

BILINGUAL INFANTS' ACCOMMODATION OF ACCENTED SPEECH

by

TAMARA MAE HUDON

B.A. (Honours), University of Ottawa, 2009

THESIS SUBMITTED TO THE FACULTY OF

GRADUATE AND POSTDOCTORAL STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR

THE DEGREE OF DOCTOR OF PHILOSOPHY

in

EXPERIMENTAL PSYCHOLOGY

UNIVERSITY OF OTTAWA

© Tamara Hudon, Ottawa, Canada 2013

Dedication

This thesis is dedicated to my husband

Jon Hudon

Thank you for holding me steady.

S.P.L.

And to

My three shining lights

Haleigh, Evan and James

Acknowledgements

Thank you very much to the parents and babies who volunteered to participate in my studies, without whom of course none of this would be possible.

I am very grateful to Christopher Fennell for accepting me as his doctoral student. I extend my sincere gratitude for bringing me into the fascinating world of wugs and kems, for the many impromptu whiteboard demonstrations, for tolerating my “emergency meetings”, but especially for allowing me the latitude to explore new methodologies and domains. Thank you also to my undergraduate supervisor Catherine Plowright who has been a mentor and role model in my research career.

I would also like to thank my committee members, Cristina Atance, Alain Desrochers and Tania Zamuner for guidance in the development of this project. The questions that you asked me in the earliest stages of this thesis helped me to move from thinking about how neat it would be to get a significant result in one of my experiments, to thinking about the broader theoretical implications of studies of this nature. Or rather, from “Can bilinguals succeed in this task” to “If bilinguals succeed in this task when monolinguals do not, what does this say about how bilinguals mentally represent language”, and importantly, “what is an accent” anyways? I would also like to thank Veronika Huta and Rachel Schmale for help in performing statistical analyses on speaker samples, as well as Dwayne Schindler for being a second set of eyes and Elizabeth Johnson for guidance in establishing the Headturn Preference Procedure at the Language Development Lab.

I will eternally hold a special place in my heart for all of my fellow LDL members, past and present. In particular, thank you to Meghan Spring for the often needed crash courses in linguistics, for letting me sleep on your floor several times, and of course for all

the fun times we've had just hanging out. Thank you to Brock Ferguson for helpful comments in the development of this project and for inspiring debates on the nature of first words. To Zhen Xu and Becky Ginsburg for being fantastic RAs and great friends. To Thyra Driver and Angeline Tsui for meticulous help in data coding and entry, and to Emily Prosser, Corinne Laliberté, and Caitlin Atkinson for helping in the recruitment and scheduling of participants. I would also like to extend a very special thank you to my colleague Sophie-Godbout Beaulieu; I am so happy that we entered this journey at the same time. Your friendship is priceless; let's get together for tea and cupcakes.

I am forever indebted to my friends and family for seeing me through this entire process. Thank you to my Mom (Rhonda Potapchuk) and Dad (Ron Potapchuk) for always having faith that I could do this, and for being so supportive along the way. Thank you especially for giving me a safe place to be imperfect. Thank you also to my fantastic in-laws Karen and Sylvain Hudon for your support, card games, and for letting Jon and I get away on so many occasions. To Megan Findlay, Derek Noon, Clay Hamelin, and Melody Larabie for being on my team. And finally, thank you to Jon, Haleigh, Evan and James to whom this thesis is dedicated. I love you guys with all my heart and thank you for being so patient while I've been at school. This has been a family project: you've all sacrificed so that I could do this and this degree is as much yours as it is mine.

Abstract

Infant word recognition is sometimes hindered by variability in the speech input. Previous research has shown that, at 9 months, monolinguals do not generalize wordforms across native- and accented-speakers (Schmale & Seidl, 2009). In the current study however, it was predicted that bilingual infants would be advantaged in accommodating for accented speech due to experience with phonetic variability across their two phonological systems. It was also predicted that this hypothesized ability would be restricted to accommodating for an accent derived from a familiar language (e.g., French-English bilinguals would accommodate for French-accented English but not Mandarin-accented English), since this type of variability would be consistent with the language sounds to which infants were regularly exposed.

Study 1 set the experimental stage by identifying native and non-native speakers with similar voices, as perceived by a group of adults. This was done in order to restrict variability across speakers to differences in accent, rather than biological differences in voice (e.g., a higher or lower pitched voice). Following speaker selection, acoustic measurements of vowels and word stress placement were taken to compare native and non-native speakers and confirmed several expected deviations between native and accented speech. Study 2 tested the hypothesis that bilingual infants would be advantaged in accommodating for these deviations when the accent is derived from a familiar phonology. Using a headturn preference procedure (HPP), 9- and 13-month-old English-learning monolinguals and French-English learning bilingual infants were tested on their ability to recognize familiarized English wordforms across a native- and French-accented speaker. Bilinguals in both age groups succeeded in generalizing wordforms across speakers, however monolingual

infants failed regardless of age. Study 3 tested whether bilinguals' success would persist when the accented speaker's first language was unfamiliar. Infants in this study failed as a group to generalize across native- and Mandarin-accented productions of English wordforms. However, bilinguals who received balanced exposure to their two languages performed better in accommodating for Mandarin accented speech than unbalanced bilinguals. This hints at a general ability to ignore irrelevant phonetic information, perhaps due to an advantage in cognitive control.

Table of Contents

Dedication	ii
Acknowledgements.....	iii
Abstract.....	v
List of Tables.....	ix
List of Figures	x
List of Appendices.....	xii
Chapter 1.....	1
Introduction.....	1
Phonetic and Phonological Development	4
Wordform Recognition	15
Word Recognition	18
Coping with Wordform Variability	21
Accented Speech as a Source of Variability	24
Defining accent.	24
Coping with accented speech.	26
The Case for a Bilingual Advantage in Accent Accommodation	32
Acquiring two phonological systems	33
Exposure to accented speech.	35
Cognates and loanwords.	36
Cognitive advantages.	37
Overview of Studies and Hypotheses	39
Chapter 2.....	44
Study 1: Pairing and Comparing.....	44
Introduction	44
Pairing Speakers.....	45
Target wordforms.	45
Acoustic recordings.	46
Speaker similarity ratings.....	47
Analysis of speaker similarity ratings.	48
Comparing Speakers.....	53
Speaker characteristics.	53
Stimuli Lengths.	53
Subphonemic comparisons.	54
Close phonetic transcriptions	54
Acoustic measurements.....	56
Summary of vowel findings.....	61
Summary of VOT findings.	63
Suprasegmental comparisons.....	64
Summary of word stress findings.....	69
Chapter 3.....	71
Study 2: Bilingual Infants' Accommodation of Phonetically Familiar Accented Speech	71
Introduction	71
Methods.....	71

Apparatus.....	71
Design.....	72
Procedure.....	73
Questionnaire measures.....	76
Bilingualism questionnaire.....	76
Accent exposure.....	76
Participants.....	77
Results.....	79
Familiarized vs. non-familiarized passages.....	79
Data cleaning and screening.....	79
Overall looking time in each half x age group x language group.....	79
Familiarity x age group x language group.....	80
GLM testing for condition effects.....	83
Influence of language and accent exposure.....	89
Variables.....	89
Curve estimation results.....	91
Logistic regression for task success by exposure to accented speech.....	93
Discussion.....	94
Chapter 4.....	100
Study 3: Bilingual Infants' Accommodation of Phonetically Unfamiliar Accented	
Speech.....	100
Introduction.....	100
Methods.....	101
Participants.....	101
Results.....	101
Familiarized vs. non-familiarized passages.....	101
Data cleaning and screening.....	101
T-tests.....	102
GLM testing for condition effects.....	102
Influence of language and accent exposure.....	103
Variables.....	103
Curve estimation results.....	104
Exploratory analysis including monolingual infants.....	107
Discussion.....	111
Chapter 5.....	113
General Discussion.....	113
Broader implications.....	117
Limitations and areas for future research.....	120
Conclusion.....	121
References.....	123
Appendix A: Passages Containing Target Words.....	140
Appendix B: Language Exposure Questionnaire.....	141

List of Tables

Table 1. Mean squared distances (squared Euclidean distances) of all English-French speaker pairs in sinewave and natural speech.

Table 2. Mean squared distances (squared Euclidean distances) of all English-Mandarin speaker pairs in sinewave and natural speech.

Table 3. Length of audio stimuli in seconds for each speaker.

Table 4. Close phonetic transcriptions of native-English, French-accented English and Mandarin-accented English target words.

Table 5. Mean difference of acoustic measures for speaking pairs; standard deviations in italics.

Table 6. Acoustic measurements of stress correlates in target words.

Table 7. Experimental design.

Table 8. Curve estimation for task success (generalization across native and French-accented English wordforms) by exposure to English.

Table 9. Linear and quadratic relationships between task success and LQM of non-native speech exposure.

Table 10. Linear and quadratic relationships between task success and exposure to English.

List of Figures

*Figure 1. **Wordform Generalization Across Accents.*** Average looking time per trial in seconds for passages containing familiarized and non-familiarized target words.

*Figure 2. **Generalization Across Accents: Performance by Group.*** Looking time towards passages containing familiarized and non-familiarized target words by group. Error bars are ± 1 SD.

*Figure 3. **Bilinguals' Performance by Condition.*** Bilingual infants' orientation towards passages containing familiarized versus non-familiarized target wordforms in each condition. Error bars are ± 1 SE.

*Figure 4. **Bilingual 9-Month-Olds.*** Bilingual 9-month-olds' performance in an accent generalization task. Error bars are ± 1 SE.

*Figure 5. **Bilingual 13-Month-Olds.*** Bilingual 13-month-olds' performance in an accent generalization task. Error bars are ± 1 SE.

*Figure 6. **Generalization Across Native and French-Accented English Speech by Exposure to English.*** Scatterplot of the square root of task success scores versus the percentage of time infants were exposed to English. English exposure was calculated as the number of hours infants' listened to caregivers (and close friends and relatives) speaking English divided by the total number of hours of language exposure.

*Figure 7. **Generalization Across Native and Mandarin-Accented English Speech by Exposure to English.*** Scatterplot of the square root of task success scores versus the percentage of time infants were exposed to English. English exposure was calculated as the number of hours infants' listened to caregivers (and close friends and relatives) speaking English divided by the total number of hours of language exposure.

Figure 8. Generalization Across Native and Mandarin-Accented English Speech by

Exposure to Non-Native Speech. Scatterplot of the square root of task success scores versus the percentage of time infants were exposed to non-native (accented) speech. Non-native speech exposure was calculated as the number of hours infants' listened to caregivers (and close friends and relatives) speaking in a second, non-native, language, divided by the total number of hours of language exposure.

Figure 9. Generalization Across Native and Mandarin-Accented English Speech by

Exposure to English with Monolinguals. Scatterplot of the square root of task success scores versus the percentage of time infants were exposed to English. English exposure was calculated as the number of hours infants' listened to caregivers (and close friends and relatives) speaking English divided by the total number of hours of language exposure. This plot includes data from monolingual infants.

List of Appendices

Appendix A: Passages containing target words.

Appendix B: Language Exposure Questionnaire.

Chapter 1

Introduction

The process of word learning can quite easily be taken for granted given the speed at which infants acquire and produce their first words. Although most infants will comprehend their first words at around 6 months (Bergelson, & Swingley, 2012; Tincoff & Jusczyk, 1999) and produce their first words by 12 months (Fenson, Dale, Reznick, Bates, Thal, & Pethick, 1994), the development of this ability involves a series of remarkable feats. During the first year, novice language learners identify which sound categories (phonemes) are meaningfully contrastive in their language environment (e.g., the distinction between /b/ in ‘bog’ and /d/ in ‘dog’; e.g., Werker & Tees, 1984). They also learn which sounds most frequently co-occur, and use this and other cues such as breath-pauses and suprasegmental information such as word stress to parse perceptually continuous speech (e.g., Johnson & Jusczyk, 2001; Pelucchi, Hay, & Saffran, 2009; Polka & Sundara, 2012). This allows infants to extract wordforms from a speech signal absent of reliable pauses between words. These wordforms are eventually mapped to the concepts to which they refer. Shortly after the first year, infants are able to use their refined phonological sensitivities to simultaneously learn even two very similar sounding labels for different objects (e.g., Yoshida, Fennell, Swingley, & Werker, 2009).

The ostensive complexity of the tasks leading to infants’ first words and the apparent ease with which babies master these tasks is even more impressive when considering infants who are raised from birth in a bilingual environment. These infants are tasked with concurrently identifying language patterns in two different languages with unique sets of regularities. Yet, evidence has suggested that bilingual-learning infants reach major developmental language milestones comparably to their monolingual peers. For example,

bilingual infants babble (Oller, Eilers, Urbano, & Cobo-Lewis, 1997), produce first words (e.g., Pettito, Katerelos, Levy, Gauna, Tetreault, & Ferraro, 2001), and follow vocabulary growth trajectories (Pearson & Fernandez, 1994) similarly to monolingual infants. Indeed, when both languages are considered, bilingual infants also have similarly sized overall vocabularies as same-aged monolinguals (Pearson & Fernandez, 1994; Pearson, Fernandez & Oller, 1995; Pettito et al., 2001). Thus, from a global perspective, monolingual and bilingual infants appear to develop language on par with one another.

Recent investigations of the aforementioned steps towards word learning have, however, revealed subtle differences between these groups. In most cases where differences have been discovered, they have initially been interpreted as a lag in bilingual development owing to a more complex learning environment. For example, some research has shown a bilingual delay in discriminating native language phonemes (Bosch & Sebastián-Gallés, 2003a; Sebastián-Gallés & Bosch, 2009), in detecting phonemic mispronunciations of words (Ramon-Casas, Swingley, Sebastián-Gallés, & Bosch, 2009), and in learning similar sounding novel labels for novel objects (Fennell, Byers-Heinlein, & Werker, 2007).

Although would-be monolingual and bilingual infants presumably begin life with the same general learning mechanisms and language learning biases (Curtin, Byers-Heinlein, & Werker, 2011), and produce first words at the same time and rate, it is likely that differential input (i.e., 1 language or 2) creates different paths to the same endpoint (Werker, Byers-Heinlein, & Fennell, 2009). Indeed it has been suggested that bilingual infants' "failures" or "delays" in language development do not necessarily reflect an inability to detect phonemic contrasts (in isolation or in words), but rather represent an attentional indifference to certain contrasts at certain points in development (Sebastián-Gallés, 2010).

As will be discussed in detail later in this thesis, relative to their monolingual peers, bilingual infants are exposed to substantial linguistic variability. Bilinguals hear two overlapping phonological systems, they habitually hear words from each of their languages that sound subtly different and yet mean the same thing (i.e., cognates and loanwords), and they are frequently exposed to accented speech which contributes to both phonological overlap and word-level variability. This type of learning environment may lead to a general tolerance for subtle phonetic or phonemic differences, consequently allowing infants to generalize words over a variety of phonetic forms. Throughout this thesis, the term tolerance will be used to characterize a tendency to ignore subtle acoustic variation.

The possibility that bilingual infants are more tolerant than bilinguals in ignoring such variability is intricately tied to experimental procedures currently used to assess infants' linguistic abilities. For decades, research with monolingual infants has profited from the idea that discrimination (of sounds or words) is evidenced by an increase in attention (e.g., through eye gaze or orientation) to a novel stimulus following a period of exposure to another stimulus (e.g., exposure to /b/ followed by /d/). If bilingual infants are not surprised by subtle auditory changes, then these methods may not be sufficiently sensitive to assess bilingual abilities (Sebastián-Gallés, 2010). In fact some evidence has already suggested that in tasks that do not require a surprise response, bilingual infants demonstrate phonemic discrimination similar to their monolingual peers (Albareda-Castellot, Pons, & Sebastián-Gallés, 2011).

This thesis will first describe the ways in which monolingual and bilingual development has been shown to both converge and diverge focusing specifically on the areas of phonetic and phonological development, wordform recognition, word recognition and mispronunciation detection. Drawing from this data, it will be argued that, due to experience

with more variable input in the learning environment, bilingual infants only appear to lag in certain abilities due to a differential behavioural response to subtle acoustic changes. That is to say that, as proposed by Sebastián-Gallés (2010), bilinguals are less likely to demonstrate an increase in attention to such changes.

In a series of three studies, it will be demonstrated that this tolerance for acoustic variability is not limitless, but rather is input dependent and therefore tied to the characteristics of the languages being learned. Furthermore, it will be argued that tolerance for variability encountered in the bilingual learning environment has beneficial consequences. Specifically, it allows bilingual infants to ignore irrelevant phonetic information to generalize wordforms across speakers with different accents; a task that monolingual infants find difficult in early language acquisition. The consequences of acquiring the ability to generalize across variable wordforms early in language development will be discussed as a potential facilitator in extracting the linguistic regularities of two languages simultaneously.

Phonetic and Phonological Development

A fundamental skill on the path to early word learning is the ability to understand which sounds are meaningful in one's language environment. Infants are born with an initially broad sensitivity to phonetic contrasts (Eimas, Siqueland, Jusczyk, & Vigorito, 1971; Streeter, 1976; Trehub, 1976). During the first year of life infants' linguistic sensitivities are tuned to the regularities of the language(s) to which they are exposed. Consequentially, infants either maintain (Best, McRoberts, Lafleur, & Silver-Isenstadt, 1995; Werker & Tees, 1984) or become increasingly proficient in discriminating between native-language sounds (Kuhl, Stevens, Hayashi, Deguchi, Kritani, & Iverson, 2006; Narayan, Werker, & Beddor, 2010; Polka, Colantonio, & Sundara, 2001; Tsao, Liu, & Kuhl,

2006). Meanwhile, non-native sound discrimination attenuates (Best & McRoberts, 2003; Bosch & Sebastián-Gallés, 2003; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Polka & Werker, 1994; Werker & Tees, 1984). The same developmental pattern has also recently been demonstrated for linguistic tone (Mattock & Burnham, 2006) and linguistic sign (Palmer, Fais, Golinkoff, & Werker, 2012). This process of “perceptual narrowing” allows infants to learn the relevant contrastive features of their language (i.e., phonemes) necessary to distinguish between words.

For example, English learning 6-8 month-olds discriminate between a voiceless, unaspirated retroflex stop consonant [ɬ], produced by curling the tongue backward in the mouth, and a dental unaspirated stop consonant [t̪] produced by touching the tongue against the back of the teeth (Werker, Gilbert, Humphrey, & Tees, 1981). This distinction is contrastive in Hindi; in Hindi two words can differ from each other based only on “place of articulation” of [t̪] (Werker et al., 1981). However, this contrast is not meaningful in English, and by 10-12 months, English-learning monolinguals perceive these two sounds as belonging to the same sound category (Werker et al., 1981). On the other hand, the difference between /m/ and /n/ is meaningful in English (e.g., the difference between “**m**ice” and “**n**ice”) despite that these sounds also differ only on place of articulation, where /m/ is produced by touching the lips together, and /n/ is produced by touching the tongue to the roof of the mouth. Convention in the field of linguistics dictates that linguistic sounds that differ from each other in a non-meaningful way in a language (i.e., “phones”) are represented in brackets, whereas meaningfully contrastive units in a language (i.e., “phonemes”) are represented in slashes. In linguistic transcription, words can be represented at the phonetic level, including all phonetic details irrespective of which sounds are meaningfully contrastive, or they can be transcribed at the phonemic level, including only

those sounds that are distinctive in the language under investigation. For example the English word “pit” could be transcribed as [p^hit] at the phonetic level or /pit/ at the phonemic level. Although English speakers aspirate following a word-initial [p] sound (indicated by the superscript *h*), aspiration of [p] is not meaningfully contrastive in English and therefore is not included in a phonemic transcription.

The phonological system of a language includes both an inventory of meaningfully contrastive sounds (phonemes) and the rules governing how those sounds interact. There is considerable debate in the linguistic literature over the age at which infants begin to acquire a phonological system (see Werker & Curtin, 2005). This is in part due to the traditional definition of phonemes as being abstract units that are used to contrast meaning (Trubetzkoy, 1939/1969). To some, this suggests that language learners must acquire a certain number of wordform-meaning links in order to learn which sound categories are meaningfully contrastive (e.g., Beckman & Edwards, 2000; Werker & Curtin, 2005). Although it is beyond the scope of this thesis to discuss the various theoretical perspectives as to what constitutes a phoneme, attention is drawn to this debate in order to clarify the terminology used throughout this thesis. Here, the terms *phone* and *phonetic* are used to describe acoustic information that is not meaningful at the lexical level (i.e., the meaning of “dog” does not differ whether it is spoken by a male or female voice). The terms *phoneme* and *phonological*, on the other hand, are used to describe meaningfully contrastive units. This perspective is taken because 1) it provides a means of describing meaningfully relevant versus irrelevant acoustic information, central to the topic at hand, and 2) although there is no current consensus on the nature of infants’ mental representation of speech sounds, infants nonetheless learn language in a phonological environment. That is to say, infants learn language from speakers who possess abstract phoneme categories and rules for combining

these categories. As such, infants appear to hone in on the meaningfully contrastive units of their language, phonemes, before having substantial productive vocabularies. Indeed, perceptual narrowing in and of itself can be thought of as evidence that the transition between phonetic and phonemic perception is continuous and begins early in language acquisition.

Perceptual narrowing is typically thought of as a linear process in monolingual infants, whose inventories of native language phonemes are typically solidified by 6-months in the case of vowels (Kuhl et al., 1992; Polka & Werker, 1994) and 12-months for consonants (e.g., Best et al., 1995; Kuhl et al., 2006; Werker & Tees, 1984). For bilingual infants however, this process appears to be more complex. For example, some studies have demonstrated a lull in bilingual infants' ability to discriminate certain *native* vowel contrasts for a period between 4 and 12 months, suggesting a U-shaped rather than linear developmental pattern (Bosch & Sebastián-Gallés, 2003a; 2003b).

Early studies examining the development of infants' phonetic/phonemic discrimination abilities were conducted using the Headturn Preference Procedure (HPP) (Kemler Nelson, Jusczyk, Mandel, Myers, Turk, & Gerken, 1995). In such discrimination tasks, infants are first familiarized with one sound (e.g., /e/) and then tested for their preference for either this sound repeated again or a new sound (e.g., /ε/). Preference is indicated by the length of time the infant will turn his or her head towards the sound and a concurrent visual stimulus (e.g., a flashing light). Depending on the length of exposure to the initial stimulus, mediated by the complexity of the stimuli (i.e., infants habituate more quickly to less complex stimuli), discrimination is indicated by longer head turns to either the old or new stimulus (Hunter & Ames, 1988). Given a very short familiarization period, the expected effect is a preference for the same stimulus as was presented during the

familiarization phase. In contrast, if the infant has already become sufficiently exposed (i.e., habituated) to the initial stimulus, then the prediction is longer orientation towards the novel stimulus (Houston-Price & Nakai, 2004; Hunter & Ames, 1988).

Using this method, Bosch and Sebastián-Gallés (2003a) demonstrated that monolingual (Spanish learning and Catalan learning) infants' ability to discriminate a Catalan vowel contrast that does not exist in Spanish (/e/ - /ɛ/) followed the predicted pattern of perceptual narrowing. At 4-months, a time when infants are still sensitive to most phonetic contrasts, regardless of language experience (Werker & Tees, 1984), both Spanish and Catalan learning infants distinguished between /e/ and /ɛ/. This was indicated by longer orientation towards /e/ following familiarization to /ɛ/ and vice versa. At 8 and 12 months, only the Catalan learning monolinguals, for which the distinction is meaningful, showed evidence of perceiving the contrast.

The bilingual results, however, were deviant from the pattern of perceptual narrowing observed with monolingual infants. Surprisingly, infants who were learning both Spanish and Catalan, and for whom the contrast is meaningful in *one* of their languages, no longer demonstrated evidence of discrimination at 8-months, but appeared to recover the ability to perceive the distinction by 12-months. The same pattern was observed when Spanish-Catalan learning infants were tested on their discrimination of /s/ - /z/, a Catalan-specific consonant contrast (Bosch & Sebastián-Gallés, 2003b), and is consistent with other research suggesting a bilingual delay in cementing phoneme categories (Sundara, Polka, & Genesee, 2006).

Bosch and Sebastián-Gallés (2003a) initially explained bilingual infants' difficulty in distinguishing the /e/ - /ɛ/ contrast as reflecting the distributional properties of these sounds in the bilingual learning environment. Firstly, the Spanish /e/ is associated with first and

second formant values (F1 and F2; height and frontness of the vowel as produced in the mouth) that are somewhere between the values associated with the Catalan /e/ and /ɛ/.

Secondly, the Spanish /e/ occurs more frequently than the Catalan vowels. Thus it is possible that all three sounds were perceived by young bilingual infants as forming a single phoneme category. By 12 months, perhaps due to continued exposure to the languages or due to the onset of word learning, these phoneme categories were sufficiently separated to demonstrate discriminatory behaviour.

The hypothesis that Spanish-Catalan bilinguals collapsed three phoneme categories into a single broad category was reasonable. This is supported by experimental evidence that infants hearing a similar distribution of sounds subsequently formed one phoneme category (Maye, Werker, & Gerken, 2002). However, a later study examining bilingual infants' discrimination of a vowel contrast that exists in both Spanish and Catalan, (/o/ - /u/), revealed the same pattern of temporary failure followed by recovery (Sebastián-Gallés & Bosch, 2009). The authors concluded that distributional learning could not account for bilinguals' apparent failures in phoneme discrimination, since /o/ and /u/ are shared between Spanish and Catalan, and therefore cannot be subsumed under a single distribution. However, it has since been noted that overlapping phonological distributions can occur not only in situations where one language contains a contrast and the other does not (i.e., /e/ - /ɛ/ in Catalan but only /e/ in Spanish), but also when a phoneme category exists in both languages but is instantiated in slightly different ways across languages (e.g., French dental /d/ vs. English alveolar /d/) (Sundara, Polka, Mulnar 2008), or due to exposure to accented speech (Fernald, 2006; Sundara, Polka, & Genesee, 2006). Therefore, distributional learning may yet play a role in bilingual infants' perception of many contrasts.

Another area to consider is the perceptual saliency of the contrast under investigation. Using the same HPP, Spanish-Catalan learning 8-month-olds have been shown to discriminate between /e/ - /u/, which exist in both Spanish and Catalan (Sebastián-Gallés & Bosch, 2009). This contrast is arguably more salient: while /e/ and /ɛ/ are both produced at mid-height towards the front of the mouth, contrasting only on closeness¹, /u/ is produced higher and further to the back of the mouth than /e/ and involves rounding the lips. Therefore /e/ - /u/ is more saliently contrastive, perhaps warranting greater attention in the HPP.

While Sebastián-Gallés and Bosch (2009) caution against using vowel proximity to explain bilingual infants' difficulties in perceiving some contrasts (given that monolingual infants are successful in discriminating between these phonemes), it is possible that more subtly different phoneme categories are insufficient to produce a preferential response in bilingual infants (Sebastián-Gallés, 2010). After all, compared to monolingual infants, bilingual infants routinely hear greater variety in vowel productions. Increased variability results from overlapping phonological distributions between languages (Bosch & Sebastián-Gallés, 2003), the presence of similar sounding words between languages (i.e., cognates) (Sebastián-Gallés & Bosch, 2009), and exposure to accented speech (Sebastián-Gallés, 2010). Thus, bilinguals may not be surprised when an /e/ sound changes to an /ɛ/ sound; at least not sufficiently so to demonstrate a preferentially based novelty effect.

This claim is supported by the finding that Spanish-Catalan bilinguals do demonstrate the ability to distinguish between /e/ and /ɛ/ when successful performance is not dependent upon a preferential response. In the Anticipatory Eye Movement (AEM) paradigm, infants learn to predict the location of a stimulus based on its association with a

¹ A close vowel is produced with the tongue as close as possible to the roof of the mouth without creating a constriction. In contrast, in producing an open vowel the tongue is positioned far from the roof of the mouth.

sound (e.g., /e/ is right and /ɛ/ is left). Using this paradigm, Albareda-Castellot and colleagues (2011) established that Spanish-Catalan bilingual 8-month-olds discriminate between /e/ and /ɛ/ similarly to their Catalan-learning monolingual peers. In contrast, Spanish-learning monolinguals, for whom /e/ and /ɛ/ is not contrastive, did not learn to predict the location of a stimulus based on the /e/ - /ɛ/ distinction. Taken together, these results suggest that, while bilingual infants may be similar to monolinguals in their ability to discriminate between phonemic contrasts, they may differ in the amount of variability that they will allow before demonstrating a behavioural response to a change in sounds. Furthermore, two other studies, using neural measures, have shown that Spanish-English learning bilinguals discriminate an English vowel contrast as easily or perhaps even superiorly to monolingual English learning infants (Shafer, Yu, & Datta, 2011; Shafer, Yu, & Garrido-Nag, 2012).

Other evidence has suggested that bilingual infants are not generally disadvantaged in learning phoneme categories. Sundara and Scutellaro (2010) tested English-Spanish learning 8-month-olds ability to discriminate /e/ – /ɛ/, which is contrastive in English (e.g., ‘bet’ vs. ‘bait’), but not in Spanish. Unlike Catalan-Spanish learning infants of the same age, English-Spanish learning infants succeeded in discriminating the contrast. Similarly, Burns, Yoshida, Hill and Werker (2007), tested French-English bilingual, and English monolingual infants’ ability to discriminate between /b/ and /p/ as realized in both English and French. This contrast exists in both languages, however the boundary between these sounds differs between them. In French, the distinction is between a prevoiced and a voiced unaspirated consonant ([b] vs. [p]), and in English the distinction is between a voiced unaspirated and voiceless aspirated consonant ([p] vs. [p^h]) (Caramazza, Keni-Komshian, Zurif, & Carbone, 1973; see also Burns et al., 2007).

To examine infants' perception of these contrasts, infants were first exposed to multiple repetitions of [p] while looking at a checkerboard screen. This continued until infants habituated to the stimulus, as indicated by a 50% decrease in looking towards the screen. Once the habituation criterion was met, infants heard two test trials: [b] and [p^h]). Boundary discrimination was indicated by longer looking time to either or both of the test trials. At 6-8 months, both monolingual English learning and French-English learning infants could discriminate the French boundary but not the English boundary². At 10-12 months, the English learning infants perceived only the English distinction (this amelioration would be expected via perceptual reorganization) whereas the bilingual infants could perceive both distinctions. This evidence is important because it reaffirms that bilinguals are not disadvantaged in learning phonemic categories, but also because it demonstrates bilingual competence in a task that requires a preferential response.

The /b/ – /p/ distinction perceived by French-English bilinguals is arguably subtle, since it is only characterized by voice onset time (VOT; the length of time between the release of a stop consonant and the beginning of vocal chord vibration), as is the /e/ – /ɛ/ contrast perceived by English-Spanish learning infants but not Spanish-Catalan learning infants. So if it is to be argued that bilingual infants differ in their response to subtle changes, then why do some bilingual infants succeed in discriminating slight phonemic differences?

One argument is that infants succeed when the learning environment provides them with cues to better distinguish between their two languages. Sundara and Scutellaro (2010) suggested that the reason for English-Spanish learning infants' success lies in the rhythmic distinction between these two languages. The characteristics of “rhythmic classes” remain to

² Burns and colleagues (2007) suggest that the French boundary is more perceptually salient than the English boundary, which may have contributed to the English-monolinguals' discrimination of this distinction but not the English distinction.

this day controversial, however there is a tradition of classifying languages into timing groups. Romance languages (e.g., French and Spanish) have been described as “syllable-timed” whereas Germanic languages (e.g., English) have been described as “stress-timed” (Abercrombie, 1967). By this definition, Spanish and Catalan are rhythmically similar (Sundara & Scutellaro, 2010), whereas Spanish and English are rhythmically distinct (e.g., Pike, 1945). Infants hearing languages that are rhythmically dissimilar may use rhythm as a cue to differentiating their languages, to tease apart phoneme distributions and ultimately to form distinct phonemic categories (Curtin, 2011; Sundara & Scutellaro, 2010). The same could also be surmised for English-French learning infants whose languages are also rhythmically distinct (e.g., Abercrombie, 1967; Lloyd James, 1940).

It is also worth noting that earlier studies of Spanish-Catalan learning infants utilized a different methodological paradigm than newer studies of Spanish-English and French-English learning infants. Although the habituation paradigm used by Burns and colleagues (2007) and Sundara and Scutellaro (2010) required an attentional response to indicate discrimination, it did not require head turns as in the earlier phoneme discrimination studies. A direct central looking response arguably does not require the same degree of “preference” as a sustained head turn and is perhaps a more sensitive test of surprise. Whether differences in methodology can account for differential success patterns awaits further research.

The functionality of consonants, as opposed to vowels, may also influence infants’ ability to discriminate between the contrasts investigated in previous research. By 10-12 months bilingual infants may have already learned that the /b/ - /p/ contrast is useful in distinguishing between words in both languages; as such it may not be as easily ignored. Comparably, vowels are not as useful for identifying words (see Cutler, Sebastián-Gallés, Soler-Vilageliu, & van Ooijen, 2000). For example, in a visual priming task, adults were

quicker to respond to target words following primes that contained the same consonants compared to the same vowels (New, Araujo, & Nazzi, 2008). This led the authors to suggest that consonants held more lexical information than vowels. On the other hand, vowels appear to hold structural information, particularly pertaining to linguistic rules. For instance, when adults were familiarized with nonsense words containing vowels ordered by a rule (e.g., the first two vowels are always the same, followed by a different vowel: AAB), they subsequently generalized this rule to new words. They did not generalize the rule when it involved the ordering of consonants. This finding has since been replicated with infants as young as 11 and 12-months (Hochmann, Benavides-Varela, Nespor, & Mehler, 2011; Pons & Toro, 2010).

In general therefore, if vowels are less important to identifying words, they may be allowed a greater degree of variability before signaling a change in meaning (and correspondingly eliciting greater attention). Additionally, and perhaps additively, evidence suggests that bilingual infants may be more accustomed to hearing variable products of vowels compared to their monolingual peers due to receiving input from bilingual parents. Bilingual adult speakers have been shown to make more vowel category errors than monolingual speakers (Bosch & Ramon-Casas, 2011). Since bilingual-learning infants tend to receive input from at least one bilingual parent (e.g., Mom is a monolingual English speaker, Dad is a French-English bilingual and they speak to each other in English which is common to them both), this suggests greater experience with more variable vowel productions. Therefore, vowels may be allowed a greater degree of latitude before eliciting a preferential response.

In summary, although some evidence has suggested that bilingual infants maintain native-language phoneme discrimination skills across infancy, other evidence suggests that

bilingual infants do not behave in the same way as their monolingual peers in phoneme discrimination tasks. Several factors have been identified as potential contributors to divergent developmental patterns including the amount of phonetic overlap between languages, rhythmic distinctiveness between languages, saliency of contrasts under investigation, the functional role of contrasts under investigation, and task demands.

Wordform Recognition

In addition to learning the sounds of their native language(s), very young infants have demonstrated the ability to parse continuous speech to extract wordforms. Wordforms are distinguished from words in that they refer only to the sound sequence of a word and not to a word-referent pairing. It is possible therefore for an infant to recognize the linguistic *form* of a word without necessarily understanding its meaning. Like many of the early phoneme discrimination studies, the earliest studies examining wordform extraction used the Headturn Preference Procedure (Kemler Nelson et al., 1995). Using this method, Jusczyk and Aslin (1995) familiarized English-learning 7.5-month-olds with two words (i.e., ‘dog’ and ‘cup’ or ‘feet’ and ‘bike’). During the testing phase infants heard repetitions of four passages, half containing the two words heard during the familiarization period and the other half containing the two wordforms not heard during familiarization. By orienting longer towards a visual stimulus while hearing passages containing familiarized wordforms, infants demonstrated the ability to recognize familiar wordforms in continuous speech. This result remained consistent when the familiarization and testing phases were reversed so that infants first heard target words in continuous speech and were then tested for a preference for lists of familiarized or novel words in isolation (Jusczyk & Aslin, 1995). In contrast 6-month-olds did not demonstrate a preference for familiarized over non-familiarized wordforms.

Building on this work, other research has used the same experimental paradigm to demonstrate that 6-month-olds can extract familiar wordforms when in proximity to commonly known words (e.g., ‘Mommy’; Bortfeld, Morgan, Golinkoff, & Rathburn, 2005); that 7.5-month-olds can extract bi-syllabic words from continuous speech (Jusczyk, Houston, & Newsome, 1999); and that 8-month-olds are proficient in recognizing words in their native language whether they are learning a stress-timed or syllable timed language (Nazzi, Mersad, Sundara, Iakimova, & Polka, 2013; Polka & Sundara, 2012; but see Nazzi, Iakimova, Bertoncini, Fredonie, & Alcantara, 2006). Furthermore, this ability extends to recognizing words in a non-native dialect (i.e., Canadian-French learners succeed when presented with the task in European French) (Polka & Sundara, 2012; but see Nazzi et al., 2013).

In another experiment, Jusczyk and Aslin (1995) tested the degree to which 7.5-month-olds encoded phonetic detail during the familiarization phase of the HPP. To do so, they familiarized infants with non-words that differed from their real-word counterparts by a single phoneme (e.g., ‘zeet’ and ‘gike’) prior to testing their preference for either passages containing similar sounding real words (e.g., ‘feet’ and ‘bike’), or novel words. Similar sounding, but phonemically different, words were not a sufficient match to produce a preferential response in 7.5-month-olds. This suggests that monolingual English learning infants can recognize familiar words in fluent speech by 7.5 months, and that they encode wordforms in sufficient detail to detect phonemic mispronunciations.

Using a slightly different methodology, Halle and de Boysson-Bardies (1996), tested wordform recognition in a sample of European French-learning infants. Rather than familiarizing infants with wordforms prior to testing their preference for either these words or non-familiarized wordforms, this study examined infants’ natural preference for frequent

versus infrequent words in their native language. They found that by 11-months, but not earlier, these infants preferred listening to words that were common in French (e.g., ‘lapin’ and ‘voiture’) compared to phonotactically similar but infrequent words. Unlike the infants in Jusczyk and Aslin (1995), 11-month-old French learning infants maintained this preference even when the common words were mispronounced by a single phoneme (e.g., ‘napin’ and ‘boiture’). In other words, these infants appeared to treat phonemically altered words as acceptable pronunciations of frequent wordforms. In contrast, Swingley (2005) found that 11-month-old Dutch learning infants preferred listening to words over non-words but did not maintain this preference when words were mispronounced at onset or offset, thus suggesting, like Jusczyk and Aslin (1995) that words are encoded in phonological detail. The divergent results between studies may reflect a language-specific difference, or simply reflect differences in methodology. Still, regardless of these differences, it is clear that within the first year monolingual infants recognize familiar wordforms.

Very little research to date has compared monolingual and bilingual wordform recognition. One study used the HPP to examine wordform recognition in monolingual (English and Welsh) and bilingual (English-Welsh) 11-month-olds (Vihman, Thierry, Lum, Keren-Portnoy, & Martin, 2007). Infants heard lists of words that were either frequent or infrequent in their natural learning environment and word recognition was indicated by longer orientation towards frequent words. Vihman and colleagues (2007) found that bilinguals recognized frequent words in each of their languages on par with their English-learning peers and before their Welsh-learning peers. These results were paralleled in a neurophysiological experiment by the same authors. Based on evidence that by 13-months, monolingual infants display characteristically different event-related potentials (ERPs) when hearing familiar versus unfamiliar words (Mills, Coffey-Corina, Neville, 1997), Vihman and

colleagues (2007) sought to determine whether the same pattern of neural discrimination would exist in their bilingual sample. The ERP data suggested that, like English-learning monolinguals, English-Welsh bilinguals recognized frequent words in both of their languages. Given that the Welsh-monolingual infants did not produce a differential ERP response to frequent versus infrequent words, it was suggested that bilinguals' exposure to English supports the ability to recognize words in Welsh (Vihman et al., 2007; but see Marchman, Fernald, & Hurtado, 2010). Thus, in this case, Welsh-English bilinguals appeared to be at an advantage over their monolingual Welsh-learning peers.

Whether bilingual infants' wordform representations are as phonologically detailed as their monolingual peers has only recently begun to be investigated. Singh and Foong (2012) examined Mandarin-English learning bilingual infants' ability to recognize wordforms using the HPP. The authors found that at 7.5-months and 11 months, bilingual infants were sensitive to phonemic changes in Mandarin tone (meaningful changes in pitch). At 9-months-of age, however, bilinguals appeared to generalize words across differences in lexical tone. These data support the U-shaped developmental pattern observed in samples of Spanish-Catalan learning children (Bosch & Sebastián-Gallés, 2003), and also support the hypothesis that bilingual learning infants may develop a temporary tolerance for phonemic substitutions that could signal a change in a word's meaning. Evidence from word recognition studies is similarly suggestive of bilingual attentional indifference to phonemic mispronunciations.

Word Recognition

Studies investigating auditory recognition of canonically and mispronounced wordforms engage infants in tasks that do not involve word comprehension and therefore do not necessarily reflect the nature of children's long-term lexical representations (see

Swingley & Aslin, 2000). Several studies have however examined the nature of early word representations using methods that harness children's knowledge of word meanings. In the preferential looking task (also referred to as "visual choice", "language guided looking", "visual fixation" or "looking while listening"), word recognition is indicated by longer and/or quicker looking towards a named object (e.g., "Look at the dog!") compared to a simultaneously displayed distractor (e.g., a picture of a shoe) (Fernald, Pinto, Swingley, Weinberg, & McRoberts, 1998; Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987). Recent research using this paradigm has shown that monolingual English-learning infants know the meanings of several common nouns even at 6-months (Bergelson & Swingley, 2012). Furthermore, the burgeoning literature in this area is converging to suggest that early words are represented in phonological detail.

Although it seems to follow that perceptual attunement to the sound categories of one's language early in life would prepare young infants for learning words in their full form, until recently the prevailing view suggested that early lexical representations were vague or holistic. According to this view, since toddlers' vocabularies are relatively small, most phonetic detail is unnecessary in order to distinguish between word meanings (e.g., Charles-Luce & Luce, 1990). For example, if a child's vocabulary consists only of 'mom' and 'dad', a child might be capable of identifying these words despite only a rough lexical representation (e.g., 'm-' and 'd-'). As the lexicon grows, more detail is necessary in order to distinguish between words and thus the lexicon "restructures" over time (e.g., Metsala & Walley, 1998; Storkel, 2002). In the last decade however, a great deal of literature has begun to suggest that children's early lexical representations include substantial phonological information.

In order to test the amount of phonological detail in toddlers' lexical representations,

Swingley and Aslin (2000) showed 18-23 month olds two pictures side by side, one of which was either named correctly (e.g., ‘doll’) or incorrectly (e.g. ‘goll’). Toddlers recognized the target word regardless of the word’s pronunciation, as indicated by greater fixation on the corresponding object. However, mispronunciation significantly reduced word recognition, suggesting that the toddlers had previously encoded words in full phonological detail. These results have been replicated using vowel mispronunciations, consonant mispronunciations varying in word position, and with participants as young as 14-months old (Bailey & Plunkett, 2002; Ballem & Plunkett, 2005; Mani, Coleman, & Plunkett, 2008; Mani & Plunkett, 2007; Ramon-Casas, Swingley, Sebastián-Gallés, & Bosch, 2009; Swingley, 2003, 2009; Swingley & Aslin, 2002; White & Morgan, 2008). Although general task performance (recognizing familiar words) improves with age, the mispronunciation effect has been shown to be stable across all ages tested to date and is not related to vocabulary size or the length of time since acquiring the word.

The mispronunciation effect found in word recognition studies has not however held across bilingual samples. Ramon-Casas and colleagues (2009) found that unlike their Catalan-learning monolingual peers, Spanish-Catalan bilingual toddlers did not detect a mispronunciation in the preferential looking task when the vowel change corresponded to the Catalan-only /e/ - /ɛ/ contrast. In this case, infants looked towards a target object equally whether it was pronounced correctly or incorrectly. The bilingual toddlers’ performance therefore mimicked a group of Spanish-learning monolinguals, for whom the contrast is not meaningfully distinctive. This was not due to a general inability to hear the distinction between these sounds, since discrimination “failures” in infancy had recovered by 12-months, and these toddlers were between 18 and 24 months. Nor, did this represent a failure to notice any vowel substitutions since mispronunciations consisting of contrasts common to

both Spanish and Catalan (i.e., /e/ - /i/ or /e/ - /a/) did hinder word recognition. Furthermore, Ramon-Casas and Bosch (2010) found that a similarly aged group of Spanish-Catalan bilinguals, were capable of perceiving the /e/- /ε/ when the target words were not cognates (words that sound similarly and share a common meaning). Taken together these results suggest that the amount or degree of phonological overlap (at the phonemic level and also the lexical level) between languages influences bilingual infants' reactions to mispronunciations.

In the last several sections it has been demonstrated that bilinguals sometimes appear not to detect, or perhaps even to ignore auditory changes that could indicate a change in word meaning. On the other hand, monolinguals rather consistently detect such changes. In the following sections, it will be shown that monolinguals sometimes overextend this sensitivity to non-meaningful variability, while bilinguals may generalize more easily across non-meaningful differences in wordforms.

Coping with Wordform Variability

Recognizing meaningful distinctions (i.e., phoneme substitutions) between wordforms is important to later word learning. However, natural language environments contain a wealth of variability that is irrelevant to word meaning. This includes speaker specific information (e.g., gender), subphonemic detail (subtle differences in the realization of phonemes, such as pronouncing an /i/ sound with the tongue positioned higher or lower in the mouth in English), and suprasegmental information (variation across segments larger than a single phoneme, such as pitch in English).

Beyond learning the phonemes of one's native language(s), infants must learn which broader linguistic components are meaningful and which can be ignored. For example, in English, pitch is not meaningfully contrastive at the lexical level. While high-low pitch

alternations within English words would sound odd to a native adult listener, such variability would not necessarily impair word recognition. On the other hand, lexical stress (which includes a combination of pitch, syllable duration and amplitude) does sometimes indicate a change in meaning (e.g., ‘insight’ and ‘incite’ differ only in placement of stress on the first or second syllable) and substituting one word for the other may lead to confusion without contextual information.

Notwithstanding the impressive ability to recognize familiar wordforms within the first several months of life, experimental evidence suggests that early wordforms are encoded in such phonetic detail as to limit generalizability across non-meaningful differences in form. When 7.5-month-old monolinguals are first presented with target words from a single speaker, they do not subsequently recognize these words in the fluent speech of an opposite gendered speaker (Houston & Jusczyk, 2000). Similarly, they do not generalize words across variations in pitch (Singh, White, & Morgan, 2008), or affective states (i.e., happy versus neutral; Singh, Morgan, & White, 2004). These sensitivities to non-meaningful variation appear to attenuate in samples of monolingual infants by 9-months for pitch, and by 10.5 months for gender and affect, suggesting that infants quickly learn to ignore irrelevant acoustic details.

Although bilingual research in this area is nearly non-existent, one study using the same method demonstrated that English-Mandarin learning bilinguals generalize across English wordforms varying in pitch at 9-months and 11-months, but not at 7.5 months (Singh & Foong, 2012). This appears to mirror previous monolingual data, however it should be noted that at 9-months, but not 7.5 or 11 months, these infants also generalized across Mandarin wordforms that contrasted phonemically in tone (for which the primary perceptual correlate is pitch; Burnham & Mattock, 2007; Sing & Foong, 2012). This leaves

open the possibility that some degree of overlap in an English-Mandarin learning environment may result in a temporary indifference to changes in pitch. This could be the case where infants sometimes hear words that sound similarly and mean the same thing across languages (i.e., cognates), but where tone is only meaningful for one of the languages. For example in English “mama” is frequently used as a diminutive meaning “mother” regardless of the pitch in which it is spoken; on the other hand the tone in which “ma” is pronounced in Mandarin makes the difference between referring to one’s mother (mā; high flat tone) or one’s horse (mǎ; falling rising tone).

Although the number of cognates between English and Mandarin have not yet been reported, there are likely to be few since English and Mandarin belong to different language families (Indo-European vs. Sino-Tibetan respectively) that evolved geographically distant from one another. Despite the fact that the number of cognates between languages can increase as contact between linguistic groups increases (Bouckaert, Lemey, Greenhill, Alekseyenko, Drummond, Gray, Suchard, & Atkinson, 2012) previous research has shown that that cognate frequency is reduced in less closely related languages than more closely related languages (Schepens, Dijkstra, Grootjen & van Heuven, 2013). On the other hand, another source of overlap that may confuse the meaningfulness of pitch would come from accented speech. For example if a native-English speaking parent is less adept at producing correct Mandarin tones; or alternatively, if a Mandarin speaker applies native-language tones to English words for which tone is not meaningful. In the next section, the qualities of accented speech are described in detail along with data suggesting that accent is a particularly problematic source of variability for monolingual infants.

Accented Speech as a Source of Variability

Defining accent.

Accented speech is speech that systematically deviates from some “standard” and can include phonemic, subphonemic and suprasegmental variation (Cristia et al., 2012; Wells, 1982). At the phonemic level, accented speakers may add, substitute, or delete phonemes in accordance with their native phonology (e.g., since /ð/ does not exist in French, a French-accented speaker may substitute /d/ for /ð/ resulting in “dis” instead of “this”). In contrast, subphonemic variation consists of pronunciation of a single phoneme that falls outside of the normal range, as compared to a standard, but that does not cross a phonemic boundary (e.g., /ɪ/ can be pronounced with the tongue higher or lower in the mouth without changing the meaning of a word). Finally, suprasegmental variation refers to variability across a segment larger than a single phoneme. This can include both meaningful and non-meaningful variation. For example, placement of lexical stress can be meaningful in English (e.g., ‘that’s perFECT’ versus ‘to PERfect’). English speakers tend to place lexical stress on the first syllable of a word, whereas French speakers tend to place stress on the final syllable of a sentence. Applying either of these “rules” to another language can result in non-native sounding speech.

Two types of accent are distinguished from one another in the psycholinguistic literature: regional or sociolectal (“within-language”; sometimes referred to as “dialectal”) accents, and non-native (“foreign”) accents (Cristia et al., 2012). Regional accents refer to the acoustic properties of one’s native-language speech derived from the speaker’s geographic or social positioning. The term regional accent is sometimes used interchangeably with “dialectal accent”. However, some dialects are not mutually

understandable, whereas two people speaking different regional varieties of a language will usually be able to communicate with each other (e.g., a conversation between an East-coast Canadian and a West-coast Canadian). A foreign accent is quite different in that it involves the application of one's native phonology to a second language. For example, the phonotactic rules of Spanish do not allow for word-initial /s/. Therefore, Spanish-accented English speakers often add /e/, or epenthesize, at the beginning of words beginning with /s/ (e.g., 'study' becomes 'estudy').

Determining what constitutes the "standard" is an issue when defining accented speech. One way to define a standard is by taking the position of the listener (Cristia et al., 2012). In this sense, the speaker can be described as speaking with an accent if his or her speech deviates systematically from the listener's speech. For example, a Canadian English speaker in Canada would be considered as speaking with an accent to a British English speaking peer and vice versa. This definition can become complicated however with preverbal listeners, particularly those who routinely hear speech that differs from the community standard (e.g., a child with two British English speaking parents who is being raised in Canada) or who are learning two languages simultaneously (e.g., a child who hears both French and English from a native-French speaker).

The question to some extent therefore becomes "what kind of speaker will this child be?" It has been suggested that children possess a "native-accent filter", allowing them to hear accented speech but produce speech in line with language community norms (Chambers, 2002). This is consistent with evidence that 20-month-old toddlers tend to tune their perceptual sensitivities to that of the local language community more so than to the accented productions of their parent(s), even when their parents are the primary source of total language exposure (Flocchia, Delle Luche, Durrant, Butler, & Goslin, 2012). Still, other

research has suggested that children hearing accented speech adopt at least some of the linguistic properties of their parents' speech (Evans, Mistry, & Moreiras, 2007). Due to the uncertainty of the child's linguistic outcomes, and the heterogeneity of listening environments in both monolingual and bilingual homes, for the purpose of this thesis standard speech is defined by the linguistic characteristics common to the listener's native language community. To clarify, if, for example, a listener routinely hears both French-accented and native-English, then "standard English" is native-English with the linguistic characteristics of the region within which the listener resides (i.e., Eastern Ontarian English), and does not necessarily refer to the linguistic varieties heard in the home-learning environment. It is important to note that the purpose of this distinction is only to provide a common term for discussing "standard" speech as opposed to "accented" speech, and does not necessarily reflect what the infant is accustomed to hearing.

Coping with accented speech.

Perhaps because accented speech is characterized by such a wide range of phonetic deviations, working across varying speech segments, accent is accommodated later than other forms of variability (i.e., pitch, affect, and gender). In a HPP, Schmale and Seidl (2009) first familiarized 9 and 13-month-old monolinguals with two wordforms produced by either a Spanish-accented or native-accented ("standard") English speaker. At test, infants heard some passages that contained familiarized wordforms, and others that did not. The passages were presented in the opposite manner as during the familiarization phase, so that infants who were familiarized with Spanish-accented tokens were tested on non-accented passages and vice versa. Nine-month-olds in this study did not demonstrate a preference for passages containing familiarized over non-familiarized tokens, however 13-month-olds preferred passages containing familiarized words. This suggested that at 9-months-old,

infants were incapable of generalizing wordforms across speakers with different accents, whereas by 13-months infants initial wordform representations were sufficiently flexible to allow for subtle differences in form. In keeping with the language of this thesis, this meant that 13-month-olds could ignore irrelevant acoustic details to recognize familiar words across speakers with different accents.

Difficulties demonstrated by younger infants were not, however, due to a dislike for accented speech (as supported by Kitamura, Panneton, & Best, in press; but see Kitamura, Panneton, Deihl, & Nottley, 2006) or a general difficulty segmenting speech in a foreign accent since 9-month-olds successfully extracted familiarized wordforms across two Spanish-accented speakers (Schmale & Seidl, 2009). This result only held however when 9-month-olds heard two Spanish-accented speakers with perceptually similar voices (as rated by adult listeners). When 9-month-olds heard two Spanish-accented speakers with dissimilar voices, their performance was again comparable to chance. This contrasts with infants' behaviour when hearing two dissimilar native speakers; in this case 9-month-olds again succeeded. Therefore it appears that infants' difficulties lie in generalizing across highly variable wordforms, such as when there is a change in speaker *and* accent. Interestingly, particularly for bilingual infants, this might be exactly the scenario that is most frequently encountered in children's natural learning environment and therefore the most ecologically valid condition.

Other research has pointed to the same developmental pattern in generalizing wordforms across speakers with differing regional accents. Schmale, Cristia, and Johnson (2012), examined the possibility that English-learning 9-month-olds inability to accommodate for Spanish-accented speech may have been due to the "substantial" differences between native- and Spanish-accented English. They noted that Spanish-accented

and native-English differ at both the subphonemic and suprasegmental levels and that generalizing wordforms across these two linguistic varieties might be particularly challenging in early infancy. In a follow-up study, they therefore tested infants' performance in generalizing across dialectal varieties of English that contrasted only on vowel implementation. Using the same paradigm, Schmale and colleagues (2010) found that 12-month-olds, but not 9-month-olds, could generalize familiarized wordforms across North Midland American and Southern Canadian varieties of English. This suggested that the subtler accent variation across two English dialects was still challenging for the younger group of infants. Although both the foreign-accent and dialectal-accent studies point to a developmental transition towards more flexible wordform representations, experimental paradigms introducing infants to novel words only allow inferences regarding how wordforms are encoded and recognized across a very short period of time. On the other hand, high-frequency word preference paradigms have shed light on the level of rigidity with which wordforms are represented in long-term memory.

As described earlier, infants prefer listening to words that are frequent rather than infrequent in their native language, as demonstrated by longer headturns towards a speaker playing the former rather than the latter word type (Halle & de Boysson-Bardies, 1996). Using a variation of this paradigm that measured preference by associating wordlists with a centrally presented visual stimulus, Best and colleagues (2009) examined American English learning 15- and 19-month-olds' preferences for frequent versus infrequent words presented in toddlers' native-variety of American English or Jamaican-accented English. Both age groups preferred frequent over infrequent words when presented in their native-variety. However, only 19-month-olds maintained this preference when listening to Jamaican-accented English productions. The authors therefore concluded that “phonological

constancy”, the ability to ignore non-meaningful phonetic variation, emerges over toddlerhood.

Together both the infant and toddler data suggest that early wordform representations are phonetically detailed, and that over time young language learners begin to ignore irrelevant acoustic information to recognize wordforms across non-meaningful phonetic variations. A limitation of both sets of studies is that they do not necessarily inform as to how or when infants come to accept multiple pronunciations for words, that is, wordform-referent pairings. However, as in the phonemic mispronunciation studies mentioned earlier, recent research tying wordforms to their associated meanings parallels evidence from tasks that do not require word-object associations. In a preferential looking paradigm, 15-month-old toddlers looked longer to a picture of a named object over a distracter object when the auditory stimuli were produced in their native-accent (Australian English) but not when they were produced in an unfamiliar accent (Jamaican English) which differed in vowel implementation at the subphonemic level (Mulak, Best, Tyler, Kitamura, & Irwin, *in press*). By 19-months however, toddlers identified known objects by their labels regardless of pronunciation, suggesting long-term word representations that are sufficiently flexible to allow for accent variability. Older toddlers have likewise demonstrated the ability to generalize a newly learned word across speakers with different accents.

Schmale and colleagues (2011) first taught 24 and 30-month-old toddlers a novel word-object pairing (e.g., “feem” + green object) produced by either a native-English speaker or a Spanish-accented English speaker. Following the training phase toddlers were presented with two test trials whereby they saw an image of two objects side by side: the trained object (e.g., green object) and a new object (e.g., orange object) that had been seen before but never named. During one test trial toddlers heard the trained object labeled (e.g.,

“look at the *feem!*”) and during the other test trial toddlers heard the novel object named (e.g., “look at the *choon!*”). Toddlers who heard native-English tokens during the training phase heard Spanish-accented English test trials and vice versa. Successful accommodation of the accent was indicated by longer looks towards the trained or novel object when respectively labeled (see mutual exclusivity; e.g., Markman & Watchtel, 1988). The data suggested that 30-month-olds accommodated for the change in accent regardless of whether the training phase consisted of native-accented or Spanish-accented productions. In contrast, 24-month-olds accommodated for the accent only if they first heard Spanish-accented training tokens. The authors interpreted these findings as indicating that brief exposure to a foreign accent may be sufficient to enable generalization across differing pronunciations of a newly learned word. Schmale and colleagues (2011) suggest that exposure to greater phonetic variability during the familiarization phase, may promote broader lexical categories and therefore more robust word representations. To elucidate this hypothesis, the authors draw from previous research showing that 1) foreign-accented speech, particularly in terms of vowel-space utilization, is more variable than in standard speech (Jongman & Wade, 2007 as cited in Schmale et al., 2011) and 2) acoustically variable word-learning conditions appear to draw infants’ attention towards relevant phonological dimensions, specifically phonemes (Rost & McMurray, 2009). The authors therefore suggest that more variable (i.e., foreign-accented) learning conditions provide toddlers with a means of determining which acoustic properties can be ignored; those aspects that are highly variable (e.g., highly varied vowels) are not attended to during the testing phase.

This analysis has since been supported by research directly investigating the role of exposure in accent generalization (Schmale, Cristia, & Seidl, 2012). Twenty-four-month-olds who first heard two minutes of short stories in Spanish-accented English without visual

referents subsequently generalized a newly learned word across native-English and Spanish-accented English. They did not however succeed in generalizing across accents if they first heard stories produced by a native-English speaker, which supports the conclusion that foreign-accent exposure was critical in facilitating performance. Similarly, White and Aslin (2011) showed that 19-month-olds who were exposed to familiar word-referent pairings where one vowel was consistently mispronounced (e.g., [a] was pronounced as [æ]: ‘dog’ -> ‘dag’), subsequently tolerated these mispronunciations in a word recognition paradigm. Toddlers who were exposed to standard pronunciations or other vowel shifts prior to test did not succeed in the task. This suggests that experience with a novel pronunciation only affects perception of that specific pronunciation; it does not result in a general tolerance for novel pronunciations. As such, mispronunciation tolerance is dependent upon specific phonetic input. The results of several studies, therefore, indicate that experience with unfamiliar accents facilitates and in some cases results in generalizing wordforms and words across variable pronunciations at an earlier age.

Bilingual infants constitute a linguistic group that likely tends to have greater experience with accented speech than same-aged monolinguals. As a result, bilingual infants, like toddlers exposed to accented speech in the studies illustrated above, could be especially proficient in accommodating for linguistic variability imposed by accent. The following section summarizes the bilingual literature elaborated in this paper to provide a full argument as to why bilingual infants might be advantaged in generalizing wordforms across speakers with different accents. This discussion will extend beyond exposure to accent itself to argue that the foundation for accent accommodation is laid through exposure to two complete phonological systems.

The Case for a Bilingual Advantage in Accent Accommodation

The studies in this thesis extend the work of Schmale and Seidl (2009) to a bilingual population, with the expectation that bilingual infants will succeed in recognizing wordforms across speakers with different accents earlier than their monolingual peers. In earlier accent accommodation studies using the HPP, 9-month-old monolingual infants did not recognize wordforms across speakers with different accents (Schmale & Seidl, 2009; Schmale et al., 2010). Wordforms between the familiarization phase and testing phase varied substantially enough at the subphonemic and suprasegmental levels, such that the typical pattern of preference for passages containing familiarized over passages containing non-familiarized words was not observed. In essence, the continuity of attention to the familiarization words was broken in the test phase when the target wordforms no longer constituted a “good enough” match. Thus, 9-month-old monolinguals attended to the subphonemic and suprasegmental variation imposed by accent and were intolerant of these inconsistencies.

In contrast, several avenues of research with bilingual infants at approximately the same age, and using the same procedure, have demonstrated a tolerance for even phonemic contrasts. For example, Spanish-Catalan learning 8-month-olds do not demonstrate a preference for /e/ sounds after substantial exposure to /ε/ or vice versa, even though the contrast is meaningful in Catalan (Bosch & Sebastián-Gallés, 2003a). Similarly, English-Mandarin learning 9-month olds accommodate for both non-phonemic variation of English pitch, and phonemic variation of Mandarin lexical tone across wordforms (Singh & Foong, 2012). Taken together, these studies suggest that despite being capable of distinguishing between even close phonetic contrasts (Albareda-Castellot et al., 2011; Burns et al., 2007; Sundara & Scutellaro, 2011), bilingual infants may not break an attentional pattern between familiarization and testing phases (e.g., by looking longer towards a novel stimulus during

the testing phase) when the acoustic change is minimal. This pattern of attentional indifference has been mirrored with bilinguals in word recognition paradigms (Ramon-Casas et al., 2009; but see Ramon-Casas & Bosch, 2010).

Although accent variation can cause phonemic changes in words, previous studies (as well as the studies in this thesis) have controlled for phonemic variability. As such the variability imposed by accent in these studies may have been even subtler than in the phoneme discrimination, wordform and word recognition tasks described above. If bilingual infants routinely ignore phonemic substitutions in wordforms/words, then they should be more tolerant variation that could not signal a meaningful change across acoustic forms. That is to say that bilingual infants may ignore subphonemic and suprasegmental variation and consequentially sustain a preferential response for familiarized words despite a change in speaker and accent. If this is the case, then what specific experience or experiences lead bilingual infants to tolerate wordform variability? Four potential candidates, which are by no means mutually exclusive, are described here in turn: experience with two distinct phonological systems, exposure to accented speech, the presence of cognates and loanwords between languages, and potential cognitive advantages afforded to bilingual learners.

Acquiring two phonological systems.

No two languages are phonologically identical. Even those languages that overlap substantially in phonetic categories have language-specific idiosyncrasies in the exact pronunciation of those sounds. For example [ɪ] occurs in both Canadian English (CE) and Canadian French (CF), however F1 (an acoustic measurement inversely related to vowel height) values for [ɪ] productions are higher for monolingual CE speakers than for monolingual CF speakers (MacLeod, Stoel-Gammon, & Wassink, 2009). A young CE-CF bilingual therefore immediately encounters an environment that consists of a higher degree

of phonetic/phonemic variability than what might be experienced in a monolingual environment. A broader range of [ɪ] sounds encountered in the learning environment may dampen a bilingual's sensitivity to subtle changes of this sound in a discrimination task (but see Sundara, Polka, & Molnar, 2008 for evidence of bilingual infants' discrimination of cross-language allophonic variants), or when encountering accented speech. This could be particularly true when the accented speech mirrors the variability of phones heard in one's environment (e.g., a native-French speaker who applies his or her native variety of [ɪ] to an English word).

Furthermore, phonetic categories across languages may differ in status as meaningful units (phonemes). Continuing with the French-English example: /i/ and /ɪ/ are phonemically contrastive in CE, whereas [i] and [ɪ] are allophones (speech sound variations) of the phoneme /i/ in CF. While previous research has shown that monolingual infants are sensitive to phonemic variation, by 11-months, attention to native-language allophones decreases (Seidl, Cristia, Bernard, & Onishi, 2009). It is unknown whether the phonemic status of the 4 sounds representing the [i] to [ɪ] vowel height continuum across CE and CF influences the relative frequency of (or variability surrounding) those sounds. However, a parallel can be tentatively drawn between the /i/ - /ɪ/ distinction in CE but not CF and the /e/ - /ɛ/ distinction that exists in Catalan but not Spanish. It is possible for example that hearing a distribution of sounds that includes the English /i/, the French /i/ and its allophone [ɪ], and the English /ɪ/ leads to a broad phonemic category includes each sound as an allophonic variant. In such case French-English bilingual infants would not be expected to discriminate between any combination of these sounds. Although possible, this seems unlikely given data suggesting that Spanish-Catalan bilingual infants can hear the difference between /e/ and /ɛ/ despite not demonstrating discriminatory *behaviour* in a HPP. The position presented in this paper

therefore is that bilinguals hear, but are attentionally indifferent to such contrasts, consequentially allowing a broad (but input-specific) range of variability in the pronunciation of wordforms. In summary, the presence of subphonemic variation, and distributions of phonetic and phonemic speech sounds that overlap imperfectly across two languages may allow bilingual infants to generalize across differently accented wordforms.

Exposure to accented speech.

In proportion to overall speech exposure, young monolingual infants may rarely be presented wordforms across two different accents. In contrast, bilingual infants may encounter this precise situation frequently in their natural learning environment. This is particularly true in two-parent situations where each parent is a native speaker of a different language (e.g., a native-English speaking father and a native-French speaking mother), but where they speak a common language in each other's presence (e.g., French). In such a case where one speaker is speaking with a non-native accent while the other speaks with a native accent, their bilingual infant would hear variable productions of the same wordforms. Furthermore, some data has suggested that even perfectly bilingual adults produce some language sounds in an intermediary way between their two languages (MacLeod et al., 2009). As such, even in an environment with perfectly bilingual caregivers, bilinguals may frequently hear wordforms with greater subphonemic variability than their monolingual peers. Although this thesis is the first to test bilingual infants' accommodation of accented wordforms, related research with monolingual toddlers has demonstrated that even brief exposure to a foreign-accent can facilitate word recognition across accents (Schmale et al., 2012). Exposure in this study consisted of Spanish-accented stories presented auditorily in the absence of visual cues; thus demonstrating that accent exposure can function to facilitate accommodation even in the absence of word-referent pairings. This implies a bottom-up

learning process, where infants learn to estimate the degree of allowable variability surrounding wordforms without making a word-referent connection. Therefore, exposure to variable wordform productions may lead to wordform representations that are sufficiently flexible to tolerate the accent to which they are exposed.

Although infants need not, as demonstrated in the study described above, process accent at a lexical level in order to become accent-tolerant, it is possible that exposure to accented speech also functions in this way. Hearing words (wordform-referent pairings) pronounced in a variety of ways may lead to a top-down understanding that a certain amount of variability can surround a word without changing its meaning. For example, an infant who dad refer to “ball” in a native-English accent, and mom refer to “ball” in French-accented English, may come to learn that words can sound differently and yet mean the same thing. Even in the absence of accented speech natural languages frequently co-evolve to include words that sound similarly.

Cognates and loanwords.

Many bilingual infants are exposed to cognates and loanwords across their two languages. Cognates are words across two languages that have a common linguistic origin (i.e., descended from the same predecessor language). This common origin leads to similar forms and usually to a similar meaning (e.g., ‘night’ in English and ‘natt’ in Norwegian), although not always (e.g., ‘dish’ in English, and ‘tisch’, which means table, in German). Loanwords occur when one language incorporates a word from another for the same concept (e.g., ‘club’; /klʌb/ in English versus /klœb/ in French). The loanword must adhere to the phonology of the recipient language and thus often changes its form, as in the vowel example above. Focusing on loanwords and on cognates that share meaning, if bilinguals are routinely exposed to two similar wordforms for the same concept, then they may be more

willing than monolinguals to accept minor phonetic/phonological changes to novel words as acceptable instances of those words. As with accented speech however, it is nonetheless possible that the presence of cognates and loanwords, simply by virtue of offering variety in the production of similar wordforms, influence perception at the wordform rather than the word level as well.

Cognitive advantages.

In the child and adult literature, bilinguals often outperform monolinguals in tasks that measure cognitive control. This includes conflict resolution (Costa, Hernandez, & Sebastián-Gallés, 2008), perspective taking (Kovacs, 2009), and attentional control as evidenced by the inhibition of habitual responses (Bialystok, 1999). Recent evidence has shown that even preverbal bilingual infants are advantaged in a task requiring cognitive control.

In a series of eye-tracking studies, Kovacs and Mehler (2009) compared monolingual and bilingual 7-month-olds' ability to inhibit a previously learned stimulus-reward association in order to learn and respond to a new combination. Infants were trained to look towards either the left or right side of a screen following the presentation of either a visual or speech cue in order to see a visually stimulating puppet. Both groups learned to look towards the correct location in anticipation of a reward. However, only the bilingual group succeeded in redirecting their anticipatory looks when the stimulus-reward contingency changed (i.e., the cue signaled the appearance of the puppet on the opposite side of the screen). The authors concluded that experience processing and representing two languages is sufficient to enhance cognitive control even prior to the onset of speech. Thus the infant data and studies of older children and adults converge to suggest a bilingual advantage in cognitive control.

Bilinguals' superior performance in such tasks has been attributed to their experience inhibiting one language while producing (or perhaps perceiving, in the case of infants) another language (Green, 1998; Meuter & Allport, 1999). It has been suggested however that the bilingual advantage is notably "confined to tasks that contain a salient but irrelevant cue" (Bialystok, Craik, Grady, Chau, Ishii, Gunjhi, & Pantev, 2005). If Bialystok et al.'s (2005) observation that bilinguals' success is predominantly limited to ignoring irrelevant cues also applies to infant speech perception, then this may have consequences in accent accommodation. Specifically, bilingual infants may be better prepared to ignore irrelevant phonetic information in order to match wordforms across speakers with different accents. Before turning to an overview of the studies presented in this thesis, it is important to note that the capacity to ignore irrelevant information does not necessarily imply an understanding of what is relevant versus irrelevant in a cognizant sense. This thesis takes the position presented in Rost and McMurray (2009), that variability in the speech input draws infants' attention towards acoustic contrasts that are meaningful because they tend to be those that are more reliably produced (low variability; e.g. phonemic contrast). On the other hand, over time, highly variable properties in the speech input (e.g., individual voices) are ignored. By this logic, exposure to acoustic forms containing a range of variability (such as in a bilingual learning environment) should allow infants to hone in on relevant features while ignoring those that are irrelevant (Rost & McMurray, 2009).

Several candidates have been proposed here as potential contributors to a bilingual advantage in accommodating for accented speech. The potential effects of these candidates may be somewhat difficult to disentangle. For instance, hearing two phonological systems, hearing accented speech, and hearing cognates would each produce a degree phonological overlap in the learning environment. Further, bilinguals may simultaneously benefit from

general cognitive advantages and hearing several sources of phonetic variability. The approach taken in the studies that follow was not to argue for a single candidate but rather to narrow the possible explanations for an advantage over two studies.

Overview of Studies and Hypotheses

The literature review provided herein has described some of the processes involved in early language acquisition including the development of phonetic and phonological categories, learning to recognize familiar wordforms/words and learning to tolerate non-meaningful variation of these wordforms/words. The development of these skills at times converges between monolinguals and bilinguals, however frequent patterns of divergence have also been described. Monolingual infants attune to native-language sound categories in a fairly linear fashion: maintaining or improving discriminatory ability of native-language contrasts, while attenuating sensitivity to non-native contrasts. They also appear sensitive to deviations of familiar wordforms and words along the phonemic dimension. This sensitivity is however sometimes overextended so that monolingual infants have periods of difficulty recognizing wordforms across different speakers, and especially across speakers with different accents.

Patterns of phonetic/phonemic discrimination in bilingual infants are sometimes consistent with monolingual data, however bilingual infants occasionally display a divergent U-shaped pattern of discriminatory behaviour across infancy; particularly for contrasts that are relatively close in perceptual space or where there is imperfect overlap between sound categories across languages. While recent evidence suggests that bilingual infants can discriminate between even close phonetic contrasts, they may not react to phonetic changes in the same way as their monolingual peers (i.e., by demonstrating a preference for a novel stimulus). Several possible contributors to this lack of behavioural response have been

described above. The argument put forth here therefore is that tolerance for phonetic changes may allow bilingual infants to succeed in wordform recognition tasks where monolinguals fail, namely in accommodating for changes in accent and speaker.

In this thesis, monolingual and bilingual 9- and 13-month-old infants were compared on their ability to generalize wordforms across speakers with different accents. Study 1 “Pairing and Comparing” set the experimental stage by identifying native and non-native accented speakers with similar voices, as perceived by a group of adults. This was done in order to minimize voice variability (e.g., biological variability imposed by, for example, the shape of the oral cavity, length of vocal chords etc.) in experimental stimuli so as to focus on accent as a primary source of variability in the infant experiments that followed (Schmale & Seidl, 2009). Furthermore, several acoustic measurements of vowels and word stress placement were taken and compared across native and non-native speakers in order to confirm expected deviations between standard and non-standard speech, and to describe phonetic deviations that may hinder infants’ capacity to recognize wordforms across speakers.

Study 2 “Bilingual Infants’ Accommodation of Phonetically Familiar Accented Speech” first tested the hypothesis that bilingual infants are advantaged in accommodating wordform variability imposed by speaker and accent. To do so, 9- and 13-month-old English-learning monolinguals and French-English learning bilinguals were tested on their ability to recognize familiarized words across a change in speaker and accent. As in Schmale and Seidl (2009) and Schmale and colleagues (2010) infants were tested in a Headturn Preference Procedure. In this case, infants were familiarized with wordform tokens in either native-Canadian English or Canadian French-accented English, and then tested for their preference between English passages containing familiarized or non-familiarized words in

the opposite accent. Given that monolingual infants typically fail in this task at 9-months, but succeed at 13-months, the prediction was that bilingual infants would succeed at generalizing across accents at 9-months whereas their monolingual peers would not succeed until 13-months. This prediction was partially confirmed; bilingual infants succeeded at both ages, however surprisingly monolinguals failed even at 13-months.

Given bilingual infants' success, two further hypotheses concerning the source of bilinguals' success were explored. Namely, that exposure to accented speech and/or infants' levels of exposure to their native languages (e.g., English-dominant, French-dominant, or relatively equal) would be related to performance in this task. In light of previous findings that exposure to an artificial "accent" produces an input-dependent pattern of accommodation (White & Aslin, 2011), where only specific vowel deviations were tolerated post-exposure, it was predicted that exposure to a French-accent in particular would facilitate success in this experiment. According to two different measures of accent however, only a very small number of infants reportedly heard French-accented English in their learning environment. Therefore hearing a French-accent *specifically* did not appear necessary to facilitate success in the task.

On the other hand, when infants who heard English-accented French were also considered, the total number of infants hearing some form of an accent was approximately 50%. The relationship between accent exposure (English-accented French or French-accented English) and task performance was not, however, significant despite testing the possibility of a non-linear relationship. In contrast, a nearly significant trend towards a relationship between bilinguals' exposure to English (inversely related to their exposure to French) and performance in the task was discovered. This relationship was linear, suggesting that greater exposure to English actually hindered performance in the task.

Two potential explanations for this relationship were identified. First, it is possible that phonetic distributions across English and French are such that overlap effects are most pronounced for infants who are hearing predominantly French. Infants in this study were tasked with generalizing wordforms across a native speaker of English and a French-accented speaker of English. Since the French-accented speaker tended to impose French-phonological characteristics onto her English productions, this might be particularly conducive to French-accent accommodation in French-dominant bilinguals. An alternative explanation however is that French-dominant bilinguals treated the task of wordform recognition differently than those with greater exposure to English. Given that the experimental stimuli consisted of English wordforms and passages, it is possible that French-dominant bilinguals were less restrictive in the amount of variability that they would allow before identifying a potential change in meaning simply due to inexperience with the English language.

In Study 3 “Bilingual Infants’ Accommodation of Phonetically Unfamiliar Accented Speech”, 9-month-old French-English bilinguals were tested on their ability to accommodate for Mandarin-Accented Canadian English. The hypothesis for this study was that bilinguals’ ability to accommodate for a change in speaker and accent in Study 2 was related to the distribution of language sounds across their two languages, and not simply owing to inexperience with English. The prediction therefore was that French-English bilinguals would not accommodate for an accent derived from an unfamiliar (Mandarin) phonology imposed upon a familiar language (English). The data from this study supported this hypothesis since bilingual infants failed to generalize wordforms across native-English and Mandarin-accented speech. However, a surprising inverted-U shaped relationship was discovered between task success and the percentage of time infants were exposed to one of

their two languages (English). These data suggested that bilingual infants who were less dominant in either language, and received more equally balanced exposure to their two languages were the most successful in accommodating for a novel accent. These results are explained by drawing from other domains of infant bilingualism research. In particular it is suggested that bilingual infants may harbour more flexible representations of events stored in memory, allowing them to generalize across conflicting cues (in this case across differing accents; Brito & Barr, 2012).

Evidence from the studies presented in this thesis support the overarching hypothesis that bilingual infants tolerate subtle variations of sounds. Experience with input that matches the characteristics of the imposing accent appears to be important in this task. Namely, experience with English and French facilitates generalization across native- and French-accented English but is not sufficient to allow infants as a group to accommodate across native- and Mandarin-accented English. On the other hand, degree of bilingualism may also be important in relaxing wordform representations more generally (i.e., to accommodate for novel phonetic deviations in foreign accented speech), as evidenced by the fact that infants who were more equally exposed to their two languages tended to perform best in generalizing wordforms across native- and Mandarin-accented English, despite the overall group failure. The following chapters present detailed descriptions of the methods and findings of the studies described herein, followed by a general discussion and presentation of limitations and areas for future research.

Chapter 2

Study 1: Pairing and Comparing

Introduction

The purpose of this Chapter is twofold; first, to pair accented and non-accented speakers who have perceptually similar voices (discussed later), and second, to describe the instances in which accented-speakers deviate from standard speech. Earlier, accent was defined as speech that systematically deviates from some “standard” and including phonemic, subphonemic and suprasegmental variation (Cristia et al., 2012; Wells, 1982). Two common ways of measuring a speaker’s deviation from “standard speech” have been used widely in the linguistic literature. The first, involves adult perceptual ratings. Here, adults are exposed to productions of accented and non-accented speech and asked to rate how accented, or perhaps how foreign, the speaker sounds (e.g., Magen, 1998; Flege, 1984; 1988). This type of measure is useful for determining degree of accent (e.g., heavy versus light), and of accent detection (e.g., accented or not), but it does not reveal exactly how a speaker deviates from standard speech.

The current study draws from the work of Schmale and Seidl (2009) and uses acoustic analysis of at the segmental and suprasegmental levels to compare accented and non-accented speakers. Using this method it is possible to more accurately describe where the “accent” lies in stimuli that will later be used to test infants’ accommodation of accented speech. Furthermore, previous work has suggested that adult second language learners perceive non-native sounds as belonging to native sound categories and therefore produce them as such (Flege, 1981). Here, native-speakers’ productions are compared to accented-speakers’ productions. Where deviations are found, we then turn to the accented-speakers’ native phonologies as described in the linguistic literature, in order to support specific

deviations as the result of second-language learning, rather than for instance an idiosyncratic pronunciation irrespective of status as a second-language learner.

In order to minimize the variability attributable to voice, previous studies examining infants' ability to generalize across speakers with different accents speakers based on the similarity of their voices (Schmale & Seidl, 2009; Schmale et al., 2010). So that it would be possible to compare the results of the current experiments with previous work, the same method for pairing native and non-native speakers was followed here. This method is described in detail below. Following selection of speaker pairs (native-English – Mandarin-accented; native-English – French-accented), acoustic analyses were conducted in order to describe subphonemic and suprasegmental differences between native and non-native speakers. Although it is possible to use perceptual ratings to determine whether an individual speaks with an accent in comparison to a standard, precise acoustic measurements allow a detailed description of differences imposed by accented speech.

Pairing Speakers

Target wordforms.

As in Schmale and Seidl (2009), the stimuli developed in this study for use in Study 2 and 3 of this thesis were adapted from the stimuli used in Jusczyk, Houston, and Newsome (1999). The original research used the four target words: “hamlet”, “doctor”, “kingdom” and “candle”. Since the goal of this study was to examine infants' accommodation of subphonemic and suprasegmental variation imposed by accented speech, and not detection of phonemic mispronunciations, it was imperative that the French and Mandarin-accented speakers' productions of the target words did not cross any phonemic boundaries (e.g., “**this**” pronounced as “**dis**” would indicate a potential change in meaning and therefore would not be an acceptable target word). The original words were pilot tested with two

native French speakers and two native Mandarin speakers. A native speaker of English asked non-native speakers to repeat the original target words several times. While French speakers' productions fell within English phonemic boundaries, it was discovered that the variations imposed by the Mandarin accent could sometimes be perceived as phonemic differences in English.

Three alternative wordforms, “ticket”, “matches”, and “kinship” were selected to replace “hamlet”, “doctor” and “kingdom” following the same pilot procedure. In selecting these alternative words it was possible to preserve some of the same vowels used in the original target words (/æ/, /ə/, and /i/), however the /ə/ of “doctor” was consistently pronounced as an /o/, which is phonemic in English (as well as in French) and therefore was not included in the current stimuli.

Acoustic recordings.

Once target words were selected, 5 native speakers of Canadian English, 5 native speakers of Canadian French, and 5 native-Mandarin speakers were recorded as potential sources of experimental stimuli. The non-native speakers self-identified as speaking English with an accent that reflected their native language. All speakers were adult females. The speakers were asked to read passages and word lists in an animated voice as if reading to a young child. They were also told that some of the passages might seem odd to them, but to read the passages exactly as they appeared on the page. Speakers read the four target passages (Appendix A) as well as four filler passages meant to familiarize the speaker with the task. The speakers were allowed to rehearse as many times as they wished. All recordings took place in a laboratory room with a Behringer B-1 condenser microphone via an Alesis io|14 audio interface into an Apple MacBook Pro. The raw passages and lists of target words produced by the speakers were exported to PRAAT as 24-bit WAV files

without any processing. The files were then normalized within PRAAT to approximately 70 dB.

Speaker similarity ratings.

Schmale and Seidl (2009) argue that since it is not known which acoustic dimensions influence speaker recognition, perceptual similarity ratings, rather than acoustic measurements, should form the basis for matching speaker pairs. The exact same method used in Schmale and Seidl (2009) was used here to pair speakers based on voice similarity. Twenty undergraduate students (16 female), who received partial credit for participation, were recruited to rate the perceptual similarity of the 15 speakers with the intention of pairing one native-English speaker and one French-accented speaker as well as one native-English speaker and one Mandarin-accented speaker.

Testing took place in a well-lit laboratory room. Participants were seated at a computer desk; the distance between the participant and the computer monitor was approximately 60cm. Participants heard audio stimuli originating from the computer over BOSE QC3 headphones. Tokens, rather than entire passages were presented in order to reduce the length of the experiment, which took approximately 20 minutes. Previous studies have also used tokens, rather than passages to conduct stimuli pairing (Schmale & Seidl, 2009).

All possible pairings of the 15 speakers saying two words in isolation (“matches” and “ticket”) were presented to the participants in both natural and sinewave speech.³ Sinewave

³ Consistent with Schmale and Seidl (2009), a single isolated token of each word was drawn from word lists and not passages. This controlled for co-articulatory information (e.g., the influence of the final sound from one word on the pronunciation of the first sound in the next word) that varies throughout passages. In contrast to the acoustic measurements displayed later, it is not possible to present adult participants with the range of co-articulatory information from the 6 tokens in passages since it would create an unduly long experiment. It is possible however to measure, for example, a range of VOTs for /k/ in candle throughout passages.

speech “preserves much of the complex temporal and prosodic information found in speech within a similar range of frequencies without retaining the distinctly biological quality of true speech” (Krentz & Corina, 2008, pp. 2). To elaborate, vocal characteristics, including indexical information such as age and gender, or biological characteristics such as the length of the speaker’s vocal chords or shape of the oral cavity, are eliminated in sinewave speech. This creates a soundwave that, for lack of better terminology, sounds “tinny” and computerized. However, sinewave speech artificially preserves the accented qualities of speech, specifically prosodic and temporal information (Krentz & Corina, 2008; Schmale & Seidl, 2009). Therefore, by subtracting the perceptual similarity ratings of sinewave speech pairings from those of natural speech pairings it is possible to measure voice similarity while reducing the confounding factor of accent (Krentz & Corina, 2008; Schmale & Seidl, 2009). All pairings were separated by 500 milliseconds of silence. The pairings of speakers, words, and speech type (natural or artificial) were quasi-randomly presented by an experimental program designed in our lab using PRAAT 4.4.30 (Boersma & Weenik, 2001), to present the stimuli in the same way as was done in Schmale and Seidl (2009). Using this program, participants rated each speaker pairing on a scale from 1 to 5 (1 meaning very dissimilar and 5 meaning very similar).

Analysis of speaker similarity ratings.

Once the similarity ratings were collected, the perceived similarity of speaker pairs for “matches” versus “ticket” tokens were compared. To do this, first the total similarity scores for each of matches, ticket and their sinewave counterparts were calculated for each participant and then averaged across participants. Paired sample t-tests revealed that similarity scores for “matches” versus “ticket” were marginally significant ($p = .053$), and there was a significant difference between the similarity scores for the sinewave counterparts

($p = .049$). For both natural and sinewave speech, raters on average judged pairings of “matches” tokens as more similar than pairings of “ticket” tokens. However, since speaker similarity is to be judged by the relative difference between similarity ratings for tokens and their sinewave counterparts, of true interest was whether there would be a significant difference between the difference scores (natural speech – sinewave speech) for matches versus ticket. A paired t-test revealed that there was no difference between participants’ evaluations of similarity between natural and sinewave tokens for matches and ticket ($p = .957$). Subsequently, similarity ratings for matches and ticket tokens were collapsed for further analyses.

The next step was to prepare the data for use in a multidimensional scaling procedure (MDS) by converting similarity ratings collected in PRAAT to dissimilarity ratings. To do so, each rating was subtracted from a number higher than the highest possible score (the highest possible score was 5, therefore each rating was subtracted from 6). Dissimilarity ratings are preferred over similarity ratings because their relationship to distances is direct and positive (Giguère, 2007). That is to say, the higher the dissimilarity rating the greater the perceived psychological distance.

The dissimilarity ratings for each speaker pair were then averaged across participants, and placed in a matrix in SPSS for MDS analysis (ALSCAL). MDS detects significant underlying dimensions from a set of variables (e.g., speakers) by analyzing a matrix of perceived dissimilarities and transforming them to represent distances (Schmale & Seidl, 2009). These distances are then reported as mean squared differences (average dissimilarity) of each speaker pair. First entered were the average dissimilarity ratings for English and French natural speech pairs and then for English and French sinewave pairs, resulting in two proximity matrices (Table 1).

To determine the most similar speaker pair, the mean squared differences from the sinewave speech are subtracted from the mean squared differences from the natural speech. Remember that the sinewave speech represents the isolated accent features. So, even if there appears to be a large distance between two speakers in natural speech, this difference could potentially be accounted for by the accent itself. As such, a speaker pair is deemed highly similar if there is a small change between the mean squared distance of a speaker pair in natural and sinewave speech. On the other hand, a speaker pair is deemed highly different if there is a relatively large change between the mean squared distance of the pair in natural and sinewave speech. In this case, the smallest change between natural and sinewave ratings was between E4 (English Speaker #1) and F2 (French Speaker #2); $1.69 - 1.83 = -.14$. In contrast, the difference for the E2-F3 pairing was 5.47, suggesting that these speakers' voices are very dissimilar. Thus, in Study 1 speaker E4 and F2 were paired. To avoid confusion in the following analyses E4 will be referred to as Native-English Speaker 1 (NE1) and F2 will be referred to as French-Accented Speaker (FA).

Table 1 Mean squared distances (squared Euclidean distances) of all speaker pairs in sinewave and natural speech.

Proximity Matrix: Natural Speech										
	E1	E2	E3	E4	E5	F1	F2	F3	F4	F5
E1	.00									
E2	1.62	.00								
E3	3.17	.66	.00							
E4	.28	.69	1.57	.00						
E5	.03	1.24	2.66	.15	.00					
F1	12.04	6.89	3.32	8.77	11.17	.00				
F2	2.23	3.48	2.74	1.69	2.18	6.86	.00			
F3	8.14	6.28	3.20	5.96	7.68	1.74	2.52	.00		
F4	2.40	6.24	6.51	2.91	2.68	13.34	1.08	6.49	.00	
F5	1.81	.98	3.23	1.61	1.55	12.93	6.56	12.03	8.30	.00
Proximity Matrix: Sinewave Speech										
	ES1	ES2	ES3	ES4	ES5	FS1	FS2	FS3	FS4	FS5
ES1	.00									
ES2	.09	.00								
ES3	3.68	2.77	.00							
ES4	2.60	1.99	4.16	.00						
ES5	.39	.23	3.43	.98	.00					
FS1	8.20	7.88	4.59	15.19	10.53	.00				
FS2	4.92	4.77	11.12	1.83	2.98	24.70	.00			
FS3	7.92	6.56	.81	7.60	7.44	5.16	16.80	.00		
FS4	1.37	.81	.60	2.19	1.18	5.87	7.13	2.79	.00	
FS5	9.39	7.73	4.15	3.38	6.61	17.30	8.92	4.47	4.72	.00

Note: E = English, F = French, # = Participant number.

Using the exact same method, proximity matrices were derived for English and Mandarin speakers (see Table 2). After subtracting the mean squared differences of sinewave speech from the mean squared distances of natural speech, speakers E1 and M3 were found to be the most similar pair (-.20), and were thus paired for the following

experiments. Henceforth E1 will be referred to as Native-English Speaker 2 (NE2) and M3 will be referred to as Mandarin-Accented Speaker (MA).

Table 2 *Mean squared distances (squared Euclidean distances) of all English-Mandarin speaker pairs in sinewave and natural speech*

Proximity Matrix: Natural Speech										
	E1	E2	E3	E4	E5	M1	M2	M3	M4	M5
E1	.00									
E2	.35	.00								
E3	1.74	.55	.00							
E4	.56	.34	1.24	.00						
E5	.04	.16	1.29	.34	.00					
M1	7.34	8.94	11.03	11.80	8.12	.00				
M2	11.22	8.75	5.89	12.16	10.61	12.04	.00			
M3	.66	.29	.63	1.26	.53	6.60	6.44	.00		
M4	5.35	5.78	6.50	8.64	5.71	1.02	6.12	3.60	.00	
M5	.16	.80	2.44	1.32	.34	5.41	11.10	.78	4.02	.00
Proximity Matrix: Sinewave Speech										
	ES1	ES2	ES3	ES4	ES5	M1	M2	M3	M4	M5
ES1	.00									
ES2	3.40	.00								
ES3	2.85	2.66	.00							
ES4	.49	4.21	5.42	.00						
ES5	.29	1.90	2.85	.54	.00					
MS1	12.35	4.37	4.56	15.71	10.44	.00				
MS2	10.85	10.35	3.04	15.86	11.59	6.19	.00			
MS3	.86	2.82	.63	2.58	1.17	7.99	5.64	.00		
MS4	2.95	.11	1.70	4.13	1.74	3.88	8.41	2.01	.00	
MS5	1.86	.25	1.71	2.76	.89	5.41	9.04	1.45	.14	.00

Note: E = English, M = Mandarin, # = Participant number

Comparing Speakers

Speaker characteristics.

The final pairings consisted of two monolingual native-English speakers and two accented-English speakers. The French-accented speaker (FA) was exposed to French from birth and attended a French language primary school where she began learning English as a second language. The Mandarin-accented speaker (MA) was exposed to Mandarin from birth and began learning English in high school.

Stimuli Lengths.

In the HPP, which is described in detail in Chapter 3, infants are first exposed to repetitions of target words during the familiarization phase and then tested for recognition of these target words in passages. As described above, repetitions of target words are presented as “word lists” where speakers were asked to repeat a target word several times, whereas in passages target words are embedded in sentences presented in a story-like format. The average length of target word lists and passages for each speaker are displayed below in Table 3.

Table 3 *Length of audio stimuli in seconds for each speaker*

Speaker	Avg. length of Word List	Range for Word List	Avg. length of Passages	Range for Passages
NE1	14.97	12.90-17.74	17.02	17.34-19.89
FA	14.81	13.16-17.31	19.37	19.32-21.12
NE2	16.19	15.31-17.31	16.23	14.78-17.41
MA	14.85	13.01-17.08	17.46	16.91-21.63

Subphonemic comparisons.

Close phonetic transcriptions

A linguist with training in phonetic description provided the close transcriptions in Table 4. Notably the native English speakers and Mandarin-accented speaker nasalized [æ] in “candle” and the stressed [ɪ] in “kinship” whereas the French-accented speaker did not. Nasalization in English is context dependent, where vowels are nasalized before nasal consonants and are otherwise oral (Malécot, 1960). In contrast, nasalization is phonemic in French indicating a change in meaning across otherwise identical words (e.g., ‘bas’ [bæ] vs. ‘banc’ [bæ̃] is the difference between ‘low’ and ‘bench’). While one might expect that this would make native-French speakers attentive to and successful in producing English nasalization of vowels, this was not the case in the current data. Possibly, the degree to which native-French speakers nasalize vowels to indicate a change in meaning is greater than the phonologically contextual nasalization of English vowels. If this is the case, and English nasalization is not as salient, then native-French speakers may not notice when English sounds are nasalized and this may be reflected in French-accented products of English. It is also possible that French-accented speakers are aware of the greater degree of nasality in their native-language, and are hyper-correcting. Notably, vowel nasalization is also non-phonemic in Mandarin, and the Mandarin speaker’s productions more closely mirrored the English speaker’s in terms of vowel nasalization. Therefore, it is possible that native-French speakers are more restrictive in nasalizing vowels.

Both French and Mandarin-accented speakers also failed to aspirate the word-initial [t] in “ticket” and the Mandarin speaker did not aspirate the word-initial [k] in “candle”. Furthermore, accented speakers did not release word-final stop-consonants. There were also differences between the native and accented productions of the word-final [l] in “candle”.

Native English speakers and the Mandarin-accented speaker produced this as a velarized (“dark”) [ɫ], whereas the French-accented speaker produced this as a non-velarized (“light”) [l]. Furthermore, while the native-speaker treated [l] as syllabic, neither of the accented speakers did so. Interestingly, the Mandarin-accented speaker inserted a [ə] vowel before the [l]. This is likely due to the monosyllabic structure of Mandarin, where each syllable of the language generally takes a CV form and constitutes a standalone word. Therefore, the Mandarin speaker inserts a vowel prior to the [l] in order to maintain consistency with her native phonology. It should be noted however, that the addition of this phoneme (which is perceptually similar to the English schwa [ə]) should be assimilated into the syllabic [l] for native English speakers (see “perceptual magnets”, Kuhl, 1991). Indeed it has been argued that syllabic [l] is actually the combination of [ə]+[l] (e.g., Oda, 2007; Wells, 1995) and is sometimes transcribed as such.

Perceived stress patterns between native and Mandarin-accented English were similar, however the French-accented speaker did not place perceptible stress on the first syllable of “kinship” or “matches. As will be described further below, this is consistent with French but not English phonology.

Table 4 *Close phonetic transcriptions of isolated target words in native and accented-English*

Word	Canadian English	Ontarian French	Mandarin Chinese
Candle	[^h kʰæ̃n·,dɫ̥]	[^h kʰæ̃n·,dl]	[^h kʰæ̃n·,døɫ]
Kinship	[^h kʰĩn·,ʃɪ̃p]	[^h kʰɪn·,ʃɪ̃p ^ˈ]	[^h kʰĩ·,ʃɪ̃p ^ˈ]
Matches	[^h mæ·,tʃəs]	[^h mæ·,tʃəs]	[^h mæ·,tʃəs]
Ticket	[^h tɪ·,kɪ̃t]	[tɪ·kɪt ^ˈ]	[tɪ·kɪt ^ˈ]

Note:

^ˈ = no audible release of phoneme

ɪ̃ = tilde under [ɪ] indicates a “creaky voice” where vocal chords are very tense

ĩ = tilde above [ɪ] or [æ] indicates nasalization

ɫ = velarized (“dark”) [ɪ]

ə = central open-mid vowel

^h = aspiration

^ˈ = primary stress

^ˌ = secondary stress

Acoustic measurements.

In order to describe the subphonemic differences between native and non-native speech, acoustic measurements were gathered for each speaker. Speaker pairs were compared on measurements of duration, F0 (pitch), F1 (height), and F2 (front-backness), for each vowel contained in the test stimuli (as in Schmale & Seidl, 2009). For F0 a higher number means higher pitch, for F1 a higher number indicates that the vowel was produced with the tongue positioned lower in the mouth, and for F2, a higher number indicates articulation closer to the front of the mouth.

Since the intention was to demonstrate differences between single speakers, using a repeated measures design was not possible since it would require groupings of more than $n = 1$ for each speaker type. As such, GLM statistics were run treating multiple tokens spoken by single speakers as independent. However, since this violates the assumption of independence of errors, given that all tokens are coming from the same speaker alpha was set at .01 to

correct for multiple exemplars from single speakers, in addition to using Scheffé's correction for family-wise error. Table 5 summarizes the key vowel findings. Importantly, where accent is driving differences it was expected that the two native-English speakers (NE1 and NE2) would produce similar measurements whereas they would differ significantly from accented speakers. Each native-speaker's measurements were therefore compared with each other as well as with each of the accented speakers. Although an ideal sample would perhaps consist of many accented versus non-accented speakers, if both native speakers differ in the same way from an accented speaker, then this lends support to the conclusion that such a difference is related to the non-native speaker's accent. Such a pattern does not preclude the possibility that an accented speaker's measure reflects a personal, idiosyncratic, pronunciation rather than the imposition of one's native phonology onto English productions. However, when such patterns are consistent with previous phonological research, this can be taken as strong support for an accent-based deviation from standard productions.

Table 5 Mean difference of acoustic measures for speaking pairs; standard deviations in *italics*

	F0 (Hz)				F1 (Hz)			
	Stressed		Unstressed		Stressed		Unstressed	
Speakers	/æ/	/ɪ/	/ɪ/	/ə/	/æ/	/ɪ/	/ɪ/	/ə/
NE2-NE1	-3.11 <i>24.37</i>	-35.60 <i>19.16</i>	-4.71 <i>30.93</i>	-7.38 <i>56.14</i>	34.93 <i>46.31</i>	38.42 <i>33.81</i>	-66.04 <i>22.25</i>	1.40 <i>89.48</i>
NE1-FA	12.16 <i>24.37</i>	5.97 <i>19.16</i>	-23.25 <i>30.93</i>	-43.21 <i>56.14</i>	-195.94* <i>46.31</i>	102.81 <i>33.81</i>	82.14* <i>22.25</i>	-7.49 <i>89.48</i>
NE2-FA	9.05 <i>24.37</i>	-29.63 <i>19.16</i>	-32.56 <i>30.93</i>	-50.59 <i>56.14</i>	-161.01 <i>46.31</i>	141.23* <i>33.81</i>	16.09 <i>22.25</i>	-6.10 <i>89.48</i>
NE2-MA	-62.99 <i>24.37</i>	-62.00 <i>19.16</i>	-53.24 <i>30.93</i>	-4.74 <i>56.14</i>	-142.26 <i>46.31</i>	14.26 <i>33.81</i>	2.51 <i>22.25</i>	-74.50* <i>89.48</i>
NE1-MA	-59.87 <i>24.37</i>	-26.41 <i>19.16</i>	-48.53 <i>30.93</i>	2.64 <i>56.14</i>	-177.19* <i>46.31</i>	-26.41 <i>33.81</i>	66.55 <i>22.25</i>	-75.90* <i>89.48</i>
	F2 (Hz)				Duration (ms)			
NE2-NE1	257.42 <i>93.66</i>	-89.74 <i>63.41</i>	36.98 <i>61.13</i>	-126.60 <i>89.48</i>	-3.17 <i>6.46</i>	32.29* <i>6.06</i>	-1.37 <i>9.31</i>	-27.99 <i>19.99</i>
NE1-FA	-31.75 <i>93.66</i>	-81.88 <i>63.41</i>	-63.02 <i>61.13</i>	86.84 <i>89.48</i>	11.71 <i>6.46</i>	-12.85 <i>6.06</i>	-23.25 <i>9.31</i>	17.39 <i>19.99</i>
NE2-FA	225.67 <i>93.66</i>	171.63 <i>63.41</i>	-26.03 <i>61.13</i>	-39.76 <i>89.48</i>	8.54 <i>6.46</i>	19.44 <i>6.06</i>	-24.62 <i>9.31</i>	-10.60 <i>19.99</i>
NE2-MA	-136.44 <i>93.66</i>	-138.68 <i>63.41</i>	-82.17 <i>61.13</i>	-239.84 <i>89.48</i>	24.30* <i>6.46</i>	-23.10* <i>6.06</i>	-11.04 <i>9.31</i>	-17.31 <i>19.99</i>
NE1-MA	-393.86* <i>93.66</i>	-48.93 <i>63.41</i>	-119.15 <i>61.13</i>	-113.24 <i>89.48</i>	27.47* <i>6.46</i>	-55.38* <i>6.06</i>	-9.69 <i>9.31</i>	10.68 <i>19.99</i>

* $p < .01$

For /æ/, 12 tokens taken from passages containing the word “matches” and “candle” (6 tokens of each word) were measured for each speaker. The overall Wilks’ Lambda was significant ($F(12, 109) = 7.80; p < .001; \eta_p^2 = .42$), which indicated that there were speaker differences in the production of this vowel. Univariate tests revealed significant differences for duration ($F(3, 44) = 7.25; p < .001; \eta_p^2 = .33$), F1 ($F(3, 44) = 9.13; p < .001; \eta_p^2 = .38$) and F2 ($F(3, 44) = 8.04; p < .001; \eta_p^2 = .35$).

In terms of duration, there was a significant difference between NE2 and MA’s productions of /æ/ ($p = .006$). Vowel duration was shorter in this case for the Mandarin-accented speaker. This was supported by the same pattern of difference for NE1 and MA ($p = .002$). There were no significant differences between the English speakers or between English and French speakers. In terms of F1, there was a significant difference between the NE1 and FA pairing ($p = .002$) but not between the NE2 and MA pairing ($p = .034$). There was a nearly significant difference between NE2 and FA ($p = .013$), which supports the NE1 and FA distinction on F1 for this vowel. The English speakers in this case produced smaller F1 values than MA or FA. Finally, for F2 measurements there were no differences found between the main pairings of speakers. Interestingly, however, there was a significant difference between NE1 and MA ($p = .002$) and FA and MA ($p = .005$). In this case, MA produced a greater average F2 measurement than the FA speaker and even greater still than NE1. Importantly, although non-significant, the English speakers did vary quite substantially from each other on this measure ($p = .07$).

For /ɪ/ (stressed), 12 tokens for each speaker were analyzed from the passages containing “kinship” and “ticket” (6 tokens of each word). The overall Wilks’ Lambda ($F(12, 109) = 9.17; p < .001; \eta_p^2 = .46$), indicated that there were speaker differences in the production of this vowel. Univariate tests revealed significant differences for duration ($F(3,$

44) = 31.73; $p < .001$; $\eta_p^2 = .68$), and F1 ($F(3, 44) = 7.13$; $p = .001$; $\eta_p^2 = .33$). Differences in F0 and F2 measurements were not significant ($ps > .01$).

There was a significant difference of duration between the NE2 and MA pairing ($p = .005$), but no difference between NE1 and FA. There was a significant difference between the two English speakers ($p < .001$). This is not surprising, since vowel length is not phonemic in English and therefore English speakers may apply variation on this dimension without signaling a change in meaning across words. Further, MA differed significantly from FA and NE1 on this measurement ($ps < .001$). This pattern of results indicates that MA produced longer durations of /ɪ/ stressed than NE2, who produced longer durations of this vowel than NE1. Finally, for F1 measurements, there was no significant difference between the two main speaker pairings of interest. There was however a significant difference between NE2 and FA ($p = .002$). Although non-significant, the general direction of this difference is supported by the comparison between NE1 and FA ($p = .037$). This therefore may suggest that the French-accented speaker is producing relatively low F1 levels on this vowel compared to the other speakers. Notably, there was a significant difference between FA and MA in the production of /ɪ/ stressed on F1 ($p = .006$). The MA speaker's value fell between that of NE1 and NE2.

For /ɪ/ (unstressed), 12 tokens for each speaker were analyzed from the passages containing “kinship” and “ticket” (6 tokens of each word). The overall Wilks' Lambda was significant ($F(12, 108) = 2.97$; $p = .001$; $\eta_p^2 = .22$). However, this was entirely carried by differences in F1 measurements ($F(3, 44) = 5.48$; $p = .003$; $\eta_p^2 = .27$). There was a significant difference in the NE1 and FA pairing ($p = .007$), but no significant difference between NE2 and MA or any other possible pairing. This difference reflects a greater F1 production for the FA speaker. Although there was no significant difference between NE2

and FA, it is worthwhile to note that NE1 and NE2 produced very similar tokens of this vowel on average in terms of F1 measurement ($p = .947$).

For /ə/ (unstressed), 6 tokens for each speaker were analyzed from the passages containing “matches”. The overall Wilks’ Lambda for speaker was significant ($F(12, 45) = 5.49$; $p < .001$; $\eta_p^2 = .55$). Once again however, this difference was entirely carried by differences in F1 measurements ($F(3, 20) = 7.40$; $p = .002$; $\eta_p^2 = .53$). There was a significant difference between the NE2 and MA pairing ($p = .009$), but no difference between NE1 and FA. There was similarly a difference between NE1 and MA ($p = .007$), and the comparison between FA and MA approached significance ($p = .017$). The pattern of results indicates that MA produced higher F1 measurements; NE1 and NE2 produced extremely similar values ($p = 1.00$).

Summary of vowel findings.

The key vowel findings in terms of stimuli pairings (NE1 and FA; NE2 and MA) are as follows. NE1 produced /ɪ/-stressed and /ɪ/-unstressed lower in the mouth than FA and produced /æ/ with the tongue positioned higher in the mouth. These findings are consistent with the theory that accent is driven by the application of first language phonology to a second language (e.g., Flege, 1980, 1981). In this case, the French-speaker appears to apply characteristics of Canadian-French phonology to English words. While Canadian-French does not contain a phonemic /æ/ it does contain /a/ (as does English) which is produced slightly lower in the mouth (MacLeod et al., 2009). Similarly, although Canadian-French does not contain a phonemic /ɪ/, it does contain a phonemic /i/. In Canadian-French [ɪ] is an allophonic variant of /i/, produced by laxing. Native speakers of Canadian-French produce [ɪ] higher in the mouth than native English speakers (MacLeod et al., 2009). Therefore it is possible that the French-accented speaker’s production of /ɪ/ was influenced by her native

phonology. The schwa /ə/ vowel was produced similarly across the French-accented and native-English speakers in this study.

None of the vowels examined here are meaningfully contrastive, or phonemic, in Mandarin, and therefore it was expected that there would be clear differences between native- and Mandarin-accented speakers' productions of all vowels. Indeed, previous studies have reported that native-Mandarin speakers produce /æ/ higher in the mouth and closer to the front of the mouth than native American English speakers (Chen, Robb, Gilbert, & Lerman, 2001). Consistent with these findings, MA produced /æ/ closer to the front of the mouth than NE1. Although the difference between MA and NE2 was not significant, it patterned in the expected direction. Unlike previous findings, MA produced /æ/ lower in the mouth than NE1. Once again, although not significant, this pattern was the same between MA and NE2. This difference could reflect an idiosyncratic difference for the particular MA speaker here, or could reflect differences between General American- and Canadian-English.

Similarly to differences in /æ/ production, Mandarin-accented speakers in previous work have produced /ɪ/ higher and closer to the front of the mouth (Chen et al., 2001). In the current study, MA generally produced /ɪ/ with the tongue closer to the front of the mouth than the native-English speakers however the only significant difference was between MA and NE1. MA also generally produced /ɪ/ higher in the mouth than the native-English speakers (with the exception of the opposite pattern between NE1 and MA on stressed-/ɪ/), although these differences were also non-significant. Although differences in /ə/ have not been evaluated elsewhere, in the current study NE2 articulated /ə/ lower in the mouth than MA. Furthermore, NE2 produced longer /æ/ vowels and shorter stressed /ɪ/ vowels than MA.

Summary of VOT findings.

In addition to vowel measurements, voice onset time (VOT) was measured for word initial /k/s (candle and kinship; 6 tokens of each) in sentences. A 4 (speaker) by 2 (token; candle or kinship) mixed measures GLM revealed a main effects of speaker ($F(3, 20) = 15.72$; $p < .001$; $\eta_p^2 = .70$) and token ($F(2, 40) = 9.17$; $p = .001$; $\eta_p^2 = .31$), but no interaction between speaker and token ($p = .09$). VOT measurements were longer for candle ($M = 73.72$; $SE = 2.68$) than for kinship ($M = 64.99$; $SE = 1.70$) across speakers. Importantly, paired post-hoc comparisons using Tukey's HSD revealed a significant difference between MA and NE2 ($p = .002$). Also of interest, MA differed significantly from NE1 and FA ($p < .001$ respectively), where MA displayed shorter VOT than the other speakers. Although Mandarin speakers generally produce longer VOT associated with a high degree of aspiration compared to English speakers, VOT for first-language Mandarin speakers producing English words has been shown to be similar to native speakers of British English (Chao & Chen, 2008; Rochet & Fei, 1991). It is possible that the Mandarin speaker here was hyper-correcting; in fact Chao and Chen (2008) also raised the possibility that L2 learners are aware of subtle differences in VOT and are capable of "correcting" their speech (2008).

Here, MA did not perceptibly aspirate the word initial /k/ in "candle". This is reflected in MA's shorter VOT overall since VOT is shorter for unaspirated stop consonants than for aspirated stops (Chao & Chen, 2008; Docherty, 1992; Lisker & Abramson, 1964). If MA was hyper-correcting to sound more "English-like" the result might possibly be the elimination of perceptible aspiration. Notably MAs VOT for /k/ in "candle" ranged from 45.43 to 58.25 with a mean = 51 milliseconds, which falls towards the low range of what would be expected of an aspirated /k/ produced by a native-English speaker. Previous research has suggested that the VOT range for native speakers of American English

producing an aspirated /k/ is 50-135 milliseconds averaging at 80 milliseconds; unaspirated /k/ the range has been shown at 0-35 milliseconds, averaging at 21 milliseconds (Lisker & Abramson, 1964). In contrast Mandarin speakers have been shown to average around 110 milliseconds for aspirated /k/. Taken together these results support the possibility of hyper-correction of VOT by MA. No other comparisons were significant.

Suprasegmental comparisons.

In addition to the acoustic measurements of VOT and vowel tokens, correlates of English stress were analyzed in order to describe suprasegmental differences between speakers. Previous studies using the headturn preference procedure have demonstrated that 7.5 month olds are sensitive to the correspondence of emphatic stress, stress intentionally placed by a speaker to highlight specific words, between familiarization and word recognition phases (Bortfeld & Morgan, 2010). Indeed, Bortfeld and Morgan (2010) have shown that although 7.5-month-olds generally prefer listening to emphatically stressed words, word recognition is facilitated when the degree of emphatic stress during familiarization is similar to the emphatic stress heard during word recognition. This evidence is particularly important to the current study; if word stress varies between accented and non-accented speech then this could make word recognition more difficult when infants are faced with one speech type in the familiarization phase and another during word recognition.

Since foreign accent is usually thought of as the application of native-language phonology to a non-native language, it seemed that this might be the case in the current sets of stimuli. English is a stress-timed language where word stress generally follows a trochaic pattern; that is, native speakers of English generally stress the first syllable in a given word (Cutler & Carter, 1987). On the other hand, French is a syllable-timed language, and native speakers of French generally place stress at the sentential rather than lexical level

(Abercrombie, 1967). Furthermore, although Mandarin speakers use duration, pitch and amplitude to distinguish between stressed and unstressed syllables similarly to English speakers, Mandarin speakers also tend to produce stressed syllables with a higher pitch than native English speakers (Zhang, Nissen, & Francis, 2009). Therefore, it was possible that English speakers could differ from their accented pairing in placement and/or magnitude of stress markers.

Several measures that correlate with English stress (duration, pitch, and amplitude) were analyzed. Duration (ms) was measured across the syllable, pitch (Hz) was measured as an average across the syllable, and amplitude (dB) was measured at the peak of each syllable (Table 6). Target pairings (candle and ticket; kinship and matches) were analyzed individually, and tokens presented in isolation and were analyzed separately from those presented in passages.

Table 6 *Acoustic measurements of stress correlates in target words*

Candle and Ticket								
Measures	NE1		FA		NE2		MA	
	Syllable 1	Syllable 2	Syllable 1	Syllable 2	Syllable 1	Syllable 2	Syllable 1	Syllable 2
Isolated Tokens								
Duration (msec)	212.6	249.9	197.2	215.5	231.7*	297.9	197.9*	241
Amplitude (dB)	77.5*	76.2	78.6*	74.5	78.7*	74.8	77.1	76.4
Pitch (Hz)	281.2*	235.2	281.5*	220.6	305*	229.3	283*	233.3
Passage Tokens								
Duration (msec)	251.6	194.4	244.7	228.7	245.6	178.4	243.4	210.2
Amplitude (dB)	73.3	72.5	74.2	73	77.2*	73.9	78.2	77.4
Pitch (Hz)	234.9	197.8	226.4	240.1	209.6*	184.2	266.6*	202
Kinship and Matches								
Measures	NE1		FA		NE2		MA	
	Syllable 1	Syllable 2	Syllable 1	Syllable 2	Syllable 1	Syllable 2	Syllable 1	Syllable 2
Isolated Tokens								
Duration (msec)	197.4*	364.2	211.7*	349.6	225.4*	414.6	198.7*	350.12
Amplitude (dB)	79.4*	73.6	84.4*	78.7	78.4*	74	76.8	76.7
Pitch (Hz)	238.9	252.9	227	255.7	298.4*	202.8	289.1*	214.6
Passage Tokens								
Duration (msec)	231.8*	303.1	257.9*	352.1	245.4	264.7	221.3	293.5
Amplitude (dB)	77.6*	74.2	81.1*	77	78.9	74	80.6	78.3
Pitch (Hz)	294.3	192	218.4	233.6	206.6*	170.5	264.2*	218.5

* Difference between Syllable 1 and 2 significant at $p < .01$ for this speaker.

Bold font indicates that magnitude of difference is significantly greater for this speaker than for their pair at $p < .01$.

GLMs were conducted for each speaker pairing (NE1 and FA; NE2 and MA); thus four analyses were conducted for each speaker pairing. Each GLM analyzed Syllable (1st vs. 2nd) as a repeated measure across measures of amplitude, duration, and pitch. Speaker Accent (native vs. accented) was treated as a between subjects factor. As in the preceding vowel measurements, the assumption of independence of errors was violated given that the measurements produced are from single speakers (native or accented). Since these results were not meant to generalize to the population (i.e., of native-English or accented-English speakers) the analyses were carried through in order to facilitate discussion of the stress markers within the experimental stimuli used in this thesis. It is unclear what type of adjustment can be made to correct for the increased probability of a Type I error; however an attempt to adjust for this error was made by setting a conservative main effect alpha of .01. Furthermore, univariate tests were assessed with alpha set at .01, which is lower than convention would suggest ($.05/3 \text{ DVS} = .016$).

For the NE1/FA pairing for “candle” and “ticket” tokens produced in isolation, there was no main effect of accent ($p = .29$), though there was a main effect of syllable ($F(3, 52) = 18.93$; $p < .001$; $\eta_p^2 = .52$). Speakers generally produced first syllables in a higher pitch ($F(1, 54) = 14.0$; $p < .001$; $\eta_p^2 = .21$), and with greater amplitude ($F(1, 54) = 38.14$; $p < .001$; $\eta_p^2 = .41$). However an interaction between syllable and accent ($F(3, 52) = 4.28$; $p = .009$; $\eta_p^2 = .20$) suggested that although both speakers produced louder first syllables, the magnitude of this effect was greater in FA ($t(27) = 6.34$; $p < .001$) than NE1 ($t(27) = 2.20$; $p = .037$). For tokens produced in passages there was no significant effect of accent, syllable, or interaction between syllable and speaker. Neither speaker produced a consistent stress cue across isolated tokens and passages, however the French-accented speaker placed a greater

magnitude of stress (amplitude) on the first syllable in isolated tokens compared to the native accented speaker.

For “kinship” and “matches” tokens produced in isolation there was no effect of accent ($p = .80$), or syllable by accent, however there was an effect of syllable ($F(3, 52) = 37.89$; $p < .001$; $\eta_p^2 = .69$). Speakers tended to produce louder first syllables ($F(1, 54) = 80.98$; $p < .001$; $\eta_p^2 = .60$), and longer second syllables ($F(1, 54) = 42.65$; $p < .001$; $\eta_p^2 = .44$). This pattern also occurred in the passages where there was a significant main effect of syllable ($F(3, 20) = 9.41$; $p < .001$; $\eta_p^2 = .59$). Speakers tended to produce louder first syllables ($F(1, 22) = 27.20$; $p < .001$; $\eta_p^2 = .56$) and longer second syllables ($F(1, 22) = 10.19$; $p = .004$; $\eta_p^2 = .32$).

For the NE2/MA pairing for “candle” and “ticket” tokens produced in isolation there was a main effect of speaker ($F(3, 52) = 10.98$; $p < .001$; $\eta_p^2 = .39$), syllable ($F(3, 52) = 70.18$; $p < .001$; $\eta_p^2 = .80$), and an interaction between syllable and accent ($F(3, 52) = 10.01$; $p < .001$; $\eta_p^2 = .37$). This speaker pairing produced longer second syllables ($F(1, 54) = 12.13$; $p < .001$; $\eta_p^2 = .18$). The speakers also produced first syllables in a higher pitch ($F(1, 54) = 100.94$; $p < .001$; $\eta_p^2 = .65$) and louder than second syllables ($F(1, 54) = 42.34$; $p < .001$; $\eta_p^2 = .44$). However, the interaction between syllable and accent for amplitude was significant ($F(1, 54) = 19.66$; $p < .001$; $\eta_p^2 = .27$), and follow up analyses suggested that while the native-accented speaker produced a louder first syllable relative to the second syllable ($t(28) = 6.78$; $p < .001$), the MA speaker did not ($t(28) = 1.75$; $p = .09$).

For the passages, the main effect of accent was not significant ($p = .18$), the main effect of syllable was significant ($F(3, 20) = 12.40$; $p < .001$; $\eta_p^2 = .65$), and interaction between these two variables was significant ($F(3, 20) = 6.42$; $p = .003$; $\eta_p^2 = .49$). The duration of syllable was not significant within passages. Pitch ($F(1, 22) = 31.97$; $p < .001$; η_p^2

= .59) and amplitude ($F(1, 22) = 8.20$; $p = .009$; $\eta_p^2 = .27$) however continued to be higher on first syllables. Although the omnibus test suggested a syllable by accent interaction, none of the univariate tests were significant with an alpha set at .01. The only consistent stress cue between isolated tokens and passages were amplitude and pitch. Notably however, the Mandarin-accented speaker did not produce significantly higher first syllable amplitude during isolated tokens. Therefore, it is possible that infants hearing Mandarin-accented “candle” and “ticket” tokens during familiarization may perform differently than those hearing familiarization and test phases that are more similar.

Finally, for “kinship” and “matches” tokens in isolation, there were effects of accent ($F(3, 52) = 4.84$; $p = .005$; $\eta_p^2 = .22$) and syllable ($F(3, 52) = 55.9$; $p < .001$; $\eta_p^2 = .76$) as well as an interaction between accent and syllable ($F(3, 52) = 8.15$; $p < .001$; $\eta_p^2 = .32$). First syllables were generally produced louder ($F(1, 54) = 26.56$; $p < .001$, $\eta_p^2 = .33$) and with a higher pitch ($F(1, 54) = 84.10$; $p < .001$; $\eta_p^2 = .60$). Second syllables were generally longer ($F(1, 54) = 91.46$; $p < .001$; $\eta_p^2 = .63$). As in the other tokens, the interaction between syllable and accent suggested that while the native-accented speaker produced louder first syllables ($t(27) = 6.97$; $p < .001$), the MA speaker did not ($t(27) = .17$; $p = .87$). Within the passages, only syllable was significant ($F(3, 20) = 20.41$; $p < .001$; $\eta_p^2 = .76$), with speakers producing louder ($F(1, 22) = 55.60$; $p < .001$; $\eta_p^2 = .72$) first syllables in higher pitch ($F(1, 22) = 15.5$; $p = .001$; $\eta_p^2 = .41$). This pattern mirrored those found in the “candle” and “ticket” tokens and suggested that the differential pattern of stress may make Mandarin accented isolated tokens more difficult to identify in passages.

Summary of word stress findings.

In general, accented speakers were more similar to their native-speaking counterparts than was predicted based on previous studies of word-stress placement and magnitude in

native French and Mandarin. Where some distinctions were made between native and accented speakers, these tended to involve the magnitude of a stress marker and not placement.

For the NE1/FA pairing, the French-accented speaker spoke first syllables of “candle” and “ticket” louder than the native speaker when the tokens were produced in isolation. Neither speaker produced consistent stress cues across isolated tokens (which are presented during the familiarization phase by one speaker) and passages (which are presented during the test phase by the opposite speaker). Consequentially, the most incongruent familiarization-test situation occurs when FA produces isolated familiarization tokens of “candle” and “ticket”, and NE1 produces the corresponding test passages. Similarly, for the NE2/MA pairing, the accented speaker did not produce first syllables significantly louder than second syllables in isolated tokens of either set of wordforms. This suggests that the most incongruent familiarization-test situation here will exist when the familiarization tokens consist of Mandarin-accented productions and passages consist of native-English productions.

With the stimuli and differences engendered by speaker accent have described in detail, it is possible to now turn to examinations of infants’ ability, or lack thereof, to accommodate these differences in wordform recognition tasks.

Chapter 3

Study 2: Bilingual Infants' Accommodation of Phonetically Familiar Accented Speech

Introduction

Previous research has indicated that monolingual infants have difficulty recognizing familiar wordforms across speakers with different accents (Schmale & Seidl, 2009; Schmale et al., 2011). The current study is the first to examine this phenomenon concurrently in monolingual and bilingual infants. Here, English learning monolingual and French-English learning bilingual 9- and 13-month-olds were tested on their ability to generalize wordforms, with which they have recently become familiar, across speakers with different accents. In this case, the “standard” speaker is a native speaker of Canadian English, whereas the accented speaker speaks English with a French-accent. Given the arguments presented in Chapter 1, it was predicted that only bilingual infants would succeed in recognizing familiar wordforms despite a change in speaker accent at 9-months, whereas both monolingual and bilingual infants would succeed at 13-months. This prediction was partially supported; bilinguals did succeed in the task at both ages, however monolinguals did not succeed at either age. The relationship between task success and language exposure is explored. The results presented in this chapter, in combination with parts of the introductory chapter, have been submitted for review in *Infancy*, a peer reviewed academic journal.

Methods

Apparatus.

Testing took place in a small dimly lit room. Each infant sat on a parent's lap facing a 68.6 cm video monitor. The distance between the infant and the monitor was 1.5 meters. Two additional 44.5 cm monitors (left and right) were mounted to the walls left and right of

the infant at a distance of 1.5 meters and a height of 1 meter from the floor. Each side monitor was accompanied by a loudspeaker that sat below the monitor on the floor. A digital video-camera was placed beneath the center monitor and covered by black cloth. The experimenter monitored infant looking times via a closed circuit television system from an adjacent room. Habit 2000, a computer program produced by the Infant Cognition Laboratory at the University of Texas at Austin, was used to order the presentation of stimuli and to collect looking time data. A computer key was pressed to begin each trial and to record looking times. Computer key presses signaled whether the infant was looking at the side monitor playing the visual stimulus. Audio stimuli were delivered at 65 dB, +/- 5 dB. As a masking control, parents wore BOSE QC3 headphones and listened to female vocal music for the duration of the experiment.

Design.

Infants were randomly assigned to hear one of two sets of target tokens (kinship/matches or candle/ticket) in one of two accent orders (accented speech: familiarization or test), which produced four total testing conditions (Table 7). In Condition 1, infants were familiarized with “candle” and “ticket” presented in native-English and were tested using French accented passages containing these words. In Condition 2, infants were familiarized with “kinship” and “matches” presented in native-English and were tested using French accented passages containing these words. In Condition 3, infants were familiarized with “candle” and “ticket” presented in French-accented English, and were tested using native-English passages containing these words. In Condition 4, infants were familiarized with “kinship” and “matches” presented in French-accented English, and were tested using native-English passages containing these words. Four 9-month-olds and three 13-month-olds

from each language group were assigned to each condition. Thus 14 participants were assigned to each condition.

Table 7 *Experimental design*

Condition	Target Tokens	Familiarization	Test Passages
1	Candle & Ticket	Native English	Accented English
2	Kinship & Matches	Native English	Accented English
3	Candle & Ticket	Accented English	Native English
4	Kinship & Matches	Accented English	Native English

Procedure.

After obtaining informed consent from a caregiver, each infant was tested in a variant of the headturn preference procedure (e.g., Jusczyk & Aslin, 1995; Kemler Nelson et al., 1995). Caregivers were instructed not to speak, point at any of the three screens, or make headturns towards the side screens. All familiarization and test trials began with an attention-getting animated red light presented on the center screen. When the infant fixated on the red light, the experimenter pressed a key that resulted in the light being extinguished and an animated green light being presented on either the left or right screen. Infants were allowed 3 seconds to make a 90 degree headturn towards the flashing light before the auditory stimulus began to play from the corresponding loudspeaker. If the infant did not make a headturn within this amount of time the attention-getter would play again on the center screen and the trial would repeat. The 3-second delay was programmed into Habit such that if the infant did look within the 3 seconds, the audio would only play after the entire 3 seconds had passed. Once the audio stimulus began, it played either until completion or until the infant looked at least 30 degrees away for at least two seconds. During the familiarization phase, the initial token (e.g., ‘ticket’ or ‘candle’) was random. After the initial trial, the two tokens repeated back and forth until the infant had accumulated 80 seconds of total looking

time. Eighty seconds (rather than the 60 seconds used in previous accent research) was used as the familiarization criterion in order to better equalize infants' exposure to both target words. The side of presentation (left/right) was random.

At test, infants heard four blocks of the four different passages for a total of 16 trials. Each block contained two passages with familiarized words (e.g., 'candle' and 'ticket') and two passages containing non-familiarized words (e.g., 'kinship' and 'matches'). The passages were presented in pseudo-random order so that each passage was only heard once during each block of trials. The initial display screen (left/right) was random and subsequent trials alternated back and forth between left and right. Each passage was played to completion.

Although test trials in the HPP are typically response contingent, playing passages to completion was intended to allow infants a better opportunity to recognize a wordform as familiar and consequentially orient towards the sound-source. Support for this variation of the HPP comes from (1) previous research showing that infants do listen and learn sound patterns from auditory presentations even when auditory presentations are not contingent on visual attention (e.g., Pelucchi et al., 2009), and (2) that infants look towards stimuli that are captivating due to either familiarity or novelty (e.g., Fennell et al., 2007; Hunter & Ames, 1988), in the absence of contingent responding. That is to say that regardless of whether infants are "controlling" the presentation of stimuli by looking towards or away from those stimuli, they can still demonstrate a novelty or familiarity preference.

Since auditory stimuli were not interrupted when infants looked away, infants could reorient towards familiarized words even if they only came to recognize them as familiar after several repetitions, or when they occurred in a particular sentential position (i.e., while listening and briefly looking away). This might be a more ecologically valid scenario since

non-native speakers typically do not cease to speak if not immediately understood. In fact, non-native speakers (like all speakers), use repair strategies that typically involve some form of repetition (Derwing & Rossiter, 2002; Reiger, 2003). Although the speakers in the current study did not receive feedback (e.g., being asked for clarification) regarding their pronunciations during recording, it is possible that their productions could become closer to standard pronunciations over several repetitions. It is also possible that greater exposure to the passages could allow infants to better accommodate for phonetic idiosyncrasies. Traditionally, infants hear a variety of target word exemplars during the familiarization phase, where tokens are presented repeatedly in isolation. However, if infants do not recognize target words within sentences early during a test trial (and consequently look away), the trial ends. As such they are precluded from hearing further exemplars. This could result in hearing an unreasonably restricted range of variability; certainly one that does not match the variability presented in the familiarization phase, and perhaps one that is not representative of natural variability in the face of non-native speech.

In order to be included in final sample, at least 12 trials needed to be successfully completed; only fully completed blocks were included in the analysis. Test trials for each participant were coded offline in a frame-by-frame analysis; 25% of participants' data were re-coded by a second highly trained observer. Both coders were blind to which test passages contained familiarized versus non-familiarized words. A preferential response consisted of longer average orientations to the display screen while hearing passages containing either familiarized or non-familiarized wordforms. All infant participants received a Language Development Laboratory t-shirt as a token of appreciation.

Questionnaire measures.

Bilingualism questionnaire.

Infants in this study were classified as monolingual if they were exposed to English at least 90% of the time since their birth, and bilingual if they were exposed to English *and* French 20-80% of the time since birth. These criteria were based on Pearson, Fernandez, Lewedag and Oller (1997), which recommends a minimum of 25-75% exposure to each language for inclusion in bilingualism studies. The current criterion is slightly more lenient in order to observe a broader range of behavioural variation. Bilinguals in the current study were simultaneous bilinguals (sometimes referred to as “crib” bilinguals), meaning that they were exposed to both languages since birth. Babies who were exposed to a language other than French or English more than 10% of the time were excluded.

To assess language exposure, the current study utilized the Language Exposure Questionnaire (Appendix B), used in previous research to classify bilingual participants (Bosch & Sebastián-Gallés, 1997; Fennell et al., 2007). Using this questionnaire, the experimenter asked the attending caregiver to estimate the number of hours their child was exposed to their native language(s) from each of his or her primary caregivers (e.g., parents, grandparents, childcare workers), and to estimate the percentage of time their child was exposed to his or her native language(s) in the community. The attending caregiver was also asked to describe the linguistic environment in which the child’s parent(s) were raised.

Accent exposure.

Accent exposure was measured in two ways. The first was by examining the linguistic environment in which the parent(s) were raised as reported by the Language Exposure Questionnaire. Parents who learned a language outside of the home environment (e.g., in school) were considered as accented. This measure will later be referred to as the

Language Questionnaire Measure (LQM) of accented speech. The second measure of accent was a subjective measure based on the perception of the attending caregiver, henceforth referred to as the Subjective Caregiver Report (SCR). The attending caregiver (almost always a participant's parent) was asked whether he or she considered themselves or their spouse (if applicable) as speaking with an accent in one or more of their spoken languages. For example, a parent might be asked both "Do you consider yourself as speaking with an English accent when you speak French?" and "Do you consider yourself as speaking with a French accent when you speak English?" For both the LQM and SCR, the percentage of time infants heard either French-accented or generally accented (e.g., English-accented French) was calculated as a proportion of all language exposure (i.e., number of accented hours/total number of hours).

Participants.

Fifty-six infants successfully completed this study. Of these, 28 were monolingual and 28 were bilingual. These groups were further divided into two age groups: 9 and 13-month olds. All infants were in apparent good health at the time of the study. All infants were born after at least 37 weeks gestation except for one bilingual 13-month-old who was born at 36 weeks. No infants had any reported developmental or physical challenges.

The 9-month-olds consisted of 16 monolinguals (12 females; age range = 8.48-10.07 months; $M = 9.27$ months; $SD = .44$) and 16 bilinguals (9 females; age range = 8.46-10.16 months; $M = 9.29$; $SD = .57$). There was no significant difference in the ages of these groups. On average, 9-month-old monolinguals were exposed to 97.76% English ($SD = 2.76$), 2.18% French ($SD = 2.84$), and .19% a language other than French or English ($SD = .75$); bilinguals were exposed to 49.21% English ($SD = 16.35$), 50.32% French ($SD = 16.26$), and .34% other ($SD = .38$). Twenty-two additional 9-month-olds were tested but were excluded from

analysis due to: physical impairment that precluded headturns in one direction (1 bilingual), fussiness/crying (15; 8 bilinguals), parental interference (1 bilingual), equipment failure (3; 1 bilingual), or experimenter error (2 monolinguals). There was no significant difference in the age of these two groups ($t(30) = .12$; $p = .90$), or in the number of completed blocks ($t(30) = .80$; $p = .43$). Eleven monolinguals and 13 bilinguals completed all 4 blocks of test trials.

The 13-month-olds consisted of 12 monolinguals (3 females; age range = 12.85-14.61; $M = 13.61$; $SD = .51$) and 12 bilinguals (2 females; age range = 11.62-14.16; $M = 13.02$; $SD = .68$). The bilingual group was on average .58 months younger than the monolingual group, and this was significant ($t(22) = 2.36$; $p = .028$; Cohen's $d = .98$). On average, 13-month-old monolinguals were exposed to 98.14% English ($SD = 2.20$), 1.29% French ($SD = 1.60$), and .56% a language other than French or English ($SD = 1.95$); bilinguals were exposed to 48% English ($SD = 21.74$), 51% French ($SD = 21.66$), and .50 Other ($SD = 1.24$). An additional twenty 13-month-olds were tested but were excluded from the analysis due to: fussiness/crying (16; 10 bilinguals), parental interference (2; 1 bilingual), illness (1 monolingual) or experimenter error (1 monolingual). All 12 monolinguals and 11 bilinguals completed all 4 blocks of test trials. There was no significant difference in the number of blocks completed between groups. Although there were no differences between language groups in the number of completed trials, there was a difference between age groups ($t(54) = -2.14$; $p = .04$; Cohen's $d = .63$); overall, 9-month-olds ($M = 3.75$; $SD = .44$) completed fewer blocks than 13-month-olds ($M = 3.96$; $SD = .20$).

Results

Familiarized vs. non-familiarized passages.

Data cleaning and screening.

For the purpose of data screening, a single “familiarity” variable was calculated by subtracting looking time (in seconds averaged over the number of trials) to passages containing non-familiarized wordforms from looking time to passages containing familiarized wordforms. Positive difference scores indicate greater looking towards passages containing familiarized target wordforms over passages containing non-familiarized counterparts while negative scores indicate the reverse. Using this difference score the data were screened for normality and the presence of outliers. Screening was conducted by cell for each language group x age group combination. For reasons outlined below, screening was also conducted for performance overall as well as within the first and second half of the experiment. The criterion for skewness was skew/std. error of skew > 2.58 ; the criterion for excessive kurtosis was likewise kurtosis/std. error of kurtosis > 2.58 . Familiarity was not excessively skewed or kurtotic. The criterion for an outlying score was ± 2.58 SD from the mean; no outliers were identified.

Overall looking time in each half x age group x language group.

The current study measured infants’ preference for passages containing familiarized vs. non-familiarized target words during 4 blocks of 4 test trials. Given the added length of the task in this study compared to the traditional response-contingent HPP, it was possible that infants could become so disinterested in the task itself that preferential behaviour (if it existed at any point) may diminish over time. Thus, overall average looking time per trial during the first half of test trials was compared to looking during the second half of trials as a function of age group (9-months vs. 13-months) and language group (monolingual vs.

bilingual). Results of the corresponding ANOVA suggested that there was a significant decline in looking between the first ($M = 9.55$; $SE = .32$) and second half ($M = 6.95$; $SE = .41$) of trials ($F(1, 52) = 61.86$; $p < .001$; $\eta_p^2 = .54$). There was no effect of age group, language group, or an interaction between these variables, suggesting that the decline in looking across block halves was consistent across groups (all $ps > .10$). On average there was a 27% decline in looking between the first and second half of the experiment. It should be noted however that this includes the data from 8 participants who only completed the first 3 blocks of trials. Although looking time is averaged over the number of completed trials, total looking time may be inflated for those participants who would not have looked at all during block 4. When their data is excluded, there is a 30% decrease in looking between block halves and this remains significant ($F(1, 43) = 81.07$; $p < .001$; $\eta_p^2 = .65$), while the difference between age and language groups remains non-significant (all $ps > .10$).

Familiarity $\times$ age group $\times$ language group.

Due to the potential for performance to change over time within a single experimental session, other infant studies using preferential paradigms have examined performance overall as well as during each half of the experimental test trials (e.g., Vouloumanos & Werker, 2007; Yoshida et al., 2009). The data here will be analyzed in the same way for three reasons: 1) observation of recorded sessions suggested that infants tended to become restless during the second half of test trials (trials 9-16; blocks 3 and 4); 2) there was a significant decrease in looking between the first two and second two blocks of trials; and 3) at least one previous study using the HPP reported performance effects that only spanned the first three blocks of trials (Jusczyk & Aslin, 1995). To clarify, the “first half” of trials refers to trials 1-8, or Blocks 1 and 2. The “second half” of trials refers to trials 9-16, or Blocks 3 and 4.

To examine performance over the entire experiment, a 2 x 2 x 2 mixed GLM was performed with familiarity (familiarized vs. non-familiarized target words in passages) as the within-subjects dependent variable and age group (9- vs. 13-months) and language group (bilingual vs. monolingual) as the between-subjects independent variables. There was no significant effect of familiarity, nor were there any 2 or 3-way interactions (all $ps > .10$). Next, using the same method, performance was examined in each half of the experiment.

For the second half of the experiment there was no main effect of familiarity, nor were there any 2-way or 3-way interactions (all $ps > .10$). For the first half of the experiment there was a main effect of familiarity ($F(1, 52) = 5.03$; $p = .03$; $\eta_p^2 = .09$). Infants looked longer towards passages containing familiarized ($M = 9.90$; $SE = .34$) over non-familiarized ($M = 9.24$; $SE = .35$) target wordforms, suggesting the ability to generalize words across speakers with different accents. There was no 2-way interaction between familiarity and age group, nor was there a 3-way interaction between familiarity, age group and language group ($ps > .10$). There was however a non-significant trend towards an interaction between familiarity and language group ($F(1, 52) = 7.90$; $p = .08$; $\eta_p^2 = .06$). This trend warranted further investigation given that the primary purpose of this study is to determine whether bilinguals are advantaged in accent accommodation.

Planned, follow up t-tests were conducted within each language group to examine whether monolinguals and/or bilinguals succeeded in wordform generalization (Figure 1). As a group, monolinguals did not demonstrate a significant difference ($t(27) = .12$; $p = .90$) in looking towards familiarized ($M = 9.63$; $SD = 2.14$) versus non-familiarized ($M = 9.6$; $SD = 1.97$) targets. Thirteen out of 28 babies (46.4%) looked longer towards familiarized wordforms in passages. In contrast, bilingual infants looked significantly longer ($t(27) = 2.85$; $p = .008$) to familiarized ($M = 10.13$; $SD = 2.82$) over non-familiarized ($M = 9.0$; $SD =$

3.14) targets. Cohen's $d = .54$, suggesting that this was a large effect. Twenty out of 28 babies (71.4%) looked longer towards passages containing familiarized targets.

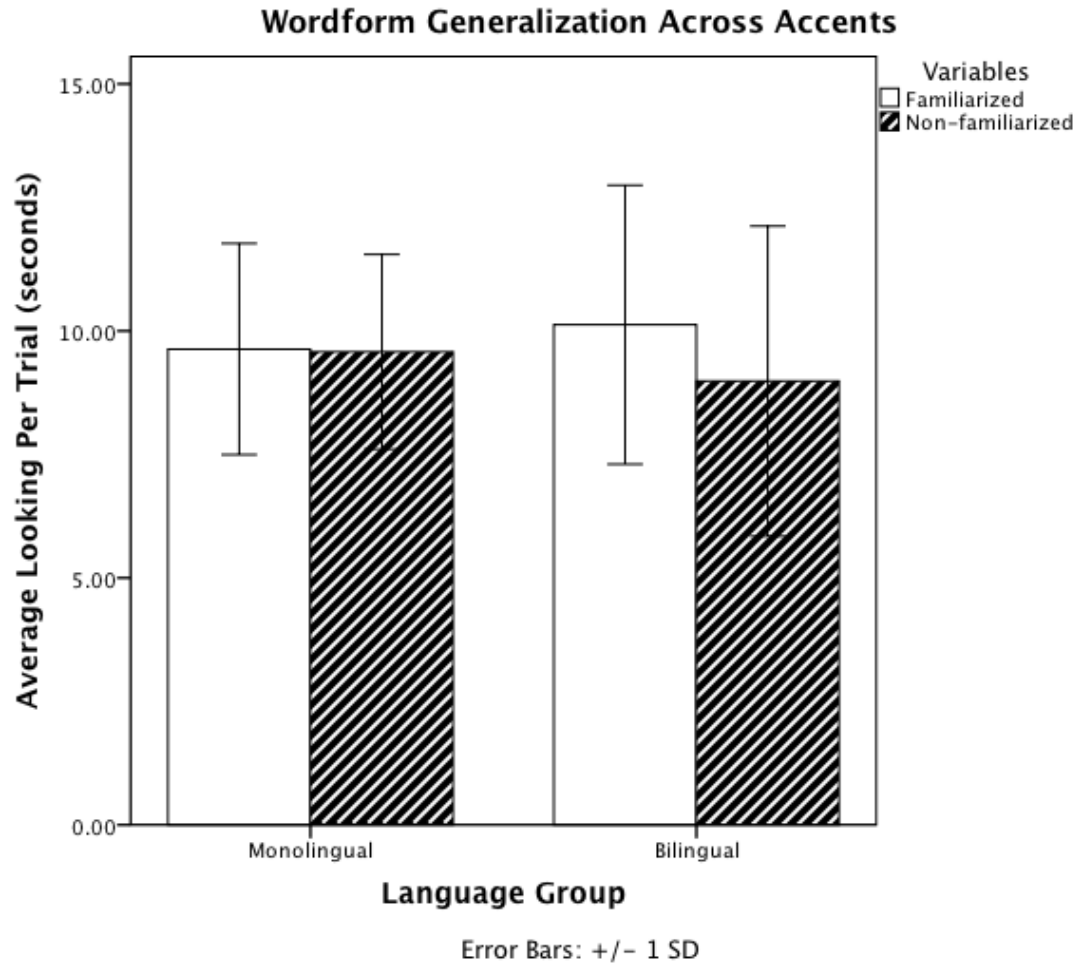


Figure 1. Average looking time per trial in seconds for passages containing familiarized and non-familiarized target words during the first half of test trials.

Although there was no significant interaction between familiarity, age group, and language group planned, follow-up t-tests within each age by language group cell were conducted to clearly identify which groups were succeeding in generalizing wordforms across accents. Neither the 9-month-old monolinguals ($t(15) = -.60$; mean difference = $-.38$; $p = .56$) nor the 13-month-old monolinguals ($t(11) = 1.20$; mean difference = $.63$; $p = .25$)

demonstrated preferential behaviour towards passages with familiarized or non-familiarized wordforms. Similarly, on their own, 9-month-old bilingual infants did not orient longer towards either passage type ($t(15) = 1.61$; mean difference = .81; $p = .13$). Thirteen-month-old bilinguals, however, oriented significantly longer towards familiarized over non-familiarized targets ($t(11) = 2.42$; mean difference = 1.59; $p = .03$). Cohen's $d = .70$, suggesting that this was a large effect. These results are displayed in Figure 2.

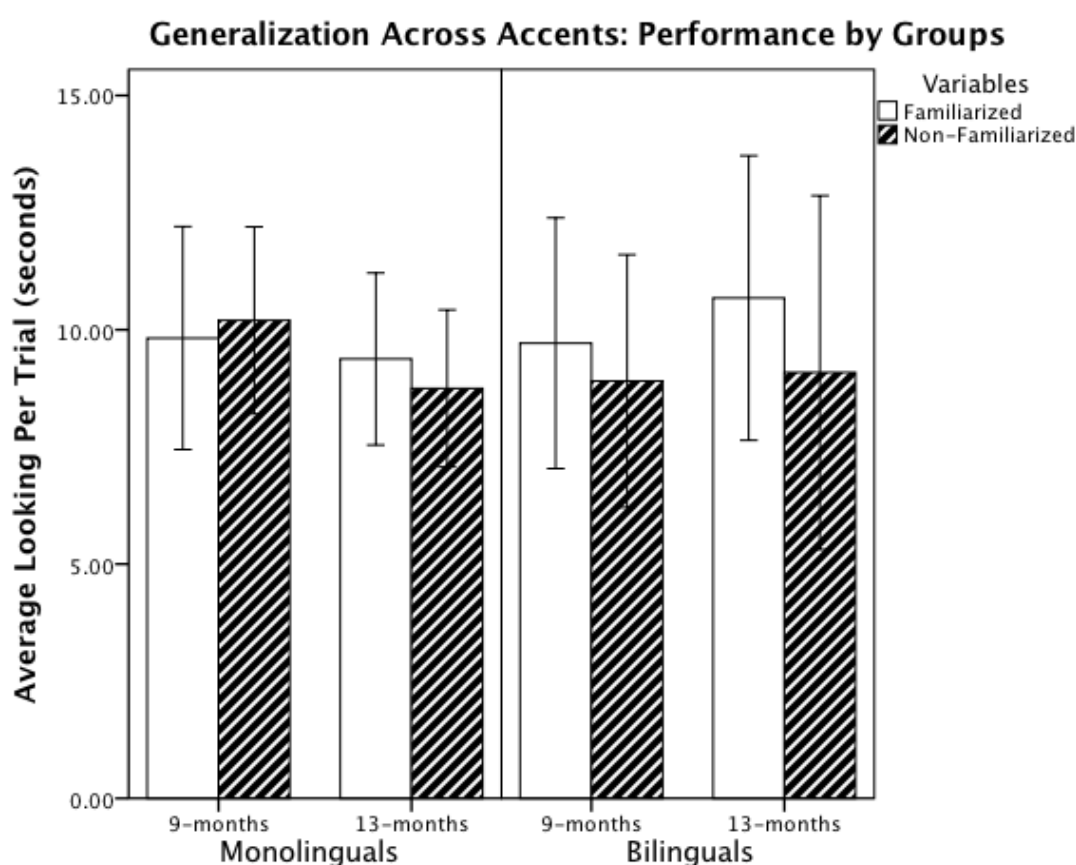


Figure 2. Looking time towards passages containing familiarized and non-familiarized target words by group. Error bars are +/- 1 SD.

GLM testing for condition effects.

In order to examine potential condition effects, 2 x 2 x 4 mixed GLMs were conducted for each linguistic group separately. Familiarity (familiarized vs. non-familiarized

during the first half of experimental trials) was the within-subject dependent variable; age group (9- vs. 13-months) and condition (1 – 4) were the between-subject independent variables.

For the monolingual group, there was no effect of familiarity, nor were there any 2 or 3-way interactions (all $ps > .10$). Therefore, difficulty with a particular condition or set of conditions was not masking an ability to accommodate for accented speech across speakers in this group. For the bilingual group, there was a main effect of familiarity ($F(1, 20) = 12.88$; $p = .002$; $\eta_p^2 = .39$). Bilingual infants oriented longer towards passages containing familiarized ($M = 10.2$; $SE = .57$) over non-familiarized ($M = 9.0$; $SE = .67$) targets. There was no 2-way interaction between familiarity and age group, nor was there a 3-way interaction between these variables and condition (all $ps > .10$). There was however an interaction between familiarity and condition ($F(3, 20) = 21.77$; $p = .01$; $\eta_p^2 = .42$). A visual representation of performance across blocks is available in Figure 3.

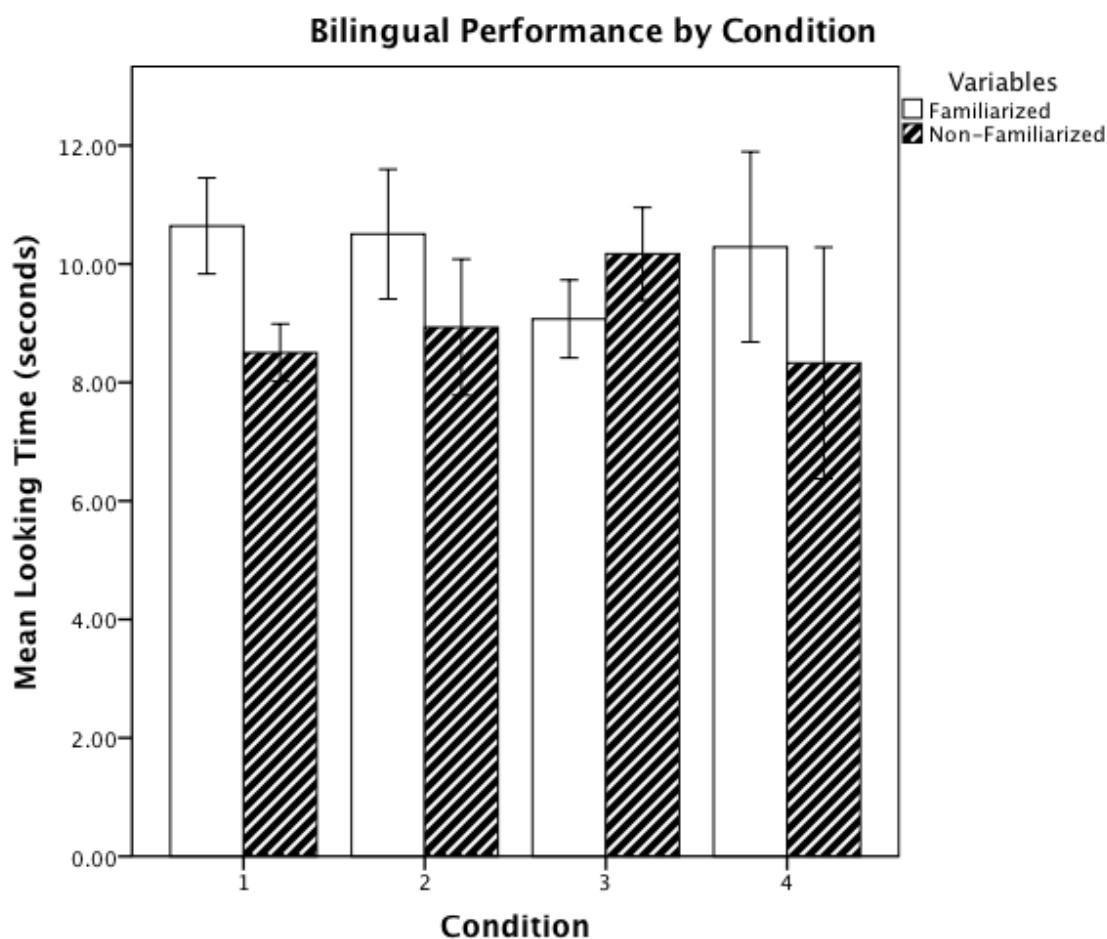


Figure 3. Bilingual infants’ orientation towards passages containing familiarized versus non-familiarized target wordforms in each condition. In Condition 1 and 2, infants heard a native-English speaker during the familiarization phase and a French-accented speaker during the test phase. In Condition 3 and 4 this was the reverse. In Condition 1 and 3 the target words were “candle” and “ticket”. In Conditions 2 and 4 the target words were “kinship” and “matches”. Error bars are +/- 1 SE.

Planned follow-up t-tests suggest preferential looking towards passages containing familiarized words in Condition 1 ($t(6) = 4.64$; mean difference = 2.14; $p = .004$; Cohen’s d

= 1.75) Condition 2 ($t(6) = 2.99$; mean difference = 1.57; $p = .02$; Cohen's $d = 1.13$) and Condition 4 ($t(6) = 2.64$; mean difference = 1.96; $p = .04$; Cohen's $d = 1$). There was no significant difference between looking at different passage types in Condition 3 ($t(6) = -1.31$; mean difference = -1.10; $p > .10$). This suggests that infants did not recognize target wordforms in passages if they were first familiarized to “candle” and “ticket” in French-accented speech and then tested with native-English passages. In all other conditions, bilingual infants succeeded in recognizing familiarized words across native-English and French-accented speakers.

A final analysis examined bilingual and monolingual 9-month-olds' performance by condition to verify whether both groups were succeeding in the task when condition effects were held constant. This was important in order to determine which age group should be tested in Study 3. Although the number of participants per condition ($n = 4$ for 9-month-olds and $n = 3$ for 13-month-olds) was too small to conduct inferential statistics, patterns of behaviour across conditions were inspected to determine whether a particular condition was perhaps more difficult for 9- and/or 13-month-olds. For 9-month-olds there was a significant effect of Familiarity ($F(1, 12) = 8.06$; $p = .02$; $\eta_p^2 = .40$), and in interaction between familiarity and condition ($F(3, 12) = 11.48$; $p = .001$; $\eta_p^2 = .74$). This interaction is displayed in Figure 4.

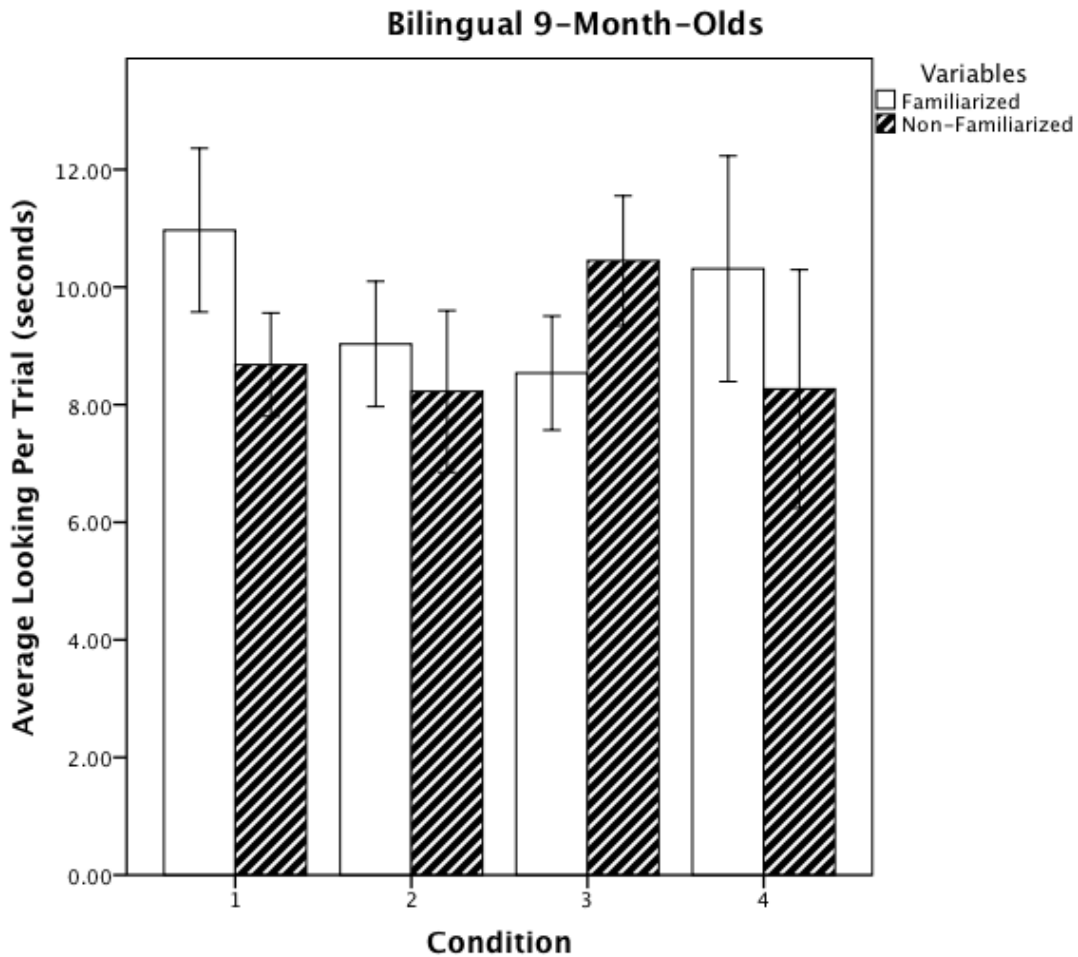


Figure 4. Bilingual 9-month-olds' performance in an accent generalization task. Error bars are +/- 1 SE.

The results of this analysis show that 9-month-olds display a general trend towards a preference for familiarized targets in passages, with the exception of in Condition 3, wherein infants tended to look longer towards passages containing non-familiarized targets.

Similarly, 13-month-olds tended to look longer to passages containing familiarized targets ($F(1, 8) = 5.32$; $p = .05$; $\eta_p^2 = .40$). There was no interaction between familiarity and

condition ($p > .10$). Once again, to acquire insight into which conditions (if any) may have been problematic, performance was examined by condition (Figure 5).

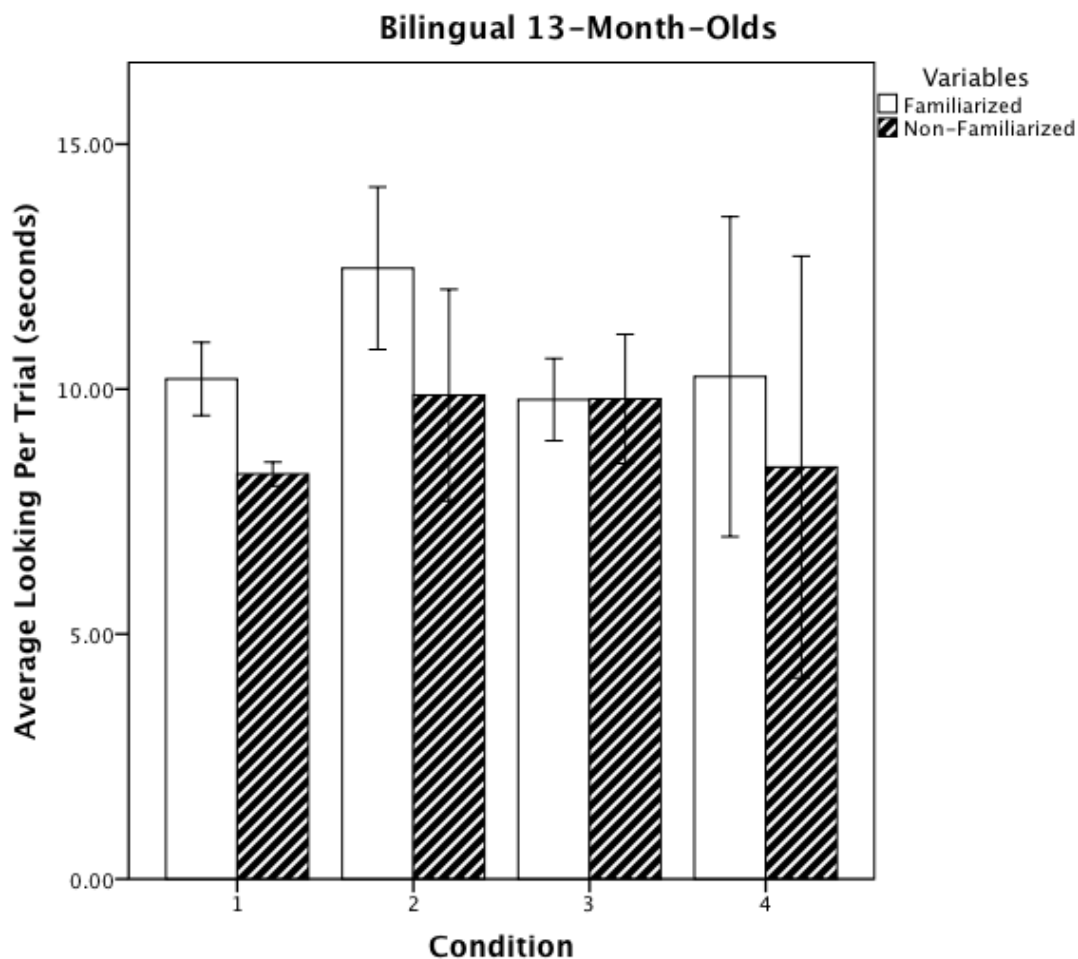


Figure 5. Bilingual 13-month-olds' performance in an accent generalization task. Error bars are ± 1 SE.

The results of this analysis suggest that similarly to the 9-month-old bilinguals, performance patterns were consistent across all conditions except for in Condition 3. In Condition 3 there was a clearly non-differential pattern of looking between passage types. Suggestions as to the origins of condition effects will be offered in the discussion section.

Influence of language and accent exposure.

Given the overall success of bilingual infants in generalizing wordforms across native and French-accented English, the next step was to identify the source of this ability. Two potential candidates were evaluated in this analysis: exposure to two languages, and exposure to accented speech. The prediction for these analyses was that there would be some degree of exposure to one language (inversely related to the other language) and to accented speech where phonological overlap would be maximized and therefore produce greatest task success. If this overlap is imperfect due to idiosyncratic pronunciations across languages (which is the assumption in this thesis), then performance will peak at the point where there is the most overlap. Performance should therefore also decline at some point prior to and following this ideal. For example, it might be the case that at 50% exposure to English (and therefore 50% exposure to French also if no other languages are heard), performance in the task would be the greatest whereas perhaps at 25% or 75% exposure to English, performance would decline. In order to test this possibility, the relationship between task success and exposure to English and accented speech was evaluated using the Curve Estimation Procedure in SPSS 21. An inverted-U shaped relationship between task success and exposure to either or both of English and accented speech was predicted.

Variables.

Exposure to English was measured using the Language Exposure Questionnaire. Although it was initially predicted that exposure to French-accented speech would be a particular facilitator of infants' success in this task, only 5 (according to the Subjective Caregiver Report (SCR)) and 10 (according to the Language Questionnaire Measure (LQM)) out of 28 infants were hearing French-accented English speech. Exposure to French-accented speech was therefore not assessed as an independent variable in the following analyses since

the number of 0 scores significantly skewed the data beyond repair through transformation⁴. In contrast 19 parents reported that their infants heard some form of accented speech (e.g., English-accented French) and 17 parents were assessed as speaking with an accent using the LQM. Therefore, the relationship between task success and exposure to accented speech in general was evaluated.

Although the SCR and LQM of exposure to accented speech were significantly correlated ($r = .44$, $n = 28$; $p = .02$), it is unclear at this stage which measure better reflects infants' exposure to non-native (accented) speech, therefore both measures were evaluated separately in the curve estimation procedure. Success in the task was measured as the difference between looking towards passages containing familiarized vs. non-familiarized targets. Since any deviation from 0 indicates a potential preference (e.g., a score of -1 is not inherently "lower" than a score of 1 because both would indicate differentiation between non-familiarized or familiarized targets), all scores were converted to positive scores. Given the trend in the current experiment towards a preferential response in the first two blocks but not the second two blocks of test trials, performance during the first half of the experiment was used as the dependent variable.

Data cleaning and screening.

Prior to the curve estimation procedure, the DV (success in the task during the first half) and IVs (LQM, SCR measures of proportion of exposure to accented speech and LEQ estimation of percentage exposure to English) were screened for normality and potential outliers. The criterion for skewness was skew/std. error of skew > 2.58 ; the criterion for excessive kurtosis was kurtosis/std. error of kurtosis > 2.58 . Task success was reasonably normally distributed, as was the percentage of time infants were exposed to English and

⁴ LQM met the criteria for logistic regression, discussed later in this thesis.

LQM. SCR was substantially positively skewed ($1.98/.44 = 4.50$) as well as platykurtic ($3.40/.86 = 3.95$). SCR was subsequently logarithmically transformed ($\log_{10}(\text{SCR} + 1)$). Transforming the variable had very little effect and therefore it was dropped from further analysis in preference of the LQM. No univariate outliers were identified for any of the variables. Scatterplots of all DV + IV combinations were screened for potential outliers and heteroscedasticity. No multivariate outliers were identified. The combination of task success and exposure to English appeared heteroscedastic. Since success in the task neared the criterion for positive skewness ($.90/.44 = 2.05$), it was subsequently square root transformed. This improved the dispersion of data across task success and exposure to English.

Curve estimation results.

The potential relationship between the square root of task success and exposure to English (inversely related to exposure to French) was examined first using Curve Estimation in SPSS 21. There was a nearly significant linear relationship between these variables ($F(1, 26) = 4.02$; $p = .06$; $R^2 = .13$). These results reflect a negative relationship between task success and exposure to English. As exposure to English decreases (and exposure to French increases), the ability to generalize wordforms across a native and French-accented English speaker improves. This relationship is depicted in Figure 6 below.

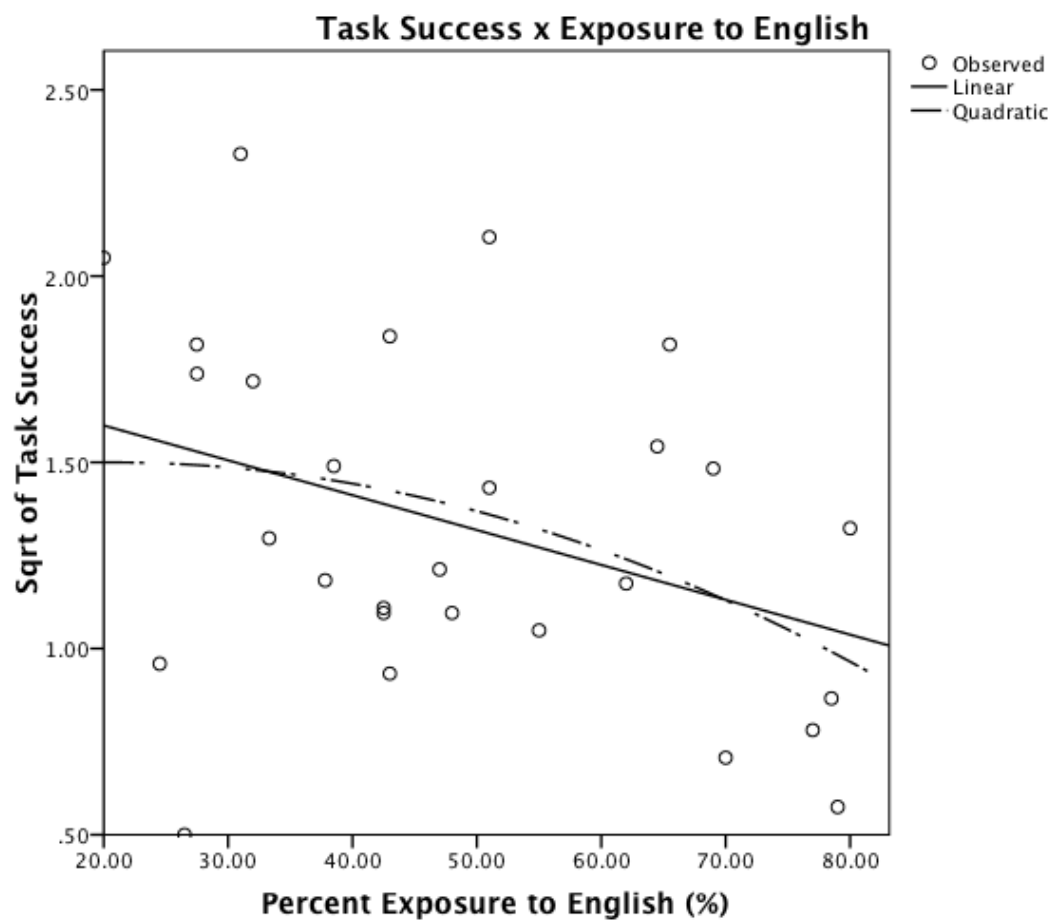


Figure 6. Scatterplot of the square root of task success scores versus the percentage of time infants were exposed to English. English exposure was calculated as the number of hours infants listened to caregivers (and close friends and relatives) speaking English divided by the total number of hours of language exposure (e.g., English + French).

There was no significant quadratic relationship ($p > .10$) between task success and exposure to English. The results of this analysis are displayed in Table 8.

Table 8. *Linear and quadratic relationships between task success and Exposure to English.*

Equation	R ²	Adj. R ²	F	df1, df2	p	Constant	b1 (std. err)	b2 (std. err)
Linear	.13	.10	4.02	1, 26	.06	1.79	-.009 (.45)	
Quadratic	.14	.08	2.09	2, 25	.15	1.44	.006 (.03)	-.0003 (.000)

Note: b1 and b2 are non-standardized regression coefficients

Next, the relationship between task success and the LQM of accented speech was evaluated. There was neither a linear nor a quadratic relationship between these variables ($ps > .10$).

Logistic regression for task success by exposure to accented speech.

Both measures of exposure to French-accented speech were substantially skewed due to the large number of infants who had never heard this variety of speech. This made these variables non-amenable to curve estimation regression. However, by creating a binary variable (heard French accent vs. did not hear French-accent) it was possible to submit the data to a logistic regression. In this case, the DV was whether or not infants were exposed to French accented speech, and the IV was task success during the first half (as measured by the difference in listening between passages containing familiarized over non-familiarized words converted to positive scores).

In order to determine whether the data were amenable to logistic regression, each DV (SCR and LQM) was screened to ensure that they met the criterion for perfect separation (the ratio of cases for the DV group with the smallest number of participants to the number of IVs had to be 10:1). Since only one IV was tested, each group of the DV had to equal at least 10. Using the SCR only five participants heard French-accented speech, thus this

criterion was not met and this variable was dropped from further analysis. Using the LQM 10 participants heard French-accented speech and 17 did not; therefore after confirming linearity between the IV and the log odds ratio probability of the DV, these data were submitted to logistic regression. Success in the task did not significantly predict whether or not infants had been exposed to French-accented speech as measured by the LQM ($X^2(1) = .33; p = .56$).

Discussion

The results of Study 2 suggest that as a group French-English learning infants were capable of generalizing wordforms across native and French-accented speakers. Bilingual infants who were familiarized with either a native or French-accented productions of target words, subsequently recognized these words in running speech when presented by a speaker with the opposite accent. This was indicated by longer overall orientation towards a visual stimulus while hearing passages containing words that the infants had previously been exposed to rather than passages containing novel words. In contrast, monolinguals did not demonstrate an attentional preference for either set of passages.

These results are consistent with previous work demonstrating that 9-month-old English learners did not generalize wordforms across a native English and Spanish-accented English speaker (Schmale & Seidl, 2009) or across speakers of different dialects of English (Schmale et al., 2010). It was surprising however that 13-month-old monolinguals failed to accommodate for accented speech in this study since previous research has suggested success at this age in a similar task (Schmale & Seidl, 2009; Schmale et al., 2010).

It is possible that the non-contingent variation of the HPP applied in the current study made the task more difficult, perhaps by making the task less interesting for infants who did not control the presentation of test stimuli. This seems unlikely however given that

monolingual and bilingual infants had similar attrition rates and completed a similar number of trials. This suggests that monolinguals' failure cannot be attributed to being less interested in the task. Furthermore, the task was designed to be simpler by virtue of giving infants greater opportunity to demonstrate learning by allowing reorientation towards the stimuli. They also had an arguably greater opportunity to become acquainted with the idiosyncratic phonetic productions associated with each speaker by receiving longer exposure to test trials. Another possible explanation however is that the French-accented stimuli in this study were more phonetically contrastive with the native English stimuli compared to the Spanish-accented or regional varieties of English in previous research. This could perhaps be tested in future research by asking adults to rate the degree of "accentedness" of French versus Spanish accented speakers. Still, despite age-related differences across current and previous studies it is telling that regardless of potential challenges imposed by the experimental design or the particular accent used in this study, bilingual infants succeeded overall.

One of the main hypotheses of this study was that bilingual infants would succeed in accommodating for accented speech in wordform recognition earlier than their monolingual peers. Clearly the data herein suggest that overall bilingual infants did generalize across speakers and accents earlier than the monolinguals. The question as to whether 9-month-olds in this study truly succeeded is somewhat less clear. Despite that there was no significant difference between 9-month-olds' and 13-month-olds' performance, the younger group did not appear to generalize across variable wordforms at a level greater than chance when all conditions were included. However it seems that with the exception of Condition 3, in which 13-month-olds also did not succeed, these infants did demonstrate the expected pattern of preference for passages containing familiarized over non-familiarized words. When controlling for the effect of condition, 9-month-olds were successful in the task.

Although it is possible that simply by chance, four out of the five 9-month-olds and two of the three 13-month-olds who did not orient longer towards passages containing target words during the first two blocks of test trials were also in Condition 3 this is very unlikely.⁵ On the other hand, there could have been something unique about either the stimuli or the participants in this condition that related to performance.

Concerning the stimuli itself, it is notable that the stress measurements taken for target words in Study 1 predicted that Condition 3 would present the most difficulty in accommodating across accents. This is because the most inconsistent word stress patterns occurred when the French-accented speaker produced the target words “candle” and “ticket” in isolation during the familiarization phase and the native-English speaker produced these words in target passages. Recall that the French-accented speaker used amplitude to emphasize the first syllable of these tokens. In fact, this speaker used a degree of amplitude that exceeded her native-accented peer. In the corresponding test passages however, neither speaker reliably used amplitude as a first-syllable cue to stress. They did both use amplitude to stress kinship and matches tokens within test passages, thus it is possible that infants in that condition were matching the stress pattern that they heard during familiarization to that which was heard during the test phase. That said, if some form of perceptual matching of word-stress across the familiarization and testing phases was responsible for this effect, then one might have expected a decrease in performance across Conditions 1 and 2 where there

⁵ Although a Chi squared test is not suitable here since 50% of cells would be smaller than 5, the probability of this particular outcome can be computed as follows. Use the combinatorial formula nCr to find the number of ways to select 6 infants from a total of 8 who looked did not look longer to familiarized words and 0 infants out of 20 who did look longer to familiarized words. Then divide that by the number of possible ways to choose 6 infants out of 28. The probability of 6 out of the 8 infants who did not look longer to familiarized words ending up in Condition 3 specifically is $1/13455$ or $p = .00007$. The probability of the 6 infants who did not look longer to familiarized words ending up in any condition together is the above probability multiplied by four, or about $1/3364$, $p = .0003$.

were no consistent stress-cues across experimental phases, and an increase in performance in Condition 4 where stress-cues between familiarization and testing were consistent. This was not however the case. Performance across Conditions 1, 2, and 4 was quite similar.

Another possibility is that generalization across familiarization and testing phases was easiest in Condition 3. If this were the case, then it is possible that infants could have habituated to passages containing recognized target words and preferred passages containing novel words. The multidimensional nature of accented speech makes it difficult to assess whether some targets may have been more easily matched across familiarization and test phases. Notably however, infants did successfully generalize across instances of “candle” and “ticket” when the accented speech came during the test phase. It cannot however be ruled out that infants who heard the accented versions of these tokens during familiarization recognized these words within passages more easily. Certainly some previous evidence suggests that such asymmetries in generalizing across accents do occur. Schmale and colleagues (2011) for example found that 24-month-olds accommodated generalized words across native and Spanish-accented English when familiarized with accented tokens and tested on native tokens, but not in the opposite condition.

Finally, it is possible that the linguistic characteristics of infants in Condition 3 made their performance differ from the others. For example if these infants were hearing a greater proportion of French, then perhaps they may have habituated to the target wordforms more quickly than their peers and consequentially oriented longer towards their novel counterparts. This possibility was tested in a 2-Way ANOVA with condition and age group as the IVs and percent exposure to French as the DV and there was no significant difference

in mean exposure to French across conditions or age groups, and no significant interaction between age and exposure to French (all $ps > .10$).⁶

Regardless of the cause of infants' inconsistent behaviour in Condition 3, when controlling for condition, both age groups demonstrated success in accommodating for French-accented speech. The natural next step was therefore to examine which environmental factors may be related to this success. Of the examined relationships, only the relationship between exposure to English and task success neared significance. As infants received less exposure to English, and more exposure to French, they were more successful in differentiating between passages containing familiarized and non-familiarized wordforms.

There are several potential explanations for this relationship. First, it is possible that the measures used to assess exposure to accented speech were not sufficient to accurately detect all cases where infants were hearing accented speech. Indeed, it was noted throughout this study that the LQM and SCR were sometimes “missing” speakers who were perceived by the experimenter as speaking with an accent. It remains therefore possible that infants who were hearing more French were also hearing French-accented speech more frequently, even though this was not suggested by the measures used herein. Although it was not feasible in the current study, future studies may profit from collecting speech samples from parents and allowing third parties to rate the presence and degree of accent in their speech.

Another explanation for the negative correlation between exposure to English and task success returns to the idea that phonetic overlap between languages makes bilingual infants more tolerant to phonetic variability. For instance, the phonetic distributions across

⁶ The range of exposure to French for 9-month-olds was 38 to 67%; Condition 1 $M = 38\%$, Condition 2 $M = 67.65\%$, Condition 3 $M = 43.75\%$ and Condition 4 $M = 51.88\%$; $SE = 8.85$. For 13-month-olds the range was 39 to 62%; Condition 1 $M = 62.40\%$, Condition 2 $M = 49\%$, Condition 3 $M = 55.34\%$; Condition 4 $M = 39.33\%$; $SE = 10.22$. None of these differences were significant.

English and French may be such that overlap effects are most pronounced in infants who are hearing predominantly French. Although it might be expected that overlap would be symmetrical, resulting in better performance for both English-dominant and French-dominant infants, this might not necessarily be the case. For example, if French is the infant's dominant language then prototypical sound categories may take on greater French-language characteristics (assuming that frequency of sounds is reflected in language-dominance). Since the French-accented speaker tended to impose French-phonological characteristics onto her English productions, this might be particularly conducive to French-accent accommodation in French-dominant bilinguals. On the other hand, perhaps French-dominant bilinguals treated the task of wordform recognition differently than those with greater exposure to English. Given that the experimental stimuli consisted of English wordforms and passages, it is possible that French-dominant bilinguals were less restrictive in the amount of variability that they would allow before identifying a potential change in meaning simply owing to a lack of experience with the language. If so, then one might also expect that a negative relationship should exist between performance in such a task and exposure to English even when the accent under investigation is phonetically unfamiliar. This will be the focus of the experiment presented in the following chapter.

Chapter 4

Study 3: Bilingual Infants' Accommodation of Phonetically Unfamiliar Accented Speech

Introduction

The results of Study 2 suggest that French-English bilingual infants were advantaged over their monolingual English learning peers in recognizing familiarized wordforms across speakers with different accents. Specifically, as a group, bilingual infants were able to generalize wordforms across a native- and a French-accented speaker of English.

Examination of the relationship between task success and exposure to English revealed that infants who were exposed to less English compared to French tended to perform better in accommodating for the accent under investigation. A question that arose from these results was whether infants with less exposure to English were less restrictive in the amount of variability that they would allow English wordforms simply owing to inexperience with the language, or whether experience with phonetic overlap between French and English (and perhaps related accents) drove success in the task.

In the current study, 9-month-old French-English bilinguals were tested on their ability to accommodate Mandarin-Accented Canadian English. The hypothesis for this study was that bilinguals' ability to accommodate for a change in speaker and accent in Study 2 was related to the distribution of language sounds across their two languages, and not simply owing to inexperience with English. The prediction therefore was that French-English bilinguals would not accommodate for an accent derived from an unfamiliar (Mandarin) phonology imposed upon a familiar language (English).

Methods

The apparatus, design, procedure and questionnaires were identical to those in Study 2. As in Study 2, a second coder, blind to the familiarity status of test trials, recoded 25% of the data. The Pearson's product-moment correlation coefficient between coders, based on the difference score between looking to passages containing familiarized vs. non-familiarized target words, was .97 ($p < .05$) suggesting high inter-rater reliability.

Participants.

The participants in this study were 16 bilingual 9-month-olds (7 females; age range = 8.57-10.28; $M = 9.31$; $SD = .59$). Participants were considered bilingual if they were exposed to English and French 20-80% of the time using the same Language Exposure Questionnaire as in Study 2. On average these infants were exposed to English 43.07% of the time ($SD = 17.16$), to French 54.64% of the time ($SD = 16.50$) and to a language other than French or English 2.26% of the time ($SD = 7.53$). None of the infants had previous exposure to Mandarin or Cantonese. All participants were born after at least 37 weeks gestation, were in apparent good health at the time of testing, and did not have a reported history of developmental or physical challenges. Fourteen of the 16 participants completed all 4 blocks of test trials. An additional 7 babies were tested but excluded due to: extreme fussiness/crying (5), equipment failure (1), or illness (1).

Results

Familiarized vs. non-familiarized passages.

Data cleaning and screening.

For the purpose of data screening, a single "familiarity" variable was calculated by subtracting looking time (in seconds averaged over the number of trials) to passages containing non-familiarized wordforms from looking time to passages containing

familiarized wordforms. Positive difference scores indicate greater looking towards passages containing familiarized target wordforms over passages containing non-familiarized counterparts while negative scores indicate the reverse. Using this difference score the data were screened for normality and the presence of outliers. The criterion for skewness was skew/std. error of skew > 2.58 ; the criterion for excessive kurtosis was likewise kurtosis/std. error of kurtosis > 2.58 . Familiarity was not excessively skewed or kurtotic. The criterion for an outlying score was ± 2.58 SD from the mean; no outliers were identified.

T-tests.

A paired samples t-test evaluated whether there was an overall difference in looking towards passages containing familiarized vs. non-familiarized target words, and revealed that this was not the case ($t(15) = -1.13$; $M_{\text{familiarized}} = 7.34$; $M_{\text{non-familiarized}} = 7.84$; $p = .28$, Cohen's $d = .28$). Given the results of Study 2, familiarity effects (if found) were expected to occur in the first half of trials. Difference scores were obtained for familiarity in the first half and in the second half and were screened for normality and outliers using the same criteria above; all data were maintained and unaltered. Two follow up t-tests evaluated whether infants demonstrated preferential looking behaviour to familiarized or non-familiarized target words in the first and second half of the experiment. Neither of these tests suggested an effect of familiarity ($t(15) = -1.75$; $M_{\text{familiarized}} = 8.23$; $M_{\text{non-familiarized}} = 9.52$; $p = .10$; Cohen's $d = .43$ and $t(15) = .28$; $M_{\text{familiarized}} = 5.92$; $M_{\text{non-familiarized}} = 5.97$; $p = .78$; Cohen's $d = .07$ respectively).

GLM testing for condition effects.

Given the non-significant trend towards a preference for passages containing non-familiarized (novel) target words during the first half, a mixed GLM with familiarity (familiarized vs. non-familiarized) as the within-subjects variable and accent (during

familiarization vs. during test) and target words (candle & ticket vs. kinship & matches) as the between subjects variables was also conducted. This was to ensure that difficulty with a particular condition was not masking success in the task. The results again suggested that there was no effect of familiarity, nor were there any 2 or 3-way interactions between variables (all $ps > .10$). Therefore, as a group, bilingual infants failed to generalize wordforms across native and Mandarin-accented English speakers.

Influence of language and accent exposure.

Despite the overall group failure, there remained the possibility that exposure to accented speech or degree of bilingualism (e.g., percentage of exposure to one language) may influence individual infants' success in the task. The same curve estimation procedure as in Study 2 was used to test for either a linear or quadratic (inverted U) relationship between success in the task and percentage of exposure to English or non-native speech ("accented").

Variables.

Exposure to English was measured using the Language Exposure Questionnaire. Although the Subjective Caregiver Report (SCR) and Language Questionnaire Measures (LQM) of exposure to accented speech were highly correlated ($r = .65$, $n = 16$, $p = .006$), it remains unclear at this stage which measure better reflects infants' exposure to non-native ("accented") speech. Therefore, both measures were evaluated separately in the curve estimation procedure. As in Study 2, success in the task was measured as the difference between looking towards passages containing familiarized vs. non-familiarized targets, and the dependent variable reflected performance during the first half of the experiment only.

Data cleaning and screening.

Prior to the curve estimation procedure, the DV (success in the task) and IVs (LQM, SCR, and percent exposure to English) were screened for normality and potential outliers. The criterion for skewness was skew/std. error of skew > 2.58 ; the criterion for excessive kurtosis was kurtosis/std. error of kurtosis > 2.58 . Task success was positively skewed ($1.48/.56 = 2.64$) and excessively platykurtic ($-.27/1.09 = 3.05$). The variable was therefore square root transformed. Following transformation, the variable was normally distributed. None of the variables contained univariate outliers. Scatterplots for each DV + IV combination were examined. One potential outlier was identified when exposure to English was plotted against success in the task. In order to test for the presence of multivariate outliers, Mahalanobis distances were calculated for each combination of DV + IV. The criterion score for identifying outliers was $X^2(2) = 9.21$. All scores were well within the allowable limit. Visual examination of scatterplots left open the possibility of either linear or quadratic relationships between task success and LQM and percentage exposure to English; the relationship between task success and SCR appeared fairly linear.

Curve estimation results.

The potential relationship between task success and exposure to English (inversely related to exposure to French) was examined first using Curve Estimation in SPSS 21. Neither a linear ($F(1, 14) = .05$; $p = .83$) nor a quadratic ($F(2, 13) = .13$; $p = .87$) equation described the data. The same was true when testing for a linear ($F(1, 14) = .37$; $p = .56$) or quadratic ($F(2, 13) = .21$; $p = .82$) relationship between the DV and SCR. Similarly there was no significant linear ($F(1, 14) = 2.35$; $p = .14$) or quadratic ($F(2, 13) = 2.36$; $p = .13$) relationship between task success and the LQM of exposure to non-native speech. Through examination of the scatterplots associated with these analyses, it was observed that only 4

out of 16 participants were exposed to English more than 50% of the time. This suggests a rather disproportionally spread sample if one predicts a curved relationship between success in the task and exposure to English (see Figure 7 below).

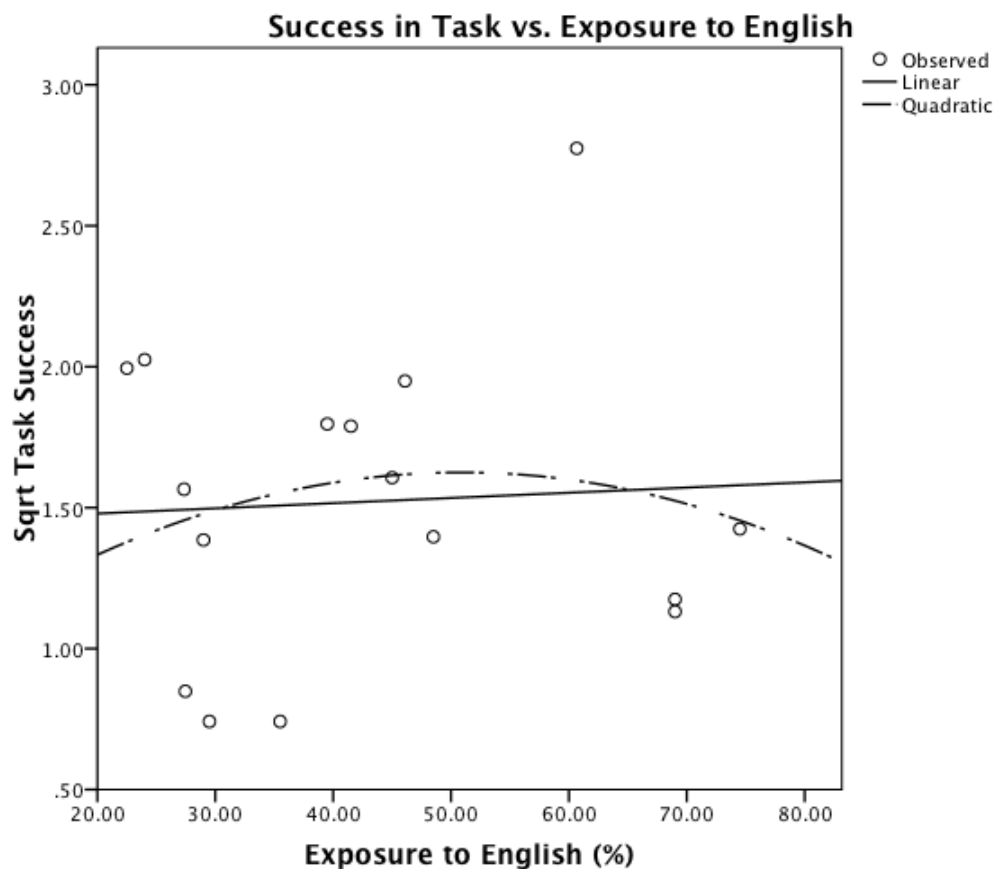


Figure 7. Scatterplot of the square root of task success scores versus the percentage of time infants were exposed to English. English exposure was calculated as the number of hours infants listened to caregivers (and close friends and relatives) speaking English divided by the total number of hours of language exposure (e.g., English + French).

Exploratory analysis including monolingual infants.

In order to examine the possibility that the quadratic relationship between task success and English exposure was non-significant owing to the lack of participants occupying the English-dominant end of the spectrum, several monolingual English learning infants were tested and included in an exploratory analysis. Seven monolingual infants (6 females) who did not meet the criteria for bilingualism were added to the curve estimation analysis. These participants were $M = 9.46$ months old (range = 8.69 – 10.39), which was not significantly different from bilingual participants ($t(21) = .56$; mean difference = .15; $p = .58$). The infants were exposed primarily to English since birth ($M = 94.79\%$), only minimally to French ($M = 5.21\%$), and never to a third language. All infants were born after at least 37 weeks gestation, and had no history of developmental or physical challenges. An additional 4 babies were tested but excluded to crying/fussiness (3), experimenter error (1), and parental interference (1). They were tested using the exact same procedure as described above.

After adding these 7 babies to the dataset, the curve estimation procedure outlined above was performed again looking for either a linear or quadratic relationship between success (the square root of the difference between looking towards passages containing familiarized or non-familiarized words) in the task and exposure to English. For consistency with the previous analysis the regression was also performed with success as the DV and each of the accent measures as the IV.

There was neither a linear ($F(1,20) = 2.08$; $p = .17$) nor quadratic ($F(2,19) = 1.00$; $p = .39$) relationship between task success and Subjective Caregiver Report of exposure to non-native speech. As demonstrated in Table 9, there was however both a significant linear relationship ($F(1,21) = 10.18$; $p = .004$; $R^2 = .33$) and an even stronger quadratic (inverted-U

shaped) relationship ($F(2,20) = 10.59$; $p = .001$; $R^2 = .51$) between non-native speech exposure as reported by the LQM and task success. Parameter estimates suggest that beyond 38% exposure to non-native speech ($b1/(2*b2)$), task success begins to decline.

Table 9 *Linear and quadratic relationships between task success and LQM of non-native speech exposure*

Equation	R^2	Adj. R^2	F	df1, df2	p	Constant	b1 (std. err)	b2 (std. err)
Linear	.32	.16	10.18	1, 21	.004	.98	2.20 (.69)	
Quadratic	.51	.31	10.59	2, 20	.001	.72	6.38 (1.62)	-8.47 (3.04)

Note: b1 and b2 are non-standardized regression coefficients

An examination of residuals confirmed that the quadratic trend better explained the data. Notably however, there were very few participants who heard > 50% non-native speech, which is not surprising given that hearing accented speech all the time would be a rarity. Still, this trend seems to indicate that exposure to overlapping distributions of native and non-native speech overlap facilitate success in accommodating for a foreign accent. A scatterplot demonstrating the relationship between task success and LQM is available below in Figure 8.

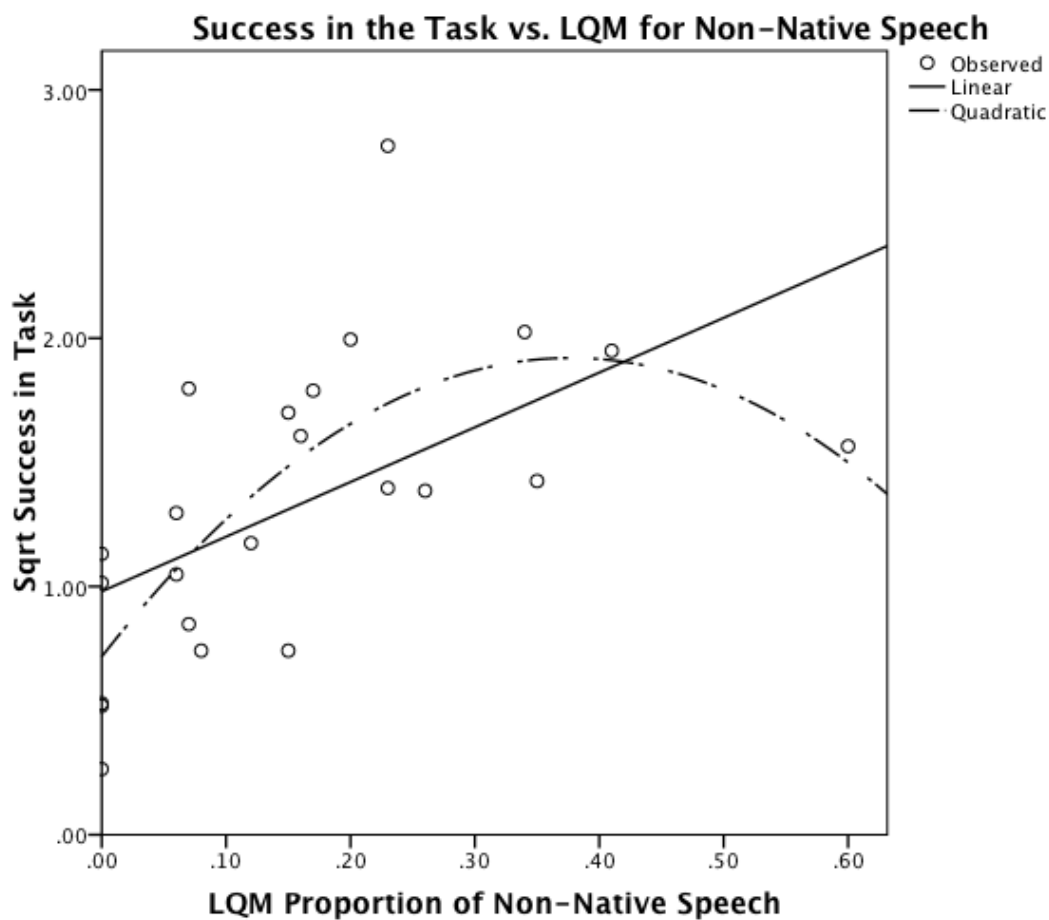


Figure 8. Scatterplot of the square root of task success scores versus the percentage of time infants were exposed to non-native (accented) speech. Non-native (“accented”) speech exposure was calculated as the number of hours infants listened to caregivers (and close friends and relatives) speaking in a second, non-native, language, divided by the total number of hours of language exposure.

Finally, the linear relationship between task success and exposure to English was significant ($F(1, 21) = 5.15$; $p = .03$; $R^2 = .20$), and suggested a negative relationship between task success and exposure to English. The quadratic relationship was also significant and provided a better fit of the data ($F(2, 20) = 5.86$, $p = .01$; $R^2 = .37$) as indicated by the greater R^2 value (Table 10). Parameter estimates suggest that beyond 48% exposure to English, success in the task begins to decrease. This quadratic relationship was reconfirmed in a visual analysis of residuals. A scatterplot of the relationship between task success and exposure to English is available in Figure 9.

Table 10. *Linear and quadratic relationships between task success and exposure to English*

Equation	R^2	Adj. R^2	F	df1, df2	p	Constant	b1 (std. err)	b2 (std. err)
Linear	.20	.16	5.15	1, 21	.03	1.88	-.009 (.004)	
Quadratic	.37	.31	5.86	2, 20	.01	.46	.048 (.025)	-.0005 (.000)

Note: b1 and b2 are non-standardized regression coefficients

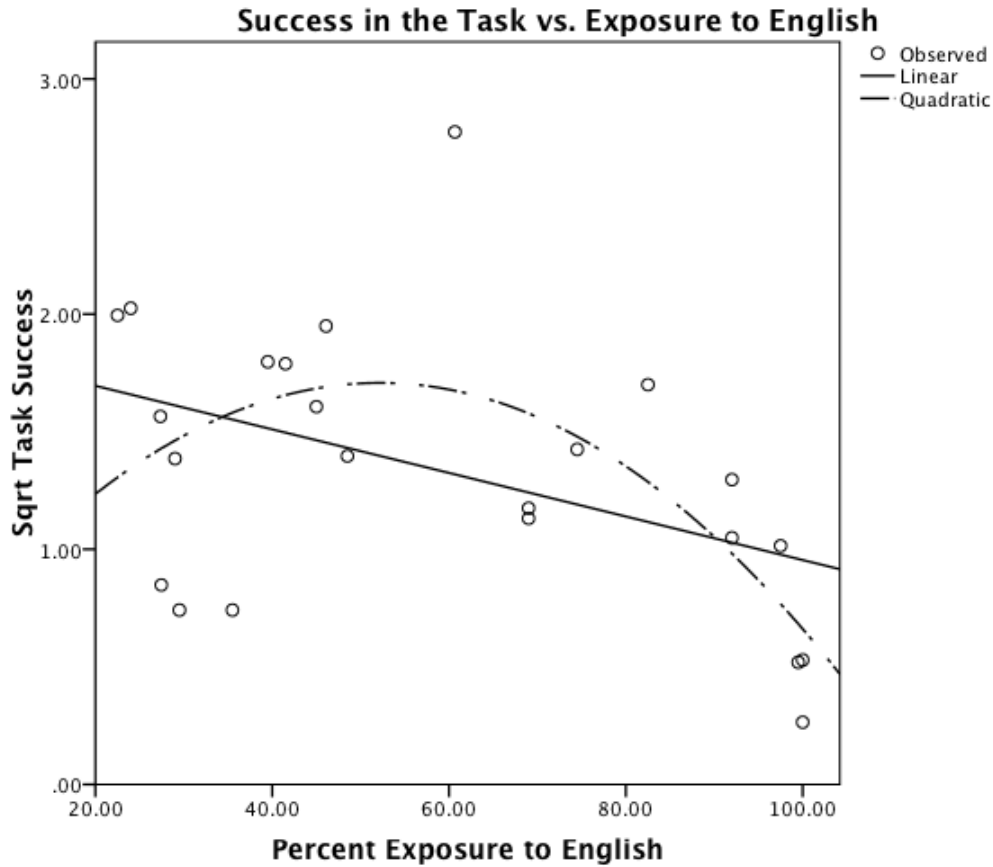


Figure 9. Scatterplot of the square root of task success scores versus the percentage of time infants were exposed to English. English exposure was calculated as the number of hours infants listened to caregivers (and close friends and relatives) divided by the total number of hours of language exposure. This plot includes data from monolingual infants.

Discussion

The results of the current study suggest that, as a group, French-English bilingual 9-month-olds do not recognize familiarized wordforms across native and Mandarin-accented productions. This overall failure, in comparison to the successes observed in Study 2, supports the hypothesis that accent tolerance is to some degree input specific. That is,

experience with the phonology of the imposing accent (either through exposure to the accent itself, or to the language from which the accent is derived) facilitates success in accommodating for that particular accent. This finding is supported by evidence that toddlers accommodate for a familiar artificially accented vowel shift but not for a novel vowel shift (White & Aslin, 2011).

A surprising finding, however, is that degree of bilingualism and exposure to accented speech was related to task performance. The quadratic nature of each of these relationships suggests that overlap between languages (i.e., English vs. French) and language varieties (native vs. non-native accents) relates to success in accommodating for a previously unheard foreign accent. This is interesting because it suggests that accent tolerance may not be as language-specific as initially proposed, and that the influence of hearing two languages and or two varieties of the same language is already affecting accent accommodation as early as 9-months. The relative flexibility of infants' early wordform representations may therefore be dependent upon the degree to which they are bilingual or exposed to multiple pronunciations of wordforms. Unfortunately, the current data do not allow inferences concerning the relative importance of accent exposure versus degree of bilingualism given that these two measures were very highly correlated in the current sample ($r(23) = -.56$; $p = .006$). In this sample, greater exposure to English also suggested less exposure to non-native speech. The implications of this relationship will be discussed in the following section in terms of limitations and future research directions.

Chapter 5

General Discussion

Previous research has suggested that young infants' ability to recognize familiar wordforms is vulnerable to variability in the input. For a time, novice language learners appear to encode wordforms in such rigid detail such that they can no longer recognize these words when they are altered at the subphonemic and suprasegmental levels. Using the headturn preference procedure, previous research has shown that despite being capable of extracting familiar words from running speech as early as 6-months (Bortfeld et al., 2005), monolingual 7.5-month-olds do not generalize wordforms across speakers of opposite genders (Houston & Jusczyk, 2000), across different pitch productions (Singh et al., 2008), or across different affective states (Singh et al., 2004). Although these types of variability are not phonemic, and therefore do not indicate a change in word meaning to mature listeners, monolingual infants appear unable to ignore these kinds of irrelevant phonetic information until between 9 and 10.5 months. Further, more complex variability such as speaker accent has been shown to disrupt monolingual infants' recognition of familiar words until 12 to 13 months (Schmale & Seidl 2009; Schmale et al., 2010).

The current studies were the first to test whether bilingual infants might be advantaged in ignoring irrelevant phonetic variation in order to generalize wordforms across speakers with different accents. Based primarily on previous research demonstrating that bilinguals sometimes ignore even phonemic contrasts (e.g., Bosch & Sebastián-Gallés, 2003a; 2003b), and that exposure to accented speech helps monolingual toddlers to generalize words across differently accented speakers (Schmale et al., 2011; Schmale et al., 2012; White & Aslin, 2011), the primary hypothesis presented in this thesis was that bilingual infants would accommodate for accented speech earlier than their monolingual

peers. It was also predicted that if bilinguals were advantaged in accommodating for accented speech, then this would be related to exposure to the specific languages (and accents) in their learning environment. If this were true, then French-English bilinguals should accommodate only for an accent that reflects the imposition of the phonology of one of their languages onto another one of their languages (i.e., French-accented English).

The results of these studies suggest that French-English bilingual infants did indeed generalize wordforms across French-accented and native English productions earlier than their monolingual peers. While both 9- and 13-month old bilinguals succeeded in this task, neither group of monolinguals did so. Follow up regression analyses suggested that the percentage of time bilingual infants were exposed to English as opposed to French was related to performance in the task, although this relationship did not reach significance ($p = .06$). Still, the trend suggested that greater exposure to English was related to poorer performance, while greater exposure to French was related to better performance. This supported the notion that accent accommodation was facilitated by exposure to the phonology of the language from which the accent was derived.

Had there been no correlation between task success and exposure to English, one might have argued that accent accommodation relied on a general cognitive advantage afforded to bilinguals. In such a case, one might expect this ability to transfer to accommodating for an accent that was rooted in an unfamiliar phonological system (i.e., Mandarin). Alternatively, bilinguals with less experience with English may have tolerated a change in wordforms simply due to inexperience with English. Imagine listening in on a conversation spoken in a foreign language; it might be possible to pick out some frequently repeated words, however the abundance of novel sounds might make it very difficult to notice subtle variation around these words. If this were the case, then one would predict a

similar pattern of overall success in accommodating for a Mandarin accent, as well as a consistent relationship between task success and exposure to English.

The results of Study 3 suggest that as a group, 9-month-old bilinguals were not able to generalize wordforms across native and Mandarin-accented English wordform productions. This supports the hypothesis that accent accommodation is input specific. However, the results of follow up regression analyses suggest that although bilingual infants failed in accommodating for the phonetically unfamiliar accent, the percentage of time infants were exposed to English was related to task success. This relationship followed an inverted U shaped pattern suggesting that bilinguals who were less dominant in either language, and therefore more “balanced” bilinguals, tended to demonstrate the greatest degree of wordform generalization as evidenced by greater discrimination between passages containing familiarized versus non-familiarized target wordforms.

Therefore, it appears that when infants hear accented speech that is consistent with the phonological systems that they are learning, then they 1) succeed as a group in accommodating for the accent and 2) are more successful if they tend to be more frequently exposed to the language of the “imposing” accent (i.e., French) than to their other language (i.e., English). On the other hand, when bilingual infants are tasked with accommodating for an accent that is phonetically unfamiliar, it tends to be more difficult overall (hence the group failure), but success peaks when infants are more or less equally exposed to two languages. Although this finding was unexpected, it is supported by bilingualism research in other domains of infant cognition. As described in Chapter 1, several avenues of research suggest that bilinguals may be advantaged in certain tasks measuring executive function, or the ability to monitor and control cognitive processes. In particular, bilinguals appear advantaged in tasks that require the ability to ignore salient but irrelevant perceptual cues.

An example of this advantage in infancy comes from work comparing monolingual and bilingual infants' ability to generalize memories for actions across conflicting perceptual cues.

Brito and Barr (2012) recently demonstrated that degree of bilingualism is related to memory flexibility in infancy. Memory flexibility refers to the ability to apply past experiences to future situations that are not perceptually equivalent to initial learning conditions. For example, an infant who learns to push the nose of a stuffed teddy bear to hear a sound would demonstrate memory flexibility if he or she later also pressed the nose on a toy dog in anticipation of the same response. Monolingual infants typically do not generalize actions across contrasting visual cues until 12-21 months depending on the degree of contrast and delay between initial learning and testing. However, based on evidence that bilingual children are sometimes advantaged in nonlinguistic tasks that require cognitive flexibility (Adi-Japha, Berberich-Artzi & Libawa, 2010; Bialystok & Senman, 2004), Brito and Barr (2012) predicted that bilinguals would be particularly adept in a task measuring memory flexibility.

To test this hypothesis, the authors first exposed monolingual and bilingual infants to a demonstration whereby an experimenter performed several actions on a hand puppet. Following a 30-minute delay, infants were presented with a novel puppet. The number of target actions the infants imitated from the previous session was recorded. Although monolingual infants did not generalize target actions across puppets of different colours and shapes, bilingual infants did. This suggests that bilinguals are able to ignore irrelevant visual cues in order to generalize learning across perceptually contrastive situations. Importantly, the authors also found that degree of bilingualism predicted memory flexibility; balanced exposure to two languages was related to greater success in the task.

If bilingual infants' superior memory flexibility applies not only to visual but also to auditory cues, this might help explain why balanced bilinguals performed best in accommodating for a phonetically unfamiliar accent. Schmale and Seidl (2009) suggest that 13-month-old infants succeed in generalizing wordforms across speakers with different accents because their mental representations of wordforms are more flexible than those of younger infants. Thus, in the face of conflicting but non-meaningful suprasegmental and subphonemic variation, 13-month-old monolinguals in their study succeeded in recognizing familiar but phonetically altered wordforms. Evidence from Brito and Barr (2012) applied to a linguistic task would suggest that bilinguals should be advantaged in recognizing words that are generally familiar but which differ based on irrelevant acoustic cues. This should be especially evident for more or less "balanced" bilinguals. Therefore, memory flexibility may have driven the relationship between degree of bilingualism and accommodation of a phonetically unfamiliar accent. Furthermore, although bilinguals in Study 2 demonstrated a different relationship between task success and degree of bilingualism, it cannot be discounted that memory flexibility added to overall success, despite that it may have been masked by an advantage afforded to bilinguals who had increased experience with the phonology of the accent under investigation.

Broader implications.

In several previous studies, bilingual infants have appeared insensitive to close phonetic contrasts (Bosch & Sebastián-Gallés, 2003a; 2003b), and incapable of learning similarly sounding words (Fennell, Byers-Heinlein, & Werker, 2007) at the same age as their monolingual peers. Until quite recently, these results have been explained by suggesting that the complexity of the bilingual learning environment leads to delays in bilingual infants' ability to succeed in such tasks (see Sebastián-Gallés, 2010). More recently, however,

evidence has suggested that bilingual infants may indeed be capable of discriminating between close phonetic contrasts (Albareda-Castellot et al., 2011) but that traditional testing methods are inefficient for detecting this ability in bilinguals. Specifically, methods that first familiarize or habituate infants with a particular stimulus (e.g., “/e/”) and then require a preferential response (e.g., turning one’s head longer to hear /ε/) give the illusion that bilingual infants cannot distinguish between these two phonemes, despite that they belong to at least one of their native languages. Evidence using reward-contingency paradigm (e.g., AEM; hear “/e/” -> look right to see puppet, hear /ε/ -> look left to see puppet), have however suggested that while bilingual infants may not respond preferentially to a change in such sounds, they can learn to predict the location of a reward based on these distinctions. This suggests that bilingual infants must be capable of discriminating between the two sounds (Albareda-Castellot et al., 2011).

Within this thesis, it has been argued (as elsewhere; Sebastián-Gallés, 2010) that bilingual infants are capable of discrimination between sound contrasts similarly to their monolingual peers, but are distinguished by their tolerance for acoustic variability. The evidence put forth by the studies herein is suggestive of a general tendency for bilingual infants to ignore irrelevant phonetic information (i.e., accent). This means that researchers designing studies to compare monolingual and bilingual language acquisition must be particularly attentive to whether results reflect differences in language learning and processing, or differences in response to task demands.

In contrast to earlier suggestions that bilingual infants are delayed in this area of language acquisition, here it is suggested that such tolerance may in fact be beneficial in a bilingual learning environment. Specifically, bilingual infants’ tolerance for subtle acoustic differences may help them solve the problem of breaking into two languages simultaneously.

This problem is widely known as the speech segmentation problem, and was discussed in the opening chapter of this thesis. Recall that the speech segmentation problem arises from the fact that there are no reliable pauses between words in running speech. Given this, infants must learn from other cues in order to “break into” the speech stream and begin extracting wordforms. One of the ways that infants do this is by tracking transitional probabilities (Pelucchi et al., 2009). For example in the sentence “**your** mother’s pretty”, the /r/ in “your” is preceded by an /o/ sound and followed by an /m/. While English-learning infants may frequently hear /o/ and /r/ adjacent to one another, they would rarely hear /r/ followed by /m/. Being sensitive to these probabilities, could allow infants determine that there is a boundary between “your” and “mother”, and evidence has suggested that this is the case (Pelucchi, Hay, & Saffran, 2009).

The fact that infants learn to segment speech within the first months of life is impressive. Now imagine a situation where you must learn the statistical regularities of not one but two languages. This is the challenge that bilingual infants face, and yet they appear capable of segmenting speech similarly to their monolingual peers, since they typically produce their first words around the same time (Pettito et al., 2001). Perhaps being tolerant of subtle variations surrounding wordforms allows infants to use the wordforms they have extracted from one language, to learn the regularities of another language. For example, imagine an infant who has learned to extract “ball” from an English speech stream. Perhaps having already extracted this word from the English speech stream later helps him or her to recognize “balle” in a French speech stream. If so, this may provide an additional mechanism for detecting the beginning and endings of words. For instance, now hearing the phrase “la balle est belle”, perhaps the infant will infer a break between “balle” and “est” since “balle” is familiar due to its similarity to “ball”. If on the other hand, “balle” and

“ball” were perceived as distinct then no such facilitation would be expected. Although this remains purely speculative at this point in time, further research could examine whether the presence of cognates across languages facilitates speech-segmentation. In the next section, other areas of limitation and suggestions for future research are discussed.

Limitations and areas for future research.

Because females have a tendency to outperform males in some linguistic measures, such as total vocabulary (Fenson et al., 1994), language researchers frequently attempt to maintain an equal proportion of males to females within experiments in order to generalize findings to the population. In the current studies the ratio between males and females was not equal. This was particularly true for the 13-month-olds in Study 2, where out of 12 participants in each group, 3 monolinguals and 2 bilinguals were female. However, other research using a similar task with similarly aged infants have shown no relationship between performance and gender (e.g., Polka & Sundara, 2012). Since the current task occurs at the wordform-level rather than the word-level it seems unlikely that gender would have had a significant effect. Notably, the proportion of males to females across linguistic groups at 13-months was roughly equivalent, and when gender was added as a factor in GLM it was non-significant. Furthermore, the number of females was greater in the monolingual 9-month-old group than in the same aged group of bilinguals. If gender were a contributing factor it should have aided rather than hindered monolingual infants.

An alternative explanation for bilingual 9-month-olds’ inability as a group to accommodate for the Mandarin accent may be that the Mandarin speaker imposed a greater degree of variability than the French-accented speaker. Since the purpose of the second study was to evaluate bilinguals’ performance in accommodating for a phonetically unfamiliar accent, Mandarin was selected for its distance from French and English in terms of

geographic origin. This distance, and resulting phonetic deviations, may make the Mandarin speaker's accent more pronounced. Furthermore, the French-accented speaker reported learning English in primary school whereas the Mandarin speaker did not learn English until high school. This also suggests that the Mandarin speaker might have had a more salient accent. Future research may wish to vary degree of accentedness, perhaps by testing stimuli with adult raters, to determine whether this may have disrupted bilinguals' performance.

Finally, measuring whether parents spoke with an accent based on self-report and by interpreting language histories proved difficult and likely unreliable. In several cases, native-English speaking experimenters perceived parents as speaking with an accent, but this was not reflected in either measure of infants' exposure to accented speech. Furthermore, the current study only aimed to assess whether infants' parent(s) spoke with an accent. Infants may have been exposed to accented speech from other parties (e.g., grandparents). Recent evidence has suggested that even exposure to accented speech via television may influence infants' perception of accents as early as 6-months (Kitamura et al., in press). Although it would be impossible to ask that everyone who comes into contact with the infant provide a speech sample, it would be possible to ask that at least the parent(s) of each child do so. Although perhaps a difficult quality for parents to quantify, survey measures should be broadened to ask questions concerning infants' contact with others who may speak with an accent.

Conclusion.

The studies presented in this thesis are the first to evaluate bilingual infants' ability to accommodate for accented speech. The results suggest both input specific, and possibly general cognitive advantages as driving forces in bilinguals' ability to generalize across wordforms that vary phonetically. These studies open the door to future research concerning

the nature of bilinguals' wordform representations. In particular it will be interesting to investigate whether bilinguals are also advantaged in accent accommodation at the lexical (label-referent) level. Furthermore, although the current study focused on examining an area of phonetic variability that may be more commonly encountered by bilinguals, it is possible that the bilingual advantage may extend to other forms of phonetic variability such as gender, pitch, or affect. Importantly, the results of these studies support the overarching hypothesis that bilingual infants are more tolerant of subtle acoustic differences. This suggests that the suitability of paradigms that require a change in attention to demonstrate perceptual ability or learning should be carefully considered in designing research to examine bilingual development.

References

- Abercrombie, D. (1967). *Elements of general phonetics*. Edinburgh: Edinburgh University Press.
- Adi-Japha, E., Berberich-Artzi, J., & Libnawi, A. (2011). Cognitive flexibility in drawings of bilingual children. *Child Development*, 81, 1356-1366. (doi: 10.1111/j.1467-8624.2010.01477.x)
- Albareda-Castellot, B., Pons, F., & Sebastián-Gallés, N. (2011). The acquisition of phonetic categories in bilingual infants: new data from an anticipatory eye movement paradigm. *Developmental Science*, 14(2), 395-401. (doi: 10.1111/j.1467-7687.2010.00989.x)
- Bailey, T.M., & Plunkett, K. (2002). Phonological specificity in early words. *Cognitive Development*, 17, 1265-1282. (doi: 10.1016/S0885-2014(02)00116-8)
- Ballem, K., & Plunkett, K. (2005). Phonological specificity in 14-month-olds. *Journal of Child Language*, 32, 159-173.
- Beckman, M. E., & Edwards, J. (2000). The ontogeny of phonological categories and the primacy of lexical learning in linguistic development. *Child Development*, 71(1), 240–249.
- Bergelson, E., & Swingley, D. (2012). At 6 to 9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences of the USA*, 109, 3253-3258. (doi: 10.1073/pnas.1113380109)
- Best, C. T. & McRoberts, G. W. (2003). Infant perception of nonnative consonant contrasts that adults assimilate in different ways. *Language & Speech Special Issue: Phonological Development*, 46, 183-216. (doi: 10.1177/00238309030460020701)
- Best, C. T., McRoberts, G.W., LaFleur, R., & Silver-Isenstadt, J. (1995). Divergent

developmental patterns for infants' perception of two nonnative consonant contrasts.

Infant Behavior and Development, 18, 339-350.

Best, C.T., Tyler, M.D., Gooding, T.N., Orlando, C.B., & Quann, C.A. (2009). Development of phonological constancy: toddlers' perception of native- and Jamaican- accented words. *Psychological Science*, 20, 539-542. (doi: 10.1111/j.1467-9280.2009.02327.x)

Bialystok, E. (1999) Cognitive complexity and attentional control in the bilingual mind. *Child Development*, 70, 636-644. (doi: 10.1111/1467-8624.00046)

Bialystok, E., Craik, F.I., Grady, C., Chau, W., Ishii, R., Gunji, A., & Pantev, C. (2005). Effect of bilingualism on cognitive control in the Simon task: evidence from MEG. *NeuroImage* 24, 40-9. (doi: 10.1016/j.neuroimage.2004.09.044)

Bialystok, E., & Senman, L. (2004). Executive processes in appearance-reality tasks: the role of inhibition of attention and symbolic representation. *Child Development*, 75, 562-579. (doi: 10.1111/j.1467-8624.2004.00693.x)

Boersma, P., & Weenink, D. (2013). *Praat: doing phonetics by computer* [Computer program]. Version 5.3.51, retrieved 30 May 2013 from <http://www.praat.org/>

Bortfeld, H., & Morgan, J.L. (2010). Is early word learning stress-full? How natural variability supports recognition. *Cognitive Psychology*, 60, 241-266. (doi: 10.1016/j.cogpsych.2010.01.002)

Bortfeld, H., Morgan, J., Golinkoff, R., & Rathbun, K. (2005). Mommy and me: Familiar names help launch babies into speech stream segmentation. *Psychological Science*, 16, 298-304.

Bosch, L., & Ramon-Casas, M. (2011). Variability in vowel production by bilingual

speakers: Can input properties hinder the early stabilization of contrastive categories?

Journal of Phonetics, 39, 514-526. (doi: 10.1016/j.wocn.2011.02.001)

Bosch, L., & Sebastián-Gallés, N. (1997). Native language recognition abilities in 4-month old infants from monolingual and bilingual environments. *Cognition*, 65, 33-69. (doi: 10.1016/S0010-0277(97)00040-1)

Bosch, L., & Sebastián-Gallés, N. (2003a). Simultaneous bilingualism and the perception of a language specific vowel contrast in the first year of life. *Language and Speech*, 46, 217-243.

Bosch, L., & Sebastián-Gallés, N. (2003b). *Language experience and the perception of a voicing contrast in fricatives: Infant and adult data*. In M. J. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 1987–1990). Barcelona: Causal Productions.

Bouckaert, R., Lemey, P., Dunn, M., Greenhill, S.J., Alekseyenko, A.V., Drummond, J., Gray, R.D., Suchard, M.A., & Atkinson, Q.D. Mapping the origins and expansion of the Indo-European language family. *Science*, 337, 957-960. (doi: 10.1126/science.1219669)

Brito, N., & Barr, R. (2012). Influence of bilingualism on memory generalization during infancy. *Developmental Science*, 15(6), 812-816. (doi: 10.1111/j.1467-7687.2012.1184.x)

Burnham, D., & K. Mattock. (2008). The perception of tones and phones. In O.S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning. In honor of James Emil Flege* (pp. 259-280). Amsterdam: John Benjamins.

Burns, T.C., Yoshida, K.A., Hill, K., & Werker, J.F. (2007). The development of phonetic representation in bilingual and monolingual infants. *Applied Psycholinguistics*, 28,

- 455–474. (doi: 10.1017.S0142716407070257)
- Caramazza, A., Yeni-Komshian, G., Zurif, E., & Carbone, E. (1973). The acquisition of a new phonological contrast: the case of stop consonants in French-English bilinguals. *Journal of the Acoustical Society of America*, 54, 421-428,
- Chambers, J.K. (2002). Dynamics of dialect convergence. *Journal of Sociolinguistics*, 6, 117-130.
- Chao, K.Y., & Chen, L.M. (2008). Cross-linguistic study of voice onset time in stop consonant productions. *Computational Linguistics and Chinese Language Processing*, 13(2), 215-232.
- Charles-Luce, J., & Luce, P.A. (1990). Similarity neighbourhoods of words in young children's receptive vocabularies. *Journal of Child Language*, 17, 205–215. (doi: 10.1017/S0305000900010023)
- Chen, Y., Robb, M., Gilbert, H. & Lerman, J. (2001) A study of sentence stress production in Mandarin speakers of American English. *Journal of the Acoustical Society of America* 109, 1681-1690.
- Costa, A., Hernandez, M., & Sebastián-Gallés, N. (2008) Bilingualism aids conflict resolution: evidence from the ANT task. *Cognition*, 106, 59–86. (doi: 10.1016/j.cognition.2006.12.013)
- Cristia, A., Seidl, A., Vaughn, C., Schmale, E., Bradlow, A., & Floccia, C. (2012). Linguistic processing of accented speech across the lifespan. *Frontiers in Psychology*, 3, 479. (doi: 10.3389/fpsyg.2012.00479)
- Curtin, S., Byers-Heinlein, K., & Werker, J. F. (2011). Bilingual beginnings as a lens for theory development: PRIMIR in focus. *Journal of Phonetics*, 39(4), 492-504. (doi: 10.1016/j.wocn.2010.12.002)

- Cutler, A., & Carter, D.M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language*, 2(3-4), 133-142. (doi: 10.1016/0885-2308(87)90004-0)
- Cutler, A., Sebastián-Gallés, N., Soler-Vilageliu, O., & van Ooijen, B. (2000). Constraints of vowels and consonants on lexical selection: cross-linguistic comparisons. *Memory and Cognition*, 28(5), 746-755.
- Derwing, T.M., & Rossiter, M.J. (2002). ESL learners' perceptions of their pronunciation needs and strategies. *System*, 30, 155-166. (doi: 10.1016/S0346-251X(02)00012-X)
- Docherty, G.J., *The Timing of Voicing in British English Obstruents*, Foris, New York, 1992.
- Eimas, P. D., Siqueland, E.R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, 171, 303–306. (doi: 10.1126/science.171.3968.303)
- Evans, B.G., Mistry, A., & Moreiras, C. (2007). An acoustic study of first- and second-generation Gujarati immigrants in Wembley: evidence for accent convergence? 16th *International Congress of Phonetic Sciences*. Saarbrücken, Germany.
- Fennell, C.T., Byers-Heinlein, K., & Werker, J.F. (2007). Using speech sounds to guide word learning: The case of bilingual infants. *Child Development*, 78, 1510–1525.
- Fenson, L., Dale, P.S., Reznick, J.S., Bates, E., Thal, D.J., & Pethick, S.J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, 59(5, Serial No. 242).
- Fernald, A. (2006). When infants hear two languages: Interpreting research on early speech perception by bilingual children. In P. McCardle & E. Hoff (Eds.), *Childhood bilingualism: Research on infancy through school age* (pp. 19–29). Clevedon, UK: Multilingual Matters Press.
- Fernald, A., Pinto, J., Swingley, D., Weinberg, A., & McRoberts, G. (1998). Rapid gains in

- speed of verbal processing by infants in the 2nd year. *Psychological Science*, 9(3), 228–231.
- Flege, J. (1980). Phonetic approximation in second language acquisition. *Language and Learning*, 30, 117-134.
- Flege, J. (1981). The phonological basis of foreign accent: a hypothesis. *TESOL Quarterly*, 15, 443-455.
- Flege, J. E. (1984). The detection of French accent by American listeners. *Journal of the Acoustic Society of America*, 76, 692-707.
- Flege, J. E. (1988). Factors affecting degree of perceived foreign accent in English sentences. *Journal of the Acoustic Society of America*, 84(1), 70-79.
- Floccia, C., Delle Luche, C., Durrant, S., Butler, J., & Goslin, J. (2012). Parent or community? Where do 20-month-olds exposed to two accents acquire their representation of words? *Cognition*, 124, 95–100. (doi: 10.1016/j.cognition.2012.03.011)
- Giguère, G. (2007). Collecting and analyzing data in multidimensional scaling experiments: A guide for psychologists using SPSS. *Tutorials in Quantitative Methods for Psychology*, 2, 26-37.
- Golinkoff, R.M., Hirsh-Pasek, K., Cauley, K.M., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, 14, 23–45. (doi: 10.1017/S030500090001271X)
- Green, D.W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism Language and Cognition*, 1, 67–81. (doi: 10.1017/S1366728998000133)
- Halle, P.A., & de Boysson-Bardies, B.D. (1996). The format of representation of recognized words in infants' early receptive lexicon. *Infant Behavior and Development*, 19, 463–

481. (doi:10.1016/S0163-6383(96)90007-7)
- Hochmann, J.R., Benavides-Varela, S., Nespor, M., & Mehler, J. (2011). Consonants and vowels: different roles in early language acquisition. *Developmental Science*, 14(6), 1445-1458. (doi: 10.1111/j.1467-7687.2011.01089.x)
- Houston, D.M., & Jusczyk, P.W. (2000). The role of talker specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance*, 26(5), 1570–1582. (doi: 10.1037/0096-1523.26.5.1570)
- Houston-Price, C., & Nakai, S. (2004). Distinguishing novelty and familiarity effects in infant preference procedures. *Infant and Child Development*, 13, 341-348. (doi: 10.1002/icd.364)
- Hunter M., & Ames E. (1988). A multifactor model of infant preferences for novel and familiar stimuli. In *Advances in Infancy Research*, Rovee-Collier C, Lipsitt L (eds), vol. 5. Ablex: Stamford; 69–95.
- Johnson, E.K., & Jusczyk, P.W. (2001). Word segmentation by 8-month-olds: when speech cues count more than statistics. *Journal of Memory and Language*, 44, 548-567. (doi: 10.1006/jmla.2000.2755)
- Jongman, A., & Wade, T. (2007). Acoustic variability and perceptual learning. In O.S. Bohn & M.J. Munro (Eds.), *Language experience in second language speech learning* (pp. 135–150). Amsterdam: John Benjamins.
- Jusczyk, P.W., & Aslin, R.N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, 29(1), 1–23. (doi: 10.1006/cogp.1995.1010)
- Jusczyk, P.W., Houston, D.M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, 39, 159-207. (doi: 10.1006/cogp.1999.0716)

- Kemler Nelson, D.G., Jusczyk, P.W., Mandel, D.R., Myers, J., Turk, A., & Gerken, L.A. (1995). The headturn preference procedure for testing auditory perception. *Infant Behavior & Development*, 18, 111-116. (doi: 10.1016/0163-6383(95)90012-8)
- Kitamura, C., Panneton, R., & Best, C.T. (in press). Development of language constancy: attention to native versus nonnative accents. *Child Development*, 2013. (doi: 10.1111/cdev.12068)
- Kitamura, C., Panneton, R., Deihl, M., & Notley, A. (2006). Attuning to the native dialect: when more means less. In *Proceedings of the 11th Australasian International Conference on Speech Science and Technology*, eds. P. Warren and C.I. Watson. Auckland: Australian Speech Science and Technology Association, 124-129.
- Kovacs, A.M. (2009) Early bilingualism enhances mechanisms of false-belief reasoning. *Developmental Science*, 12, 48–54. (doi: 10.1111/j.1467-7687.2008.00742.x)
- Kovacs, A.M., & Mehler, J. (2009). Cognitive gains in 7-month-old bilingual infants. *PNAS*, 100(16), 6556-6560. (doi: 10.1073/pnas.0811323106)
- Krentz, U.C., & Corina, D.P. (2008). Preference for language in early infancy: The human language bias is not speech specific. *Developmental Science*, 11, 1–9. (doi: 10.1111/j.1467-7687.2007.00652.x)
- Kuhl, P.K. (1991). Human adults and infants exhibit a “perceptual magnet effect” for speech sounds, monkeys do not. *Perception and Psychophysics*, 50, 93-107.
- Kuhl, P.K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, 9, F13–F21. (doi: 10.1111/j.1467-7687.2006.00468.x)
- Kuhl, P.K., Williams, K.A., Lacerda, F., Stevens, K.N., & Lindblom, B. (1992). Linguistic

- experience alters phonetic perception in infants by 6 months of age. *Science*, 255, 606-608. (doi: 10.1126/science.1736364)
- Lisker, L., & Abramson, A.S (1964). A cross-language study of voicing in initial stops: acoustical measurements. *Word*, 20, 384-422.
- Lloyd James, A. (1940). *Speech Signals in Telephony*. London: Pitman & Sons.
- MacLeod, A.A.N., Stoel-Gammon, C., & Wassink, A.B. (2009). Production of high vowels in Canadian English and Canadian French: a comparison of early bilingual and monolingual speakers. *Journal of Phonetics*, 37(4), 374-387. (doi: 10.1016/j.wocn.2009.07.001)
- Magen, H.S. (1998). The perception of foreign-accented speech. *Journal of Phonetics*, 26(4), 381-400. (doi: 10.1006/jpho.1998.0081)
- Malécot, A. (1960). Vowel nasality as a distinctive feature in American English. *Language*, 36, 222-229.
- Mani, N., Coleman, J., & Plunkett, K. (2008). Phonological specificity of vowel contrasts at 18 months. *Language and Speech*, 51,3–21.
- Mani, N., & Plunkett, K. (2007). Phonological specificity of vowels and consonants in early lexical representations. *Journal of Memory and Language*, 57, 252–272. (doi: 10.1016/j.jml.2007.03.005)
- Marchman, V.A., Fernald, A., & Hurtado, N. (2010). How vocabulary size in two languages relates to efficiency in spoken word recognition by young Spanish-English bilinguals. *Journal of Child Language*, 37(4), 817-840. (doi:10.1017/S0305000909990055)
- Markman, E.M., & Wachtel, G.F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, 20(2), 121-157. (doi: 10.1016/0010-0285(88)90017-5)

- Mattock, K., & Burnham, D. (2006). Chinese and English infants' tone perception: Evidence for perceptual reorganization. *Infancy*, 10, 241–265. (doi: 10.1207/s15327078in1003_3)
- Maye, J., Werker, J.F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, 82(3), 122-134. (doi: 10.1016/S0010-0277(01)00157-3)
- Metsala, J.L., & Walley, A.C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89–120). Hillsdale, NJ: Erlbaum.
- Meuter, R.F.I., Allport, A. (1999) Bilingual language switching in naming: asymmetrical costs of language selection. *Journal of Memory and Language*, 40, 25-40. (doi: 10.1006/jmla.1998.2602)
- Mills, D. L., Coffey-Corina, S., & Neville, H. J. (1997). Language comprehension and cerebral specialization from 13 to 20 months. *Developmental Neuropsychology*, 13, 397–445.
- Mulak, K.E., Best, C.T., Tyler, M.D., Kitamura, C., & Irwin, J. (in press). Development of phonological constancy: 19-month-olds, but not 15-month-olds, identify familiar words spoken in a non-native regional accent. *Child Development*. (doi: 10.1111/cdev.12087)
- Narayan, C., Werker, J.F., & Beddor, P. (2010). The interaction between acoustic salience and language experience in developmental speech perception: Evidence from nasal place discrimination. *Developmental Science*, 13, 407–420. (doi: 10.1111/j.1467-7687.2009.00898.x)

- Nazzi, T., Iakimova, G., Bertoncini, J., Fredonie, S., & Alcantara, C. (2006). Early segmentation of fluent speech by infants acquiring French: Emerging evidence for crosslinguistic differences. *Journal of Memory and Language*, 54, 283-299. (doi: 10.1016/j.jml.2005.10.004)
- Nazzi, T., Mersad, K., Sundara, M., Iakimova, G., & Polka, L. (2013). Early word segmentation in infants acquiring Parisian French: task-depend and dialect-specific aspects. *Journal of Child Language*, 10, 1-34.
- New, B., Araujo, V., & Nazzi, T. (2008). Differential processing of consonants and vowels in lexical access through reading. *Psychological Science*, 19, 1223-1227. (doi: 10.1111/j.1467-9280.2008.02228.x)
- Oda, T. (2007). Further Evidence on the Underlying Schwa of Syllabic Consonants in Present-Day English: Against the Form without Schwa. *Proceedings of JEL 2007*, 113-118.
- Oller, D.K., Eilers, R.E., Urbano, R., & Cobo-Lewis, A.B. (1997). Development of precursors to speech in infants exposed to two languages. *Journal of Child Language*, 24, 407-425.
- Palmer, S.B., Fais, L., Golinkoff, R.M., & Werker, J.F. (2012). Perceptual narrowing of linguistic sign occurs in the 1st year of life. *Child Development*, 83(2), 243-253. (doi: 10.1111/j.1467-8624.2011.01715.x)
- Pearson, B.Z., & Fernandez, S.C. (1994). Patterns of interaction in the lexical growth in two languages of bilingual infants and toddlers. *Language Learning*, 44, 617-653. (doi:10.1111/j.1467-1770.1994.tb00633.x)
- Pearson, B.Z., Fernandez, S., & Oller, D.K. (1995). Cross-language synonyms in the lexicons of bilingual infants: one language or two? *Journal of Child Language*, 22,

345-368.

- Pearson, B.Z., Fernandez, S.C., Lewedeg, V., & Oller, D.K. (1997). The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics*, 18, 41-58. (doi: 10.1017/S0142716400009863)
- Pelucchi, B., Hay, J.F., & Saffran, J.R. (2009). Statistical learning in a natural language by 8-month old infants. *Child Development*, 80, 674-685. (doi: 10.1111/j.1467-8624.2009.01290.x)
- Petitto, L.A., Katerelos, M., Levy, B.G., Gauna, K., Tetreault, K., & Ferraro, V. (2001). Bilingual signed and spoken language acquisition from birth: implications for the mechanisms underlying early bilingual language acquisition. *Journal of Child Language*, 28, 453–496. (doi:10.1017/S0305000901004718)
- Pike, K. (1945). *The Intonation of American English*. Ann-Arbor: University of Michigan Press.
- Polka, L., Colantonio, C., & Sundara, M. (2001). A crosslanguage comparison of /d/–/ð/ perception: Evidence for a new developmental pattern. *Journal of the Acoustical Society of America*, 109, 2190–2201. (doi: 10.1121/1.1362689)
- Polka, L., & Sundara, M. (2012). Word segmentation in monolingual infants acquiring Canadian English and Canadian French: native language, cross-dialect, and cross-language comparisons. *Infancy*, 17(2), 198-232. (doi: 10.1111/j.1532-7078.2011.00075.x)
- Polka, L., & Werker, J. F. (1994). Developmental changes in perception of non-native vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 421-435.
- Pons, F., & Toro, J.M. (2010). Structural generalizations over consonants and vowels in 11-

- month-old infants. *Cognition*, 116(3), 361-367. (doi: 10.1016/j.cognition.2010.05.013)
- Ramon-Casas, M., & Bosch, L. (2010). Are Non-Cognate Words Phonologically Better Specified than Cognates in the Early Lexicon of Bilingual Children? In *Selected Proceedings of the 4th Conference on Laboratory Approaches to Spanish Phonology*, ed. Marta Ortega-Llebaria, 31-36. Somerville, MA: Cascadilla.
- Ramon-Casas, M., Swingley, D., Sebastián-Gallés, N., & Bosch, L. (2009). Vowel categorization during word recognition in bilingual toddlers. *Cognitive Psychology*, 59, 96–121. (doi: 10.1016/j.cogpsych.2009.02.002)
- Reiger, C.L. (2003). Repetitions as self-repair strategies in English and German conversations. *Journal of Pragmatics*, 35, 47-59. (doi: 10.1016/S0378-2166(01)00060-1)
- Rochet, B.L. & Fei, Y. (1991). Effect of consonant and vowel context on Mandarin Chinese VOT: production and perception. *Canadian Acoustics*, 19(4), 105-106.
- Rost, G., & McMurray, B. (2009) Speaker variability augments phonological processing in early word learning. *Developmental Science*, 12(2), 339-349.
- Schepens, J., Dijkstra, T. Grootjen, F., & van Heuven, W.J.B. (2013). Cross-language distributions of high frequency and phonetically similar cognates. *PLoS ONE*, 8(5), e63006. (doi:10.1371/journal.pone.0063006)
- Schmale, R., Cristia, A., & Seidl, A. (2012). Toddlers recognize words in an unfamiliar accent after brief exposure. *Developmental Science*, 15(6), 732-738. (doi: 10.1111/j.1467-7687.2012.01175.x)
- Schmale, R., Cristia, A., Seidl, A., & Johnson, E.K. (2010). Infants' word segmentation across dialects. *Infancy*, 15, 650-662. (doi: 10.1111/j.1532-7078.2010.00032.x)
- Schmale, R., Hollich, G., & Seidl, A. (2011). Contending with foreign accent in early word

- learning. *Journal of Child Language*, 38, 1096-1108. (doi: 10.1017/S0305000910000619)
- Schmale, R., & Seidl, A., (2009). Accommodating variability in voice and foreign accent: flexibility of early word representations. *Developmental Science*, 12, 583-601. (doi: 10.1111/j.1467-7687.2009.00809.x)
- Sebastián-Gallés, N. (2010). Bilingual language acquisition: where does the difference lie? *Human Development*, 53, 245-255. (doi:10.1159/000321282)
- Sebastián-Gallés, N., & Bosch, L. (2009). Developmental shift in the discrimination of vowel contrasts in bilingual infants: Is the distributional account all there is to it? *Developmental Science*, 12, 874–887. (doi: 10.1111/j.1467-7687.2009.00829.x)
- Seidl, A., Cristia, A., Bernard, A., & Onishi, K.H. (2009). Allophonic and phonemic contrasts in infants' learning of sound patterns. *Language Learning and Development*, 5, 1991-202. (doi: 10.1080/15475440902754326)
- Shafer, V. L., Yu, Y. H., & Datta, H. (2011). The development of English vowel perception in monolingual and bilingual infants: neurophysiological correlates. *Journal of Phonetics*, 39(4), 527-545. (doi: 10.1016/j.wocn.2010.11.010)
- Shafer, V. L., Yu, Y. H., & Garrido-Nag, K. (2012). Neural mismatch indices of vowel discrimination in monolingually and bilingually exposed infants: Does attention matter? *Neuroscience Letters*, 526(1), 10-14. (doi: 10.1016/j.neulet.2012.07.064)
- Singh, L., & Foong, J. (2012). Influences of lexical tone and pitch on word recognition in bilingual infants. *Cognition*, 124, 128-142. (doi: 10.1016/j.cognition.2012.05.008)
- Singh, L., Morgan, J.L., & White, K.S. (2004). Preference and processing: the role of speech affect in early spoken word recognition. *Journal of Memory and Language*, 51, 173–189. (doi: 10.1016/j.jml.2004.04.004)

- Singh, L., White, K.S., & Morgan, J.L. (2008). Building a word-form lexicon in the face of variable input: Influences of pitch and amplitude on early spoken word recognition. *Language Learning and Development*, 4, 157 - 178. (doi: 10.1080/15475440801922131)
- Storkel, H.L. (2002). Restructuring of similarity neighbourhoods in the developing mental lexicon. *Journal of Child Language*, 29(2), 251–274. (doi: 10.1017/S0305000902005032)
- Streeter, L.A. (1976). Language perception of two-month old infants shows effects of both innate mechanisms and experience. *Nature*, 250, 39–41. (doi: doi:10.1038/259039a0)
- Sundara, M., Polka, L., & Genesee, F. (2006). Language-experience facilitates discrimination of /d– ð/ in monolingual and bilingual acquisition of English. *Cognition*, 100(2), 369–388. (doi: 10.1016/j.cognition.2005.04.007)
- Sundara, M., Polka, L., & Molnar, M. (2008). Development of coronal stop perception: bilingual infants keep pace with their monolingual peers. *Cognition*, 108, 232–242. (doi:10.1016/j.cognition.2007.12.013)
- Sundara, M., & Scutellaro, A. (2011). Rhythmic distance between languages affects the development of speech perception in bilingual infants. *Journal of Phonetics*, 39(4), 505-513. (doi: 0.1016/j.wocn.2010.08.006)
- Swingle, D. (2003). Phonetic detail in the developing lexicon. *Language and Speech*, 46, 265–294. (doi: 10.1177/00238309030460021001)
- Swingle, D. (2005). 11-month-olds' knowledge of how familiar words sound. *Developmental Science*, 8, 432–443. (doi: 10.1111/j.1467-7687.2005.00432.x)
- Swingle, D. (2009). Onsets and codas in 1.5-year-olds' word recognition. *Journal of*

- Memory and Language*, 60, 252-269. (doi:10.1016/j.jml.2008.11.003)
- Swingle, D., & Aslin, R.N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, 76, 147–166. (doi: 10.1016/S0010-0277(00)00081-0)
- Swingle, D., & Aslin, R.N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, 13, 480–484. (doi: 10.1111/1467-9280.00485)
- Tincoff, R., & Jusczyk, P.W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science*, 10(2), 172-175. (doi: 10.1111/1467-9280.00127)
- Trehub, S.E. (1976). The discrimination of foreign speech contrasts by infants and adults. *Child Development*, 47, 466–472.
- Trubetzkoy, N. S. (1969). *Principles of phonology*. (C. Baltaxe, Trans.). Berkeley: University of California Press. (Original work published in 1939).
- Tsao, F.M., Liu, H.M., Kuhl, P. K. (2006). Perception of native and non-native affricate-fricative contrasts: Cross-language tests on adults and infants. *Journal of the Acoustical Society of America*, 120, 2285-2294.
- Vihman, M.M., Thierry, G., Lum, J., Keren-Portnoy, T., & Martin, P. (2007). Onset of word form recognition in English, Welsh, and English-Welsh bilingual infants. *Applied Psycholinguistics*, 28, 475–493. (doi:10.1017/S0142716407070269)
- Vouloumanos, A., & Werker, J.F. (2007). Listening to language at birth: evidence for a bias for speech in neonates. *Developmental Science*, 10(2), 159-171. (doi: 10.1111/j.1467-7687.2007.00549.x)
- Wells, J.C. (1982). *Accents of English: An introduction*. Melbourne: Cambridge University Press.

- Wells, J.C. (1995). "New Syllabic Consonants in English," *Studies in General and English Phonetics: Essay in Honour of Professor J.D. O'Connor*, ed. By J. Windsor Lewis, 401-412, London: Routledge.
- Werker, J.F., Byers-Heinlein, K., & Fennell, C.T. (2009). Bilingual beginnings to learning words. *Philosophical Transactions of the Royal Society B*, 364, 3649-3663. (doi:10.1098/rstb.2009.0105)
- Werker, J.F., & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language Learning and Development*, 1(2), 197-234.
- Werker, J.F., Gilbert, J.H.V., Humphrey, K., & Tees, R.C. (1981). Developmental aspects of cross-language speech perception. *Child Development*, 52, 349-355.
- Werker, J.F., & Tees, R.C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7, 49-63. (doi: 10.1016/S0163-6383(02)00093-0)
- White, K.S., & Aslin, R.N. (2011). Adaptation to novel accents by toddlers. *Developmental Science*, 14(2), 372-384. (doi: 10.1111/j.1467-7687.2010.00986.x)
- White, K.S., & Morgan, J.L. (2008). Sub-segmental detail in infants' early lexical representations. *Journal of Memory and Language*, 59, 114-132.
- Yoshida, K.A., Fennell, C.T., Swingle, D., & Werker, J.F. (2009). Fourteen-month-old infants learn similar sounding words. *Developmental Science*, 12, 412-418. (doi: 10.1111/j.1467-7687.2008.00789.x)
- Zhang, Y., Nissen, S.L., & Francis, A.L. (2008). Acoustic characteristics of English lexical stress produced by native Mandarin speakers. *Journal of the Acoustical Society of America*, 123(6), 4498-4513. (doi: 10.1121/1.2902165).

Appendix A: Passages Containing Target Words

Candle Passage	The candle in the kitchen was almost melted. So Annie bought another candle at the stationary store. She came home and put away the old candle. Fran gave that candle to you later. Then she made a place for the new big candle. Your candle is very pretty and smells nice too.
Ticket Passage	Your ticket lies just over the hill. Far away from here near the sea is an old ticket. People from the ticket like to fish. Another ticket is in the country. People from that ticket really like to farm. They grow so much that theirs is a very big ticket.
Kinship Passage	Your kinship is in a faraway place. The prince used to sail to that kinship when he came home from school. One day he saw a ghost in this old kinship. The kinship started to worry him. So he went to another kinship. Now in the big kinship he is happy.
Matches Passage	The matches saw you the other day. They're much younger than the old matches. I think your matches are very nice. He showed other matches your pretty picture. Those matches thought you grew a lot. Maybe someday you'll be big matches.

Appendix B: Language Exposure Questionnaire

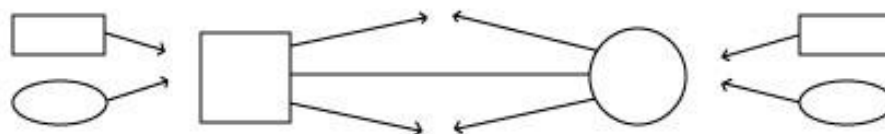
Language Exposure Questionnaire

Personal Data

NAME OF PATIENT _____
 NAME OF BABY _____
 DATE OF BIRTH _____ AGE: _____ MONTHS _____ DAYS
 DATE OF EXPERIMENT _____

Language Environment

Languages spoken by family



Are there other people living in your home? yes ☐ no ☐
 If yes please include them in the table below.

Caretakers/Friends Of Your Baby

(e.g. parents, grandparents, siblings, playmates, etc.)

Who	Languages Spoken	Since When?	Days/Week	Hours/Day

Daycare

Since _____ Hours/Day _____ Languages Spoken _____

AMOUNT OF CONTACT WITH PEOPLE WHO SPEAK DIFFERENT LANGUAGES THAN THOSE SPOKEN AT HOME

Please estimate the percentage of time your baby is exposed to each language,
 including all sources except for TV or Radio: _____%
 _____%

Accent Exposure:

Mom _____ (yes, no, or n/a for each language) Describe: _____

Dad _____ (yes, no, or n/a for each language) Describe: _____