. Results show that the model does a fair job of predicting the data (Fig. 4). The visual representation of the model output can be seen in Fig. 5.

Table 3

Mean and standard deviation for f0 of male and female speakers when reading to or greeting stimuli of different genders and face ID

Task	Adult Gender	Baby Photo Gender	Face ID	f0 (Hz)	
				Mean	SD
Greeting	Male	Male	1	144.82	35.22
			2	143.93	37.02
			3	137.93	35.20
		Female	1	133.85	37.35
			2	141.41	35.92
			3	139.00	41.94

	Female	Male	1	241.80	29.79
			2	241.73	28.86
			3	228.70	28.19
		Female	1	231.5	27.57
			2	224.05	33.64
			3	234.81	25.74
Reading	Male	Male	1	117.57	18.63
			2	117.27	19.31
			3	117.71	18.64
		Female	1	118.18	18.89
			2	119.41	19.14
			3	118.17	19.04
	Female	Male	1	211.65	21.85
			2	211.23	19.93
			3	210.82	21.79
		Female	1	211.95	19.24
			2	209.92	21.08
			3	211.60	20.27

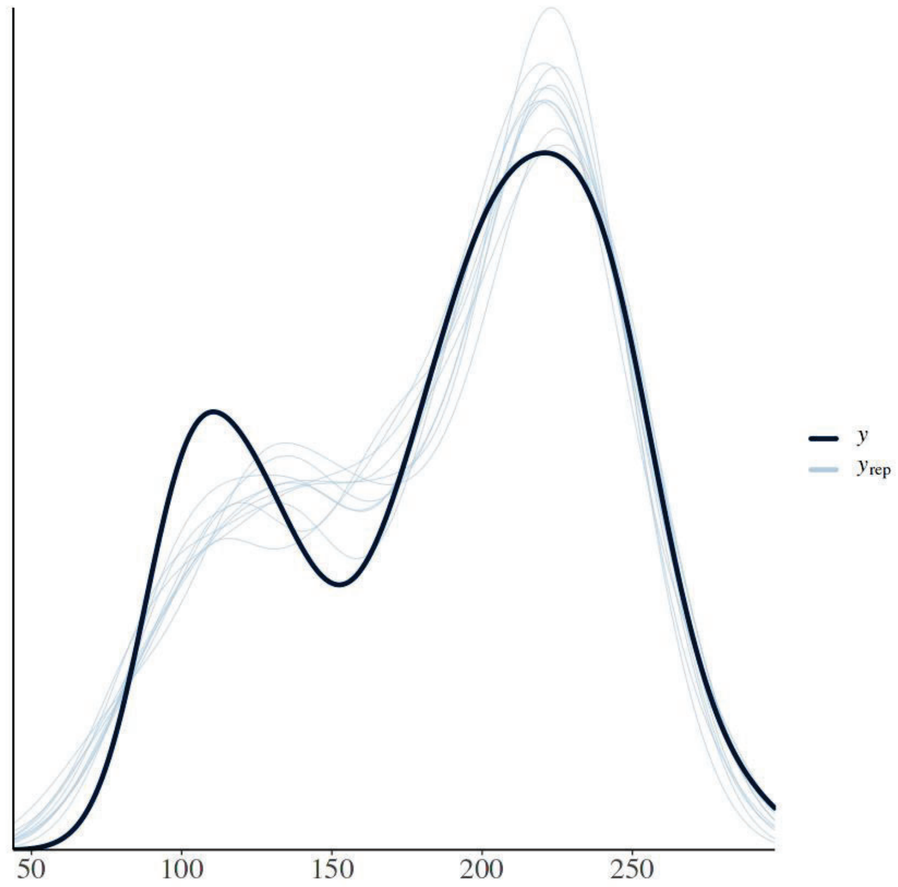


Figure 4: A posterior predictive check of the f0 model where the light blue line represents predicted data and the dark blue line indicates actual data. F0 values are represented on the x axis.

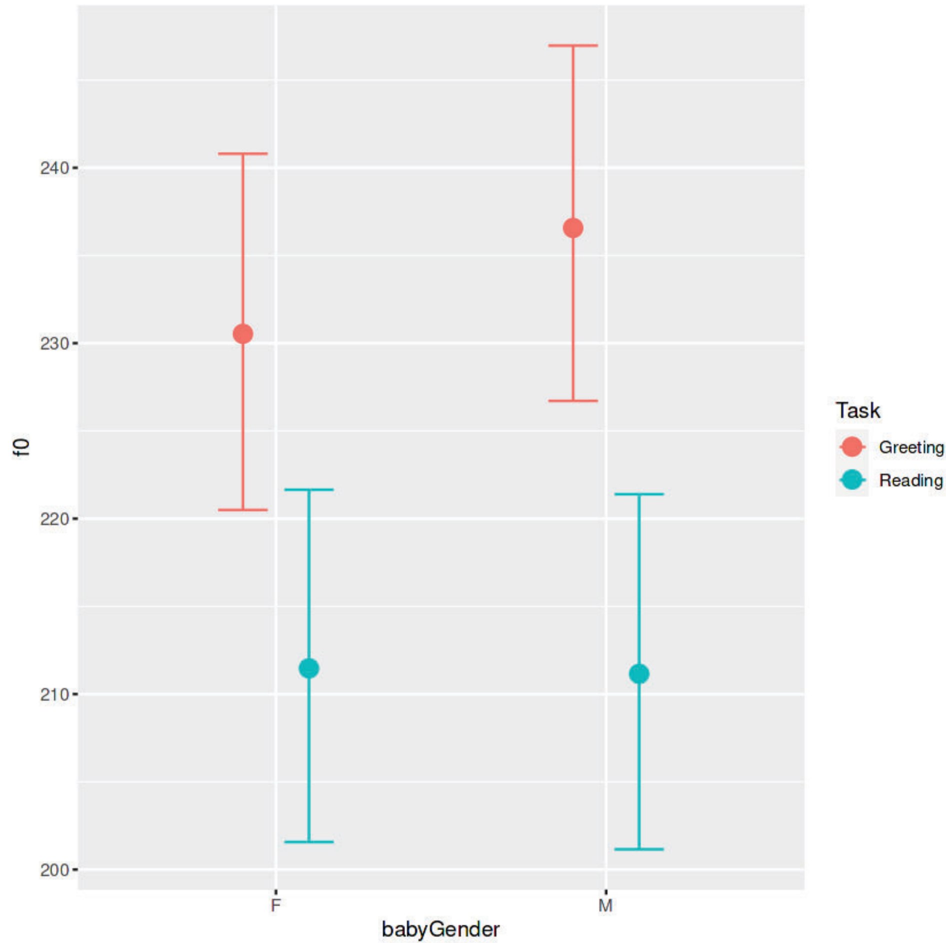


Figure 5: Conditional effects plot from the mixed effects Bayesian model output which reflects f0 values for adult speakers when they are greeting or reading stimuli of different genders.

Population level effects reported in Table 4 showed significant effects of baby gender (9.41, 95% CI: 0.81, 17.93), speaker gender (-90.32, 95% CI: -106.36, -73.94) and the task (-9.77, 95% CI: -18.31, -1.25). The f0 of the subject's speech was 9.41 Hz higher when speaking to boy photos than to girl photos. Moreover, the task also influenced the f0 values as reading to the stimuli was 9.77 Hz lower than when the stimuli were greeted.

This result can be further seen in Figure 6, which shows that the mean f0 for greeting the boy stimulus was significantly higher than greeting the girl stimulus for both male and female speakers. In addition, adult males were still shown to produce speech 90.32 Hz lower than females on average. The order of the photos presented to adults did not produce any significant effect (1.85, 95% CI: -0.02, 3.71).

Table 4

Population level effects of the model measuring f0 values in relation to stimuli gender, speaker gender, presentation order, the task at hand and the interaction between the task and stimuli gender

Parameter	Estimate	Est. Error	Lower 95% Ci	Upper 95% CI
Intercept	213.44	8.20	195.60	229.99
Baby Gender (M)	9.41	4.35	0.81	17.93
Speaker Gender (M)	-90.32	8.18	-106.36	-73.94
Order	1.85	0.97	-0.02	3.71
Task (reading)	-9.77	4.31	-18.31	-1.25
Interaction between baby gender (M) and task (reading and greeting)	-9.43	6.12	-21.40	2.62

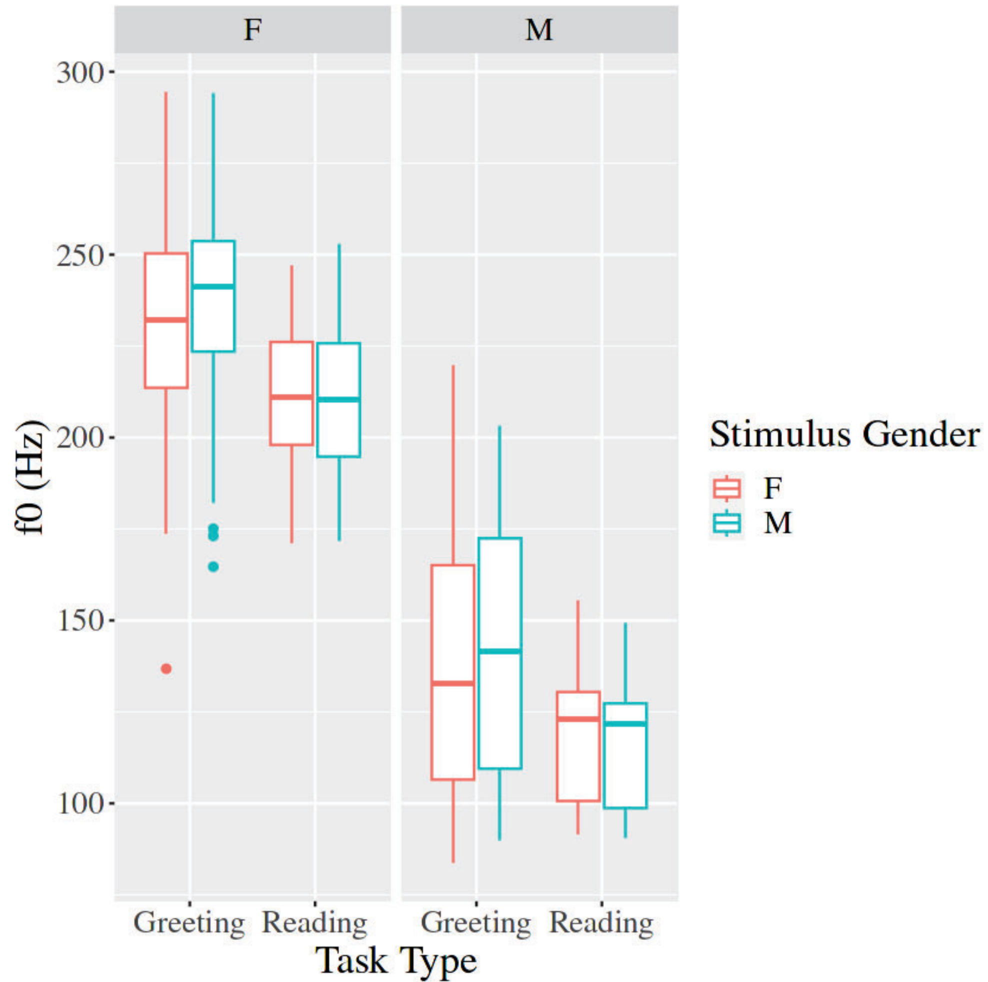


Figure 6: Differences in f0 when male and female adults greet and read to stimuli of different genders**.** The y axis shows relative f0 (Hz) and the x axis differentiates between the task (either greeting or reading to the stumuli). In addition, the graph is split into two sections in which female adults are represented on the left side of the graph and males are represented on the right side. The boxplots in orange indicate female stimuli and the blue represents male stimuli. This graph shows that both males and females had higher f0 when greeting boy stimuli they did girl stimuli and that the mean f0 was higher for greeting than for reading.

3.3 AVERAGE INTENSITY RESULTS

In addition to m1 values and f0, intensity was also analyzed for all the audio recordings. Descriptive statistics for intensity values when adults read to or greeted different images of babies were calculated and reported in Table 5. This model (4 chains, 2000 iterations) also showed that all R^s = 1.00 which showed that the model converged. A posterior predictive check of the model showed that it did not perfectly fit with the current data (Fig. 7). The visual representation of the model output is showcased in Fig. 8.

Table 5

Mean and standard deviation for intensity values when adults read to or greeted stimuli of different gender and face ID

Task	**Adult Gender**	**Baby Photo Gender**	**Face ID**	**Intensity (dB)**	
				Mean	SD
Greeting	Male	Male	1	53.78	6.40
			2	58.08	6.45
			3	58.86	6.53
		Female	1	57.98	6.33
			2	56.66	6.88
			3	57.15	6.92
	Female	Male	1	57.80	4.56
			2	59.69	4.23
			3	59.51	3.56
		Female	1	57.74	6.30
			2	57.55	4.42

			3	58.74	4.99
Reading	Male	Male	1	53.78	5.08
			2	53.34	5.60
			3	53.84	5.70
		Female	1	53.63	5.21
			2	53.84	5.70
			3	53.57	5.57
	Female	Male	1	55.12	4.20
			2	55.48	4.32
			3	55.49	4.29
		Female	1	55.50	4.48
			2	55.19	4.37
			3	55.52	4.39

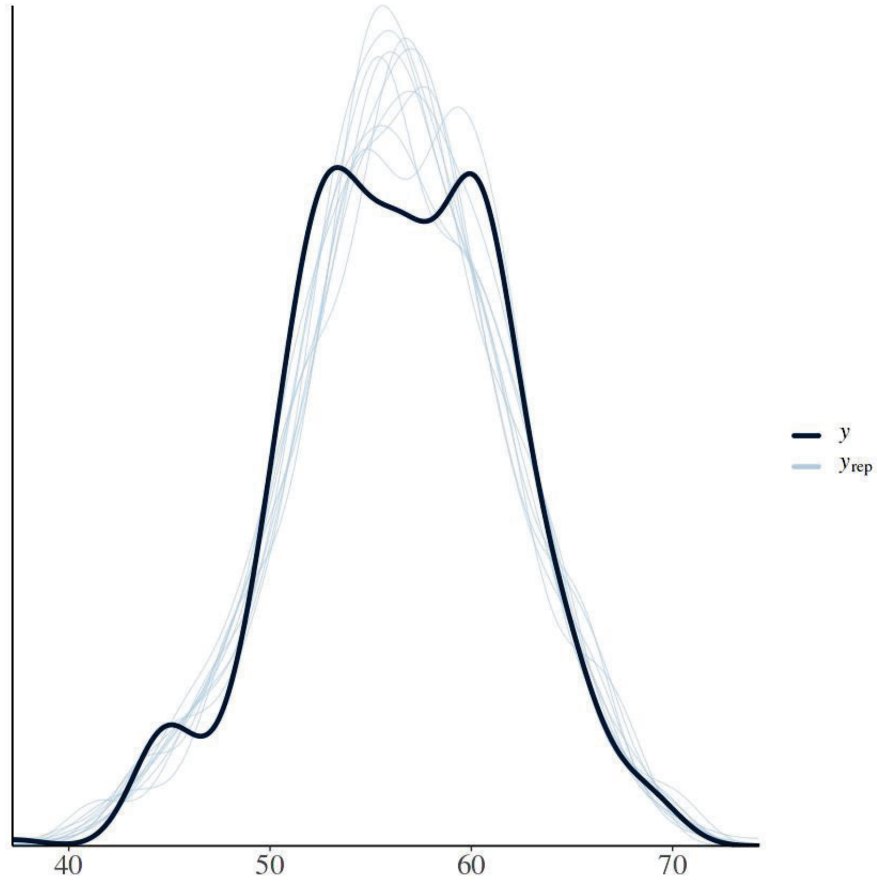


Figure 7: A posterior predictive check of the intensity model where the light blue line represents predicted data and the dark blue line indicates actual data. Intensity values are represented on the x axis.

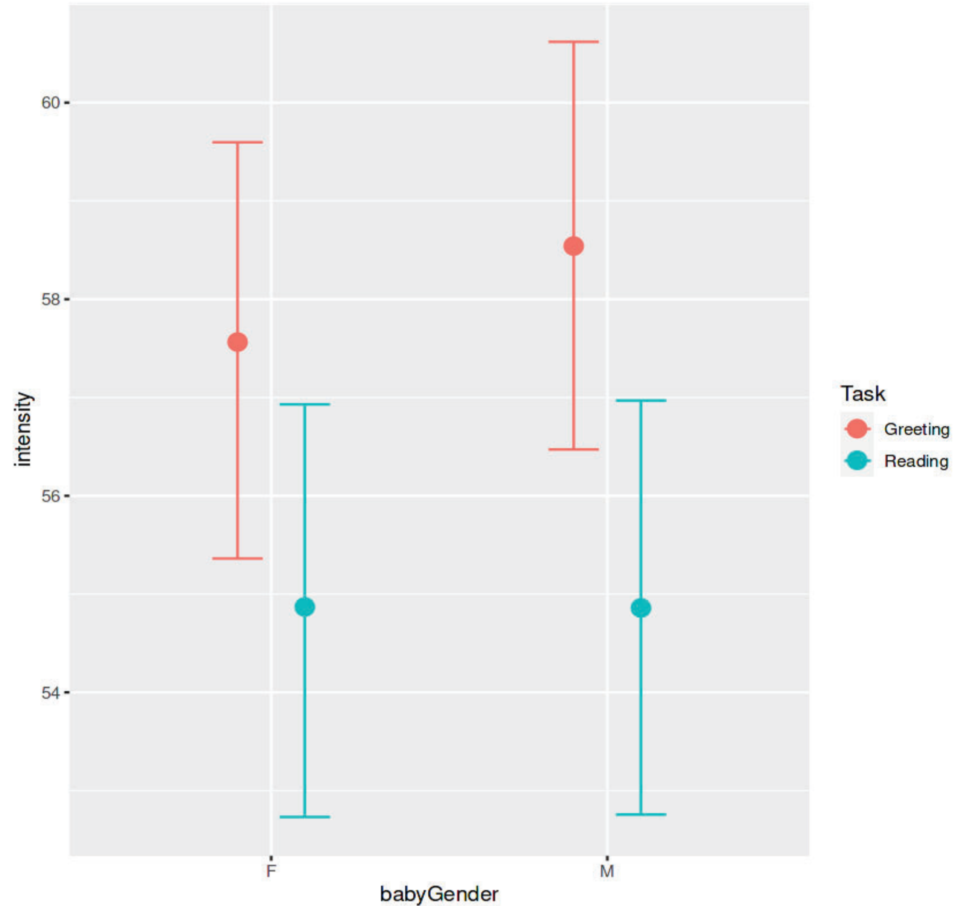


Figure 8: Conditional effects plot from the mixed effects Bayesian model output which reflects intensity values for adult speakers when they are greeting or reading to different gendered baby photos.

Population level effects reported in Table 6 indicate that baby gender (0.83, 95% CI: -0.71, 2.55) and speaker gender (0.03, 95% CI: -2.99, 2.99) were not significant as the credible intervals crossed zero. In contrast, the presentation order of the photos (0.15, 95% CI: 0.02, 0.28) and the interaction between baby gender and the task (-0.82, 95$ CI: -1.47, -0.14) were marginally positively significant. Further significance was found in whether adults were greeting or reading to the baby photos (-3.03, 95% CI: -3.50, -2.55).

This indicates that on average, adults read to the stimuli 3.03 dB lower than when they were greeting them (Fig. 9).

Table 6

Population level effects of the model measuring intensity values in relation to stimuli gender, speaker gender, presentation order, the task at hand and the interaction between the task and stimuli gender

Parameter	Estimate	Est. Error	Lower 95% CI	Upper 95% CI
Intercept	57.28	1.06	55.14	59.35
Baby Gender (M)	0.83	0.81	-0.71	2.55
Speaker Gender (M)	0.03	1.54	-2.99	2.99
Order	0.15	0.06	0.02	0.28
Task (reading)	-3.03	0.24	-3.50	-2.55
Interaction: baby gender (M) & task (reading)	-0.82	0.34	-1.47	-0.14

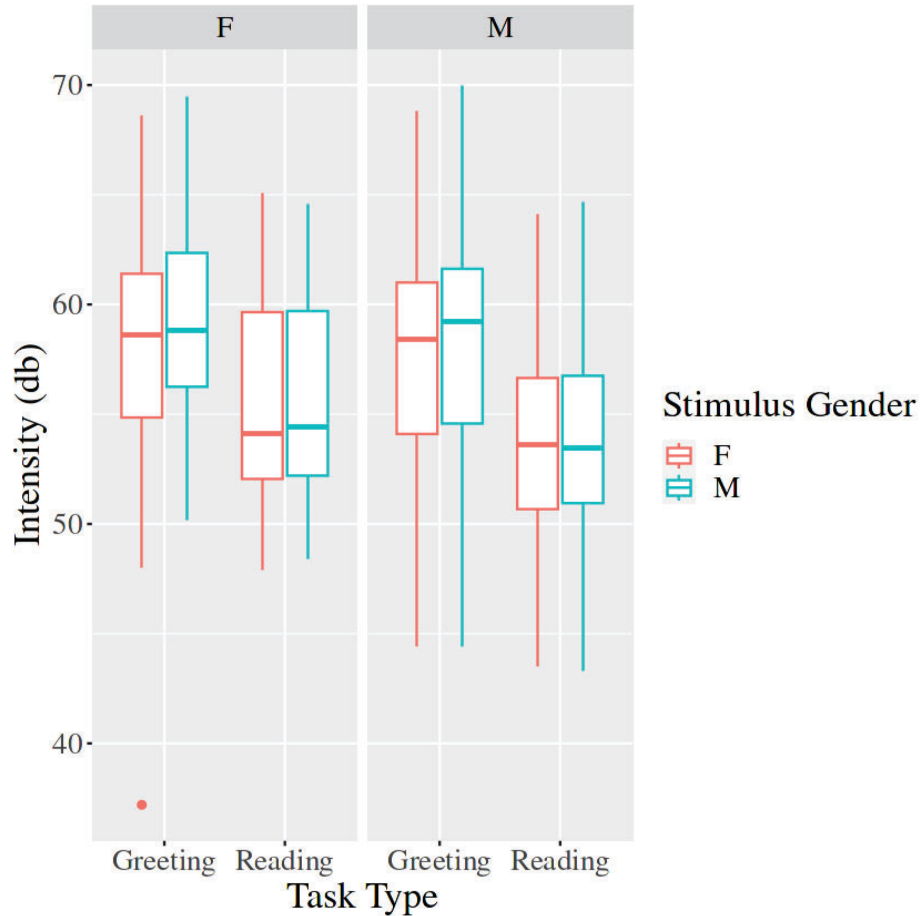


Figure 9: Differences in intensity when adults read to or greeted stimuli of different genders. The y axis shows relative intensity, and the x axis differentiates between the task (either greeting or reading to the stimuli). In addition, the graph is split into two sections in which female adults are represented on the left side of the graph and males are represented on the right side. The boxplots in orange indicate female stimuli and the blue represents male stimuli. This graph shows that both males and females had higher intensity when greeting boy photos than they did girl photos and that the mean intensity was higher for greeting than for reading.

CHAPTER 4: DISCUSSION

The current study was designed to investigate whether adults speak differently to boy babies than they do to girl babies, and whether those differences, if they exist, could be detected at a phonetic level of speech production. This research aims to targets the understanding of the emergence of gender specific speech patterns related to input in infants in early childhood, in which children's speech production patterns begin to develop. In addition, the current study provides another layer of discussion to the ongoing nature versus nurture debate.

Although the results of this study show no robust differences in the three acoustic parameters measured (m1, f0 and intensity) in adult speech to stimuli of different genders, they do provide some interesting insight. The results indicated a difference of 10.38 Hz for spectral mean difference in /s/ production when adults are talking to boy versus girl stimuli, although this finding was not significant. Although, the average m1 values for male and female adults was significantly different in which males produced lower average m1 values (1657.03 Hz less) than females. This suggests that males have a more posterior articulation of /s/ which is consistent with previous literature (see Fox & Nissen, 2005).

Results also showed significance for fundamental frequency differences (f0), indicating a 9.4 Hz difference when adults greeted boy stimuli compared to greeting girl stimuli. This difference is rather large, given the just noticeable difference (JND) which is around 0.3-0.5% (Ghitza & Goldstein, 1983) of the frequency of the sound. This suggests that adults were significantly more enthusiastic when greeting the boy stimuli. Interestingly, this finding is the opposite of the expected result. In a previous study by Kitamura & Burnham (2003) it was demonstrated that mothers used a greater f0 when

addressing girls than when addressing boys. This opposition may have arisen due to the differing subject pool used in the two studies. The current study used university students in which the majority were not parents, as opposed to the study by Kitamura & Burnham (2003) which used mothers. Since mothers have been known to engage more with their daughters than their sons (Clearfield & Nelson, 2006), it may be possible that the greater engagement affects their fundamental frequency in a way that elicits higher f0 values when addressing daughters than sons. Given that the opposite result was found in the current study, further research should be conducted to clarify these opposing results.

The results also revealed a higher intensity value of 0.83 dB when adults spoke to boy stimuli than when talking to girl stimuli. This result seems to be trending to the notion that adults spoke with greater intensity (greater volume) to the images of boy babies than the images of girl babies as the standard JND is 1.0 dB (McShefferty et al., 2015). This suggests that adults tended to greet boy stimuli louder than they greeted girl stimuli, even though this does not quite reach the JND threshold.

The results also show a significant difference when comparing reading and greeting for both f0 and intensity analyses. The fundamental frequency (f0) of adults' speech output sample was 9.77 Hz higher when greeting the stimuli as compared to reading to them, when analyzed spectrographically. Similarly, the intensity of adults' speech was 3.03 dB higher when greeting the stimuli than reading to them. This suggests that adults were more enthusiastic while greeting the stimuli than reading to them, which may be a result of spontaneous nature of the speech task. Adults were able to elicit a clearer difference in their speech to the stimulus that was spontaneous in nature, perhaps because spontaneous speech to baby photos is a little more naturalistic than reading to them. Further, greetings are designed to gain an infant’s attention so more variation in

pitch and loudness is expected. Reading to photos does not allow for any interaction or response, which is an important aspect of reading to actual babies. Overall, the results indicate small effects for f0 and intensity, but the values for intensity are not enough to be perceptually different.

A recent study by Dailey et al. (2022) compared word and utterance use in mothers and fathers while talking to babies aged 6 to 18 months old in their homes. They found no difference in the use of noun type, noun token, and utterance type when examining the speech produced by mothers and fathers when speaking to boys versus girls. This finding, collectively with the results of the current study, suggests that the differences in input when adults talk to boy versus girl babies may not be as salient as expected. It remains a possibility that input differences could occur later in a child's development after they become a true conversational partner. Ongoing research is being conducted in the lab on whether there are differences in the ways adults address photos of children of different genders to test this hypothesis.

4.1 LIMITATIONS AND FUTURE CONSIDERATIONS

Several elements in the current study may provide an explanation for the lack of substantial differences in the gender-specific input gathered from this experiment. First, since participants were required to read the same passage six times, that repetition may have induced a fatigue effect that could have diminished any results, had there been any evidence in the acoustic recordings. Second, the study was not conducted in a naturalistic environment in which adults were interacting with real babies. Due to the recent pandemic, it was not possible to recruit baby participants and conduct testing in their homes. The artificial interaction with baby photos may have prevented us from identifying robust gendered differences in adult speech when they addressed babies of

different genders. Further, it may be possible that a lab setting may not reflect a real word setting when it comes to language differences. Although the pandemic introduced various limitations, the current research can still be utilized as a tool to help determine what aspects of speech are influenced by gender. Lastly, the participants in the current study comprised a convenience sample of university students. University students are not a representative sample of the population. A more even distribution of adults of different backgrounds, including parents, would be a more representative sample. Although there were three subjects in the current study who were parents, the sample size was too small to concretely say anything about the speech patterns of parents to baby photos of different genders.

4.2 CONCLUSION

Another possible explanation for these results is that there may be another mechanism, outside of the speech input children receive from adults, that may account for gender differences in speech in early childhood. This mechanism highlights the possibility of role-modelling and imitation from a same-sex adult in a child's environment. In other words, it may be that the gender specific speech patterns that children acquire are a direct result of children modelling the behaviours of the same sex adults in their environment.

One example to consider is pitch, which is one of the most recognizable features in gendered speech. One study found that during an imitation task, both boys and girls aged 6 to 10 years old spontaneously lowered their pitch when imitating masculine characters and raised their pitch when imitating feminine characters (Cartei et al., 2019). This finding suggests that young children are aware of gender differences in speech and can accurately imitate those differences (Cartei et al., 2019). The fact that gender-specific

speech patterns are able to be identified before puberty indicates that sex dimorphism may not be the sole cause for this finding (Munson & Koeppe, 2022). Instead, these differences may be learned or adopted through social learning such as imitation and role modeling.

The emergence of gender differences identified in early childhood speech may be a product of gendered play behaviour that is encouraged or observed in a child's immediate environment. Adults can directly influence the behaviours of children through the toys they purchase and by the behaviours they expose their children to, which may cause the development of gender stereotypes (Mesman & Groeneveld, 2017). These gender stereotypes may encourage or limit specific perspectives or behaviours adopted by young children in their early development.

Looking at these gendered patterns in childrens' behaviour may be an important factor to consider in a child's development, as these behavioural differences may help to influence and shape differences in speech production. Adults may unknowingly reinforce certain behaviours and speech patterns when interacting with children of different genders. In turn, children may learn to adopt these gendered behaviours which may give rise to the gender specific speech differences that emerge in early childhood. It is possible that this mechanism is more powerful than the one focused in this study, that being the mechanism of gender-differentiated speech input to children.

Future research should be conducted in which parents are recorded speaking to their babies in their home environment. This would provide a more naturalistic setting in which it may be more plausible that adults would produce the results that I expected to find in this study. Interacting with real babies may give more power to the study because babies can react to the stories that parents read them, whereas talking to a photo is one-

sided and does not allow for social or communicative reactions to storytelling. In addition, since the spontaneous task of greeting baby photos showed more significance, doing a follow up study using a spontaneous speech task may provide more insightful results.

In conclusion, gender specific patterns in children's speech may be a product of one of the mechanisms discussed previously or a combination of both. In the same ways nature and nurture may often go hand in hand, so too may the mechanisms in which we ascribe to the emergence of gendered speech in children.

REFERENCES

Albuquerque, L., Valente, A. R. S., Teixeira, A., Oliveira, C., & Figueiredo, D. (2021). Age and gender effects in European Portuguese spontaneous speech. *Loquens*, *8*. https://doi.org/10.3989/loquens.2021.077

Bablekou, Z. (2009). Dominant and submissive language in children's conversational acts: A gender comparison. *European Childhood Education Research Journal*, *17*(3), 283-296. https://doi.org/10.1080/13502930903101354

Barbier, B., Boë, L., Captier, C., & Laboissière, R. (2015). Human vocal tract growth: A longitudinal study of the development of various anatomical structures. *Interspeech 2015 -16th Annual Conference of the International Speech Communication Association*, International Speech Communication Association, Dresden, Germany.

Bauer, P. J., & Coyne, M. J. (1997). When the name says it all: Preschoolers recognition and use of the gendered nature of common proper names. *Social Development*, *6*(3), 271-291. https://doi.org/10.1111/j.1467-9507.1997.tb00106.x

Berggren, K. N., Hung, M., Dixon, M. M., Bounsanga, J., Crockett, B., Foye, M. D., Gu, Y., Campbell, C., Butterfield, R. J., & Johnson, N. E. (2018). Orofacial strength, dysarthria, and dysphagia in congenital myotonic dystrophy. *Muscle Nerve*, *58*(3), 413-417. https://doi.org/10.1002/mus.26176

Boersma, P., & Weenink, D. (2001). PRAAT, a system for doing phonetics by computer. *Glot international*. *5*. 341-345. https://www.researchgate.net/publication/208032992

Bürkner, P. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, *10*(1), 395–411. doi:10.32614/RJ-2018-017

Cangi, M. E., Isildar, A., Tekin, A., & Sarac, A. B. (2020). A preliminary study of normative speech rate values of Turkish speaking adults. *ENT Updates*, *10*, 381-389. doi: 10.32448/entupdates.769051

Cartei, V., Garnham, A., Oakhill, J., Banerjee, R., Roberts, L., & Reby, D. (2019). Children can control the expression of masculinity and femininity through the voice. *Royal Society Open Science*, *6*. https://doi.org/10.1098/rsos.190656

Chen, J., Quiang, H., & Wu, X. (2016). Frequency importance function of the speech intelligibility index for Mandarin Chinese. *Speech Communication*, *83*, 94-103. https://doi.org/10.1016/j.specom.2016.07.009

Clearfield, M. W., & Nelson, N. M. (2006). Sex differences in mothers' speech and play behaviour with 6-,9-, and 14-month-old infants. *Sex Roles*, *54*, 127-137. Doi:

10.1007/s11199-005-8874-1

Cockrell, M. D. (2020). Provider-patient interaction: Exploring elderspeak in simulated preclinical chiropractic student encounters. *Gerontology & Geriatric Medicine*, *6*, 1-6 . https://doi.org/10.1177/23337214209234

Dailey, S., & Bergelson, E. (2022). Talking to talkers: Infants' talk status, but not their gender, is related to language input. *Child Development*, 1-19. https://doi.org/10.1111/cdev.13872

De Decker, P. M., & Nycz, J. R. (2012). Are tense [æ]s really tense? The mapping between articulation and acoustics. *Lingua*, *122*(7), 810-821. https://doi.org/10.1016/j.lingua.2012.01.003

Denham, S. A., Bassett, H. H., & Wyatt, T. M. (2010). Gender differences in the socialization of preschoolers' emotional competence. *New Directions for Child and Adolescent Development*, *2010*(128), 29-49. https://doi.org/10.1002/cd.267

Deterding, D. (2006). The North Wind versus a Wolf: Short texts for the description and measurement of English pronunciation. *Journal of the International Phonetic Association*, *36*(2), 187-196. https://www.jstor.org/stable/44526411

Eagly, A. H., & Wood, W. (2013). The nature-nurture debates: 25 years of challenges in understanding the psychology of gender. *Perspectives on Psychological Science*, *8*(3), 340-357. doi: 10.1177/1745691613484767

Eckhaus, E., & Ben-Hador, B. (2019). Gossip and gender differences: A content analysis approach. *Journal of Gender Studies*, *28*(1), 97-108. https://doi.org/10.1080/09589236.2017.1411789

Fitzpatrick, M. A., Mulac, A., & India, K. (1995). Gender preferential language use in spouse and stranger interaction. *Journal of Language and Social Psychology*, *14*, 18-39. doi:10.1177/0261927X95141002

Fivush, R., Brotman, M. A., Buckner, J. P., & Goodman, H. S. (2000). Gender differences in parent-child emotion narratives. *Sex Roles*, *42*, 233-253. https://doi.org/10.1023/A:1007091207068

Foulkes, P., & Docherty, G. (2006). The social life of phonetics and phonology. *Journal of Phonetics*, *34*, 409-438. doi:10.1016/j.wocn.2005.08.002

Fox., R. A., & Nissen, S. L. (2005). Sex-related acoustic changes in voiceless English fricatives. *Journal of Speech, Language, and Hearing Research*, *48*(4), 753-765. https://doi.org/10.1044/1092-4388(2005/052)

Fung, P., Schertz, J., & Johnson, E. K. (2021). The development of gendered speech in children: Insights from adult L1 and L2 perceptions. *JASA Express Letters*, *1*(1),

1-6. https://doi.org/10.1121/10.0003322

Ghitza, O., & Goldstein, J. L. (1983). JNDs for the spectral envelope parameters in natural speech. *HEARING — Physiological Bases and Psychophysics*, Berlin, Heidelberg.

Government of Alberta (n.d.). *Alberta's top baby names*. Retrieved April 27, 2021 from https://www.alberta.ca/top-baby-names.aspx

Government of British Columbia (n.d.). *Most popular baby names in British Columbia*. Retrieved April 27, 2021 from https://connect.health.gov.bc.ca/babynames?year=2019

Government of Saskatchewan (2021). *Liam and Olivia still Saskatchewan's most popular baby names in 2020*. Retrieved April 27, 2021 from https://www.saskatchewan.ca/government/news-and-media/2021/january/29/liam-and-olivia-still-saskatchewans-most-popular-baby-names-in-2020

Hannah, A., & Murachver, T. (1999). Gender and conversational style as predictors of conversational behavior. *Journal of Language and Social Psychology*, *18*(2), 153-174. https://doi.org/10.1177/0261927X99018002002

Hespanhol, L., Vallio, C. S., Costa, L. M., & Sarahiotto, B. T. (2019). Understanding and interpreting confidence and credible intervals around effect estimates. *Brazilian Journal of Physical Therapy*, *23*(4), 290-301. doi: 10.1016/j.bjpt.2018.12.006

Hughes, S. M., Farley, S. D., & Rhodes, B. C. (2010). Vocal and physiological changes in response to the physical attractiveness of conversational partners. Journal of Nonverbal Behavior, 34, 155-167. https://doi.org/10.1007/s10919-010-0087-9

Hustad, K. C., Mahr, T., Natzke, P. E. M., & Rathouz, P. J. (2020). Development of speech intelligibility between 30 and 47 months in typically developing children: A cross-sectional study of growth. *Journal of Speech, Language, and Hearing Research*, *63*, 1675-1687. https://doi.org/10.1044/2020_JSLHR-20-00008

Ikeda, Y. & Masataka, N. (2002). A variable that may affect individual differences in the child-directed speech of Japanese women. *Japanese Psychological Research*, *41*(4), 203-208. https://doi.org/10.1111/1468-5884.00120

Jacewicz, E., Fox, R. A., O'Neill, C., & Salmons, J. (2009). Articulation rate across dialect, age, and gender. *Language Variation and Change*, *21*(2), 233-256. doi: 10.1017/S0954394509990093

Jacobson, J. L., Boersma, D. C. Fields, R. B., & Olson, K. L. (1983). Paralinguistic features of adult speech to infants and small children. *Society for Research in Child Development*, *54*(2), 436-442. https://www.jstor.org/stable/1129704

Janssen, A., & Murachver, T. (2004). The relationship between gender and topic in gender-preferential language use. *Written Communication*, *21*(4), 344-367. doi:10.1177/0741088304270028

Jiang, C., Whitehill, T. L., McPherson, B., & Ng, M. L. (2016). Spectral moment analysis of affricates produced by Mandarin-speaking pre-adolescents with repaired cleft palate. *International Journal of Pediatric Otorhinolaryngology*, *84*, 137-142. https://doi.org/10.1016/j.ijporl.2016.01.029

Johnson, J., & Young, K. (2002). Gendered voices in children's television advertising. *Critical Studies in Media Communication*, *19*(4), 461-480. doi: 10.1080/07393180216572

Kennison, S. M., & Byrd-Craven, J. (2015). Gender differences in beliefs about infant-directed speech: The role of family dynamics. *Child Development Research*, *2015*, 1-6. https://doi.org/10.1155/2015/871759

Kitamura, C., & Burnham, D. (2003). Pitch and communicative intent in mother's speech: Adjustments for age and sex in the first year. *Infancy*, *4*(1), 85-110. https://doi.org/10.1207/s15327078in0401_5

Knoll, M. A., & Costall, A. (2015). Characterising F(0) contour shape in infant- and foreigner-directed speech. *Speech Communication*, *66*, 231-243. https://doi.org/10.1016/j.specom.2014.10.007

Kotami, K., & Yoshimi, T. (2018). Machine-learning classification of pronunciation difficulty for learners of English as a Foreign Language. *Research in Corpus Linguistics*, *6*, 1-8. doi: 10.32714/ricl.06.01

Kuchirko, Y. A., Schatz, J. L., Fletcher, K. K., & Tamis-Lemonda, C. S. (2020). Do, say, learn: the functions of mothers' speech to infants. *Journal of Child Language*, *47*, 64-84. doi:10.1017/S0305000919000308

Kwon, H. (2010). Gender difference in speech intelligibility using speech intelligibility tests and acoustic analyses. *The Journal of Advanced Prosthodontics*, *2*, 71-76. doi:10.4047/jap.2010.2.3.71

Labov, W. (1990). The interaction of sex and social class in the course of linguistic change. *Language, Variation and Change*, *2*, 205-254. https://doi.org/10.1017/S0954394500000338

Ladefoged, P., & Johnson, K. (2015). *A course in phonetics* (7th ed.). Cengage Learning.

Ladegaard, H. J., & Bleses, D. (2003). Gender differences in young children's speech: The acquisition of sociolinguistic competence. *International Journal of Applied Linguistics*, *13*, 222-233. https://doi.org/10.1111/1473-4192.00045

Lathrop-Marshall, H., Keyser, M. M. B., Jhingree, S., Giduz, N., Bocklage, C., Couldwell, S., Edwards, H., Glesener, T., Moss, K., Frazier-Bowers, S., Philips, C., Turvey, T., Blakey, G., White, R., Mielke, J., Zajac, D., & Jacox, L. A. (2022). Orthognathic speech pathology: Impacts of class III malocclusion on speech. *European Journal of Orthodontics*, *44*, 340-351. https://doi.org/10.1093/ejo/cjab067

Leaper, C., & Ayres, M. M. (2007). A meta-analytic review of gender variations in adults' language use: Talkativeness, affiliative speech, and assertive speech. *PSPR*, *11*(4), 328-363. doi: 10.1177/1088868307302221

Leaper, C., & Smith, T. E. (2004). A meta-analytic review of gender variations in children's language use: Talkativeness, affiliative speech, and assertive speech. *Developmental Psychology*, *40*(6), 993-1027. doi: 10.1037/0012-1649.40.6.993

Leavell, A. S., Tamis-LeMonda, C. S., Ruble, D. N., Zosuls, K. M., & Cabrera, N. J. (2012). African American, White and Latino fathers' activities with their sons and daughters in early childhood. *Sex Roles*, *66*, 53-65. doi 10.1007/s11199-011-0080-8

Lee, A., & Doherty, R. (2017). Speaking rate and articulation rate of native speakers of Irish English. *Speech, Language and Hearing*, *20*(4), 206-211. https://doi.org/10.1080/2050571X.2017.1290337

Li, F., Rendall, D., Vasey, P. L., Kinsman, M., Ward-Sutherland, A., & Diano, G. (2016). The development of sex/gender-specific /s/ and its relationship to gender identity in children and adolescents. *Journal of Phonetics*, *57*, 59-70. https://doi.org/10.1016/j.wocn.2016.05.004

Mack, S., & Munson, B. (2012). The influence of /s/ quality on ratings of men's sexual orientation: Explicit and implicit measures of the 'gay lisp' stereotype. *Journal of Phonetics*, *40*(1), 198-212. https://doi.org/10.1016/j.wocn.2011.10.002

Masataka, N. (2002). Pitch modification when interacting with elders: Japanese women with and without experience with infants. *Journal of Child Language*, *29*, 939-951. doi :10.1017/S0305000902005378

McLeod, S., Crowe, K., & Shahaeian, A. (2015). Intelligibility in context scale: Normative and validation data for English-speaking preschoolers. *Language, Speech, and Hearing Services in School*, *46*, 266-276. Doi: 0.1044/2015_LSHSS-14-0120

McShefferty, D., Whitmer, W. M., & Akeroyd, M. A. (2015). The just-noticeable difference in speech-to-noise ratio. *Trends in Amplification*, *19*, 1-9. doi:10.1177/2331216515572316

Mesman, J. & Groeneveld, M. G. (2017). Gendered parenting in early childhood: Subtle

but unmistakable if you know where to look. *Child Development Perspectives*, *12*(2), 22-27. https://doi.org/10.1111/cdep.12250

Morris, R. J., McCrea, C. R., & Herring, K. D. (2008). Voice onset time differences between adult males and females: Isolated syllables. *Journal of Phonetics*, *36*(2), 308-317. https://doi.org/10.1016/j.wocn.2007.06.003

Munson, B., Lackas, N., & Koeppe, K. (2022). Individual differences in development of gendered speech in preschool children: Evidence from a longitudinal study. *Journal of Speech, Language, and Hearing Research*, *65*, 1311-1330. https://doi.org/10.1044/2021_JSLHR-21-00465

Nalborczyk, L., Batailler, C., Loevenbruck, H., Vilain, A., & Bürkner, P.-C. (2019). An introduction to Bayesian multilevel models using brms: A case study of gender effects on vowel variability in standard Indonesian. *Journal of Speech, Language and Hearing Research*, *62*(5), 1225-1242. https://doi.org/10.1044/2018_JSLHR-S-18-0006

Nicollas, R., Garrel, R., Ouaknine, M., Giovanni, A., Nazarian, B., & Triglia, J. M. (2008). Normal voice in children between 6 and 12 years of age: Database and nonlinear analysis. *Journal of Voice*, *22*, 671-675. https://doi.org/10.1016/j.jvoice.2007.01.009

Oberauer, K. (2022). The Importance of Random Slopes in Mixed Models for Bayesian Hypothesis Testing. *Psychological Science*, *33*(4), 648-665. https://doi.org/10.1177/09567976211046884

Pépiot, E. (2014). Male and female speech: A study of mean f0, f0 range, phonation type and speech rate in Parisian French and American English speakers. *Speech Prosody*, *7*, 305-309. https://shs.hal.science/halshs-00999332

Pexels. (n.d.). The best free stock photos, royalty free images & videos shared by creators. [Stock images]. https://www.pexels.com/

Prodi, N., Visentin, C., Borella, E., Mammarella, I. C., & Di Domenico, A. (2019). Noise, age, and gender effects on speech intelligibility and sentence comprehension for 11- to 13-year old children in real classrooms. *Frontiers in Psychology*, *10.* doi: 10.3389/fpsyg.2019.02166

Pruden, S. M., & Levine, S. C. (2017). Parents' spatial language mediates a sex difference in preschoolers' spatial-language use. *Psychological Science*, *28*(11), 1583–1596 doi: 10.1177/0956797617711968

Rodriguez-Cuadrado, S., Baus, C., & Costa, A. (2018). Foreigner talk through word reduction in native/non-native spoken interactions. *Bilingualism*, *21*(2), 419-426. https://doi.org/10.1017/S1366728917000402

Rothermich, K., Harris, H. L., Sewell, K., & Bobb, S. C. (2019). Listener impressions of foreigner-directed speech: A systematic review. *Speech Communication*, *112*, 22-29. https://doi.org/10.1016/j.specom.2019.07.002

Rubin, D. B. (1984). Bayesianly justifiable and relevant frequency calculations for the applied statistician. *The Annals of Statistics*, *12*(4), 1151–1172. http://www.jstor.org/stable/2240995

Seguino, S. (2007). Plus Ça Change? Evidence on global trends in gender norms and stereotypes. *Feminist Economics*, *13*(2), 1-28. https://doi.org/10.1080/13545700601184880

Shaw, C. A., & Gordon, J. K. (2021). Understanding elderspeak: An evolutionary concept analysis. *Innovation in Aging*, *5*(3), 1-18. https://doi.org/10.1093/geroni/igab023

Simpson, A. P. (2009). Phonetic differences between male and female speech. *Language and Linguistics Compass*, *3*(2), 621-640. doi:10.1111/j.1749-818x.2009.00125.x

Steensma, T. D., Kreukels, B. P. C., de Vries, A. L. C., & Cohen-Kettenis, P. T. (2013). Gender identity development in adolescence. *Hormones and Behaviour*, *64*(2), 288-297. https://doi.org/10.1016/j.yhbeh.2013.02.020

Stennes, L. M., Burch, M. M., Sen, M. G., & Bauer, P. J. (2005). A longitudinal study of gendered vocabulary and communicative action in young children. *Developmental Psychology*, *41*(1), 75-88. doi:10.1037/0012-1649.41.1.75

Suire, A., Tognetti, A., Durand, V., Raymond, M., & Barkat-Defradas, M. (2020). Speech acoustic features: A comparison of gay men, heterosexual men, and heterosexual women. *Archives of Sexual Behaviour*, *49*, 2575-2583. https://doi.org/10.1007/s10508-020-01665-3

Tenders, A., Rispoli, M., Sethilselvan, A., & Loucks, T. A. (2019). Early speech rate development: A longitudinal study. *Journal of Speech, Language, and Hearing Research*, *62*, 4370-4381. https://doi.org/10.1044/2019_JSLHR-19-00145

Tripp, A., & Munson, B. (2022). Perceiving gender while perceiving language: Integrating psycholinguistics and gender theory. *WIREs Cognitive Science*, *13*(2). https://doi.org/10.1002/wcs.1583

Uther, M., Knoll, M. A., & Burnham, D. (2007). Do you speak E-NG-L-I-SH? A comparison of foreigner- and infant-directed speech. *Speech Communication*, *49*, 2-7. doi:10.1016/j.specom.2006.10.003

Walker, J. F., & Archibald, M. D. (2006). Articulation rate in preschool children: A 3-year longitudinal study. *International Journal of Language and Communication Disorders*, *41*(5), 541-565. doi: 10.1080/10428190500343043

Weirich, M., & Simpson, A. P. (2018). Gender identity is indexed and perceived in speech. *PLos One*, *13*(12). https://doi.org/10.1371/journal.pone.0209226

Whiteside, S. P., Hanson, A., & Cowell, P. E. (2004). Hormones and temporal components of speech: Sex differences and effects of menstual cyclicity on speech. *Neuroscience Letters*, *367*(1), 44-47. https://doi.org/10.1016/j.neulet.2004.05.076

Whiteside, S. P., & Marshall, J. (2001). Developmental trends in voice onset time: Some evidence for sex differences. *Phonetica*, *58*, 196-210. https://doi.org/10.1159/000056199

Yoho, S. E., Borrie, S. A., Barrett, T. S., & Whittaker, D. B. (2019). Are there sex effects for speech intelligibility in American English? Examining the influence of talker, listener, and methodology. *Attention, Perception, & Psychophysics*, *81*, 558-570. https://doi.org/10.3758/s13414-018-1635-3

Zimman, L. (2018). Transgender voices: Insights on identity, embodiment, and the gender of the voice. *Language and Linguistics Compass*, *12*(8). https://doi.org/10.1111/lnc3.12284

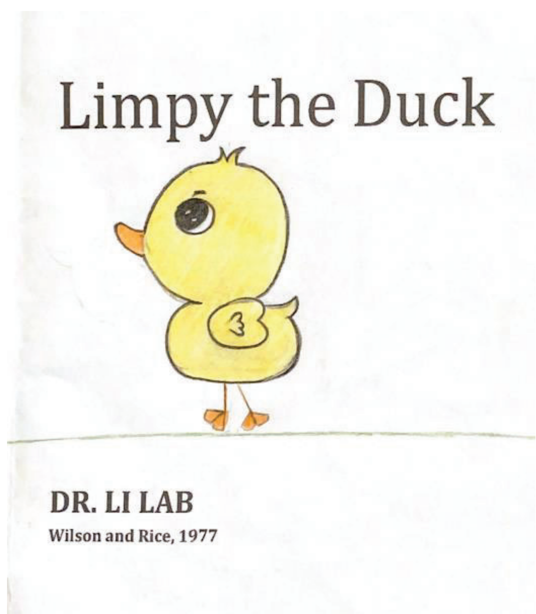
Zosuls, K. M., Ruble, D. N., Tamis-LeMonda, C. S., & Shrout, P. E. (2009). The acquisition of gender labels in infancy: Implications for gender-typed play. *Developmental Psychology*, *45*(3), 688-701. doi:10.1037/a0014053

Appendix A

Passages used in the current study

Limpy the Duck from Berggren et al. (2018)

Limpy is a fuzzy, yellow, baby duck. He belongs to a fisherman. The fisherman lives in a little house by the bay. Every morning children go swimming in the bay. About 10:00, Limpy waddles out to the road to wait for the children. When he hears them coming he begins a loud, excited quacking. The children always bring bread or corn for Limpy. He will nip at their fingers or peck at their bare toes until he is fed. Limpy never follows the children down to the shore. He likes to swim in his own little pond. It is much safer.



The North Wind and the Sun (short version) from Kotami & Yoshimi (2018)

The North Wind and the Sun were disputing which was the stronger, when a traveller came along wrapped in a warm cloak.

They agreed that the one who first succeeded in making the traveller take his cloak off should be considered stronger than the other.

Then the North Wind blew as hard as he could, but the more he blew the more closely did the traveller fold his cloak around him; and at last the North Wind gave up the attempt.

Then the Sun shone out warmly, and immediately the traveller took off his cloak.

And so the North Wind was obliged to confess that the Sun was the stronger of the two.

The North Wind and the Sun

DR. LI'S LAB

Aesop Fables

Appendix B

Baby photos, names and facts used in the current study

Lucas

Olivia

Liam

Sophia

Noah

Emma

Lucas loves playing with dinosaurs.
Sophia loves being pushed in the stroller.
Noah loves trains.
Olivia loves kittens.
Emma loves teddy bears.
Liam loves being outside with his dad.

Appendix C

Face ID of photos

Baby face ID #1
Boy = Lucas
Girl = Olivia

Baby face ID #2
Boy = Liam
Girl = Sophia

Baby face ID #3
Boy = Noah
Girl = Emma