

**FUNCTIONAL SPECIALIZATION OF PHONOLOGICAL PERCEPTION:  
HOW BILINGUALISM MODIFIES NEURAL ORGANIZATION**

DR. SHIRLEY M. CHEUNG

This thesis is submitted in fulfillment of the requirements for the degree of  
Doctor of Philosophy

Lancaster University  
Department of Psychology

September 2019



### Declaration

I declare that this thesis is entirely my own work completed under the supervision of Professor Gert Westermann and Dr Silke Brandt, and that it has not been submitted, in whole or in part, for the award of a higher degree elsewhere.

A handwritten signature in black ink, appearing to be 'J. Westermann', written above a horizontal line.

Signature

18.07.2019

Date

## Acknowledgments

I would first like to thank my supervisors Professor Gert Westermann and Dr Silke Brandt for their continuous support and guidance throughout my studies. I especially want to thank Gert for teaching me how to look at things on the brighter side and that no obstacle should be a reason to despair. I really could not have asked for a more encouraging supervisor, and I am so grateful that you have taught me how to always find the positive out of any situation. Also a big thank you to Dr Eugenio Parise for working with me during the design of our first fNIRS study and for taking the time to teach me the basics of MATLAB.

The work presented in this thesis would not have been possible without all of the families that participated in my studies, as well as the Leverhulme Trust for funding my research and training. Additionally, I greatly thank the staff at the Cognitive Neuroimaging Centre (CoNiC) at Nanyang Technological University for their generous hospitality and help during my three-month visit for data collection in Singapore.

I would not have made it to this point without all of the friends that I've made during my time in Lancaster. I am particularly deeply grateful for my friendships with Elisa Roberti, Esther Asubiojo, Irina Tache, Christina Winters, and Marina Loucaides. Thank you for your sisterhood and great advice throughout the years. Thank you for listening to me, understanding me, loving me, celebrating with me, and most importantly, believing in me. I want to also give a very special mention to my greatest friend from day one, Marina, and that I feel so lucky to have met you and done this side-by-side with you. I am so grateful to have gotten to know all of you and to have each of you as my closest friends. I am so fortunate to have you all in my life.

The greatest thank you goes to my family for watching me from afar and cheering me on. You all have never failed to believe in me even during the times I

found it almost impossible to believe in myself. This alone has kept me going until the end, and I will always be thankful to you for that. I especially want to thank my Mother for her support during my final year. You have always told me to go after my wildest dreams and that you would support me every step of the way. During the final year I finally realized what you had meant, as I saw you were doing exactly just that. Thank you for your unconditional support. You are the best Mother and role model I could have asked for.

Lastly, I would like to express my utmost gratitude for my undergraduate supervisor, mentor, and friend, Professor Pui Fong Kan, for giving me my first opportunity to explore the world of scientific research. Your selfless guidance and support throughout my undergraduate thesis up until now has allowed me to manifest my biggest dream in submitting this very thesis for a PhD.

## **Abstract**

It is commonly known that infants undergo perceptual narrowing within the first year of life, where the universal perception of all speech sounds becomes attuned to the native language by 12 months of age. Most of this knowledge has come from studies of monolingual infants, however, over half of the world's children are born into families that use a second language, and the effects of acquiring two phonological systems on perceptual development has not been thoroughly investigated. It is possible that the onset of perceptual narrowing for bilingual infants might occur later than monolinguals due to having to learn information from two separate and possibly overlapping phonological systems. As a result, bilinguals might retain flexibility in non-native phonemic processing while their monolingual peers have already attuned to the native language. This occurrence is referred to as the Perceptual Wedge Hypothesis, and to investigate this, we used functional near-infrared spectroscopy (fNIRS) to explore the brain regions associated with native and non-native phonemic processing in English monolingual and Mandarin-English bilingual infants and adults. The stimuli used for the current series of studies comprised of minimal word-pair contrasts from English (consonant), Mandarin (lexical tone), and Hindi (dental/retroflex) languages.

Chapter 3 assessed monolingual and bilingual adults to provide an initial framework of neural activation to the three contrasts and whether language experience functionally modulated cortical structures in adulthood. From actively listening to the Hindi contrast, monolinguals and bilinguals recruited different brain regions to process the non-native contrast. It is possible that language experience influences the recruitment of executive brain regions to manipulate perceptual information.

In Chapter 4, we tested a sample of younger monolingual infants between 5 and 7 months of age to assess brain activation patterns during universal phonological

perception. Our results showed that there were no differences in brain activation patterns across all languages, and that neural activity was uniformly localized in the left inferior parietal cortex (sensorimotor interface) and the right superior temporal cortex. These neurophysiological findings complement previous behavioral research by demonstrating that infants at 5-7 months of age perceive phonological information on a universal, acoustic basis. The sensorimotor aspect of speech perception at that age is discussed.

Following this, Chapter 5 investigated 10-12-month-old monolingual and bilingual infants to assess whether bilingualism prolongs the onset of perceptual narrowing (i.e. the Perceptual Wedge Hypothesis). We expected that only bilinguals would exhibit flexibility in speech perception by remaining able to discriminate the non-native phonemic contrast at this age. We were able to confirm our hypothesis by finding that bilinguals showed significant differential activation to Hindi in the left inferior frontal cortex.

Lastly, Chapter 6 compared brain activation between younger and older monolingual infants from Chapters 4 and 5 to assess developmental changes in brain activation during phonological processing across the first year of life.

The exploratory nature of the work presented in this thesis shows how acquiring two phonological systems from birth affects phonemic perception across infancy and in adulthood. Whereas monolingual and bilingual infants use the same cognitive resources to acquire language, bilinguals need to allocate their limited resources across two language systems. The present research stands as a demonstration for the complexities of dual language acquisition that bilingual infants may face, and how it can affect the recruitment of cortical regions during the perception of phonological units.

## Table of Contents

Declaration.....	ii
Acknowledgments.....	iii
Abstract.....	v
Table of Contents.....	vii
List of Tables .....	x
List of Figures .....	xi
Terminology.....	xiv
Abbreviations.....	xv
Dedication .....	xvi
Epigraph.....	xvii
<b>CHAPTER 1: Introduction.....</b>	<b>1</b>
Bilingual language acquisition: Significance and why it needs to be studied .....	2
The emergence of categorical perception of phonetic boundaries.....	3
Universal speech perception .....	5
Perceptual narrowing is driven by language-specific experience.....	8
Hemispheric lateralization of language .....	11
Characteristics of Mandarin lexical tone .....	14
Perception of lexical tones .....	16
Neural organization of native-language phonemic category perception .....	20
Speech perception is a sensorimotor task .....	21
The language processing neural network.....	23
Native-language neural commitment (NLNC) hypothesis .....	24
Phonological acquisition in bilinguals develops differently from monolinguals .....	27

Bilingualism facilitates perceptual plasticity .....	29
Research objectives of the thesis .....	31
<b>CHAPTER 2: fNIRS Methodology .....</b>	<b>36</b>
The discovery of fNIRS and the first study on infants .....	37
The hemodynamic response function (HRF) .....	38
Basic principles of fNIRS .....	40
Advantages and limitations of using fNIRS in infant language research .....	42
Variation in the HRF response.....	44
Stimulus complexity .....	45
Stimulus presentation / Experimental paradigm .....	46
Infant fNIRS paradigms.....	48
Data pre-processing .....	52
Determining time windows for statistical analysis .....	54
The current series of experiments .....	56
<b>CHAPTER 3: Monolingual and bilingual adults show language-specific responses to native and tonal contrasts but respond differently to the non-native contrast .....</b>	<b>58</b>
<b>CHAPTER 4: Infants show enhanced sensorimotor activation and acoustic processing of native and non-native speech during universal language perception .....</b>	<b>82</b>
<b>CHAPTER 5: Bilingual 10-12-month-old infants show greater sensitivity to non- native phonemic information in Broca's Area.....</b>	<b>109</b>
<b>CHAPTER 6: Changes in the organization of cortical activity between younger and older monolingual infants reveal a possible neural signature of universal language processing in the left inferior parietal cortex .....</b>	<b>140</b>



## **CHAPTER 7: The influence of bilingualism on phonological perception: A**

<b>discussion .....</b>	<b>155</b>
Objectives .....	156
The influence of language experience on cortical activation during phonemic processing across infants and in adulthood.....	156
Lexical tone processing in monolinguals and bilinguals whose two languages conflict in the use of pitch.....	159
Comparing the current work to the findings of Petitto et al. (2012).....	161
All infants.....	163
Younger vs. Older monolingual infants.....	164
Monolingual vs. Bilingual infants.....	164
Extending the work of Petitto et al. (2012).....	166
Limitations .....	167
Conclusion .....	170
<b>References.....</b>	<b>173</b>

## **List of Tables**

### **Chapter 3**

- 3.1. Frequency and intensity measures for all token stimuli.
- 3.2. HbO means for each bilateral ROI by language group.

### **Chapter 5**

- 5.1. Mean HbO values for the Mandarin ROIs in the left and right hemispheres, by condition and group.
- 5.2. Mean HbO values for the Hindi ROIs in the left and right hemispheres, by condition and group.

## **List of Figures**

### **Chapter 1**

- 1.1. Unimodal (solid line) versus Bimodal (dashed line) frequency distributions of the [da - ta] continuum. Reproduced from Maye, Werker, and Gerken (2002).
- 1.2. The steps in perceptual development that guide acquisition of the native language. Reproduced from Werker and Hensch (2015).
- 1.3. F0 contours of the four Mandarin lexical tones of [ma] spoken by a female speaker. Reproduced from Moore & Jongman (1997).
- 1.4. The dual-stream model of the functional anatomy of language. Reproduced from Hickok & Poeppel (2007).

### **Chapter 2**

- 2.1. Depiction of a typical canonical hemodynamic response function (HRF) observed in a neonate. Reproduced from Issard & Gervain (2018).
- 2.2. The effect of source-detector separation on the vertical penetration of light. Retrieved from nirx.net.
- 2.3. Testing procedure from Peña et al. (2003).
- 2.4. The grand averaged time courses of the HRF across all channels for the phonemic and prosodic conditions. Reproduced from Arimitsu et al. (2011).
- 2.5. An image taken from the HomER2 user interface showing raw data from three channels in the left hemisphere of one infant participant.

### **Chapter 3**

- 3.1. Diagram of channel placement in the left and right hemispheres over the approximate underlying cortical structures.
- 3.2. Illustration of the procedure. The block order was identical for all language conditions.
- 3.3. Anatomical illustration of the channel placement over the left and right hemispheres. The three ROIs in each hemisphere are circled in red (anterior to posterior: IFC, STG, IPC).
- 3.4. The grand averaged time course of the HbO (red) and HbR (blue) hemodynamic properties in right-hemispheric Channel 16 to English, Mandarin, and Hindi across all adult participants.

3.5. Mean HbO (red) and HbR (blue) difference values to the English consonant contrast in monolingual and bilingual adults, split across Hemisphere (columns) and ROI (rows).

3.6. Mean HbO (red) and HbR (blue) difference values to the Mandarin lexical tone/pitch contrast in monolingual and bilingual adults, split across Hemisphere (columns) and ROI (rows).

3.7. Mean HbO (red) and HbR (blue) difference values to the Hindi dental/retroflex contrast in monolingual and bilingual adults, split across Hemisphere (columns) and ROI (rows).

## **Chapter 4**

4.1. (Top) Image of an infant wearing the fNIRS headband showing the optode configuration during a testing session. (Bottom) Anatomical diagram of channel placement over the approximate underlying cortical structures.

4.2. Illustration of the procedure. The order of presentation for the auditory blocks was counterbalanced across participants.

4.3. HbO time courses for each auditory condition from Channel 14 in the right STG. The vertical line signifies the onset of auditory stimulation.

4.4. (Top) Channel placement over the left and right hemispheres. (Bottom) Projection of the channels over the approximate underlying cortical structures. The three ROIs in each hemisphere are circled in red (anterior to posterior: IFC, STG, IPC).

4.5. Mean HbO values for the English alternating (dark grey) and non-alternating (grey) conditions. Columns indicate hemispheres and rows indicate ROIs.

4.6. Mean HbO values for the Mandarin alternating (dark grey) and non-alternating (grey) conditions. Columns indicate hemispheres and rows indicate ROIs.

4.7. Mean HbO values for the Hindi alternating (dark grey) and non-alternating (grey) conditions. Columns indicate hemispheres and rows indicate ROIs.

4.8. The dual-stream model of the functional anatomy of language. Reproduced from Hickok & Poeppel (2007).

## **Chapter 5**

5.1. (Top) Image of an infant wearing the fNIRS headband showing the optode configuration during a testing session. (Bottom) Anatomical diagram of channel placement over the approximate underlying cortical structures.

5.2. Illustration of the procedure. The order of presentation for the auditory blocks was counterbalanced across participants.

5.3. Monolingual (left) and Bilingual (right) HbO time courses for each auditory condition from Channel 14 in the right STG. The vertical line indicates the onset of auditory stimulation.

5.4. (Top) Channel placement on the left and right hemispheres. (Bottom) Approximate projections onto the cortical regions of interest (circled in red).

5.5. Combined monolingual and bilingual mean HbO values for the English non-alternating and alternating conditions, split across hemispheres and ROIs,  $*p < .05$ ,  $**p < .01$ .

5.6. Mean HbO values for the Mandarin non-alternating and alternating conditions in monolinguals and bilinguals, split across hemispheres and ROIs,  $*p < .05$ ,  $**p < .01$ .

5.7. Mean HbO values for the Hindi non-alternating and alternating conditions in monolinguals and bilinguals, split across hemispheres and ROIs,  $*p < .05$ ,  $**p < .01$ .

## Chapter 6

6.1. Mean HbO values for English in younger and older English monolingual infants, split across hemispheres and ROIs. As there was no effect of condition, the alternating and non-alternating values were averaged together.  $*p < .05$ ,  $**p < .01$ .

6.2. Mean HbO values for Mandarin in younger and older English monolingual infants, split across hemispheres and ROIs. As there was no effect of condition, the alternating and non-alternating values were averaged together.  $*p < .05$ ,  $**p < .01$ .

6.3. Mean HbO values for Hindi in younger and older English monolingual infants, split across hemispheres and ROIs. As there was no effect of condition, the alternating and non-alternating values were averaged together.  $*p < .05$ ,  $**p < .01$ .

6.4. The grand averaged time courses of HbO (red line) and HbR (blue line) for Hindi (Alt – non-Alt) in younger and older infants and across all channels in the left and right hemispheres.

## Terminology

The following are definitions of the linguistic terminology and abbreviations that will be used throughout the thesis.

**Phonetics** is the physiological and acoustic study of human sounds without direct regard to their systemic patterns in a language. It is simply the study of all speech sounds used in all languages.

**Phones**, which are universal units of phonetics and thus not specific to any language, are encased in “[ ]” (e.g. [p]).

**Phonemics** or **phonology** involves the study of phonemes and their distribution and status in a particular language.

**Phonemes** are the smallest, distinctive units of sound that distinguish meaning, for example, the [b] and [p] in the English /bat/ vs. /pat/. Each language has a distinct and specific repertoire of phonemes. Phonemes that are relative to a specific language are encased in “/” (e.g. the English phoneme /p/; the French phoneme /p/).

**Allophones** are variations of a phoneme that do not distinguish meaning between words. They are dependent on their position in a word or the adjacent phone. For example, the English phoneme /p/ is pronounced with an aspirated [p<sup>h</sup>] in *pot*, but with an unaspirated [p] in *spot*.

In summary, **phonetics** is universal, whereas **phonemics** is language-specific.

## Abbreviations

<b>Alt</b>	alternating condition
<b>DPF</b>	differential pathlength factor
<b>EEG</b>	electroencephalography
<b>ERP</b>	event-related potential
<b>F0</b>	fundamental frequency
<b>fMRI</b>	functional magnetic resonance imaging
<b>fNIRS</b>	functional near-infrared spectroscopy
<b>HbO</b>	oxygenated hemoglobin
<b>HbR</b>	deoxygenated hemoglobin
<b>HbT</b>	total hemoglobin
<b>HRF</b>	hemodynamic response function
<b>IFC</b>	inferior frontal cortex
<b>IPC</b>	inferior parietal cortex
<b>IS</b>	Interactive Specialization
<b>L1</b>	first language
<b>L2</b>	secondary language
<b>MEG</b>	magnetoencephalography
<b>MMN</b>	mismatch negativity
<b>MMR</b>	mismatch response
<b>NLNC</b>	Native Language Neural Commitment Hypothesis
<b>non-Alt</b>	non-alternating condition
<b>Spt</b>	Sylvian parietal temporal
<b>STG</b>	superior temporal gyrus
<b>STS</b>	superior temporal sulcus

## **Dedication**

This work is dedicated to me – to serve as a reminder for times when I feel doubtful or discouraged, to look back and see what I can accomplish when I simply choose to believe in myself.



**Epigraph**

“Once we believe in ourselves, we can risk curiosity, wonder,  
spontaneous delight, or any experience that reveals the human spirit.”

–E. E. Cummings

# CHAPTER 1

## **Introduction**

## ***BILINGUAL LANGUAGE ACQUISITION: SIGNIFICANCE AND WHY IT NEEDS TO BE STUDIED***

Given that more than half of the world's population is born into families that use a second language (Grosjean, 2010), it is important to study bilingualism and to understand the theoretical implications of how linguistic experience can modulate cognition and the neural architecture of the brain. Although growing up bilingual or multilingual is normative in many countries, research on language development is still very much based on monolingualism. This body of research is not an accurate representation of how humans, on a global level, learn and use language in diverse and culturally integrated environments. Indeed, there is much more to be understood about the abilities of the young language-learner in acquiring two separate yet overlapping language systems each with its own phonology, syntax, lexicon, and pragmatic attributes, while using the same cognitive resources as another individual who is only learning one language system.

One of the earliest stages of language acquisition involves the learning of the sound system, or phonology, of the native language. Bilingual phonological development takes on a different trajectory than monolinguals (Ferjan Ramírez, Ramírez, Clarke, Taulu, & Kuhl, 2016; Garcia-Sierra et al., 2011), as emerging research on bilingual phonological acquisition has suggested that bilingual infants' phonological perception remain flexible and open for a longer period of time than monolinguals (Petitto et al., 2012; Singh, Loh, & Xiao, 2017). Contributing to the sparse literature on bilingual phonological processing, the current thesis used functional near-infrared spectroscopy (fNIRS) to investigate the brain basis underlying the processing and discrimination of familiar (native) and unfamiliar (non-native) phonemic contrasts in

monolingual and bilingual adults, as well as developing 5-7- and 10-12-month-old infants.

### ***THE EMERGENCE OF CATEGORICAL PERCEPTION OF PHONETIC BOUNDARIES***

Every spoken language has a unique collection of phones within an acoustic space, and all languages differ in the number of phonemic categories that they use. Over the course of development, acoustic differences that are contrastive in the native language become more salient to perceive (e.g. [t - d]), whereas non-contrastive sounds such as allophones are more difficult to distinguish (e.g. [t - t<sup>h</sup>]) (Kuhl et al., 2006). This categorical perception of language-specific phonemic categories allows the language learner to map any token to the phonology of their native language (Myers, 2014). A wealth of experimental studies has supported this developmental process from language-universal to language-specific reorganization of the language learner's phonemic repertoire. The first published study on the ability of very young infants to categorically perceive and discriminate speech sounds was conducted by Eimas, Siqueland, Jusczyk, and Vigorito (1971), where the high-amplitude sucking technique (HAS) was used to measure categorical perception. One- and four-month-old infants successfully discriminated the categorical boundary between [b] and [p], which was indicated by a rebound in sucking rate following a change in stimulus compared to baseline. Categorical perception in 1-4-month-old English-learning infants was also found for unfamiliar French nasalized vowel contrast /pa - pã/, Czech consonants /ža - řa/ (Trehub, 1976), and natural speech vowel tokens [pa - pi] and [ta - ti] (Trehub, 1973). The evidence for very young infants to discriminate unfamiliar contrasts in the absence of specific native language experience, which adults experience considerable

difficulty (Werker, Gilbert, Humphrey, & Tees, 1981; Werker & Tees, 1984), suggests that native categorical boundaries are acquired in development and sustained in adulthood.

Categorical perception allows for the discrimination of sounds that are distinct enough to fall across phonetic categories, whereas sounds that fall within a category, such as allophones, are harder to discriminate (Best & McRoberts, 2003; Eimas, Siqueland, Jusczyk, & Vigorito, 1971; Liberman, Harris, Hoffman, & Griffith, 1957). The discovery of the location of phonetic boundaries in the acoustic space is necessary for successful phonetic discrimination (Myers, 2014). The emergence of native categorical perception is driven by the statistical distribution of phonemic tokens within a language (Maye, Werker, & Gerken, 2002). It is known that the phonetic tokens of a language are not evenly distributed in the acoustic space, therefore the listener utilizes the distributional information to statistically determine phonetic boundaries between tokens from two different distributions (or categories; e.g. [b] vs. [p]) and assimilate all other tokens that fall close a distribution into the nearest category (Best, 1994). The influence of the statistical distribution of phonetic tokens on categorical perception was demonstrated by a study that tested 6- and 8-month-old infants on a [da - ta] contrast that varied along an eight-step continuum (Maye, Werker, & Gerken, 2002). Infants were familiarized with either a unimodal or bimodal distribution, in which tokens from the center of the continuum were presented more frequently in the unimodal condition and tokens from the endpoints of the continuum were presented more in the bimodal condition (see Figure 1.1). Infants who were exposed to a unimodal distribution were expected to merge the two phonetic units into a single category, whereas a bimodal exposure would separate the two. Following a 2.3-minute familiarization phase from either frequency distribution, infants were tested on their discrimination of [da] and [ta]

(i.e. Tokens 1 and 8 on Figure 1.1). The findings showed that infants from the unimodal distribution did not discriminate the contrast, whereas infants from the bimodal distribution showed discrimination characterized by a novelty preference to the non-alternating, medial trials (i.e. Tokens 3 and 6).

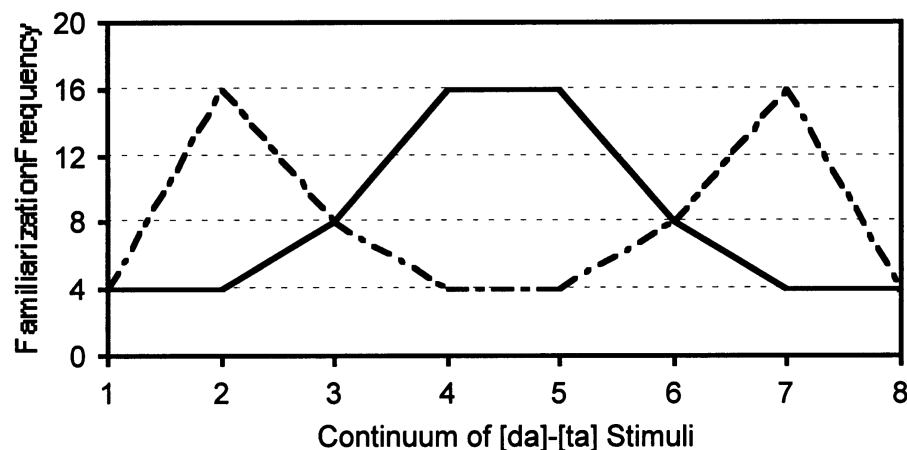


Figure 1.1. Unimodal (solid line) versus Bimodal (dashed line) frequency distributions of the [da - ta] continuum. Reproduced from Maye, Werker, and Gerken (2002).

### ***UNIVERSAL SPEECH PERCEPTION***

It is widely established that monolingual infants before 6-8 months of age have the ability to discriminate a wide range of phonetic contrasts in many, if not all, languages in the world (Bertoncini, Bijeljac-Babic, Blumstein, & Mehler, 1987; Eimas, Siqueland, Jusczyk, & Vigorito, 1971; Streeter, 1976; Trehub, 1976; Werker, Gilbert, Humphrey, & Tees, 1981; Werker & Tees, 1984). Simply put, all infants start perceiving speech on a universal basis as they have not yet acquired knowledge of their native phonological systems, and therefore they are able to distinguish contrasting speech elements from any language solely based on acoustic sensitivity. With age and

native-language experience, infants' universal sensitivities shift towards language-specific phonemic perception.

Our knowledge of universal phonetic perception is largely based upon behavioral and some ERP studies concerned with contrast discrimination abilities over the course of the first twelve months of life. For example, the English boundary between the phonemes /l/ and /r/ generate two separate semantic representations between “light” and “right”. However, the English phonemes /l/ and /r/ share the same boundary in the Japanese language, yet Japanese infants at 6-8 months of age were able to distinguish between the two categories that was only contrastive in the English language (Kuhl et al. 2006). Evidence from previous neurophysiological studies of universal phonetic perception has revealed findings in accordance with behavioral results on universal sensitivities between birth and 6-8 months of age. For example, by using a more temporally sensitive measure (i.e. electroencephalography or EEG), Rivera-Gaxiola, Silva-Pereyra, and Kuhl (2005) found that 7-month-old infants discriminated both native English and non-native Spanish consonant contrasts through examining event-related potentials (ERPs) in the form of mismatch negativities (MMNs) in a double oddball paradigm. The paradigm tested two deviant stimuli against a standard stimulus within the same experimental conditions, where MMNs were automatically elicited by the deviant or oddball stimuli (Näätänen et al., 1997). In an earlier study, the same response was observed in 6-month-old Finnish infants, where MMNs were elicited to a native Finnish vowel /*ö*/ and a non-native Estonian vowel /*õ*/ against a standard /*e*/ (Cheour et al., 1998).

An MEG investigation on brain oscillatory activity in the theta band revealed that 6-month-old infants had higher theta power for frequently presented stimuli for both native Finnish and non-native Mandarin consonant contrasts in a passive listening

oddball paradigm (Bosseler et al., 2013). It has been shown that in infants and adults, relative theta power increases when attention increases (Klimesch, 1999; Stroganova, Orekhova, & Posikera, 1998). Results from Bosseler et al. indicated that 6-month-old infants attended to frequently occurring acoustic speech signals regardless of language nativity, and that these oscillations were driven by the distributional frequency of speech events. At 12 months of age, the pattern of oscillatory activity reversed, so that older infants attended to native over non-native speech information regardless of the frequency of presentation. Finally, adults' oscillatory patterns were significantly higher for non-native phonemic units regardless of frequency of presentation. These findings suggest that 12-month-old infants and adults no longer attended preferentially to highly frequent material as they did at 6 months of age; and as seen in adults, processing non-native syllables require greater attention and cognitive effort. This was a result of having learned and established native phonemic categories, which in turn restricts their ability to learn new phonetic material.

Previous work has shown that language-specific perception of vowels emerges at 6 months of age (Cheour et al., 1998; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Polka & Werker, 1994), whereas it is seen between 6 and 9 months for consonants (Anderson, Morgan, & White, 2003). Vowels and consonants have different learning trajectories because each plays a different role in speech recognition and analysis. It has been hypothesized that the main role of vowels is to aid in the classification of the rhythmic class of languages (e.g. Byers-Heinlein, Burns, & Werker, 2010; Nazzi, Bertoncini, & Mehler, 1998). Moreover, vowels are longer in duration than consonants, which makes them more salient to perceive as they are the main carriers of intonation and stress (Nespor, Peña, & Mehler, 2003). In contrast, because there is a higher number of consonants than vowels in most language systems, consonants are relatively



more informative and therefore concern the identification of lexical elements (Nespor et al., 2003). As vowels are the main carriers of prosody and are heard more frequently in speech than consonants, they are learned more quickly as attention increases for highly frequent information (Maye, Werker, & Gerken, 2002).

### ***PERCEPTUAL NARROWING IS DRIVEN BY LANGUAGE-SPECIFIC EXPERIENCE***

Universal speech perception declines as a result of listening experience (i.e. learning the phonemic distribution of the native language), tuning the perceptual system to favor native phonological information from the environment (Bosseler et al., 2013; Maye, Werker, Gerken, 2002). Native speech perception is modulated by individual experience, in which stimuli that once were easily discriminable before 6-8 months of age become more difficult to discern by 10-12 months (Flom, 2014; Werker & Tees, 1984). This shift is often referred to as perceptual narrowing, and it is characterized by the maintenance and enhancement in the perception of native phonological information, coupled with a decrease in sensitivity to non-native information (Kuhl et al., 2006; Maurer & Werker, 2014). Ultimately, infants become native language listeners through the process of perceptual narrowing.

The onset and offset of perceptual sensitivities is driven by critical/sensitive periods, which are windows in development where a system is open to modification from external input from the environment (Werker & Hensch, 2015, Figure 1.2). The opening of these periods can be driven by biological and brain maturation factors (Peña, Werker, & Dehaene-Lambertz, 2012; Werker & Hensch, 2015), whereas time points outside of the critical/sensitive window are resistant to external influences. Perceptual narrowing is portrayed as the closing of a these periods, driven by experience, when

the system has accrued and learned enough perceptual information from its external input.

Critical and Sensitive periods explain two different developmental phenomena. While both concepts have been interchangeable in previous literature (e.g. Werker & Hensch, 2015), a distinction should be made between the two. Sensitive periods generally refer to a limited time window in development during which the effects of experience on the brain are unusually strong. Critical periods, however, are defined as a special class of sensitive periods where behaviors and their neural substrates do not develop normally if appropriate stimulation is not received during a restricted period of time (Knudsen, 2004).

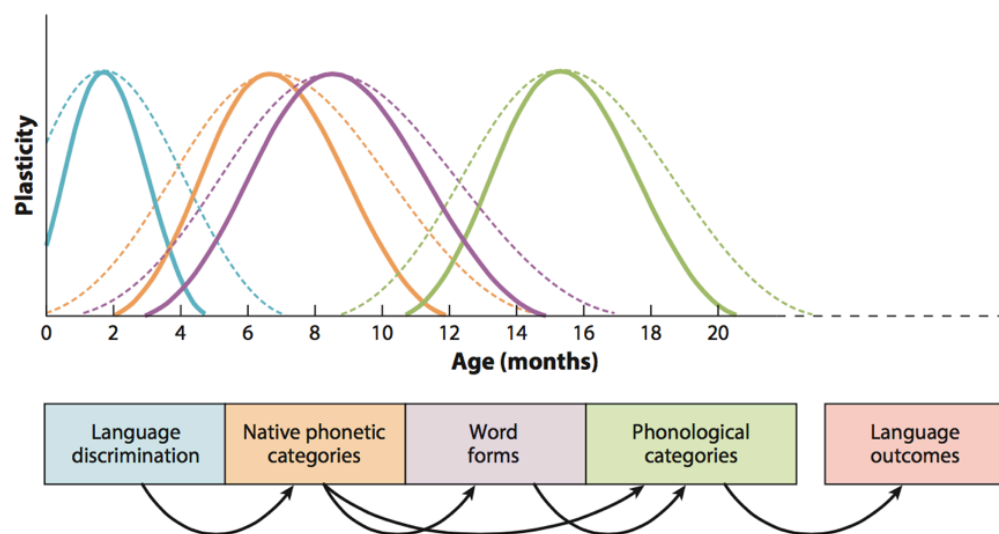


Figure 1.2. The steps in perceptual development that guide acquisition of the native language. The solid lines represent the opening and closing of sensitive periods for each step, and the dashed lines signify how these critical periods can be altered by external factors (e.g. sensory deprivation, language experience). Reproduced from Werker and Hensch (2015).

The process of perceptual narrowing is not unique to speech perception. It has also been found, for example, in facial recognition where infants' recognition ability improves for same-race faces but deteriorates for other-race faces (Kelly et al., 2007, 2009) and in musical rhythm perception where infants lose the ability to distinguish complex rhythms from unfamiliar musical styles (Hannon & Trehub, 2005a,b). Taken together, perceptual narrowing promotes a more focused and rapid learning of frequently occurring stimuli closely followed by the decline in the perception to less frequent and non-native sensory information. From an evolutionary perspective, some researchers have argued that perceptual narrowing occurs as an adaptive response to native social groups (Pascalis et al., 2014).

One of the most well-known and frequently cited study on perceptual narrowing was conducted by Werker and Tees (1984). In a series of three experiments, English infants aged 6-8, 8-10, and 10-12 months were tested using a head-turn procedure in which the infants were conditioned to turn their heads away from the experimenter and towards a loud speaker when they perceived a change in the speech sound category. The stimuli used were Salish /k<sup>2</sup> - q<sup>2</sup>/ and Hindi /t - ɖ/ manner-of-articulation contrasts. All infants at 6-8 months successfully discriminated both non-native contrasts, whereas a gradual decline in discrimination was observed with age, where 10-12-month-olds were no longer able to discriminate either contrast.

Since this first study, our understanding of the process of specialization to the native phonetic repertoire has been extended and refined. We now know that not all sensitivity to non-native phonetic contrasts is lost by the first year of life. Some non-native phonemic contrasts, such as click consonants in Zulu, a Bantu language spoken in central and southern Africa, remain discriminable across infancy and adulthood (Best, McRoberts, & Sithole, 1988). The authors explained that Zulu click contrasts

were psychoacoustically robust and therefore discriminability would be high for non-native contrasts that are unlikely to assimilate to native English categories (Best, 1994; Burnham, 1986). Rather, non-native listeners would discriminate the contrast solely on the basis of its acoustic features. In other words, it was suggested from Best et al. (1988) that phonemic perception is the process of assimilating non-native speech sounds to native categories. But if speech sounds are robust enough to resist assimilation, then the sounds are assumed to be perceived non-linguistically.

### ***HEMISPHERIC LATERALIZATION OF LANGUAGE***

If distributional information from spoken language alone is adequate to guide the learning of phonetic boundaries in the acoustic space (Maye, Werker, & Gerken, 2002), it should also be able to modify neural activation patterns as native phonemic boundaries are being learned. The brain is functionally specialized to process segmental and suprasegmental properties of speech respectively in the left and right hemispheres. Segmental properties define the rigid features of language, which involve the fast, structural, and lexical characteristics of natural speech (e.g. phones) (Minagawa-Kawai, Cristià, & Dupoux, 2011). Suprasegmental properties encompass the slower, more flexible features such as changes in prosody and how language is spoken (e.g. tone of voice).

In adults, previous research has established that the left hemisphere shoulders much of the work on phonetic discrimination (or segmental properties in general), whereas the role of the right hemisphere processes suprasegmental information, such as changes in pitch (Zatorre, Evans, Meyer, & Gjedde, 1992). In infants, functional neuroimaging studies have shown that the brain is already specialized to process speech

information asymmetrically: segmental properties of language, such as syntactic structure and unfiltered speech samples were generally processed in the left hemisphere (Gervain, Berent, & Werker, 2012; Kotilahti et al., 2010; Minagawa-Kawai et al., 2010), whereas suprasegmental properties, such as prosodic contrasts, low-pass filtered speech, and slow acoustic modulated sounds were processed in the right hemisphere (Arimitsu et al., 2011; Homae, Watanabe, Nakano, Asakawa, & Taga, 2006; Homae, Watanabe, Nakano, & Taga, 2012; Telkemeyer et al., 2009). Moreover, newborns already exhibit robust, leftward neural activation to normal speech but not to backward, unpredictable, or non-linguistic auditory information (Dehaene-Lambertz, Dehaene, & Hertz-Pannier, 2002; Dehaene-Lambertz et al., 2010; Gervain, Macagno, Cogoi, Peña, & Mehler, 2008; Kotilahti et al., 2010; Peña et al., 2003), indicating a perceptual system tuned for speech processing at birth. However, there were a few studies that did not observe left-lateralized responses to speech in infants (e.g. Ferjan Ramírez, Ramírez, Clarke, Taulu, & Kuhl 2016; Mercure et al., 2019; Petitto et al., 2012). These results could have been attributed to the smaller lexicon in infants younger than 12 months of age (Ferjan Ramírez, Ramírez, Clarke, Taulu, & Kuhl 2016; Mills, Plunkett, Prat, & Schafer, 2005), the faster maturation and stronger cerebral blood flow of the right hemisphere in young infants (Chiron et al., 1997; Leroy et al., 2011; Roche-Labarbe et al., 2012), and the right hemisphere as a novelty detector (Goldberg, Podell, & Lovell, 1994).

An interesting aspect of language is lexical tone, where changes in pitch can denote segmental and suprasegmental functions depending on whether a language is tonal or not. Tonal languages (e.g. Cantonese, Mandarin, Thai) use fluctuations in pitch to distinguish one word from another, whereas non-tonal languages (e.g. English, French, Dutch) use pitch changes for non-linguistic purposes such as communicative

intent (Hallé, Chang, & Best, 2004). A tonal contrast therefore may be perceived phonemically or prosodically depending on the listener's language background. Earlier studies have used behavioral dichotic listening paradigms to investigate hemispheric lateralization of Mandarin lexical tone in native (Mandarin monolingual, English-Mandarin bilingual) and non-native (Norwegian, American) listeners (Wang, Behne, Jongman, & Sereno, 2004; Wang, Jongman, & Sereno, 2001). Participants were presented with two different tones simultaneously in each ear, and they were instructed to identify the stimuli that they heard in each ear on an answer sheet. The results revealed that native listeners showed a significant left-hemispheric advantage by generating a higher number of correct responses in the contralateral right ear, whereas non-native listeners did not show evidence for a hemispheric advantage. In further support of these findings, a neuroimaging PET study revealed that Mandarin listeners recruited additional areas in the left hemisphere that were not activated in non-tonal English listeners during the discrimination of Mandarin lexical tone contrasts (Klein, Zatorre, Milner, & Zhao, 2001). Moreover, only the English listeners recruited the right inferior frontal region during lexical tone discrimination.

Differences in hemispheric processing were also observed in an fNIRS brain imaging study that tested non-tonal Japanese neonates with synthesized phonemic (/itta/ vs. /itte/) and prosodic (/itta?/ vs. /itta/) contrasts. The phonemic condition was a vowel substitution, whereas the prosodic contrast was a change in rising and falling pitch contour used in interrogative and affirmative statements (Arimitsu et al., 2011). Results showed a perceptual asymmetry for segmental and suprasegmental properties of speech, where the newborns exhibited a significant rightward lateralization of cortical activity in the temporal regions when presented with the prosodic contrast and bilateral activation with a stronger signal in the left parietal region to the phonemic

contrast. An earlier study that used the same experimental materials from Arimitsu et al. (2011) found that 85% of right-handed Japanese-speaking adults showed a robust, left-lateralized pattern of activation when they were presented with the phonemic contrast relative to the pitch contrast (Furuya & Mori, 2003). These findings suggest that the neonatal brain is functionally specialized to process language at birth but with more sensitivity for suprasegmental features as they are easier to process (Arimitsu et al., 2011). As age and listening experience increases, lateralization for segmental features shifts leftward and is found to occur as early as 10 months of age (Sato, Sogabe, & Mazuka, 2010).

### ***CHARACTERISTICS OF MANDARIN LEXICAL TONE***

Mandarin lexical tone is a property of spoken language that can be processed phonemically or prosodically in different hemispheres of the brain. Over half of the world speaks a tonal language (Yip, 2002). These languages use lexical tones, which are changes in fundamental frequency (F0) or pitch over a syllable word. The direction of the pitch height (high, middle, low) and contour (rising, falling, flat) assigns lexical meaning to a word. The tonal system of Mandarin Chinese is comprised of four tones that can be instantiated onto a syllable word (see Figure 1.3). Tone 1 (T1) is characterized as high level, T2 as rising, T3 as low-dipping, and T4 as falling. For example, the Mandarin word “ma” spoken with T1 means *mother*, with T2 means *hemp*, and with T3 and T4 means *horse* and *scold*, respectively (Jongman, Wang, Moore, & Sereno, 2006). In addition to F0 height and contour, each of the Mandarin lexical tones differs in temporal properties. As depicted on Figure 1.3, Tones 2 and 3 tend to be the longest, whereas Tone 4 appears to be the shortest, though the relative duration of any tone may change as a function of its position in a sentence (Nordenhake & Svantesson,

1983). Unlike segmental phonemes that change rapidly as speech is articulated such as consonants and vowels, lexical tones are recognized as a form of prosody, that is, a slower, suprasegmental form of phoneme called a tone phoneme or “toneme” (Chao, 1965).

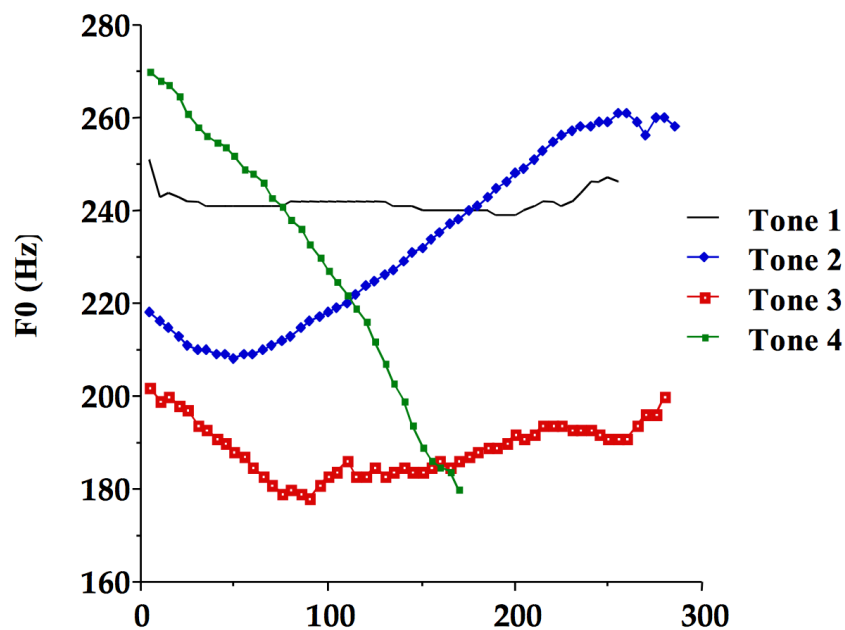


Figure 1.3. F0 contours of the four Mandarin lexical tones of [ma] spoken by a female speaker. Reproduced from Moore & Jongman (1997).

Listeners processing non-tonal languages would not need to learn the function of lexical tones. Lexical tones and changes in pitch are fundamentally identical, but their pragmatic use is determined by the rules that govern the language itself. For example, pitch changes have no lexical relevance in non-tonal languages, yet they are still significant in non-linguistic contexts (e.g. musical tones, emotional expression). Changes in pitch contours in non-tonal speech help speakers and listeners to express and detect communicative intentions and tone of voice, rather than to distinguish one word from another (Hallé, Chang, & Best, 2004). For example, a change in pitch



contour over a word can serve as a non-linguistic, pragmatic function such as altering a phrase from an interrogative to a declarative statement (e.g. “She is driving the *cà*r?” vs. “She is driving the *cà*r.”). Therefore, pitch perception is not exclusive only to tonal language learners; in fact, non-tone speakers can detect pitch changes as it is a natural application of psychoacoustic pitch perception, and they might treat these kinds of discrimination tasks pragmatically but not lexically (Burnham et al., 1996).

### ***PERCEPTION OF LEXICAL TONES***

As lexical tones are carried over syllable vowels, both typically share similar properties with each other. Previous studies have shown that the language-specific perception of lexical tones follows a similar developmental trajectory to vowels emerging between 4 and 6 months (Polka & Werker, 1994; Yeung, Chen, & Werker, 2013). Evidence has shown that native tonal language listeners performed better than non-native listeners when assessed on lexical tone discrimination (Burnham et al., 1996; Hallé, Chang & Best, 2004; Wang, Behne, Jongman, & Sereno, 2004; Wang, Jongman, & Sereno, 2001). Moreover, tone language users showed left-lateralized MMNs during lexical tone processing, suggesting that using a lexical system that implements F0 contours changes the way the perceptual system encodes pitch (Xi, Zhang, Shu, Zhang, & Li, 2010). PET brain imaging of Thai, Chinese, and English adults on their discrimination of Thai lexical tones, non-linguistic pitch changes, and Thai consonant contrasts revealed that native Thai lexical tone discrimination by Thai speakers activated Broca’s Area, whereas homologous low-pass filtered speech stimuli (i.e. non-linguistic pitch discrimination) did not elicit the same pattern of activation (Gandour et al., 2000). Interestingly, native Chinese listeners who underwent the same procedure did not exhibit activation in Broca’s Area, further supporting that the

lateralization effect when listening to Thai tonal contrasts was specific to Thai listeners only. On the other hand, children and adults who spoke non-tonal languages perceived pitch changes with greater accuracy when they were superimposed over a musical note and in low-pass filtered speech, but not over normal speech (Burnham et al., 1996). When pitch contours in speech were attenuated below normal levels, represented as filtered speech and musical notes, Burnham et al. suggested that it might have allowed non-tonal speakers to focus more on global pitch patterns such as intonation and not segmental, phonemic features.

Regarding lexical tone perception in infants, Mattock and Burnham (2006) investigated the discrimination of two pairs of Thai lexical tone contrasts (rising-falling and rising-low) in 6- and 9-month-old infants learning English or Chinese (Mandarin, Cantonese). Using the conditioned head-turn procedure, Chinese infants discriminated the lexical tone contrast equally well at both ages, whereas discriminability declined from 6 to 9 months in English infants. These results were further supported in a subsequent study that tested non-tone learning English and French infants at 4, 6, and 9 months using the same stimuli from Mattock and Burnham (2006). However, the study used a different paradigm that was the stimulus alternating preference procedure, where infants were familiarized to one tone type until a maximum duration of 30 seconds of looking time was accrued (Mattock, Molnar, Polka, & Burnham, 2008). Then the infants were tested on novel stimulation blocks with alternating tones and familiar blocks with non-alternating tones. Discrimination of the contrast was indicated by longer looking times to the novel alternating conditions. The results were in line with those from the previous study (i.e. Mattock & Burnham, 2006) in finding that non-tone learning French and English infants were able to discriminate lexical tone contrasts at 4 and 6 months, but not at 9 months. These two studies demonstrated that perceptual

reorganization of lexical tone as a result of language experience occurs by 9 months of age, but more importantly, that infants exposed to tonal languages remained sensitive to other non-native tonal contrasts.

As non-tonal speakers can discriminate changes in pitch into adulthood (Hallé, Change, & Best, 2004), to what extent can the perceptual reorganization of lexical tone be assessed? It has been suggested that the salience of a tonal contrast might have an effect on the perceptual reorganization of lexical tone in non-tone learning infants. As previously mentioned, non-tonal infants were able to discriminate Thai tonal contrasts at 4 and 6 months, but the ability was lost by 9 months (Mattock, Molnar, Polka, & Burnham, 2008). However, in a more recent study, Liu and Kager (2014) tested Dutch infants aged between 5 and 18 months on a salient Mandarin T1-T4 (flat-falling) contrast. Infants at all ages discriminated the contrast, supporting the view that acoustic salience enables non-tonal listeners to perceive these contrasts based solely on their psychoacoustic features (Burnham et al., 1996). In a subsequent study, the T1-T4 contrast was acoustically manipulated to resemble a contracted and subtler distinction. In a new sample of 5-18-month-old infants, the results revealed that only infants at 5-6 and 17-18 months discriminated the contracted contrast. In conclusion, lexical tone perception in non-tone learning infants resembled a U-shaped developmental pattern. Perceptual attenuation of lexical tone was observed at 8-9 months of age and rebounded at 17-18 months, and in some cases, as early as 14 months (Hay, Graf Estes, Wang, & Saffran, 2015). Liu and Kager (2014) speculated that the rebound might have been driven by the non-tone-learning infants' growing knowledge of their native intonation system. It is possible that non-tonal infants have learned to analyze pitch changes with pragmatic meaning, as it has been shown that English-learning infants are sensitive to prosodic information in speech (Nazzi, Nelson, Jusczyk, & Jusczyk, 2000; Seidl &

Cristià, 2008). Furthermore, it is possible that non-tonal infants would no longer treat the pitch differences as an integral part for word learning. They would not be “deaf” to changes in pitch, instead they would learn to treat those changes acoustically rather than lexically (Hallé, Chang, & Best, 2004; Hay, Graf Estes, Wang, & Saffran, 2015).

Lastly, it is important to note that context plays an important role in lexical tone discrimination (and language processing in general). Context would allow the listener to use referential and contextual cues to ascribe linguistic or non-linguistic meaning of speech. For example, testing for lexical tone discrimination in bilinguals who are learning contrasting tonal and non-tonal languages (e.g. Mandarin-English bilinguals) can be complex, as changes in pitch contour can be perceived lexically or acoustically, depending on language context. A recent study was conducted to examine English and Mandarin monolingual and bilingual 6-, 9-, 12-, and 13-month-old infants on their perception of a salient (T1-T3) and a subtle (T2-T3) Mandarin lexical tone contrast (Singh et al., 2018). At six months, Mandarin infants successfully discriminated only the salient contrast, but English infants did not. At nine months, Mandarin infants discriminated both salient and subtle contrasts, whereas English infants discriminated only the salient contrast. From 12 months and onwards, English and Mandarin monolinguals could discriminate both contrasts. Surprisingly, bilinguals failed to show discrimination at each of the four age groups. The authors suggested that the pattern of results were indicative of the varied role that pitch plays in spoken language. Additionally, the ambiguity may have been compounded in bilingual infants who were acquiring two language systems that use pitch in contrasting ways. It was suggested that traditional phonological discrimination paradigms may not be ideal to test bilingual populations as they are likely to rely on referential and contextual support as they navigate a more complex phonological space than monolinguals. Moreover, Singh et

al. (2018) demonstrated that traditional theories of perceptual narrowing do not readily apply for lexical tones. To date, there are no published studies that used neuroimaging techniques to examine the development of Mandarin lexical tone perception in monolingual and bilingual non-tone-learning and tone-learning infants during the first twelve months of life.

### ***NEURAL ORGANIZATION OF NATIVE-LANGUAGE PHONEMIC CATEGORY PERCEPTION***

It has been widely accepted from studies on adults that the bilateral posterior-superior temporal cortices constitute the main component of constructing sound-based representations of speech (Hickok & Poeppel, 2000). Encasing the primary auditory cortex, the bilateral superior temporal regions is involved with analyzing all incoming phonetic and non-phonetic auditory information before transmitting it to associative areas predominantly in the left hemisphere for subsequent linguistic processing (e.g. semantic analysis) (Zatorre, Evans, Meyer, & Gjedde, 1992). Associative areas such as the left anterior and middle superior temporal sulci were found to activate during the discrimination of familiar consonant-vowel (CV) syllables (e.g. /ba/, /da/) compared to their equivalent, non-phonemic spectrally modified equivalents (i.e. formant manipulation), suggesting that these regions play a role in phonemic perception (Liebenthal, Binder, Spitzer, Possing, & Medler, 2005). Moreover, a meta-analysis of over 100 functional imaging studies concluded that the left mid-superior temporal gyrus was consistently involved in phonemic processing (DeWitt & Rauschecker, 2012). This gradient pattern of activation across adjacent temporal regions reflects the non-uniform structure of phonetic categories (Myers, 2014). These findings suggest that fine-grained

acoustic distinctions that are characteristic of crossing phonetic boundaries take place predominantly in the left temporal regions.

While acoustic and phonetic analyses activate temporal cortices, evidence also suggests that categorical speech perception and production involves the left inferior prefrontal cortex. In an fMRI study, the left inferior prefrontal cortex was found to activate when acoustic changes in the auditory stimuli (e.g. changes in voice onset time) crossed a categorical boundary between [da] and [ta] than when it was within a category (Myers, Blumstein, Walsh, & Eliassen, 2009). The authors suggested from the findings that the left inferior prefrontal cortex is involved in the computation of category representations. For example, in order to determine that the English /d/ in “dart” and the /t/ in “tart” are from separate phonological categories, the auditory features that have been identified by temporal lobe mechanisms must be related to articulatory representations in the inferior frontal cortex. Furthermore, in another fMRI study, English-speaking participants were tested on their perception of the categorical boundary between /ba/ and /da/ along a 10-step continuum (Lee, Turkeltaub, Granger, & Raizada, 2012), and it was found that the left inferior frontal and the pre-motor areas were activated when the acoustic changes of the stimuli crossed the perceptual category that was individually predetermined by each participant before fMRI scanning.

### ***SPEECH PERCEPTION IS A SENSORIMOTOR TASK***

The involvement of pre-motor areas in speech perception has been increasingly evidenced, suggesting that speech perception is a sensorimotor task (Bruderer, Danielson, Kandhadai, & Werker, 2015; Liberman & Mattingly, 1985; Westermann & Miranda, 2004; Wilson, Saygin, Sereno, & Iacoboni, 2004). The sensorimotor

interface, known as Sylvian parietal temporal (Spt), lies within the the Sylvian fissure at the boundary between the posterior superior temporal gyrus and the inferior parietal cortex (Hickok & Poeppel, 2007). The role of the Spt is involved in coding sensory speech input and regulating the fine motor control of articulatory movements for accurate speech reproduction (Hickok, Buchsbaum, Humphries, Muftuler, 2003). Interestingly, the Spt is not speech specific, meaning that its sensorimotor responses are equally robust when participants listened to piano melodies during the sensory task and rehearsed or hummed the auditory stimuli for the motor task (Hickok et al., 2003).

Regarding developmental research, an MEG study conducted on newborns, 6-, and 12-month-olds revealed an emerging sensorimotor network in the left inferior frontal region was seen as early as 6 months of age (Imada et al., 2006). The infants were tested on their discrimination of [ta] and [pa], and the activation seen at 6 months was indicative of the developing connection between auditory and motor areas during the time in which the onset of canonical babbling occurs. Note, however, that Imada and colleagues did not collect data from the right hemisphere. Interpreting results from the left hemisphere would be more robust if contralateral activation from the right hemisphere was compared. Further evidence in support of the sensorimotor mechanism during speech perception was seen in a study of 6-month-olds, who were given teething toys that either prevented tongue-tip movement (flat teether) or allowed free movement (gummy teether) (Bruderer, Danielson, Kandhadai, & Werker, 2015). The infants were tested on their discrimination of a non-native Hindi dental-retroflex contrast, where the production typically involves the movement and placement of the tongue-tip. Using eye-tracking during an alternating and non-alternating sound presentation paradigm, Bruderer et al. found that using the flat teether impeded infants' discrimination of the dental-retroflex contrast, whereas infants tested with the gummy teether and no teether

discriminated the contrast. This study successfully demonstrated that impeding oral/motor articulators alone affects auditory/sensory perception.

### ***THE LANGUAGE PROCESSING NEURAL NETWORK***

Taken together, the bilateral superior temporal, left inferior prefrontal, pre-motor, and inferior parietal cortices all form the dorsolateral pathway of the dual-stream model of speech processing (Figure 1.4; Hickok & Poeppel, 2007). The dorsal pathway lies dominantly in the left hemisphere and is responsible for speech perception, whereas speech comprehension tasks involve the bilateral ventral pathway. In support of the functional connectivity of the nonhomologous and nonadjacent regions that make up the dorsal stream, functional resting state data have revealed a significant low frequency correlation between Broca's Area (left inferior prefrontal cortex) and Wernicke's Area (left posterior superior temporal gyrus), in which the magnitude of the correlation increased during a listening task of continuous speech from a recorded story (Hampson, Peterson, Skudlarski, Gatenby, & Gore, 2002). Moreover, the left pre-motor cortex was found to have a significant correlation with Broca's Area at rest during the listening task, supporting the sensorimotor aspect of speech perception. The findings from Hampson et al. (2002) clearly depict the functional connectivity of areas that were long implicated to be involved in language processing and support the activity of the dorsolateral network of speech perception at resting and active states.



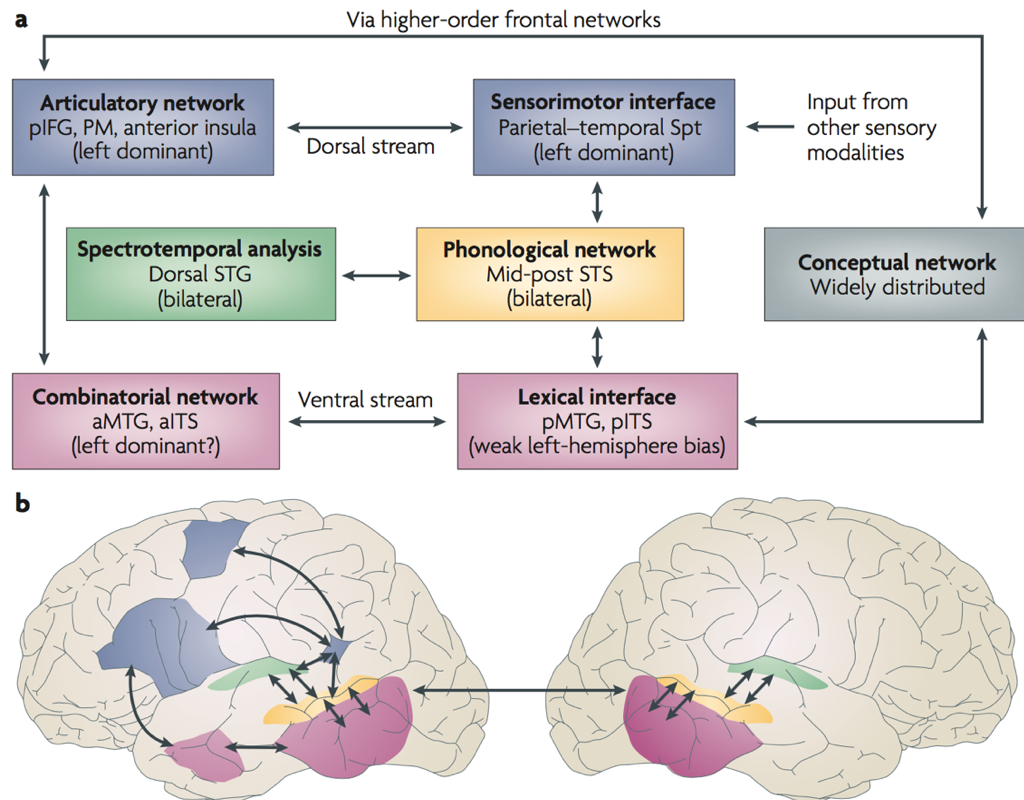


Figure 1.4. The dual-stream model of the functional anatomy of language. Reproduced from Hickok & Poeppel (2007).

### ***NATIVE-LANGUAGE NEURAL COMMITMENT (NLNC) HYPOTHESIS***

So far, native language acquisition has been reviewed with regard to the transition from initial acoustic perception to language-specific processing by which infants learn the phonemic distributions of their native language or languages. In addition, neural activity and the cortical regions that belong to the phonological-processing network were identified in infants and adults. The process of acquiring native phonology can be explained through the process of native language neural commitment, in which native phonological perception is closely linked with linguistic experience.

Linguistic experience shapes the neural circuitry to attune towards the frequent and salient information that constitute the native language. In order for native language acquisition to take place, the initial, universal stage of phonetic perception must give way to native language neural commitment. This perceptual narrowing process is explained by the native language neural commitment (NLNC) hypothesis, in which it facilitates higher language learning by shaping the brain through native linguistic experience during a critical period in language development (Kuhl, 2004; Kuhl & Rivera-Gaxiola, 2008). The NLNC works in two directions: (1) learning increases for more complex language patterns (i.e. words) that are compatible with the learned phonetic structure of the native language system, and (2) attention to non-native patterns is reduced and the learning of them is no longer facilitated. As a result, neural networks become committed and therefore more sensitive to the patterns of the native language and less committed to the alternative, thus enabling rapid and successful language learning in later ages.

Previous behavioral (Kuhl, Conboy, Padden, Nelson, & Pruitt, 2005) and ERP studies (Kuhl et al., 2008; Rivera-Gaxiola, Klarman, Garcia-Sierra, & Kuhl, 2005) have highlighted that infants' phonetic performance at an early age predicts later language outcomes. Consistent across these studies, it was found that excellent native phonemic discrimination at 7 and 11 months predicted higher language abilities such as word production, sentence complexity, and mean length of utterances between 18 and 30 months. In contrast, those who were more proficient in non-native phonemic discrimination within the first year were slower to develop native language skills at later ages. It was suggested that those infants remained at an earlier phase of development, reflecting an uncommitted neural network (Kuhl et al., 2008). Neural commitment to the native language was measured in a study using EEG, in which two

time windows (P150-250 and N250-550) were examined during the discrimination of a native English and non-native Spanish contrast in 11-month-old infants (Rivera-Gaxiola, Klarman, Garcia-Sierra, & Kuhl, 2005). All infants produced significantly larger N250-550 amplitudes only to the native contrast, whereas no differences were observed for the non-native contrast. The infants' responses to the non-native contrast were then divided into two subcategories, where the first group displayed larger P150-250 amplitudes and the second group displayed larger N250-550 amplitudes after stimulus. The results revealed that infants who responded at P150-250 to the non-native contrast at 11 months had higher scores in word production at 18, 22, 25, 27, and 30 months of age, suggesting that the earlier P150-250 component reflects a less specialized and acoustic processing of the non-native stimuli, whereas the N250-550 component resembles a more mature phonological analysis.

Another study extending previous work (Kuhl et al., 2008; Rivera-Gaxiola, Klarman, Garcia-Sierra, & Kuhl, 2005) to a bilingual population examined early brain measures of phonemic discrimination and later language abilities (Garcia-Sierra et al., 2011). This longitudinal study used EEG to examine 6-9- and 10-12-month-old Spanish-English bilingual infants' neural discrimination of both native Spanish and English phonemic contrasts, as well as examined whether word production scores in both languages at 15 months were related to early ERP responses and language exposure. Neural discrimination was correlated with vocabulary scores taken from the CDI: Words and Gestures for Spanish and English at 15 months of age (Fenson, 2007; Jackson-Maldonado et al., 2003). EEG recordings in the form of MMNs revealed that bilingual infants discriminated both native contrasts only at 10-12 months of age but not at 6-9 months. This pattern of development is divergent from monolinguals, in which they would already show native language neural discrimination between 6-9

months (Cheour et al., 1998). The results from Garcia-Sierra et al. (2011) support the NLNC hypothesis whereby neural commitment to a language is dependent on the amount of exposure to that language. Bilinguals are exposed to the same amount of total language input as monolinguals, however, that amount is divided between two languages. As such, bilinguals receive less exposure in each language and therefore may take longer to neurally commit to both of their languages. As a result, bilinguals may remain in the universal stages of language perception for a longer period of time until a sufficient amount of experience has been accumulated, thus protracting the onset of perceptual narrowing (Kuhl et al., 2008). This way allows the bilingual an advantage in mapping two languages simultaneously. Additionally, Garcia-Sierra et al. (2011) found that neural commitment, bilingual language exposure, and word production were significantly interrelated. Fifteen-month-old bilinguals who were English-dominant in word production scores showed better discrimination of the English contrast and had strong English exposure at home at 10-12 months of age. The same pattern was also found for Spanish-dominant bilingual infants.

### ***PHONOLOGICAL ACQUISITION IN BILINGUALS DEVELOPS DIFFERENTLY FROM MONOLINGUALS***

Indeed, language exposure influences later listening preferences in infants. English monolingual and English-Tagalog bilingual newborns less than five days old were tested on their ability to recognize and discriminate the two languages at the prosodic level (Byers-Heinlein, Burns, & Werker 2010). Using a high-amplitude sucking-preference procedure (HAS), the newborns were played alternating English and Tagalog sentences for ten minutes. Indicated by the higher average number of sucks, monolingual newborns showed a preference for their native language, whereas

bilinguals showed equal preferences for both languages. A follow-up study assessing discrimination of English and Tagalog in bilinguals confirmed that the infants were able to discriminate both of their native languages, and that it was not due to confusion between the two. The authors suggested that prenatal listening experience influenced listening preference in monolingual and bilingual infants at birth. More importantly, they demonstrated that infants are born already familiarized with the prosodic patterns of their native languages and that bilinguals could distinguish between the two.

At the phonemic level, bilingual infants were able to discriminate phonemic contrasts in each of their two languages at 10-12 months of age. Using a habituation procedure where infants were habituated to a medial token [pa] and tested on the English phonemic boundary /p<sup>h</sup>a/ and the French phonemic boundary /ba/, Burns, Yoshida, Hill, and Werker (2007) found that English-French bilingual infants at 10-12 months and 14-20 months of age were sensitive to the phonemic boundaries of each of their two languages, indicated by a recovery in looking time. Moreover, in an infant-controlled visual habituation procedure, bilingual French and English 10-12-month-olds readily discriminated native dental French and alveolar English variants of [d] (Sundara, Polka, & Molnar, 2008).

At the neural level, 11-month-old bilingual infants showed distinct differences from monolinguals in the perception of native phonemic contrasts. Ferjan Ramírez, Ramírez, Clarke, Taulu, and Kuhl (2016) employed whole-head magnetoencephalography (MEG) in a double oddball paradigm that used a medial token [ta] against the Spanish /da/ and English /t<sup>h</sup>a/ to assess native phonemic perception in the form of MMRs in English monolingual and Spanish-English bilingual 11-month-olds. MMR, or mismatch response, consists of an early component typically occurring between 100 and ~260 ms post stimulus, and a late component with a typical

latency of ~260-460 ms. Their findings revealed that both groups were equally sensitive to the English contrast, but bilinguals showed a stronger response to Spanish contrasts compared to monolinguals. However, when the infants' responses were split into early and late MMR time windows, bilinguals showed a significantly larger English MMR in the early time window, whereas monolinguals showed a significantly larger English MMR in the late time window. As early MMR components signify a less mature and universal encoding of information and late MMR for specialized phonetic analyses, the authors suggested that bilinguals neurally discriminated the contrasts within each of their two native languages only at the acoustic level, whereas monolinguals discriminated the English contrast at the phonemic level. These results further support that bilingual infants' sensitivity to the phonemic boundaries of both their languages suggests a slower transition from the acoustic (universal) to phonemic (language-specific) analysis of native speech (Kuhl et al., 2008). This can result from dealing with a complex workload, in which the increased amount of phonological information that bilingual infants need to learn facilitates a higher-functioning and flexible perceptual system.

### ***BILINGUALISM FACILITATES PERCEPTUAL PLASTICITY***

According to the NLNC hypothesis, it is plausible that bilinguals undergo a protracted development in acquiring native phonology. Although it appears that bilingual infants are knowledgeable with the phonemic boundaries in both of their native languages by 12 months of age (Burns, Yoshida, Hill, & Werker, 2007; Sundara, Polka, & Molnar, 2008), studies that have examined neural activation patterns have suggested that the development of native phonological systems in bilingual infants differ from that of monolingual infants, which was characterized by a slower transition

from acoustic to phonemic analysis (Ferjan Ramírez, Ramírez, Clarke, Taulu, & Kuhl, 2016; Garcia-Sierra et al., 2011). Testing bilinguals' perception of a third contrast that is outside of their native phonological systems would help elucidate whether bilinguals retain perceptual plasticity in universal phonetic discrimination, or a slower commitment to native phonology, for a longer period of time than monolinguals. This protocol, however, has only been done in a few studies.

In a visual infant-controlled habituation paradigm, 11-month-old English monolingual and English-Mandarin bilingual infants were assessed on the discrimination of a non-native Hindi dental-retroflex contrast and a native English consonant contrast, as well as own-race and other-race faces (Singh, Loh, & Xiao, 2017). It was found that bilinguals showed perceptual plasticity in the language domain by discriminating the non-native contrast, whereas monolinguals did not. The results, however, did not extend across to the perceptual domain of face recognition.

To date, there has been only one study that used neuroimaging to examine brain activation patterns in the bilateral superior temporal gyrus (STG) and inferior frontal cortex (IFC) during non-native phonemic perception in English monolingual infants and a heterogeneous sample of bilingual infants who had been receiving exposure to French, Spanish, or Chinese in addition to English. Petitto and colleagues (2012) used fNIRS to test 4-6- and 10-12-month-olds on a native English /ba - pa/ contrast and a non-native Hindi dental-retroflex /ʈa - ɽa/ contrast in an event-related oddball paradigm. Upon initial whole-brain analyses of the activation patterns, their findings showed that there was no significant difference between the left and right hemispheres across all experimental conditions. When analyzed by region of interest (STG, IFC), their findings revealed a greater right-hemispheric activation in the STG by all infants (i.e. 4-6- and 10-12-month-old monolinguals and bilinguals). In the left STG, infants of all

ages and language groups showed similar brain activation patterns. The authors suggested that activation in the left STG to all auditory stimuli is observed early and remains stable across the first year of life. In the inferior frontal regions, the right IFC showed a decrease in activation from the younger to older age groups. The authors interpreted the finding as a developmental shift in lateral dominance, characteristic of perceptual narrowing, where activity in the language network in the left hemisphere increases with linguistic experience (Sato, Sogabe, & Mazuka, 2010). Regarding the left IFC in isolation, all babies showed differences in activation between the native and non-native contrasts. English monolingual infants showed a much greater pattern of activation to their native language, whereas bilingual babies had similar activity levels for both native and non-native contrasts. This finding allowed the authors to find support for the Perceptual Wedge Hypothesis, in which exposure to more than one language “wedges” open the closing doors of native language commitment and perceptual attunement. As a result, the Perceptual Wedge allows language sensitivity to be held open for longer due to the increased neural and computational demands of bilingual language processing (Petitto et al., 2012). Simply put, Petitto and colleagues have shown through fNIRS brain imaging that bilinguals’ perceptual systems remain open at the time where monolinguals’ have already attuned to the native language.

### ***RESEARCH OBJECTIVES OF THE THESIS***

While Petitto et al. (2012) compared monolingual English infants with a heterogeneous sample of bilinguals, there is, to date, no published study that used fNIRS brain imaging to examine the neuroanatomical correlates of native and non-native phonological development in homogeneous samples of monolingual and bilingual infants. Behavioral and ERP studies lack in unveiling the specific neural



structures associated with phonemic processing. Although ERPs have been shown to be a reliable method in studying phonological processing and development (Cheour et al., 1998; Garcia-Sierra et al., 2011; Rivera-Gaxiola, Klarman, Garcia-Sierra, & Kuhl, 2005), this measure is not suitable for adequate spatial localization. As the language processing neural network spans over multiple lobules where distinct regions and hemispheres are specialized in the higher order processing of speech (e.g. acoustic vs. linguistic processing) (Hickok & Poeppel, 2007; Zatorre, Evans, Meyer, & Gjedde, 1992), cortical localization becomes paramount in localizing activation during phonological processing. The current series of experiments will address three open questions that have not been adequately assessed in the current literature.

First, as fNIRS is a relatively new neuroimaging technique, the body of research within the developmental population is steadily increasing. There has only been one fNIRS study examining the changes in cortical activation to native and non-native phonemic perception across two time points in development (Petitto et al., 2012). However, Petitto et al. measured only from the bilateral superior temporal gyrus (STG) and inferior frontal cortex (IFC) to native and non-native phonemic information in English monolingual infants and a heterogeneous sample of bilingual infants who had been receiving exposure to French, Spanish, or Chinese in addition to English. Testing a homogeneous sample of bilingual infants can eliminate potential confounds resulting from language distance. For example, the language distance between Mandarin and English is greater than the distance between Spanish and English. Moreover, examining all areas that encompass the dorsal pathway of speech processing (i.e. STG, IFC *and* IPC) would be advantageous in studying all functional aspects related to phonological processing in the speech perception network. The first objective was to follow the evolving brain activation patterns of phonological development across the first year of

life and in adulthood, specifically measuring all areas that make up the dorsal pathway and a homogenous sample of Mandarin-English bilingual infants.

Second, much research has been conducted using behavioral paradigms on lexical tone perception in tone-learning and non-tone-learning infants (Liu & Kager, 2014; Mattock & Burnham, 2006; Mattock, Molnar, Polka, & Burnham, 2008). Although there is one published study that used fNIRS to study lexical tone perception in bilingual adults (Zinszer, Chen, Wu, Shu, & Li, 2015), it has not been thoroughly assessed using fNIRS in infants. It remains in question whether fNIRS is a suitable technique in localizing hemodynamic activation in young infants during lexical tone processing. The second objective of the current work examined the neural correlates of lexical tone perception in bilingual infants and adults learning languages that contrast in the use of pitch contours.

Third, previous research assessing bilingual phonemic perception have only tested bilingual infants on phonemic contrasts from both of their native languages (e.g. Burns, Yoshida, Hill, & Werker, 2007; Ferjan Ramírez, Ramírez, Clarke, Taulu, & Kuhl, 2016; Garcia-Sierra et al., 2011; Sundara, Polka, & Molnar, 2008) but seldom on a third contrast belonging outside of the bilinguals' phonemic repertoire. A direct way to test the effect bilingualism has on perceptual flexibility is to present a non-native phonemic contrast to monolingual and bilingual infants at the age when monolinguals' phonological perception abilities start to attune to the native language (i.e. 10-12 months). If bilinguals remain sensitive to non-native contrasts when monolinguals do not, it can be inferred that the onset perceptual narrowing is protracted in bilinguals, thus providing supporting evidence for the Perceptual Wedge Hypothesis. The final objective of the current work was to replicate the findings of young monolingual and bilingual infants from Petitto et al. (2012) and extend previous findings for the

Perceptual Wedge Hypothesis. The current work was unique in its exploration in replicating and extending the findings of young monolingual and bilingual infants from Petitto et al. through assessing bilinguals' discrimination of a non-native Hindi dental-retroflex contrast on a homogeneous sample of Mandarin-English bilingual infants and implemented a more traditional block stimulation paradigm rather than an oddball paradigm used in Petitto et al.

As such, these open questions therefore provide strong theoretical motivation for the present series of experiments to locate similarities or differences in cortical activation during the processing of native and non-native speech sounds in the first year of life. And next, to investigate how evidence taken from previous behavioral and electrophysiological studies would translate to distinct areas on the cortical surface of the infant brain. In pursuit of this challenge, it would be beneficial to demonstrate how recordings of infant and adult hemodynamic activity relate to findings from previous behavioral and ERP studies. In order to achieve the two research aims that require the localization of active cortical areas, fNIRS was chosen as a viable, non-invasive neuroimaging method yielding both excellent spatial resolution and tolerance to movement artifacts. More information about the advantages, experimental design, and data analysis methods of the technique is found in the next methodological chapter. The current series of experiments is unique in its exploration of identifying the brain regions associated with the processing of a native English consonant contrast, and a Mandarin Chinese lexical tone contrast in non-native listeners as well as listeners learning two languages that differ in their use of lexical tone. Most importantly, the perception of a non-native Hindi dental-retroflex contrast was also tested. Discrimination of such contrast would unveil a phonological system that is not yet specialized to the phonemic distribution of the native language. The current thesis aims to contribute to the

increasing knowledge of bilingual language processing and its development, as well as the effectiveness of using fNIRS to study infant speech perception.

## CHAPTER 2

### **fNIRS Methodology**

***THE DISCOVERY OF FUNCTIONAL NEAR-INFRARED SPECTROSCOPY (fNIRS) AND THE FIRST STUDY ON INFANTS***

The discovery of a possible optical window to the human body was made when a cellular biologist, Frans Jöbsis-vanderVliet, was having a family dinner in 1976. He held a flat beef bone up towards a light and observed the shadow of his finger passing behind the diffuse red light coming through the bone. It was then that he realized tissues of the human body, particularly beneath the skull, could be monitored safely and non-invasively through this method (Jöbsis-vanderVliet, 1999). Following the discovery, the first prototype of a bedside machine employing laser diodes and a novel light source was completed a few years later in 1980. Since then, fNIRS has evolved to become an effective method for the non-invasive study of changes in blood-oxygen concentration levels at the surface of the cortex.

Up until the discovery of fNIRS, it was difficult to measure brain activity in awake infants using fMRI brain imaging. As fMRI requires little to no head movement for data acquisition, infants and children usually need to be sedated or asleep to undergo testing (Born, Rostrup, Leth, Peitersen, & Lou, 1996; Born et al., 1998; Martin et al., 1999). The first study that successfully measured cerebral hemodynamic changes using fNIRS on awake infants was conducted by Meek and colleagues (1998) just over twenty years ago. Brain activation in the occipital area of infants from two days to three months of age was measured in response to a visual stimulus of a flickering black and white checkerboard. Their results showed a significant increase in total hemoglobin in the occipital region of the test infants, whereas no significant changes were found in the frontoparietal region of the control infants. This study was the first to demonstrate that fNIRS is an adequate technique to safely localize hemodynamic responses without the use of invasive substances in awake, and not sedated or sleeping, infants.

### ***THE HEMODYNAMIC RESPONSE FUNCTION (HRF)***

fNIRS indirectly measures neural activation by monitoring changes in blood-oxygen concentration, which are the fluctuations in blood-oxygen levels at the surface of the cortex. When a pool of neurons is active, oxygen is recruited to the area where the hemodynamic response is localized within 1-3 mm of this neural activity (Shmuel, Yacoub, Chaimow, Logothetis, & Ugurbil, 2007). Therefore, the increase in oxygen is indirectly correlated to neural activity (Gervain et al., 2011). When a group of neurons activate, energy is required to meet their metabolic demands. Thus, an increase in cerebral blood flow followed by changes in blood-oxygenation is observed in that region. Through emitting two wavelengths of infrared light, fNIRS can measure three different properties of hemoglobin, that is, oxygenated (HbO), deoxygenated (HbR), and total (HbT) hemoglobin, where HbO is most responsive and comparable to active, neural metabolic demands. The link between blood-oxygenation and observed changes in the intensity of a wavelength of light is made using the modified Beer-Lambert law, where  $A$  is the attenuation,  $I_0$  and  $I$  are the initial and final intensities of light,  $c$  is the concentration or density of the medium,  $\varepsilon$  is the molar extinction coefficient of the medium for a light of wavelength  $\lambda$ ,  $d$  is the distance between source and detector,  $DPF$  is the differential pathlength factor, and  $G$  is an unknown term due to scattering losses (Gervain et al., 2011; Minagawa-Kawai, Mori, Hebden, & Dupoux, 2008).

$$A = -\log(I_0 / I) = (c * \varepsilon * d * DPF) + G \quad (1)$$

Simply put, as light travels through a medium, the concentration of the medium can be obtained by taking into account the scattering and absorption of light within the

medium, knowing the distance that the light will travel (i.e. source-detector separation/channel distance), the DPF, the intensities (i.e. wavelengths) of light that is emitted into the medium, and the intensity of light that leaves the medium. Change in concentration over time is seen as a hemodynamic response function (HRF). Typically, the HRF entails the neurovascular coupling of an initial increase in HbR followed immediately by a simultaneous sustained increase in HbO and a decrease in HbR before returning to baseline (Aslin, Shukla, & Emberson, 2015) (see Figure 2.1). This response is sampled much more slowly (in seconds) than EEG (milliseconds) but faster than fMRI. More importantly, the HRF can be detected using near-infrared light.

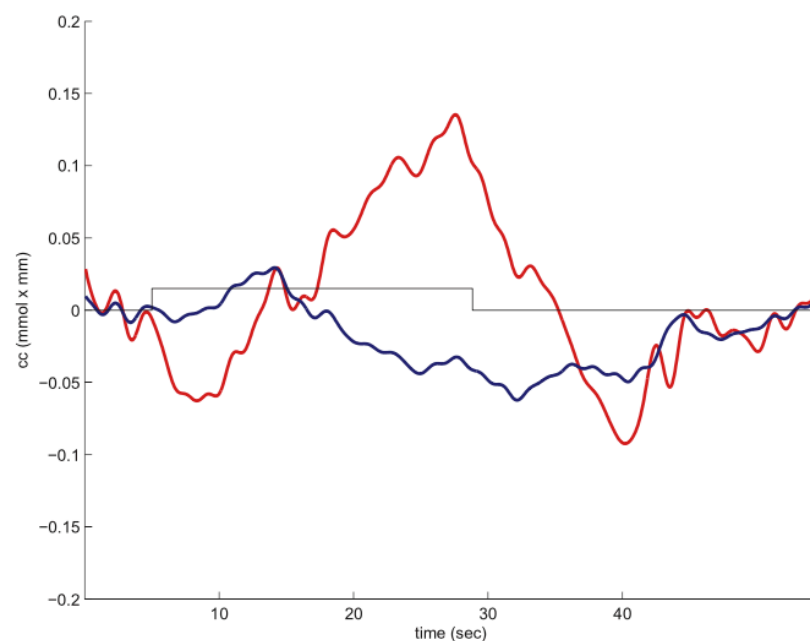


Figure 2.1. Depiction of a typical canonical hemodynamic response function (HRF) observed in a neonate. Red line: HbO, blue line: HbR, raised block: stimulation. Reproduced from Issard & Gervain (2018).



It is important to consider that the HRF is affected by additional components external to the functional response specifically elicited from stimulation. These external signals include cardiac rhythms, respiratory pulses, measurement noise, and motion artifacts. Additionally, there is considerable diffusion of light before it reaches the cortex because the light will have to first travel through any hair that may obstruct the path of light, then through the scalp, skull, and cerebral spinal fluid, before it reaches the brain. These components can be filtered out through a series of data pre-processing steps which will be discussed in the following sections.

### ***BASIC PRINCIPLES OF fNIRS***

Functional near-infrared spectroscopy is a non-invasive, neuroimaging technique that employs infrared light to monitor changes in blood-oxygen concentration on the cortical surface. Infrared light is used because it can penetrate human tissue more deeply as it has a longer wavelength than light from the visible spectrum, and HbO and HbR are strong absorbers of infrared light. Two wavelengths from the infrared spectrum (between 650-1000 nm) are shone into the skin via laser or LED light-emitting optodes (sources), absorbed and scattered by the underlying tissues, and the remainder of the light is then measured by detector optodes from the skin surface.

The space between a source and a detector is identified as a channel, where brain activation is measured. It is important to set an appropriate distance ( $d$ ) between each source and detector to estimate the DPF, which is the non-linear trajectory of light traveling through the medium, or the ratio of the mean path traveled by light in the tissue and  $d$  (Gervain et al., 2011; Minagawa-Kawai, Mori, Hebden, & Dupoux, 2008)

(the DPF is typically automatically estimated by fNIRS software systems). The distance between the source and detector will affect the light's depth of penetration into the brain: a greater source-detector separation will allow the light to travel a greater vertical distance, whereas a shorter source-detector separation would allow a shallower penetration (Figure 2.2). Increasing the distance between source and detector decreases the spatial resolution as the dispersion of infrared light is greatly increased. Moreover, it also decreases the signal-to-noise ratio (i.e. signal strength) and possibly masks the functional response (Gervain et al., 2011). The trade-off between depth of penetration (or source-detector separation) and spatial resolution requires an optimal separation between sources and detectors. Typically, the distance between sources and detectors is 30 mm for adults and 20 mm for infants (Taga, Homae, & Watanabe, 2007).

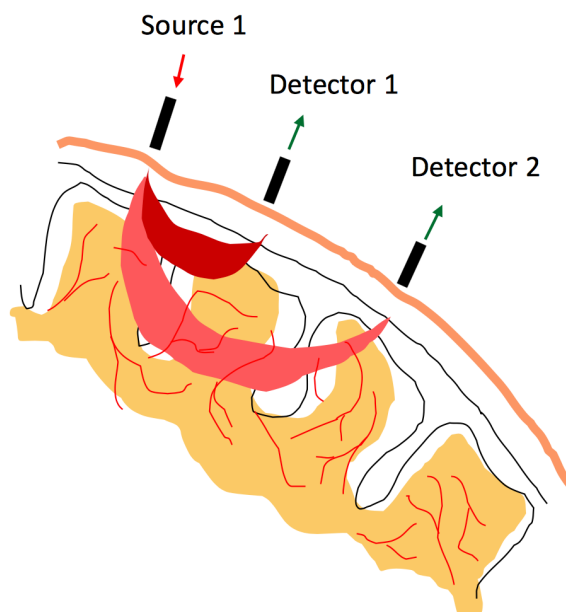


Figure 2.2. The effect of source-detector separation on the vertical penetration of light. A shorter separation (Source 1–Detector 1) allows for higher spatial resolution but a shallower depth of penetration, whereas a greater separation (Source 1–Detector 2) allows for the light to reach a greater depth but has less spatial resolution. Retrieved from nirx.net.

## ***ADVANTAGES AND LIMITATIONS OF USING fNIRS IN INFANT LANGUAGE RESEARCH***

Compared to adults, the skull is considerably thinner in infants and young children, which permits a more transparent view of the cortical surface. Therefore, fNIRS is an optimal method for the study of the developmental population as it also hosts a range of advantages relative to other brain imaging modalities. Due to the near-silent operation of fNIRS, auditory stimuli can be presented in a quiet acoustic field. Moreover, fNIRS yields good spatial and temporal precision, in which brain activity can be localized with greater accuracy than EEG and sampled much faster than fMRI (Lloyd-Fox, Blasi & Elwell, 2010). Most importantly, this technique does not require rigid head stabilization as in fMRI and thus provides a higher level of comfort during testing, as infants and vulnerable populations that have inhibited motor control would have some freedom in head movements during data acquisition. These experimental conditions are optimal for language discrimination studies on young infants, as auditory stimuli can be played in a quiet acoustic environment with infants not having to be fully restrained in their head movements. Additional advantages of the fNIRS technique include the portability of the system, affordability, ease of application, and its versatility. Due to its compact size (in relation to fMRI), the whole system including the console, LED and/or fiber optic cables, computer, and headgear can be easily transported to different rooms in a hospital and even remote locations to study cognition in infants (e.g. Lloyd-Fox et al., 2014). Moreover, the application of infant and adult headgear is simple: for the present experiments, a customized rubber-padded headband fitted with source and detector optodes was used on infants, and an elastic cap (Easycap) for adults. The application process seldom requires the use of gel or other substances to separate hair.

Most fNIRS studies conduct data analysis on HbO, as it is the property of hemoglobin that is most responsive to stimuli (e.g. Petitto et al., 2012; Sato, Sogabe, & Mazuka, 2010). Whereas fMRI can only measure HbR, fNIRS provides a more robust depiction of the hemoglobic properties of the brain. However, some studies have reasoned that HbT may provide a better signal-to-noise ratio as the calculation is additive of the absorbance measurements at two wavelengths (i.e. HbO & HbT) (Minagawa-Kawai, Mori, Naoi, & Kojima, 2007). Nevertheless, previous work has shown that HbR (Wilcox, Bortfeld, Woods, & Wruck, 2005) and HbT (Furuya & Mori, 2003; Minagawa-Kawai, Mori, Naoi, & Kojima, 2007; Peña et al., 2003; Watanabe et al., 1998) can also be reliable measures to study hemodynamic activity. Lastly, it is also possible to combine multiple techniques with fNIRS. For example, EEG and fNIRS permits both high-quality temporal *and* spatial resolution to study phonotactic processing in 6-month-olds (Obrig et al., 2017), and TMS was used in conjunction with fNIRS to study functional connectivity in the brain and the causal relationships between non-adjacent neural networks in adults (Kozel et al., 2009).

On the other hand, fNIRS has some disadvantages. First, fNIRS has less spatial resolution than fMRI; the spatial resolution of fNIRS is determined by the source-detector separation (i.e. 20-30 mm), whereas fMRI provides a spatial resolution of 3-4 mm (Glover, 2011). And as the placement of the optodes are guided by the 10-20 referencing system for EEG, identifying the exact cortical regions underlying each channel may not be entirely accurate. However, the 10-20 referencing system has been shown to be quite reliable for the co-registration of underlying cortical structures (Herwig, Satrapi, & Schönfeldt-Lecuona, 2003).

Second, the infrared light being emitted into the underlying tissue cannot safely reach deep-brain structures, as these areas require higher intensities of light to be

reached and may cause tissue damage. Thus, hemodynamic activity in these deep-brain regions cannot be measured. Conveniently, however, language processing can be quantified using neurophysiological methods that do not reach deep brain structures (i.e. EEG, MEG), suggesting that language processing occurs at the superficial layers of the cortex which fNIRS is capable of measuring (Cheour et al., 1998; Ferjan Ramierz, Ramírez, Clarke, Taulu, & Kuhl, 2016; Garcia-Sierra et al., 2011).

A third disadvantage is that the hemodynamic currents in the brain operate at a much slower speed than electrical activity on the scalp measured by EEG, therefore fNIRS is unable to provide temporal resolution in the millisecond range like EEG. As speech is comprised of fast temporal changes of auditory information, fNIRS would not be adequate for detecting the processing of rapid changes in speech. As a result, fNIRS experimental paradigms typically include longer stimulation periods (between 5 and 30 seconds) which are separated by long silent or baseline periods (between 15 and 30 seconds) (e.g. Bouchon, Nazzi, & Gervain, 2015; Peña et al., 2003). In summary, despite these shortcomings, fNIRS possesses many methodological and practical benefits relative to other techniques in measuring brain activation of infants and young children during language processing, such as its ease of application, unrestricting equipment, and near-silent operation.

### ***VARIATION IN THE HRF RESPONSE***

The HRFs elicited by infants differ from those of adults. Non-canonical responses, such as inverted responses with an increased or steady level of HbR and a decrease in HbO, are often observed more commonly in newborns and young infants. A common cause for this type of HRF results from the immature vascular system in

infants. However, with age, the HRF will increasingly take on the canonical shape (Issard & Gervain, 2018). Additionally, it is important to note that brain maturation is not homogenous across different regions; therefore, the hemodynamic response may vary from one region to another. In an fMRI study conducted by Leroy and colleagues (2011), it was discovered that the right temporal structures of asleep infants aged between 2.6 – 16.3 weeks matured much more quickly than the left, indicated by resting-state brain activity. Moreover, previous research has found that the resting cerebral blood flow is stronger in the right hemisphere than in the left in infants (Chiron et al., 1997; Roche-Labarbe et al., 2012).

### *Stimulus complexity*

A second factor that might have an impact on the variation of the HRF response is stimulus complexity. Previous research has suggested that infants actively seek and internalize information with intermediate levels of complexity and avoid allocating cognitive resources to overly predictable or overly surprising events (Kidd, Piantadosi, & Aslin, 2012). It has been hypothesized that these results may parallel the characteristics of the HRF, in which the shape of the response is canonical and highest in amplitude when stimuli are presented with an intermediate range of complexity, whereas atypical and inverted responses would be observed for stimuli that are too complex or too simple (Issard & Gervain, 2017).

For example, Issard & Gervain (2017) conducted an fNIRS study to investigate the effect of stimulus complexity on the HRF and observed canonical and inverted responses in neonates when stimulus complexity was manipulated. It was found that normal and moderately time-compressed speech (intermediate complexity) elicited a

canonical response in newborns, whereas highly time-compressed speech (high complexity) produced an inverted response. Thus, a qualitative change in the *pattern* of activation might be observed as a function of stimulus complexity, but not a quantitative change in amplitude of the response. It is important to remember that the HRF is representative of cognitive effort. Populations in the nascent stages of cognition, such as infants, may generalize their input to reduce the memory demands of highly variable or highly complex information. As a result, a change in stimulus complexity might trigger a qualitative (and not quantitative) change in response function. Selecting the appropriate level of stimulus complexity for a certain stage in cognitive development is crucial to uncovering a robust and canonical HRF.

#### *Stimulus presentation / Experimental paradigm*

The way in which stimuli are presented may also affect neural activation patterns. A common way to test infants' responses to novel stimuli is through examining the impact of familiarity (Hunter & Ames, 1998). It is supposed that when initially exposed to new information, infants would show a preference for the stimulus as they are in the process of encoding and learning that information. When the memorization or learning process has reached a sufficiency criterion, infants would then show less preference for the familiar stimulus and show an interest in novel information (Houston-Price & Nakai, 2004). However, between the initial exposure and the learning of the stimulus, it is hypothesized that infants would show *no preference* to either the familiar or novel stimulus as they would be equally attracted to both (Hunter & Ames, 1998). Thus, incorporating an adequate amount of trials during stimulus presentation, in which infants would have enough opportunities to learn and

encode the information, might have an impact in observing a novelty preference in the form of a canonical HRF.

Accordingly, just like the effect of stimulus complexity has on the HRF, stimulus repetition also may have an effect on the functional response. In other words, a repeated stimulus may either enhance or suppress the functional response over time. Repetition enhancement involves a significant increase in the amplitude of neural activity that can be seen when a novel stimulus is repeatedly presented, whereas repetition suppression refers to a significant decrease in neural activity with novel stimulus repetition. Both effects can indicate that the listener has learned and encoded the information. For example, repetition enhancement effects have resulted from repeating complex stimuli in which the listener eventually develops a memory trace through learning the information (Gervain, Macagno, Cogoi, Peña, & Mehler, 2008), whereas cases of repetition suppression have been suggested to arise from the reduced cognitive demand of processing a familiar (i.e. repeated) stimulus (Kobayashi et al., 2011).

A recent within-subjects study that investigated repetition effects revealed that the complexity of the stimulus may determine the direction of the effect (i.e. enhancement or suppression) (Bouchon, Nazzi, & Gervain, 2015). Asleep neonates were presented with 12 sequential blocks of 6 artificial ABB grammar sequences (e.g. “mulele”, “junana”) and then another 12 sequential blocks of more complex ABC sequences (e.g. “mulevi”, “junary”) during one testing session. It was found that brain activation to the ABC patterns increased over time in the left fronto-temporal cortex. For the ABB sequences, although an effect of repetition suppression was not found, a slight decline in activation was observed from an enhanced to a stable response. The authors suggested that further decrease in the complexity of the ABB sequences might



be needed to observe a repetition suppression effect. These studies suggest that a link is present between stimulus complexity and familiarity effects where familiarity effects are more likely to be observed with relatively simple and non-complex stimuli. In the following, several infant fNIRS experimental paradigms are reviewed in which canonical responses were evoked from various stimulus presentation procedures, suggesting that these paradigms were designed with appropriate levels of stimulus complexity and repetition.

### ***INFANT fNIRS PARADIGMS***

Peña and colleagues (2003) investigated cortical activation of asleep newborns aged between 2 to 5 days old in response to forward speech, backward speech, and silence. Their stimulus presentation was designed in a classic block-order fashion, in which 10 consecutive blocks of each auditory condition were presented for 15 seconds each and separated by a silent baseline of 25 or 35 seconds (Figure 2.3). The forward speech, backward speech and silent conditions were randomly presented for each infant. Through examining total hemoglobin responses, the authors found that HbT concentration was significantly greater in the left hemisphere when newborns were presented with forward speech, whereas no significant differences were observed in the backward speech and silent conditions. Importantly, significant increases in HbT were observed, indicating that canonical responses were acquired from the newborns.

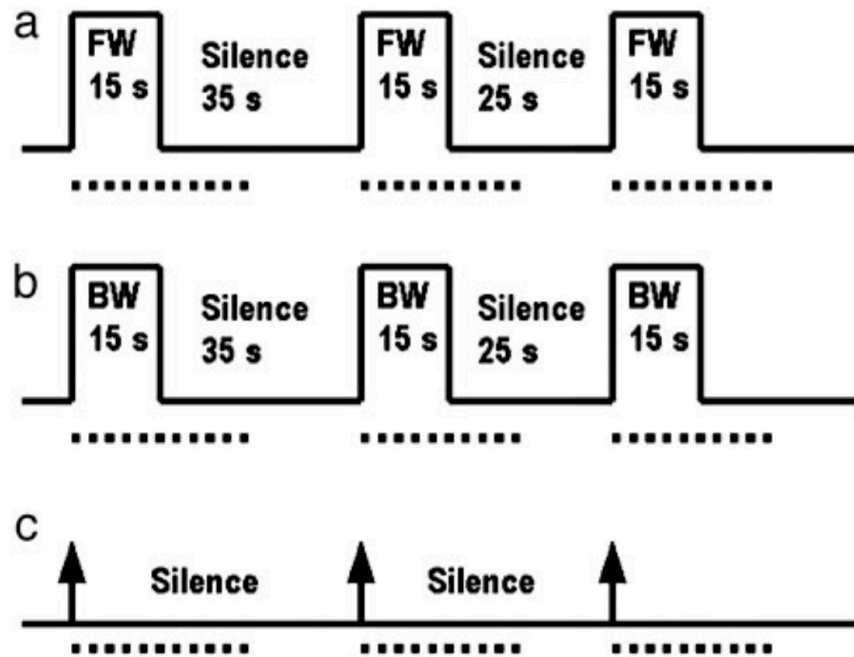


Figure 2.3. Testing procedure from Peña et al. (2003). Newborns were tested on 10 blocks of each condition: (a) forward speech, (b) backward speech, and (c) silence. The non-auditory baselines varied at 25 or 35 seconds. The periods selected for statistical analyses began at the stimulus onset and ended 30 seconds later, as indicated by the dotted line.

Another paradigm that is commonly used in infant phonemic discrimination studies is the alternating block design, where non-silent blocks serve as a baseline between auditory stimulation periods (Arimitsu et al., 2011; Issard & Gervain, 2017; Sato, Sogabe, & Mazuka, 2010). The use of a non-silent baseline permits the extraction of cortical response components specific to the differences presented in the test blocks (Arimitsu et al., 2011). In an fNIRS study examining phonemic (/iita/ vs. /itte/) and prosodic (/iita/ vs. /itta?/; falling vs. rising pitch) discrimination in Japanese, the presentation of the auditory stimuli alternated between baseline blocks (AAAA) and test blocks (BABA) for at least seven repetitions each. As newborns do not yet have

well-developed memory systems, it was reasonable to use a technique that did not require the infants to form internal or memorized representations of AAAA and BABA blocks between long silence periods to compare the subtle phonemic distinctions between the two. Canonical responses to the test conditions were observed, in which it was found that the phonemic condition showed a significant increase in HbO in the left hemisphere, whereas the right hemisphere showed significant HbO increases in response to the prosodic condition (Figure 2.4, Arimitsu et al., 2011).

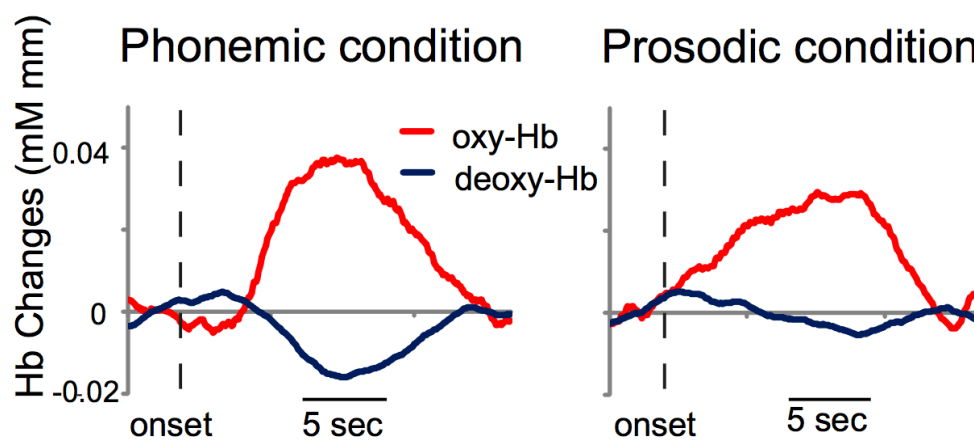


Figure 2.4. The grand averaged time courses of the HRF across all channels in the left and right hemispheres for the phonemic and prosodic conditions. The vertical dashed line represents the onset of the test block and the horizontal solid line represents the time window used for statistical analyses. Reproduced from Arimitsu et al. (2011).

Lastly, although the hemodynamic response unfolds at a much slower rate relative to electrical impulses measured as event-related potentials (ERPs), it is still possible to use event-related paradigms in fNIRS procedures (e.g. Homae, Watanabe, Nakano, Asakawa, & Taga, 2006). In these types of designs, stimulation periods as

short as 500 ms and baseline periods as short as 4.5 s have previously been used (Petitto et al., 2012). In these cases, the HRF would typically overlap with the responses from the adjacent stimuli. Although event-related paradigms are not a common method of stimulus presentation relative to block paradigms in fNIRS research, resolving the issue of overlapping response functions can be done using General Linear Modeling (GLM) for statistical analyses (Issard & Gervain, 2018).

Indeed, implementing the right paradigm is paramount to elicit response components specific to the change in stimulus. The paradigm from Peña et al. (2003) utilized repeated infant-directed speech samples between silent baselines, where the objective of that study was to compare neural activation of newborns in response to three different types of auditory conditions. Subsequently, the alternating-block paradigm described in Arimitsu et al. (2011) used a non-silent baseline to test phonemic and prosodic discrimination in newborns. The paradigm was designed in this way to extract only neural responses specific to the detection (or lack thereof) of subtle changes in Japanese minimal contrasts. Through directly comparing these two studies, speech samples presented in Peña et al. were longer in duration (15 s) than stimuli at the word-level in Arimitsu et al. (360 ms). Unlike Arimitsu et al., Peña et al. did not require the newborns to discriminate between baseline and test blocks. Taken together, fNIRS studies in speech perception with longer stimuli have used the classic block paradigm (e.g. Bouchon, Nazzi, & Gervain, 2015; May, Gervian, Carreiras, & Werker, 2017; Peña et al., 2003), whereas studies assessing phonemic discrimination of minimal pairs tend to typically present their stimuli in an alternating-block fashion (e.g. Arimitsu et al., 2011; Minagawa-Kawai, Mori, & Sato, 2005; Sato, Sogabe, & Mazuka, 2010).

Regarding neurophysiological findings, such as those from fNIRS studies, an important distinction must be made between perceptual discrimination and differential

brain activation. In most cases, perceptual discrimination leads to differential brain activation, however, a lack of differential brain activation does not always indicate a lack of perceptual discrimination between two conditions. In theory, a difference in activation would be observed when a change in stimuli was detected. But if differential brain processing is not found, the absence of cortical activity could be explained by the lower temporal resolution of fNIRS, or that the differential activation took place in regions that were not examined. Another possibility for the lack of differential activation could be attributed to the way in which the stimuli were presented. If the experimental paradigm did not extract neural responses specific to the change between two conditions (e.g. Arimitsu et al., 2011), then it is possible that the response was not accurately measured.

### ***DATA PRE-PROCESSING***

As fNIRS is a relatively novel technique, there is no standardized protocol to prepare the data for analysis. This section reviews the various data preparation steps that have been used in previous studies and the steps that the current series of experiments have adapted. First, it is expected that motion artifacts would be present in the data acquired from infants and young children. Motion artifacts are caused by head movements and characterized by rapid shifts in the coupling between optical fibers and the scalp, resulting in a period of high-frequency noise in the recorded NIRS data (Cooper et al., 2012). Motion artifacts are generally easy to identify through visual inspection (see Figure 2.5); however, it is common practice to use automated methods for a more objective approach of motion artifact detection. Some studies have defined motion artifacts as concentration changes that are greater than 0.1 mmol\*mm over a period of 0.2 seconds (Bouchon, Nazzi, & Gervain, 2015; Gervain, Macagno, Cogoi,

Peña, & Mehler, 2008; May, Gervain, Carreiras, & Werker, 2017) or two consecutive samples (Peña et al., 2003; Sato, Sogabe, & Mazuka, 2010), and in some cases, signal variations over 0.7 mmol\*mm between successive samples (Arimitsu et al., 2011). As blocks that are contaminated with motion artifacts are discarded and excluded from the final analyses, it is important to incorporate enough trial repetitions in developmental paradigms as these artifacts are common in the data of infants and young children. Common criteria for the inclusion of data in the final analyses are for each channel to have a minimum of three valid trials per condition (Lloyd-Fox et al., 2013). Additional physiological components of the raw data such as respiratory and cardiac rhythms are also filtered out using high and low bandpass filtering. Common values used typically range between 0.01 and 1 Hz to remove slow drifts in blood-oxygen concentration and heart rate (Bouchon, Nazzi, & Gervain, 2015; Gervain, Macagno, Cogoi, Peña, & Mehler, 2008; Peña et al., 2003).

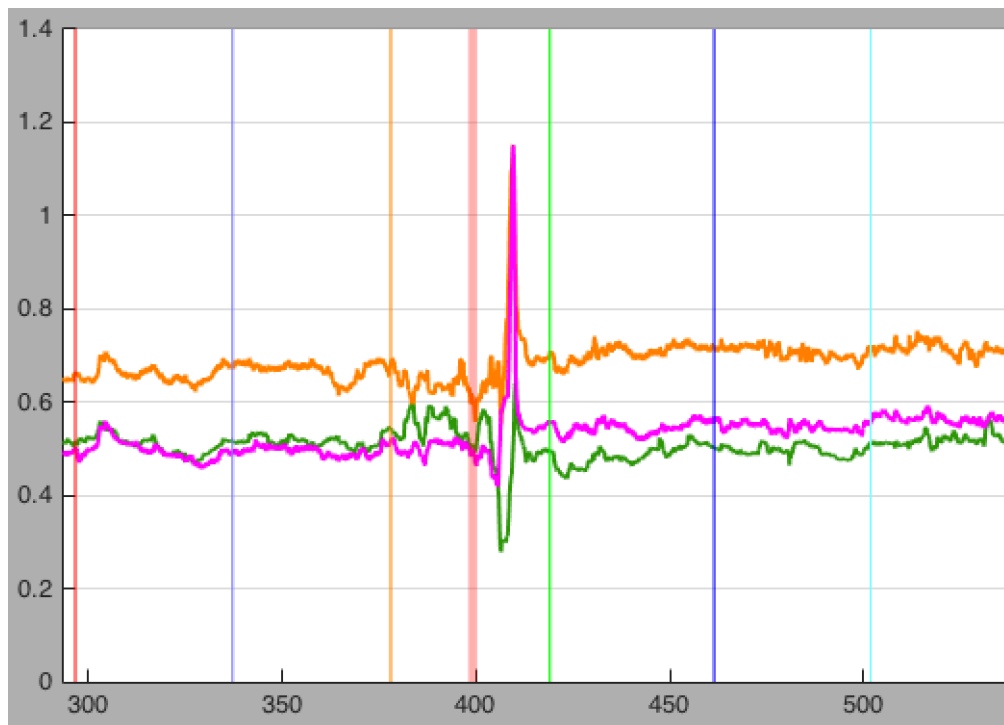


Figure 2.5. An image taken from the HomER2 user interface showing raw data from three channels in the left hemisphere of one infant participant. The x-axis represents time in seconds and the y-axis represents HbO concentration. The colored vertical lines signify the start of each experimental block. The motion artifact can be easily spotted just after the 400 s mark.

### ***DETERMINING TIME WINDOWS FOR STATISTICAL ANALYSIS***

Upon completion of motion artifact removal and the filtering of physiological noise and rhythms, the next step is to determine time windows for statistical analyses. This step is varied across studies as there are also no standardized procedures for setting time windows. Further, the length of the windows might differ depending on the experimental design. For example, in Peña et al. (2003), where 15-second speech samples were played between silent baseline periods, the authors determined a 30-second time window starting at the onset of auditory stimulation and calculated the

mean change in concentration of HbT within the window for each condition and each channel. These values were then used in repeated measures ANOVAs. Bouchon et al. (2015) had also set a 30-second time window at the onset of auditory stimulation, which the authors stated that it would capture the full time course of the HRF in each block. The stimulus presentation consisted of ~9 seconds of auditory stimulation followed by a silent baseline period of 20 or 25 seconds. Separate ANOVAs were conducted on the mean values of HbO and HbR time windows separately in that study.

It is possible that averaging the values within a wider time window may mask the peak amplitude of the signal. An alternative option might be to set shorter time windows centered around the peaks of the HRF. However, the peak latency of functional responses is not homogenous across the brain, meaning some channels may have different peak latencies than others. Arimitsu et al. (2011) objectively implemented a shorter, five-second time window for analysis through calculating the peak latency for all conditions (i.e. phonemic and prosodic) by averaging the Hb time course for all channels and all participants. They found that the peak occurred at 11.1 seconds after the onset of stimulation. Thus, a 5-second time window was centered on that value as the target period. Another 5-second window prior to stimulus onset was determined as the baseline period. Then, t-tests were conducted on these values to assess any significant differences between the baseline and target periods, as well as laterality effects.

Studies that used alternating block paradigms (e.g. Minagawa-Kawai, Mori, & Sato, 2005; Sato, Sogabe, & Mazuka, 2010) have set the time windows for analysis at the same length or slightly longer than their test blocks. Shorter time windows are typically used in alternating block paradigms; as the baseline blocks are non-silent, peak latencies for the test blocks (i.e. stimulus change) are reduced as the temporal



cortices are already receiving auditory stimulation before the test blocks. In contrast, classic block paradigms (e.g. Gervain, Macagno, Cogoi, Peña, & Mehler, 2008; May, Gervain, Carreiras, & Werker, 2017; Peña et al., 2003) average hemodynamic responses over a longer time period. Peak latencies are increased in these paradigms, as the listener would begin processing auditory stimulation after a longer silent baseline period.

### ***THE CURRENT SERIES OF EXPERIMENTS***

Compared with the fNIRS paradigms reviewed above, the current series of studies used novel stimulus presentation procedures, in which aspects from the alternating block and standard block designs were incorporated. This was done to explore whether phonemic discrimination could be measured using a modified block-paradigm and not the previous ones described. Subsequently, the response patterns elicited from the new paradigms would also be assessed for any stimulus complexity and stimulus repetition effects. In all three studies, phonemic contrasts from English, Mandarin Chinese, and Hindi were presented in non-alternating baseline (i.e. A-A-A-A- or B-B-B-B-) and alternating test (i.e. B-A-B-A- or A-B-A-B-) blocks. The procedure of Study 1 (Chapter 3) consisted of three blocks for each language condition – the first block was always a non-alternating block, followed by an alternating block and a silent baseline of 12 seconds. This paradigm was adapted from Arimitsu et al. (2011), in which baseline (non-Alt) and test (Alt) blocks were adjacent to each other. However, silence periods were always inserted after each test block and before the next non-Alt block of a new language condition in Study 1. It was predicted that adult HRF patterns in Study 1 would show a decrease in activation during the non-Alt blocks and a rebound in activity after the onset of the stimulus change in the Alt blocks. In Studies

2 and 3 (Chapters 4 and 5), 25-second silent periods were added between the non-Alt and Alt conditions for each language, resembling a paradigm seen in a speech discrimination studies (e.g. Peña et al., 2003) but not so far in phonemic discrimination studies (e.g. Arimitsu et al., 2011; Issard & Gervain, 2017; Sato, Sogabe, & Mazuka, 2010). It was predicted that if discrimination was successful, greater brain activation would be observed for the Alt blocks, as well as significant differences in activation between the non-Alt and Alt blocks within each language. The next chapters discuss each study and its findings in greater detail.

## CHAPTER 3

**Monolingual and bilingual adults show language-specific responses to native and tonal contrasts but respond differently to the non-native contrast**

## Introduction

Adults are generally more proficient in discriminating native phonemic contrasts over non-native contrasts as they are neurally committed to their native language(s) (Kuhl, 2004; Kuhl & Rivera-Gaxiola, 2008). Previous research has assessed bilinguals' discriminatory abilities on phonemic contrasts native to both of their languages but not often on a third, non-native contrast. To date, there are only a few studies that directly examined non-native phonemic discrimination in bilinguals (e.g. Archila-Suerte, Zevin, Ramos, & Hernandez, 2013; Golestani & Zatorre, 2004; Petitto et al., 2012). Previous research has suggested that language experience influences neural activation patterns in language perception tasks (Dehaene et al., 1997; Kovelman, Shalinsky, Berens, & Petitto, 2008; Perani et al., 1996). Through examining cortical responses to non-native contrasts that do not belong in the adult listener's phonemic repertoire, we can investigate whether the adult brain responds to non-native contrasts differently and whether language background has an effect on phonemic perception. More importantly, it would allow us to observe the adult brain's specialization for one's native language(s) by examining cortical responses to various types of non-native contrasts (e.g. Mandarin lexical tone/pitch vs. Hindi dental/retroflex).

Neuroanatomically, the experience of acquiring and using a second language alters the cortical structure of the adult human brain. For example, it was discovered that early and late English-Italian bilingual adults who had acquired Italian before 5 years and between 10-15 years of age, respectively, had a significantly greater volume of grey matter in the inferior parietal cortex relative to English monolinguals. Further, the density of grey matter increased with second-language proficiency and decreased as age of acquisition increased (Mechelli et al., 2004).

In language comprehension tasks with bilingual adults, cortical activation was influenced by the age of acquisition and the degree of language proficiency (Perani et al., 1996). Using PET, the authors studied brain activation of Italian-English bilingual adults with moderate proficiency in English (L2) as they listened to stories in each language. When listening to Italian (L1), the participants showed a robust activation of the left-hemispheric language processing network, whereas listening to L2 displayed reduced bilateral activation residing in the temporal areas. Although cortical activation can be influenced by the age of acquisition of the second language and the degree of language proficiency (Perani et al., 1996), a subsequent study showed that overall language proficiency has a greater influence than age of acquisition on the cortical representation of the second language (Perani et al., 1998). In that study, spoken story processing in both native languages by early and late Italian-English bilinguals (equally exposed to two languages from birth vs. exposed to the second language after 10 years of age) with high dual-language proficiency revealed a single and common neural network consisting of bilateral temporal areas, hippocampus, the left superior temporal sulcus, and left inferior parietal cortex. In other words, simultaneous and late bilinguals with high proficiency in both native languages showed comparable processing patterns between the two.

In a sample of moderately late bilinguals (exposed to the second language after 7 years of age) differential activation was observed in response to each of their two languages (Dehaene et al., 1997). Listening to L1 (French) robustly activated regions in the left temporal lobe with a weaker but similar pattern of activity in the right hemisphere, as well as a consistent activation in the left inferior frontal sulcus. In contrast, listening to L2 (English) activated a diffuse pattern of activity across all participants that was mainly centralized in the left temporal lobe. Although the

participants revealed left-hemispheric activity while listening to L2, its magnitude was smaller than for L1. In addition, listening to L2 activated small subregions in the right temporal lobe. These findings provide supporting evidence whereby language experience shapes the neural architecture of language processing by showing that dedicated language networks remain active in the left hemisphere, whereas second languages acquired after 7 years of age recruit additional substrates outside of the typical language processing regions.

### The present study

Motivated by previous research, the current study used fNIRS to compare cortical activation between English monolingual and Mandarin-English bilingual adult listeners to native and non-native phonemic contrasts to investigate how linguistic experience can guide specific patterns of activation in response to perceptual information. The following hypotheses were tested: first, monolingual and bilingual adults should show no differential activation to the English contrast that was native to both groups. Further, all participants were expected to show left-lateralized responses to the English contrast, mainly localized in the left inferior frontal cortex. Second, monolinguals and bilinguals should exhibit hemispheric differences in the processing of the Mandarin lexical tone contrast, where bilinguals would show left-lateralized activation and monolinguals would show right-lateralized activation. And third, both language groups were expected to show little to no cortical activation in response to the non-native Hindi contrast, consistent with the evidence on native-language neural commitment in adulthood (e.g. Werker & Tees, 1984). However, because of the participants' different language backgrounds, it was predicted that monolinguals and bilinguals would recruit different cortical regions for the processing of the non-native

contrast. The findings from the present study help set the framework for the subsequent chapters on infant phonemic perception, and extend the evidence for the influence of bilingualism on the cortical network of speech processing across infancy and adulthood.

## Method

### *Participants*

Changes HbO were measured in a total of 26 participants (18-29 years of age, mean age = 22.14 years, SD = 2.63). Fourteen participants were English monolinguals and twelve were Mandarin-English bilinguals. The monolingual participants reported that they had never received formal instruction of a second language or had previous experience in Mandarin or Hindi. The bilingual participants were included in the analysis only if they started formally learning English as a second language before the age of 7. The average age of English acquisition for the bilingual sample was 5.73 years (range = 0-7 years, SD = 1.88). Eleven out of twelve bilingual participants reported to use, on average, not more than 60% Mandarin in their daily language use. An additional 15 participants did not meet the conditions for the final analyses due to inadequate channel connectivity ( $n = 7$ ), excessive motion artifacts ( $n = 6$ ), and inattention during the task ( $n = 2$ ). All participants reported that they had no learning or hearing disadvantages and gave written and informed consent prior to the start of the experiment.

### *Stimuli*

Three minimal pair consonant-vowel-consonant (CVC) contrasts in English, Mandarin Chinese, and Hindi were used for the stimuli. Female speakers from each

language naturally produced five tokens for each word in adult-directed speech. A VOT consonant contrast tab - dab (phonemically represented as /tæb/ - /dæb/) was used for English. Out of the four Mandarin Chinese lexical tones, the rising (T2) and a low-level (T3) tone were instantiated on the word /taw/, meaning “naughty” and “to ask”, respectively. These tones were selected because the two form a subtle contrast and are therefore more difficult to discriminate for non-native or non-tonal listeners (i.e. English monolinguals) but lexically accessible, thus discriminable, for native, tonal listeners (i.e. Mandarin-English bilinguals). The Hindi stop contrast was a voiceless, unaspirated retroflex /ɖa:l/ and dental /ɗa:l/ consonant that meant “branch” and “lentil”, respectively. Retroflex consonants are produced with the tongue curled and placed at the area between the roof of the oral cavity (hard palate) and the alveolar ridge (area between the teeth and hard palate). The articulation of dental consonants involves the front of the tongue tip placed on the backside of the teeth and then released. To retain the natural characteristics of each language, all tokens were unfiltered and equalized to an intensity of 70 dB using Praat (Boersma & Weenink, 2012). They were also averaged to durations of 524 ms for English, 541 ms for Mandarin, and 879 ms for Hindi. The greatest difference in duration was between the English and Hindi tokens at 355 ms. The unequal durations were not expected to have a considerable effect on the NIRS signal, as it is not highly sensitive to temporal resolution, unlike ERPs. Pitch (Hz) and intensity (dB) information for each token stimulus is shown in Table 3.1.



Table 3.1. Frequency and intensity measures for all token stimuli.

	English				Mandarin				Hindi			
	/dæb/		/tæb/		/taw2/		/taw3/		/ɖa:l/		/ɖa:l/	
	Hz	(dB)	Hz	(dB)	Hz	(dB)	Hz	(dB)	Hz	(dB)	Hz	(dB)
1	244	(76.01)	249	(76.03)	238	(75.55)	213	(75.40)	194	(75.50)	193	(75.45)
2	229	(76.15)	229	(75.95)	235	(75.49)	219	(75.35)	193	(75.86)	216	(75.37)
3	244	(76.34)	236	(76.01)	231	(75.57)	194	(75.41)	182	(75.61)	206	(75.46)
4	242	(76.15)	229	(76.17)	235	(75.57)	194	(75.55)	213	(75.54)	207	(75.38)
5	237	(76.41)	252	(76.06)	227	(75.57)	210	(75.52)	226	(75.43)	216	(75.41)

### *fNIRS recording*

Hemodynamic responses were sampled at a rate of 7.81 Hz using a multichannel NIRx NIRScout system emitting two continuous wavelengths at 760 and 850 nm. Eight near-infrared light source emitters and eight detectors were arranged in two staggered 2 x 4 arrays positioned bilaterally on each side of the head, resulting in 9 channels per hemisphere (Figure 3.1). The source-detector separation was 30 mm, and the adult cap was centered at Cz on each participant's head and secured by a Velcro chinstrap. If a channel signal was poor due to the obstruction of hair, as indicated by system calibration, the corresponding optodes were taken out from the cap and the hair underneath was parted to ensure all optodes were in clear contact with the scalp for adequate channel connectivity and NIRS recording.

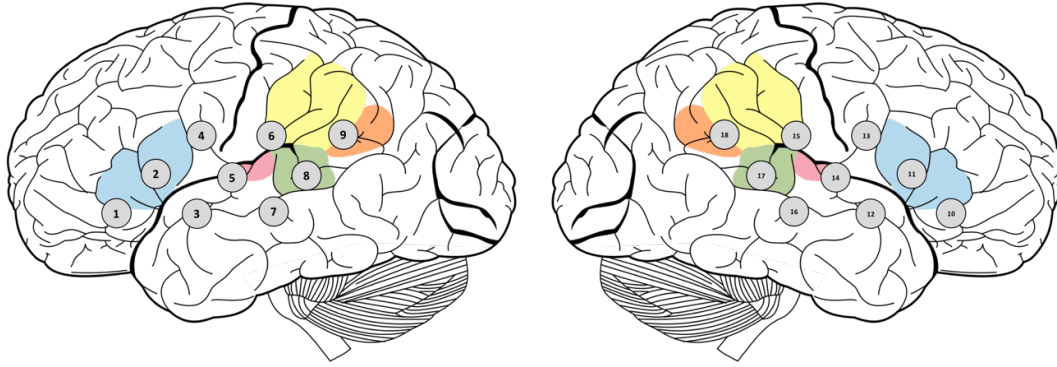


Figure 3.1. Diagram of channel placement in the left and right hemispheres over the approximate underlying cortical structures.

### *Procedure*

The experiment was conducted in a noise-isolated laboratory room. After receiving informed consent, participants completed a questionnaire that asked about some general demographic information and their language background and usage. Then, their head measurements were taken and the appropriate fNIRS cap was fitted using the 10-20 referencing system. The participants were seated approximately one meter away from a computer monitor and concealed front-facing loudspeakers where sound stimuli were administered at 70 dB. The participants underwent an active listening task where they were instructed to pay attention to the stimuli to detect any slight changes to the auditory tokens. A silent black-and-white animation of continuous slow-moving shapes played on the monitor as a visual filler. The video did not synchronize with the auditory stimuli and was not associated with the study.

The experiment had three language conditions (English, Mandarin, Hindi), in which the order of presentation was counterbalanced. Each condition contained three blocks. The order was made up of non-alternating (familiarization), alternating (test), and silence (baseline) blocks, which were always presented in this order. For each

minimal-pair, one word served as the “standard” or baseline word, and the other served as the “target” or change word. The standard and target words were counterbalanced across participants. The non-alternating block contained 10 randomized presentations of the standard tokens (e.g. A-A-A-A-... or B-B-B-B-...), whereas the alternating block had 10 presentations between the standard and target tokens (e.g. B-A-B-A-... or A-B-A-B-...). Intervals between each token were jittered between 1000 and 1500 ms to avoid phase-locked neural responses (Benavides-Varela, Hochmann, Macagno, Nespor, & Mehler, 2012; Gervain et al., 2008). The silence period lasted for 12 seconds to allow the hemodynamic response to return to baseline before the start of the next language condition. The three languages alternated in a fixed order, which was counterbalanced across participants, and were repeated five times for a total of 15 conditions, that is, 45 blocks. Languages were never mixed within a block or condition. The procedure is illustrated in Figure 3.2. The timing and presentation order of the stimuli was controlled via MATLAB, and the fNIRS computer software, NIRStar 14.0, was used for data acquisition. The total duration of the experiment was approximately 13 minutes. Participants were given a full debrief at the end of the experiment.

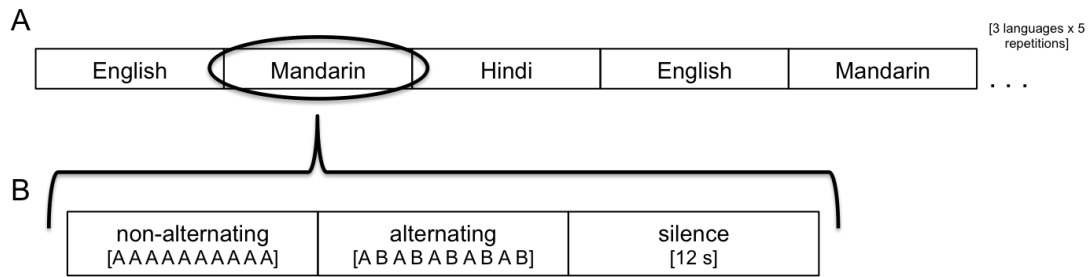


Figure 3.2. Illustration of the procedure. (A) The three languages were presented in a fixed order for 5 times for a total of 15 conditions. (B) Each condition contained three blocks (non-alternating, alternating, silence). The block order was identical for all language conditions.

#### *fNIRS data pre-processing*

Optical data collected from the two wavelengths were transformed into oxygenated (HbO), deoxygenated (HbR), and total (HbT) hemoglobin signals using the modified Beer-Lambert Law. Because the precise optical path length of the light traveling through brain tissue was unknown, the unit of HbO signals was molar concentration multiplied by length (Homae, Watanabe, Nakano, & Taga, 2007). Standard fNIRS data pre-processing, such as motion artifact (MA) removal, was carried out in the HomER2 (Hemodynamic Evoked Response data analysis GUI, version 2.1) user interface (Huppert, Diamond, Franceschini & Boas, 2009). MAs are disturbances in the raw signal that resemble sharp spikes, and they usually occur from sudden and rapid head movements or shifts in the point of contact between an optode and the scalp. MAs were defined in HomER2 as changes in the raw signal exceeding 0.1 mmol x mm over 0.1 s. Automatic identification and reconciliation of MAs involved applying wavelet-based motion artifact removal to reconstruct the signal after identifying and eliminating outlying coefficients of the Gaussian distribution from the raw signal.

(Brigadoi et al., 2014; Molavi & Dumont, 2012). However, if any block (i.e. non-Alt, Alt, or silent) contained irreconcilable MAs, the entire language condition was excluded from the analysis. Lastly, to eliminate high-frequency trends such as cardiac rhythms, the data were bandpass-filtered between 0.01 and 0.7 Hz.

In order to be included in the final analysis, it was required that each channel had to contain valid data in all three language conditions, as well as a minimum of three valid stimulation blocks per condition. If any block (i.e. non-alternating, alternating, or silent) within a language condition was invalid, the entire condition was excluded from the analysis.

#### *Defining cortical regions of interest (ROIs)*

Three anatomical ROIs were selected based on the co-registration of the underlying cortical structures using the international 10-20 referencing system. The optode configurations bilaterally overlaid the inferior frontal cortex (IFC), superior temporal gyrus (STG), and inferior parietal cortex (IPC). Channels 1, 2, and 4 in the left hemisphere were placed over the left IFC. The left IFC is also referred to as Broca's Area, known as a key language-processing center activating under tasks that require cognitive control, speech production, and phonetic working memory (see Novick, Trueswell, & Thompson-Schill, 2010, for a review). Channels 3, 5, 7, and 8 projected onto the T3 and T4 positions of the STG, that contains the primary auditory cortex, and channels 6 and 9 corresponded to the IPC, or the sensorimotor cortex. This configuration overlaid identical regions for the contralateral channels in the right hemisphere (Figure 3.3).

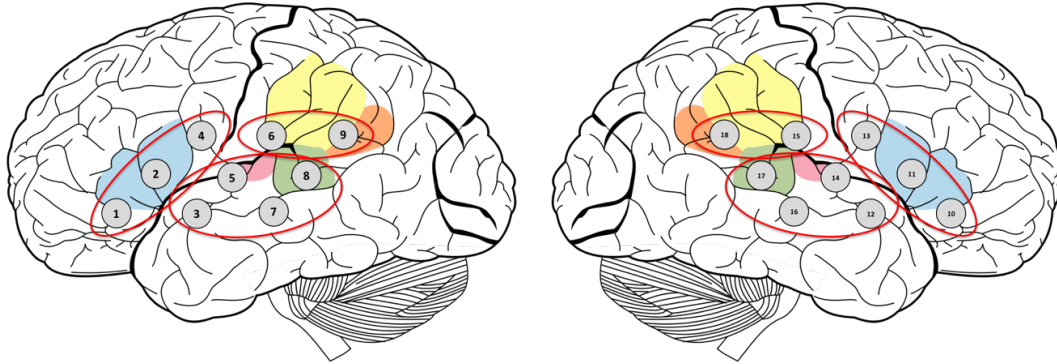


Figure 3.3. Anatomical illustration of the channel placement over the left and right hemispheres. The three ROIs in each hemisphere are circled in red (anterior to posterior: IFC, STG, IPC).

### *Statistical analyses*

Baseline correction was taken from the last 5 seconds of the silence block preceding the non-alternating block. Then, two ten-second time windows for analysis were set between 0-10 s and 20-30 s after stimulus onset, or at the start of the non-Alt blocks and Alt blocks, respectively. The adult hemodynamic response typically peaks earlier than infant responses, and differences in adult peak latencies range between 2-4 seconds (Thierry, Boulanouar, Kherif, Ranjeva & Démonet, 1999). Therefore, these time windows best incorporated peak activation from all channels. Next, a difference measure for each channel was calculated by averaging the amplitude in each time window and subtracting the non-alternating average amplitude value from the alternating value. The difference measure thus indicates the difference in amplitude between the alternating sequence and the non-alternating sequence. Positive difference values signify an increase in the hemodynamic response for the alternating sequence, associated with the detection of the contrast, and negative difference values might suggest a repetition suppression effect or the possibility that the contrast was not

perceived. The difference measures for all channels within an ROI were averaged per participant and used in the final analysis. As significant increases in HbO and significant decreases in HbR indicate cortical activation from baseline, both properties were used for the following analyses.

## Results

### *Confirmatory analyses*

In an overall analysis, separate 3 x 3 x 2 x 2 mixed-design ANOVAs were conducted on the difference values in HbO and HbR with Language (English, Mandarin Chinese, Hindi), ROI (IFC, STG, IPC), and Hemisphere (left, right) as the within-subject factors, and Group (monolingual, bilingual) as the between-subject factor. No significance was found for Language in either HbO or HbR blood properties. However, there was a significant interaction between ROI and Group in HbO,  $F(2, 48) = 4.12$ ,  $MS_e = 486$ ,  $p = .022$ ,  $\eta_p^2 = .15$ , indicating that in monolingual adults, cortical activation to all phonemic contrasts was relatively highest in the posterior regions in the bilateral IPC. In bilingual adults, on the other hand, activation was relatively highest in the anterior regions that is the bilateral IFC (see Table 3.2).

Table 3.2. HbO means for each bilateral ROI by language group.

	ROI	Mean difference value	(SD)
Monolingual	IFC	-4.90	(1.44)
	STG	-3.82	(1.54)
	IPC	-2.52	(1.95)
Bilingual	IFC	-0.54	(1.55)
	STG	-2.63	(1.67)
	IPC	-5.23	(2.10)

### *Exploratory analyses*

Although the effect of Language was not significant, there were visible differences in some specific channels that show that the hemodynamic responses of both monolingual and bilingual adults might differ across languages. Figure 3.4 shows the grand averaged hemodynamic responses of all participants to each language in a channel associated with the right STG. The figure shows an increase then a decline in HbO during the English and Mandarin familiarization/non-alternating blocks, and then a rebound in HbO activity following the onset of the English and Mandarin contrast/alternating blocks. The non-native Hindi contrast, however, did not elicit a rebound in HbO in the alternating block. To further investigate language-specific brain activation patterns of monolingual and bilingual adults, each language was examined separately using a series of mixed, repeated-measures ANOVAs. The following reports the results for each language.



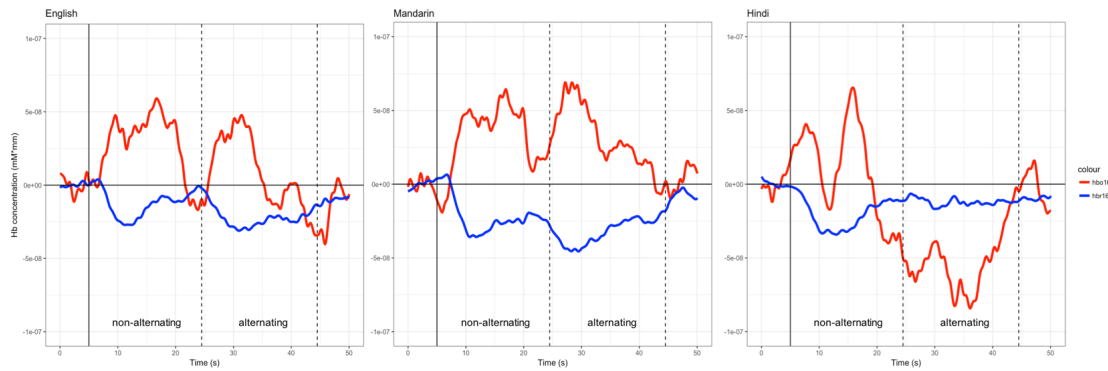


Figure 3.4. The grand averaged time course of the HbO (red) and HbR (blue) hemodynamic properties in right-hemispheric Channel 16 to English, Mandarin, and Hindi across all adult participants.

### *English*

A mixed, repeated-measures 3 x 2 x 2 ANOVA was conducted on the difference values with ROI and Hemisphere as within-subject factors and Group as the between-subject factor. No significant main effects or interactions were found in either HbO or HbR properties between monolingual and bilingual adults. These results suggest that monolingual and bilingual adults showed no differences in brain activation patterns to the native consonant contrast. Moreover, the difference measures in HbO were below zero, possibly suggesting a repetition suppression effect or the inability to detect the contrast. The mean difference values in both hemoglobin properties are depicted in Figure 3.5.

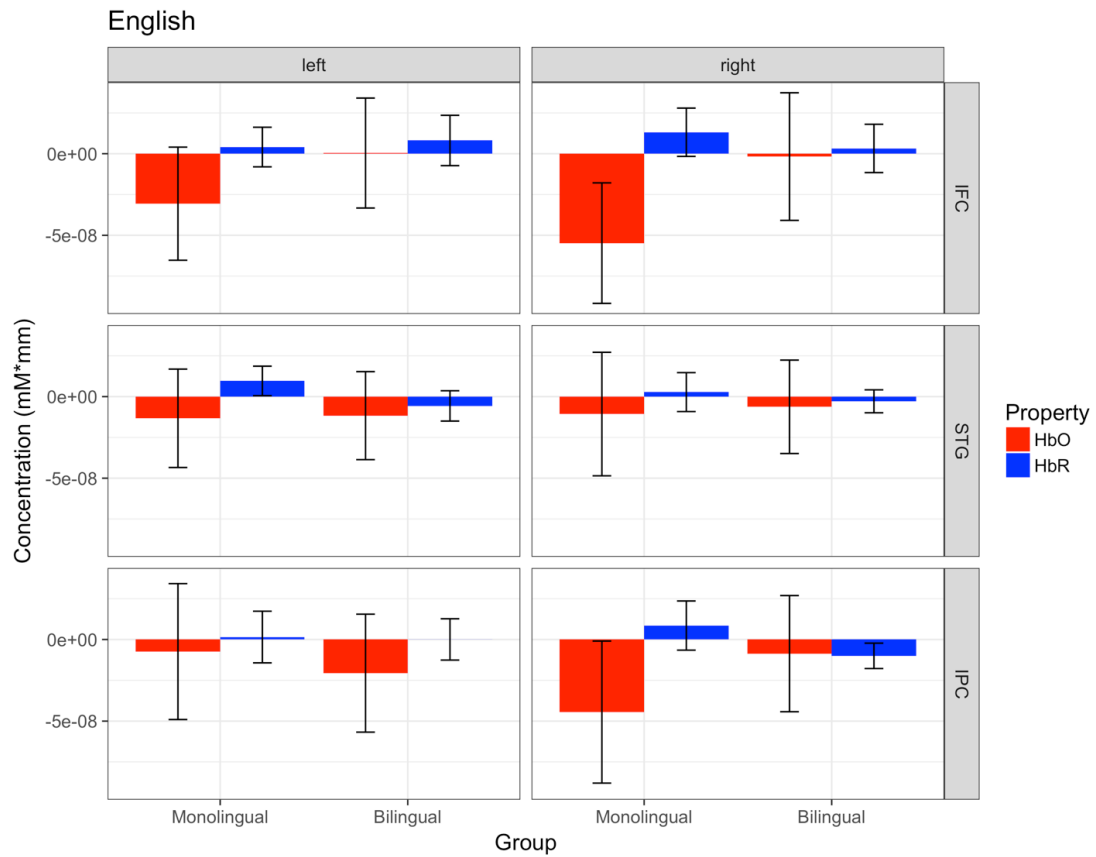


Figure 3.5. Mean HbO (red) and HbR (blue) difference values to the English consonant contrast in monolingual and bilingual adults, split across Hemisphere (columns) and ROI (rows).

### *Mandarin*

A mixed, repeated-measures 3 x 2 x 2 ANOVA with ROI and Hemisphere as within-subject factors and Group as the between-subject factor found no significant main effects or interactions in HbR. However, there was a significant three-way interaction between ROI, Hemisphere, and Group in HbO, showing that bilingual adults exhibited greater left-hemispheric activation in the IFC, or Broca's Area, than monolinguals,  $F(2, 48) = 2.35$ ,  $MS_e = 187$ ,  $p = .046$ ,  $\eta_p^2 = .09$ . There was also no significant main effect of Hemisphere in HbO,  $F(1, 24) = 3.00$ ,  $MS_e = 189$ ,  $p = .096$ ,

$\eta_p^2 = .11$ , indicating that both language groups showed no overall hemispheric differences in the processing of the Mandarin tonal/pitch contrast. These findings indicate that Mandarin-English bilingual adults utilized unilateral and bilateral mechanisms in the IFC (phonemic processing) and STG (auditory processing) while processing a native, lexical tone contrast, whereas monolinguals did not appear to detect the tonal contrast at all.

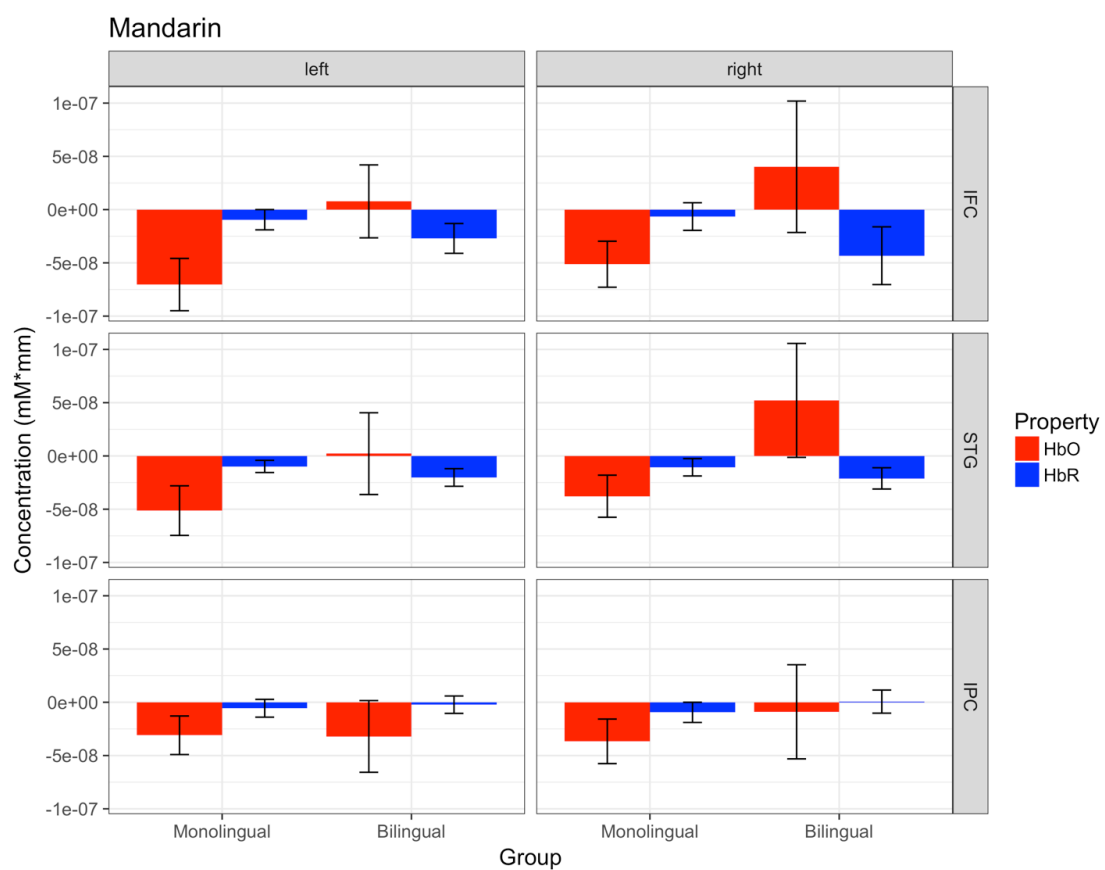


Figure 3.6. Mean HbO (red) and HbR (blue) difference values to the Mandarin lexical tone/pitch contrast in monolingual and bilingual adults, split across Hemisphere (columns) and ROI (rows).

### *Hindi*

A mixed, 3 x 2 x 2 repeated-measures ANOVA was carried out on HbO, and the analyses reported that there was a significant interaction between ROI and Group,  $F(2, 48) = 4.07$ ,  $MS_e = 396$ ,  $p = .023$ ,  $\eta_p^2 = .15$ , in which bilingual adults had significantly higher HbO difference values compared to monolingual adults in the bilateral IFC following the presentation of the non-native contrast,  $F(2, 22) = 4.69$ ,  $MS_e = 214$ ,  $p = .020$ ,  $\eta_p^2 = .30$ . On the other hand, monolingual adults appeared to have higher HbO values compared to bilinguals in the bilateral IPC. No significance was found in HbR. Through visually observing the mean difference values for HbO and HbR in Figure 3.7, it appears that both monolingual and bilingual adults did not show any HbO activation to the Hindi contrast, as all HbO difference values were negative. Statistical analyses show, however, that upon actively listening for auditory differences in the the non-native, manner-of-articulation contrast, monolingual adults significantly exhibited greater relative activation in the left and right IPC, whereas bilinguals showed a greater recruitment of the bilateral IFC in processing the same contrast.

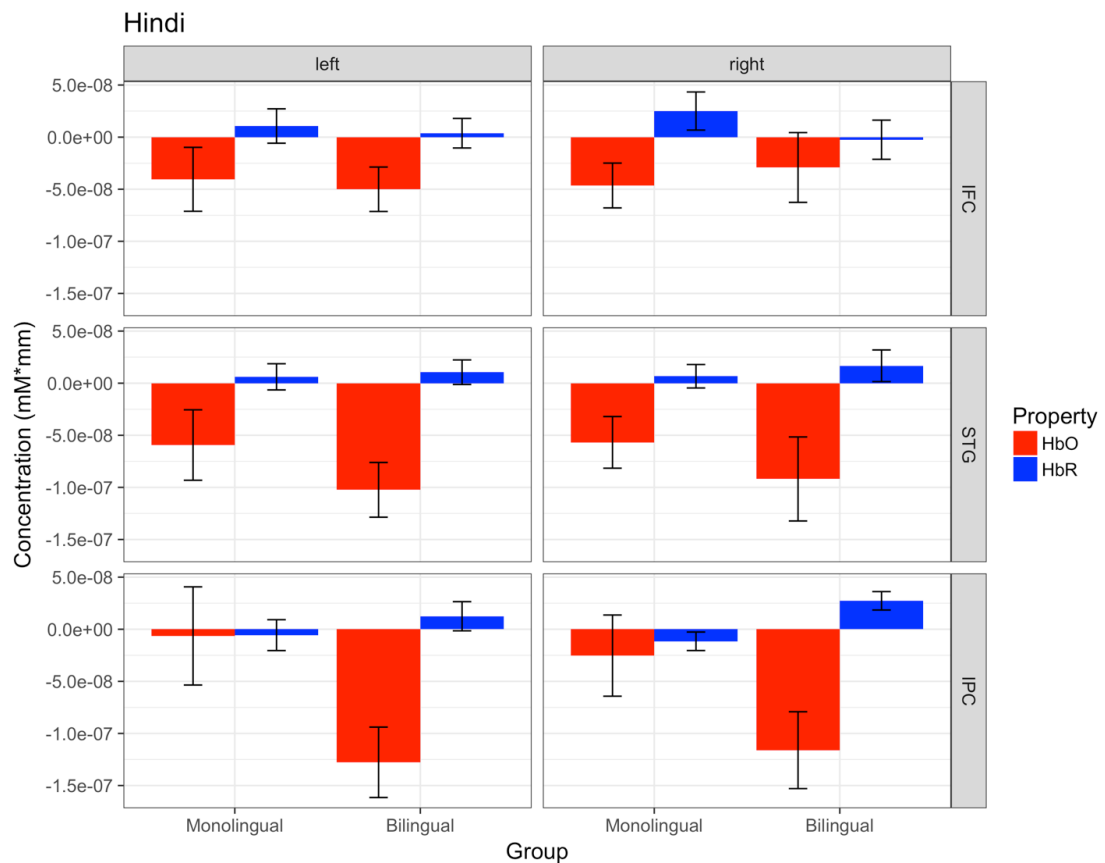


Figure 3.7. Mean HbO (red) and HbR (blue) difference values to the Hindi dental/retroflex contrast in monolingual and bilingual adults, split across Hemisphere (columns) and ROI (rows).

## Discussion

Taken together, the confirmatory analyses found no significant differences in activation between English, Mandarin, or Hindi. However, it was shown that cortical activation to all languages was relatively highest in the posterior regions in monolinguals and the anterior regions in bilinguals. This finding suggests that monolingual adults recruited attentional mechanisms located in the bilateral IPC to analyze phonological information, whereas bilingual adults utilized phonological and articulatory mechanisms located in the bilateral IFC.

Exploratory analyses suggested that there were no significant differences in HbO or HbR activity in the processing of the English contrast, suggesting that English monolingual and Mandarin-English bilingual adults processed the native English contrast similarly. This finding partially supports our hypothesis, as both language groups were expected to show comparable brain responses to the native English contrast. However, it was unexpected to observe no significant main effects or interactions in hemodynamic activity in the regions associated with native phonemic processing, such as the bilateral STG and the left IFC (Broca's Area). As previous evidence has shown that adults are excellent in discriminating native phonemic contrasts (e.g. Dehaene-Lambertz, 1997; Werker & Tees, 1984), it is assumed that the native English stimuli used in the present study was discriminable by all adult participants. One possible explanation could be that the discrimination of the English contrast was far too simple for adults. As repeated stimuli take less effort to process (the contrast was presented five times), it is possible that cognitive demand for the processing of the native contrast was reduced (e.g. Dehaene-Lambertz et al., 2006; Kobayashi et al., 2011).

For the Mandarin language condition, it was anticipated that non-tonal listeners (monolinguals) would show rightward lateralization in the processing of the tonal/pitch contrast, whereas tonal listeners (bilinguals) would show left-lateralized activation in areas responsible for phonemic processing. The exploratory results for Mandarin were partially in accordance with our predictions, where bilingual adults significantly recruited the left IFC in the processing of the native lexical tone contrast. Additionally, although not significant ( $p = .096$ ), both groups showed a relatively greater activation to the contrast in the right hemisphere, suggesting that monolinguals and bilinguals might have recruited cortical regions functionally specialized for the suprasegmental

processing of speech, such as changes in pitch (e.g. Arimitsu et al., 2011). These findings are in line with previous evidence for the role of the bilateral STG in analyzing all incoming auditory information before transmitting it to associative areas, in this case the left IFC, for subsequent linguistic processing (i.e. phonemic analysis) only in bilinguals (Zatorre, Evans, Meyer, & Gjedde, 1992; Zatorre, Meyer, Gjedde, & Evans, 1996).

One important aspect to consider is that the adults had access to lexical representations of the native word stimuli that were being presented. This could have activated additional cortical regions unrelated to phonological processing, such as the bilateral anterior STG (DeWitt & Rauschecker, 2012). To prevent the possibility of acquiring neurophysiological data that are ambiguous to the task that is being measured, the current study would have benefitted from the use of CV non-word stimuli.

Lastly, although it appears that the Hindi contrast was not perceived (indicated by the decreases in HbO activity and negative HbO values in our exploratory results see Figures 3.4 and 3.7), monolingual and bilingual adults exhibited relatively differential processing patterns to the non-native language condition that was observed in separate regions of the brain. It is important to note that the participants were instructed to pay attention to the stimuli, as subtle and/or salient changes would be presented. From actively listening to the Hindi contrast, monolinguals recruited posterior areas in the bilateral IPC, whereas bilinguals exhibited bilateral activation in the anterior areas in the IFC. The IFC is commonly known as the articulatory/phonemic processing center, where it has been shown to activate under phonemic discrimination tasks, as well as speech production and articulation. This finding is supported by another fMRI study that has also shown how early Spanish-English bilinguals with varying ages of second language acquisition and proficiency levels recruited bilateral

superior temporal and inferior frontal regions during non-native phonemic perception (Archila, Ramos, Zevin & Hernandez, 2010).

The IPC is known for its role in the sensorimotor integration of speech (Hickok & Buchsbaum, Humphries, & Muftuler, 2003; Hickok & Poeppel, 2007), as well as the allocation of attention (Behrmann, Geng, & Shomstein, 2004; Liu, Slotnick, Serences, & Yantis, 2003). Thus, it is possible that during the presentation of Hindi, monolingual adults substantially shifted their attention towards the stimulus that was relatively difficult to discriminate by recruiting attentional mechanisms located in the bilateral IPC. This finding has also been observed in monolingual Japanese listeners, where they showed prolonged activity in the bilateral IPC during the processing of the non-native English /r – l/ contrast (Zhang, Kuhl, Imada, Kotani, & Tohkura, 2005). On the other hand, bilingual adults recruited regions that activate when an acoustic change crossed a categorical boundary (Myers, Blumstein, Walsh, & Eliassen, 2009), suggesting that they utilized phonological and articulatory mechanisms to analyze subtle acoustic changes in linguistic auditory stimuli. The differential activation seen in the processing of the non-native Hindi contrast suggests that language experience influences the recruitment of executive brain regions to manipulate perceptual information.

It is important to note that there was a high dropout rate (37%) for the adult participants. This could have been attributed to the type of hair on the participants. Chinese adults generally have thicker and darker hair than Caucasians, which might have reduced the signal-to-noise ratio. However, system calibration was administered at the start of each testing session, and the experiment would not have started unless channel connectivity was good/excellent for each participant. Another factor to consider in accounting for the high dropout rate was skull shape between individuals of Chinese and Caucasian descent. It has been observed that Chinese skulls are generally



rounder and smaller than Caucasian skulls, which are more oval than Chinese heads. Moreover, the anterior (forehead) and posterior (back) areas of Chinese skulls are flatter, whereas they are more rounded in Caucasian skulls (Ball et al., 2010). As the headgear used in this study was designed for the Western anthropological head shape, it may not have been suitable for Chinese heads. Precautions should be taken in the future when using headgear on individuals with Chinese and Caucasian descent in the same study.

One limitation of this study is that we did not assess varying levels of language proficiency or age of acquisition in bilingual adults. As previous research has shown that either factor plays a role in the language processing and neural organization of bilinguals (Perani et al., 1998), this study would benefit greatly from supplementary investigations of these factors to expand our knowledge of how the availability of cognitive processes at the age of second language exposure and attained L2 proficiency would have an effect on the recruitment of brain regions to encode non-native phonemic information (e.g. Archila-Suerte, Zevin, Bunta, & Hernandez, 2012).

A second limitation to consider is that our sample of bilingual adults attained second language (English) proficiency at an average age of 5.74 years. According to the theories of perceptual narrowing, which state that all infants undergo attenuation by 12 months of age, bilingual adults would have perceptually narrowed as monolinguals. Therefore, the brain data of our adult bilingual sample might not be representative of simultaneous bilingual adults, as there is a possibility that sequential bilinguals might show different activation patterns in processing the first and second language. In a previous investigation (Wei et al., 2015), it was found that in participants who acquired L2 at an average age of 9.47 years (range = 0-21), earlier L2 exposure was associated with larger volumes in the right parietal cortex. Moreover, the cortical area of the right

superior parietal lobule increased as age of acquisition of L2 decreased. The authors of the study suggested that the structure of the human brain is reworked by the experience of acquiring a non-native language. In terms of linguistic proficiency measured using a behavioral paradigm, it appeared that participants who acquired L2 before 16 years of age exhibited native-like performance in a sentence judgment task, whereas participants who acquired L2 after 17 years showed a significant decline in performance (Birdsong & Molis, 2001). These results implicate that while sequential bilinguals may perform equally well as monolinguals in overt linguistic tasks, brain imaging is able to unveil differences in cortical structure as an effect of age of acquisition.

## CHAPTER 4

**Infants show enhanced sensorimotor activation and acoustic processing of native  
and non-native speech during universal language perception**

## Introduction

The universal ability to perceive phonetic boundaries is driven by the sensitivity of the infant's perceptual system in detecting acoustic differences that contrast in speech. However, by twelve months of age, language experience attunes the perceptual system to favor frequently occurring phonological information in the environment (Maye, Werker, & Gerken, 2002). This shift is commonly known as perceptual narrowing, which is seen as a decrease in perception of non-native information with an increase in native phonemic discrimination (Flom, 2014). This phenomenon has been studied in detail by a large body of research (Burns, Yoshida, Hill, & Werker, 2007; Kuhl et al., 2006; Mattock & Burnham, 2006; Mattock, Molnar, Polka, & Burnham, 2008; Werker & Tees, 1984), but what is less studied is the period before perceptual reorganization. More specifically, it has not been clearly established which cortical structures are involved in the language-universal perception of native and non-native phonetic units.

Our knowledge of infant speech perception is widely based upon behavioral studies and a few ERP studies concerned with phonetic discrimination abilities over the course of the first twelve months of life. These studies have demonstrated that infants before 6 to 8 months of age have the capacity to discriminate all phonetic contrasts. However, behavioral and ERP studies lack in unveiling the specific neural structures associated with phonetic processing. As such, the gap in evidence therefore provides strong theoretical motivation for the present study to, first, locate any similarities or differences in cortical activation during the processing of native and non-native speech sounds. And second, to investigate how evidence taken from previous behavioral and electrophysiological studies translates to distinct areas on the cortical surface of the infant brain. In pursuit of this challenge, it would be beneficial to demonstrate how

using more spatially-accurate recordings of infant brain activity would relate to the findings from previous behavioral and ERP studies. In order to achieve the two research aims or challenges that require the localization of active cortical areas, a viable and non-invasive neuroimaging method, such as fNIRS, that yields both excellent spatial resolution and is tolerant to movement artifacts is needed to localize active cortical regions.

### ***BEHAVIORAL EVIDENCE ON UNIVERSAL PHONEMIC PROCESSING***

Behavioral paradigms have yielded dissimilar results in identifying whether young infants can discriminate between certain types of contrasts. For example, a recent study conducted by Sundara and colleagues (2018) challenged a previous finding that suggested subtle native-language contrasts (i.e. [na]-[ɲa]) and not salient contrasts (i.e. [ma]-[na]) might not be discriminated until later at 12 months of age (Narayan, Werker, & Beddor, 2010). This was demonstrated by Narayan et al. (2010) testing Filipino and English infants on the two types of contrasts using a visual habituation paradigm. Although the Filipino language includes the syllable-initial alveolar consonants /na/ and /ɲa/, the Filipino-learning infants showed discrimination at 10-12 months but not at 6-8 months. Noting that the null results at 6-8 months may have been caused by methodological constraints, by which the procedure did not implement an infant-controlled habituation paradigm, Sundara et al. (2018) conducted a replication study where infants controlled the length of test trials. Following habituation, the infant's attention was directed towards an attention getter. Then, test trials consisted of the presentation of a checkerboard accompanied by auditory stimuli. The duration of listening (i.e. looking time) was recorded manually on-line and trials ended once the

infant looked away from the checkerboard for more than one second or at the end of the 19-second trial.

Findings from Sundara et al. (2018) revealed that, indeed, 6-month-olds were successful in discriminating subtle contrasts by firstly discriminating the same Filipino /na/-/ɲa/ contrasts as well as additional subtle dental-retroflex nasal (/ɳ/-/ɳ/) and lateral (/l/-/ɭ/) contrasts in Tamil. Sundara et al. (2018) argued that infants are initially sensitive to all subtle and salient, native and non-native phonetic contrasts, and that language experience is not *necessary* for discrimination (see Kuhl & Miller, 1975, 1978). Rather, such experience serves to modify initial, universal sensitivities by maintaining, reducing, or facilitating them. The findings from Sundara et al. (2018) question the claims made by previous reports on contrasting fricatives (e.g. [sa]-[za], [fa]-[θa], [fi]-[θi]) and how they may resist discrimination during the early stages of phonetic perception and would require the child to learn and gain more experience to be discriminated at a later age (Eilers, Wilson, & Moore, 1977; Eilers & Minifie, 1975). Selecting the appropriate behavioral infant-preference paradigm to study phonetic discrimination is crucial to the outcomes and their interpretation. Therefore, researchers must have a clear understanding of the effect they are testing for and how it can be interpreted by their data. If the outcome is unclear, it is suggested that researchers should seek the significant preference in both directions (Houston-Price & Nakai, 2004).

In summary, there is strong behavioral evidence suggesting that infants do not show differences in the discrimination of native and non-native phonetic contrasts before 6-8 months of age. What remains in question is whether the outcome translates to neural activation. Therefore, implementing neurophysiological measures would

provide a deeper and more objective look into the kinds of neural mechanisms infants utilize during the initial, universal stages of phonological processing.

### ***NEUROPHYSIOLOGICAL EVIDENCE OF UNIVERSAL PHONEMIC PROCESSING***

Studies examining the neural substrates of phonetic perception have revealed findings that are in agreement with previous behavioral research. Using a more sensitive measure (i.e. electroencephalography), Rivera-Gaxiola, Silva-Pereyra, and Kuhl (2005) found that 7-month-old infants discriminated both native English and non-native Spanish syllable contrasts through examining event-related potentials in the form of mismatch negativities (MMNs). MMNs are automatically elicited by deviant or oddball stimuli (Näätänen et al., 1997). In an earlier study, the same response was observed in 6-month-old Finnish infants, where MMNs were elicited to a native Finnish vowel and a non-native Estonian vowel (Cheour et al., 1998).

Further, an MEG investigation on brain oscillatory activity in the theta band revealed that 6-month-old infants had higher theta power for frequently presented stimuli for both native Finnish and non-native Mandarin phonemic contrasts (Bosseler et al., 2013). It has been shown that in infants and adults, relative theta power increases when attention increases (Klimesch, 1999; Stroganova, Orekhova, & Posikera, 1998), thus Bosseler et al. suggested that infants attend to frequently occurring acoustic speech signals at 6 months, and that these oscillations are driven by the distributional frequency of speech events. In sum, ERPs have been shown to be a reliable method in studying phonological processing and development. However, the measure is still not equipped with adequate spatial localization. As the language processing neural network spans over multiple lobules where distinct regions and hemispheres are specialized in the higher order processing of speech (e.g. acoustic processing in the STG vs. linguistic

processing in the left IFC) (Hickok & Poeppel, 2007; Zatorre, Evans, Meyer, & Gjedde, 1992), brain imaging becomes paramount in localizing activation during phonological processing tasks.

Some, but not many, brain imaging studies have identified a few regions that were active during syllable discrimination tasks. For example, an MEG study revealed that the left inferior frontal and superior temporal regions were active in 6-month-olds in the discrimination of [ta] and [pa] (Imada et al., 2006). Note, however, that only data from the left hemisphere was collected. A recent neuroimaging measure, fNIRS, that is increasingly becoming popular in studying infant development has been used to examine vowel discrimination in Japanese 3-4-month-old infants (Minagawa-Kawai, Naoi, Nishijima, Kojima, & Dupoux, 2007). It was found that both native and non-native contrasts were discriminated, which was indicated by bilateral hemodynamic activation in the left and right temporal areas. Another fNIRS study that tested a group of 4-6-month-old monolingual infants showed that, overall, higher activation was localized in the right STG and in the left IFC when discriminating native English and non-native Hindi consonant contrasts (Petitto et al., 2012). As the current neuroimaging evidence suggests, brain activation during phonemic discrimination is not always consistent across studies. This could be a result of inconsistent experimental procedures across studies, stimulus presentation, and the accuracy of co-registering underlying cortical structures.

Taken together, identifying brain regions involved in phonetic discrimination would be beneficial in determining the effects of language experience on speech perception. Due to the advancement in a wide range of methods to measure neural activity in infants, neurophysiological evidence that complements behavioral results would help enhance the current understanding of one of the first steps in language



acquisition. The goal of the present study aimed to deepen the understanding of phonological development through using fNIRS to elucidate neural mechanisms employed during the processing of native and non-native phonemic information. It is driven by two main research questions: (1) Are there differences in activation patterns to native and non-native speech? (2) If so, where do these differences, or lack thereof, manifest on the cortex of the infant brain? We hypothesized that 5-7-month-old infants would process native and non-native phonemic contrasts equally, and that fNIRS would capture the regions that are activated during phonological processing. We expected to observe activation patterns in the left and right temporal areas, as well as Broca's Area in the left inferior frontal cortex. The current study implemented a block design for stimulus presentation and measured the left and right inferior frontal, superior temporal, and inferior parietal regions in response to English, Mandarin, and Hindi phonemic contrasts.

## Method

### *Participants*

A total of 24 infants were included in the final analysis (11 girls, mean age = 6.51 months, SD = 20.22 days, min = 5.60 months, max = 7.83 months). An additional five infants were excluded due to fussiness ( $n = 2$ ), insufficient placement of the headgear ( $n = 2$ ), and excessive motion artifacts in the raw signal ( $n = 1$ ). All infants were healthy and born full-term with no auditory or cognitive disabilities. All infants were from English monolingual backgrounds, and their caregivers reported that they had no experience in listening to Mandarin or Hindi. Families were given a thorough description of the study and what the testing procedure would entail. Then, caregivers

were given the opportunity to ask questions regarding the study and gave informed consent before testing began.

### *Stimuli*

Three minimal pair CVC contrasts in English, Mandarin, and Hindi were used. Native female speakers from each language naturally produced five tokens for each word in adult-directed speech. The stimuli consisted of an English consonant contrast /dæb – tæb/, a Mandarin lexical tone contrast /taw2 - taw3/, and a Hindi dental-retroflex contrast /ɖa:l - ɖa:l/. All tokens were unfiltered and equalized to an intensity of 75 dB and adjusted to a duration of 650 ms using Praat (Boersma & Weenink, 2012). Detailed information about the stimuli can be referred to in Chapter 3.

### *fNIRS recording*

Hemodynamic responses were sampled at a rate of 7.81 Hz using a multichannel NIRx NIRScout system emitting two continuous wavelengths at 760 and 850 nm. Eight infrared emitters and eight detectors were arranged in two staggered 2 x 4 arrays on a flexible headband secured in the back by Velcro strips. They were positioned bilaterally on each side of the head resulting in 9 channels per hemisphere (Figure 4.1). The source-detector separation was 20 mm, and the placement of the headband was guided by the international 10-20 system for EEG.

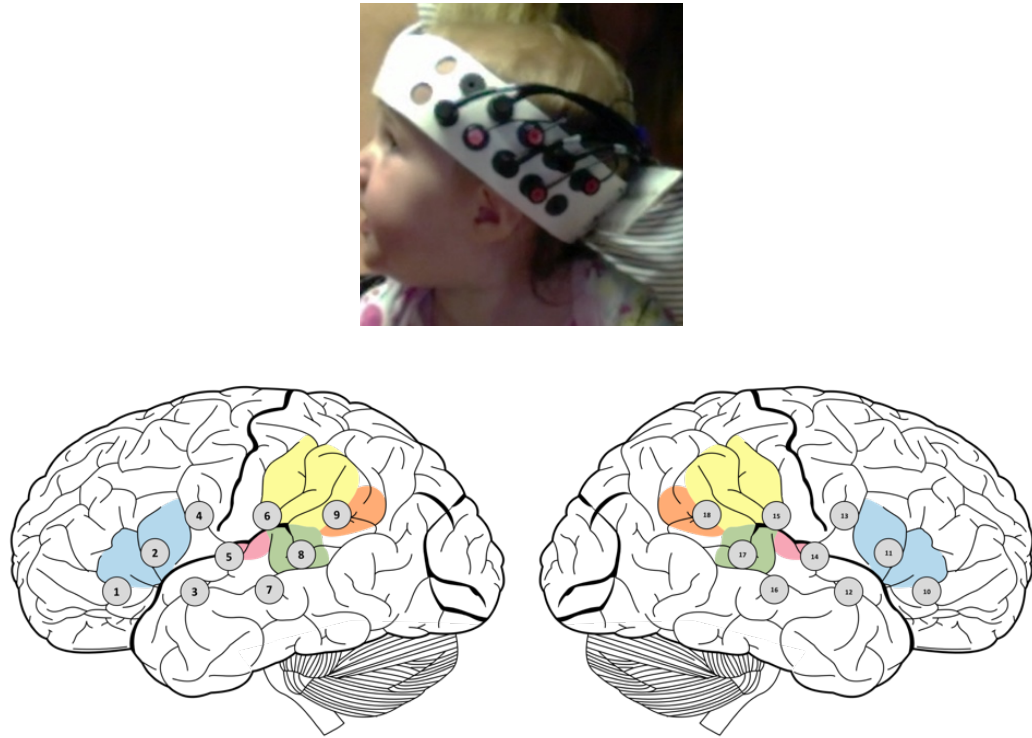


Figure 4.1. (Top) Image of an infant wearing the fNIRS headband showing the optode configuration during a testing session. The red optodes signify light source emitters and the black optodes represent infrared light detectors. (Bottom) Anatomical diagram of channel placement over the approximate underlying cortical structures.

### *Procedure*

The experiment was conducted in a noise-isolated laboratory room where the infant sat on their caregiver's lap facing approximately 1 meter away from a computer monitor. Then, the experimenter carefully aligned the fNIRS headband on the infant. Sound stimuli were administered through concealed front-facing loudspeakers at approximately 70 dB. The timing and presentation of the stimuli was controlled via MATLAB, and NIRStar 15.0 was used for data acquisition.

Stimuli from the three languages were presented, with the order of languages counterbalanced across all participants. For each language, one word from the minimal-

pair served as the “standard” or baseline word, and the other served as the “target” or change word. The standard and target words were also counterbalanced across participants. Each language had a non-alternating and alternating condition (e.g. English alternating = Ealt; Mandarin non-alternating = Mnon; Hindi non-alternating = Hnon), where the non-alternating condition contained 10 randomized presentations of the standard tokens (e.g. A-A-A-A... or B-B-B-B...), and the alternating condition had 10 randomized presentations alternating between the standard and target tokens (e.g. B-A-B-A... or A-B-A-B...). Intervals between each token were jittered between 1000 and 1500 ms to avoid phase-locked neural responses (Benavides-Varela, Hochmann, Macagno, Nespor, & Mehler, 2012; Gervain, Macagno, Cogoi, Peña, & Mehler, 2008). Twenty-five-second silence periods always followed each auditory block to allow for the hemodynamic response to return to baseline before the start of the next auditory block. The three languages alternated in a fixed order and were repeated five times for a total of 30 blocks of auditory stimulation. A diagram of the procedure is shown in Figure 4.2. A silent video of slowly moving black and white shapes played on the monitor as a visual filler while the experimenter waved silent toys to keep the infant entertained throughout the study. The video did not synchronize with the auditory stimuli and was not associated with the study. The duration of the testing session was approximately 20 minutes. Parents and caregivers were given a full debrief at the end of the experiment.

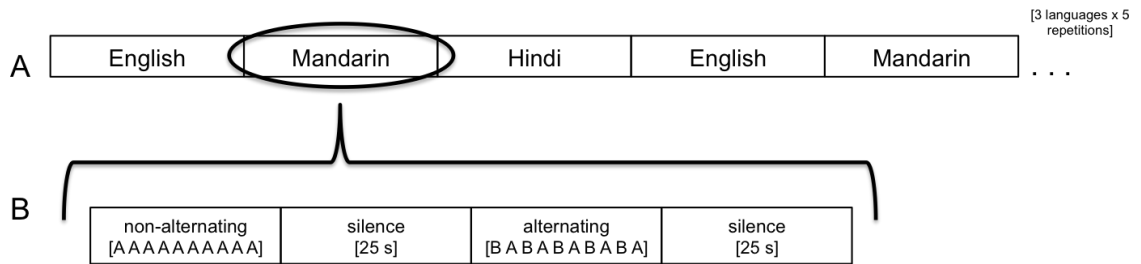


Figure 4.2. Illustration of the procedure. (A) The three language conditions were repeated in a fixed order for a total 15 repetitions. (B) Each language condition contained a non-alternating and an alternating auditory block each followed by a 25-second silence period. The order of presentation for the auditory blocks was counterbalanced across participants.

#### *fNIRS pre-processing*

Optical data collected from two wavelengths were transformed into HbO, HbR, and HbT. HomER2 was used to carry out motion artifact correction and filtering procedures (Huppert, Diamond, Franceschini & Boas, 2009). MAs were identified as changes in the raw signal that exceeded 0.1 mmol x mm over 0.1 s, and wavelet-based motion artifact removal was applied to reconstruct the signal after identifying and eliminating the MAs (Brigadoi et al., 2014; Molavi & Dumont, 2012). The data were also bandpass-filtered between 0.01 and 0.7 Hz. Baseline correction was taken from a five-second interval preceding the onset of auditory stimulation where HbO responses for each condition (auditory stimulation + silence period) were averaged across all repetitions per participant. Finally, the total time course of each condition (40 s) was plotted for visual inspection of the peaks of activation. Figure 4.3 depicts an example of the HbO hemodynamic response function (HRF) for each condition from one channel.

In order to be included in the final analysis, it was required that each channel had to contain valid data in all three language conditions, as well as a minimum of three valid trials per condition. If any auditory block was invalid, the entire language condition was excluded from the analysis.

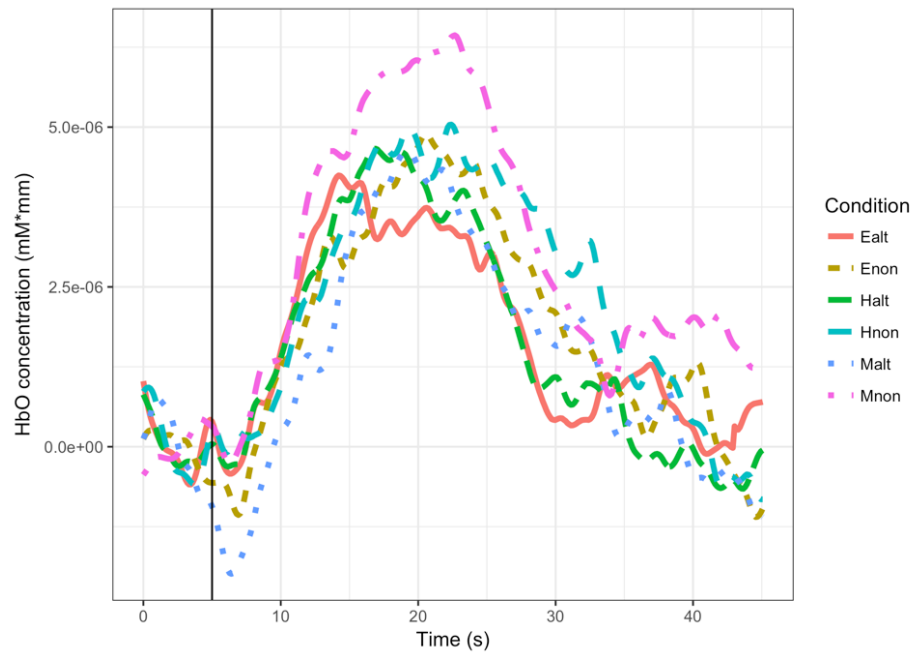


Figure 4.3. HbO time courses for each auditory condition (E = English; H = Hindi; M = Mandarin; alt = Alternating; non = Non-alternating) from Channel 14 in the right STG. The vertical line signifies the onset of auditory stimulation. This channel was selected for a visual example based on the clear responses elicited in each condition.

#### *Defining cortical regions of interest (ROIs)*

Three main regions of interest were selected based on the co-registration of the underlying cortical structures using the international 10-20 system. Channels were grouped into ROIs according to the 10-20 referencing system and then averaged within each region. The channels lay bilaterally over the IFC (LH: channels 1, 2, 4; RH: 10,

11, 13), STG (LH: channels 3, 5, 7, 8; RH: 12, 14, 16, 17), and IPC (LH: channels 6, 9; RH: 15, 18) (see Figure 4.4).

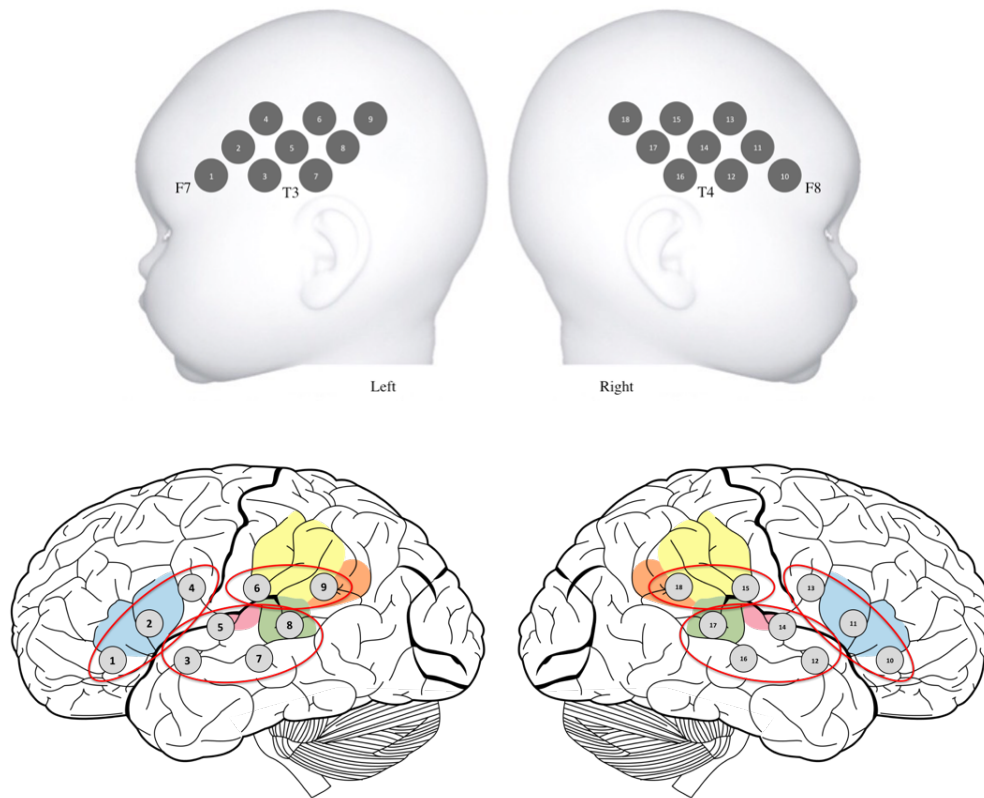


Figure 4.4. (Top) Channel placement over the left and right hemispheres. (Bottom) Projection of the channels over the approximate underlying cortical structures. The three ROIs in each hemisphere are circled in red (anterior to posterior: IFC, STG, IPC).

### *Statistical analyses*

A 20-second time window was set five seconds after the onset of the auditory stimulation blocks (non-alternating and alternating). The length of the time window ensured that all peak values from each condition were incorporated. This time window was then applied to all channels. Previous fMRI research on adults has shown that event-related BOLD responses have different peak latencies (2-4 s differences) during

phonological processing (Thierry, Boulanouar, Kherif, Ranjeva & Démonet, 1999). However, peak latencies in infants are longer than adult peak latencies, whereby infants' vascular systems may not be mature enough to elicit adult-like responses (Issard & Gervain, 2018). Therefore, setting a broad time window of 20 seconds would ensure that all response peaks would be incorporated.

Data points within each time window were averaged, and prior to statistical analyses, boxplots were generated from the averaged values to observe the distribution of peak activity and for the identification of outliers. Two participants were removed from the final analysis, because their values were outliers for over 10% of their data (condition x channel;  $6 \times 18 = 108$  data points), or for more than eleven instances. The outlying values of participants who were not excluded from the final analysis were not removed.

## Results

Hemodynamic activity in the bilateral inferior frontal, superior temporal, and inferior parietal lobes were recorded in English-learning 5- to 7-month-olds in response to English, Mandarin, and Hindi phonemic contrasts. We expected the participants to exhibit equal or similar responses for all languages, because at 5-7 months of age, their perceptual system is encoding phonetic information at the universal, acoustic level prior to perceptual narrowing.

### *Confirmatory analyses*

In an overall analysis, a  $3 \times 3 \times 2 \times 2$  repeated-measures ANOVA was conducted on HbO with Language (English, Mandarin, Hindi), ROI (IFC, STG, IPC), Condition (Alt, non-Alt), and Hemisphere (left, right) as within-subjects factors. Results revealed



that the main effect of Language was insignificant,  $F(2, 46) = .20$ ,  $MS_e = 656$ ,  $p = .823$ ,  $\eta_p^2 = .01$ . However, there was a significant interaction between ROI and Hemisphere,  $F(2, 46) = 8.95$ ,  $MS_e = 24642$ ,  $p = .001$ ,  $\eta_p^2 = .28$ , where subsequent analyses showed that brain activation was significantly lateralized to the right STG,  $F(1, 23) = 13.11$ ,  $MS_e = 7250$ ,  $p = .001$ ,  $\eta_p^2 = .36$ .

### ***Exploratory analyses***

The following discusses the results for English, Mandarin, and Hindi in more detail. HbO activation patterns were examined in 3 x 2 x 2 repeated-measures ANOVAs with HbO values as the dependent variable and ROI, Condition, and Hemisphere as the within-subject factors.

### ***English***

There was a significant main effect of ROI with the bilateral STG showing the highest activation,  $F(2, 46) = 9.98$ ,  $MS_e = 15741$ ,  $p < .001$ ,  $\eta_p^2 = .30$ . Additionally, there was a significant interaction between ROI and Hemisphere,  $F(2, 46) = 5.53$ ,  $MS_e = 7222$ ,  $p = .007$ ,  $\eta_p^2 = .19$ , where simple main effects analyses found that the STG showed the greatest activation in the right hemisphere,  $F(2, 46) = 14.19$ ,  $MS_e = 9048$ ,  $p < .001$ ,  $\eta_p^2 = .38$ . Pairwise comparisons (Bonferroni corrected) found significant differences between STG-IFC (mean difference 38.79,  $p < .000$ ) and STG-IPC (mean difference 20.96,  $p = .035$ ). In the left hemisphere, a marginally significant simple main effect indicated that the IPC had the highest amount of activation,  $F(2, 46) = 3.03$ ,  $MS_e = 2433$ ,  $p = .058$ ,  $\eta_p^2 = .12$ .

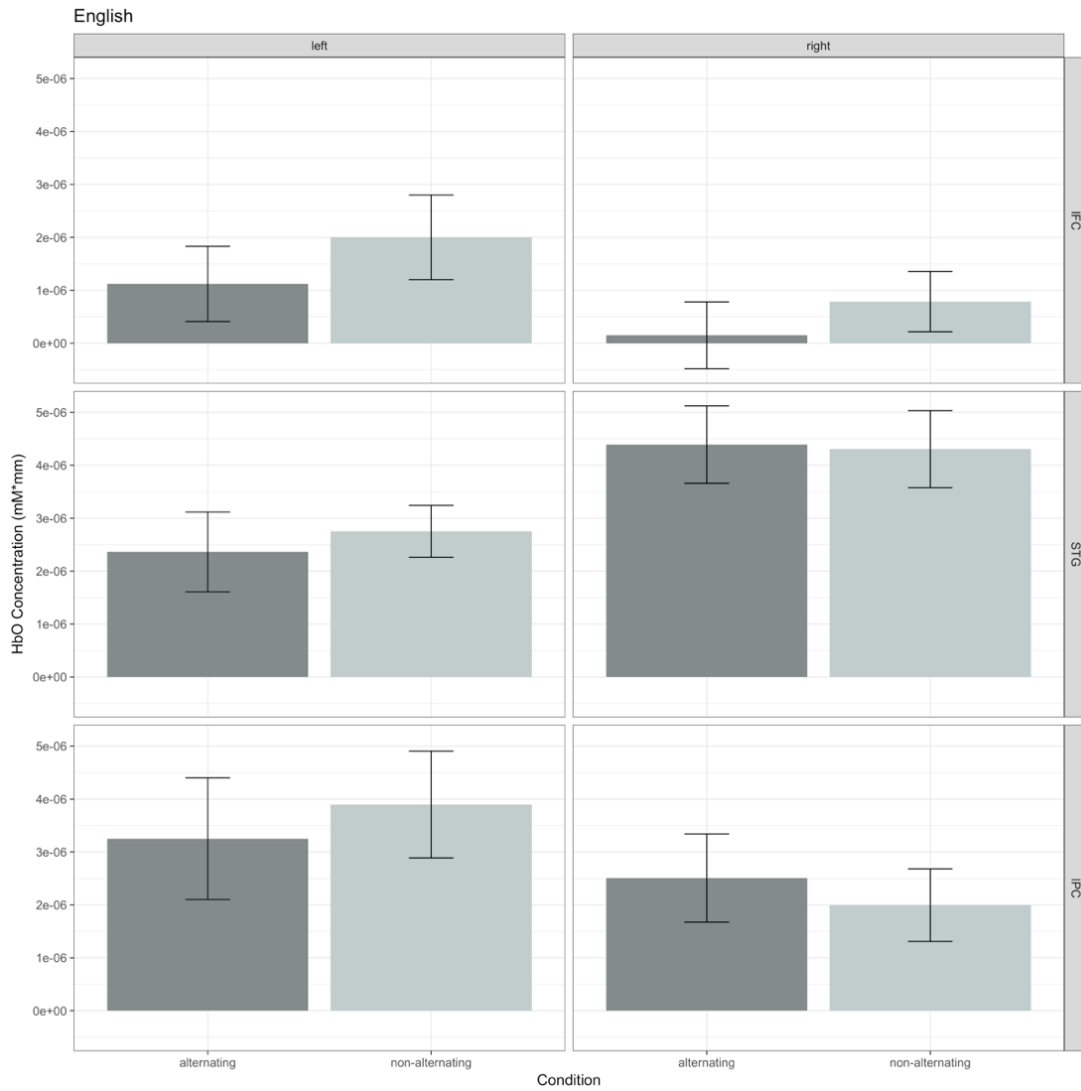


Figure 4.5. Mean HbO values for the English alternating (dark grey) and non-alternating (grey) conditions. Columns indicate hemispheres and rows indicate ROIs.

### *Mandarin*

The results revealed a significant main effect of ROI, with the greatest activation in the bilateral STG,  $F(1.54, 35.47) = 16.03$ ,  $MS_e = 25610$ ,  $p < .001$ ,  $\eta_p^2 = .41$ . An interaction was also found between ROI and Hemisphere,  $F(2, 46) = 6.00$ ,  $MS_e = 7420$ ,  $p = .005$ ,  $\eta_p^2 = .21$ , where subsequent analyses indicated three simple main effects. First, in the left hemisphere, the IPC showed the greatest activation,  $F(2, 46) = 4.68$ ,  $MS_e = 3310$ ,  $p = .014$ ,  $\eta_p^2 = .17$ . Pairwise comparisons (Bonferroni corrected)

revealed a significant difference between IPC-IFC (mean difference 23.42,  $p = .015$ ). Second, in the right hemisphere, the STG elicited the greatest activation,  $F(2, 46) = 19.50$ ,  $MS_e = 10273$ ,  $p < .001$ ,  $\eta_p^2 = .46$ . Pairwise comparisons showed significant differences between STG-IFC (mean difference = 41.34,  $p < .001$ ) and STG-IPC (mean difference = 22.20,  $p = .007$ ). Third, and in the opposite direction, brain activation in the STG was significantly lateralized to the right hemisphere,  $F(1, 23) = 9.59$ ,  $MS_e = 11974$ ,  $p = .005$ ,  $\eta_p^2 = .29$ .

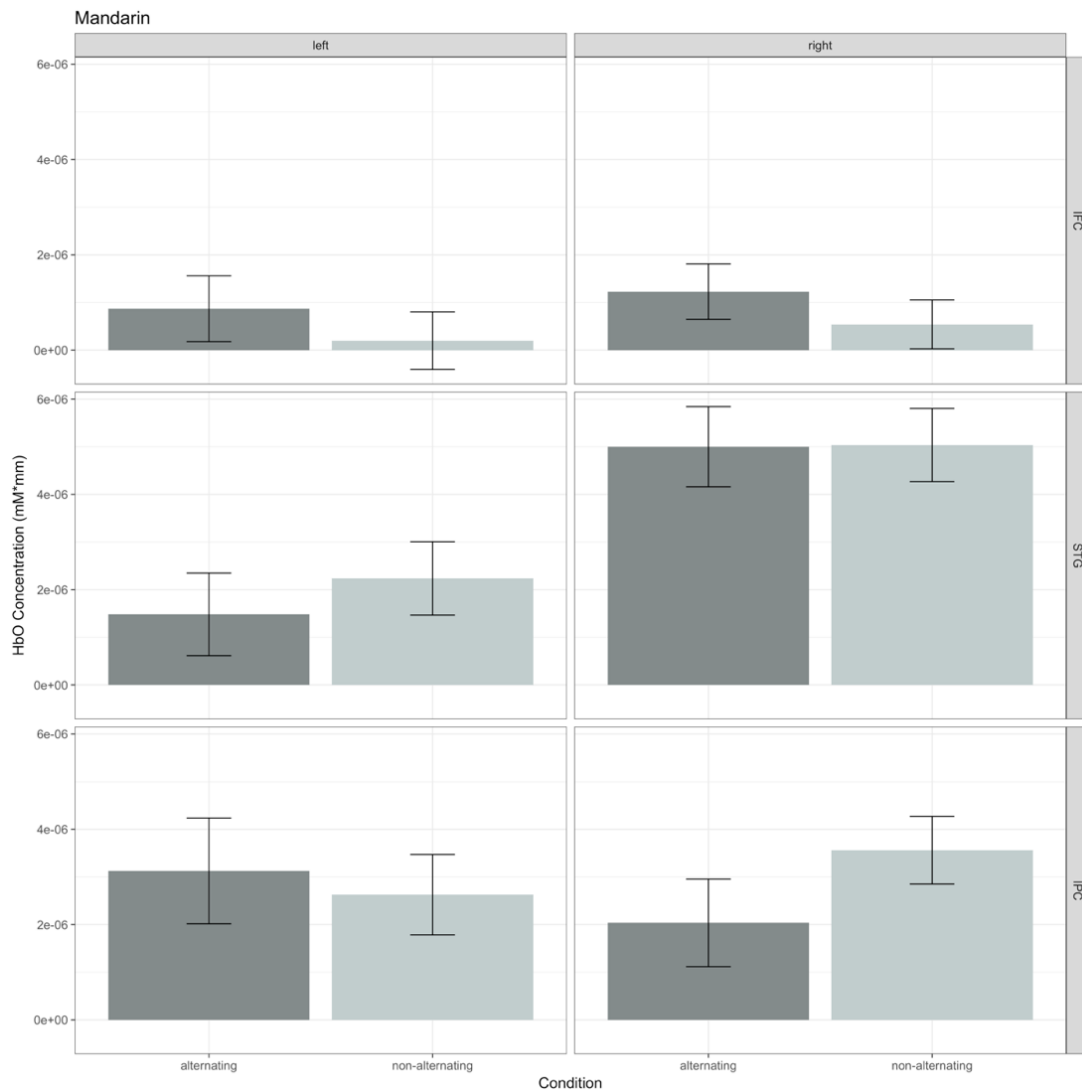


Figure 4.6. Mean HbO values for the Mandarin alternating (dark grey) and non-alternating (grey) conditions. Columns indicate hemispheres and rows indicate ROIs.

### *Hindi*

There was a significant main effect of ROI, with the greatest activation in the bilateral IPC,  $F(1.49, 34.34) = 11.68$ ,  $MS_e = 35326$ ,  $p < .001$ ,  $\eta_p^2 = .34$ . There was also a significant interaction between ROI and Hemisphere,  $F(1.49, 34.25) = 6.99$ ,  $MS_e = 14654$ ,  $p = .006$ ,  $\eta_p^2 = .23$ , where subsequent analyses indicated that IPC in the left hemisphere showed the greatest amount of activation,  $F(1.41, 32.47) = 7.46$ ,  $MS_e = 12996$ ,  $p = .005$ ,  $\eta_p^2 = .25$ . Pairwise comparisons indicated a significant difference between IPC-IFC (mean difference = 38.77,  $p = .014$ ). Although the effect of Condition was not significant, the means show, however, that activation was higher for the Alt condition ( $M = 50.50$ ,  $SD = 13.08$ ) than the non-Alt condition ( $M = 41.93$ ,  $SD = 12.20$ ) in the left IPC (see Figure 4.7). Further, in the right hemisphere, STG had the greatest activation,  $F(2, 46) = 13.95$ ,  $MS_e = 9467$ ,  $p < .001$ ,  $\eta_p^2 = .38$ . Pairwise comparisons indicated a significant difference between STG-IFC (mean difference = 39.70,  $p < .001$ ). The current findings showed that during the processing of Hindi stimuli, the left IPC showed the greatest increase in HbO activation relative to all other languages, and the means indicated a greater response to the Alt over the non-Alt condition.

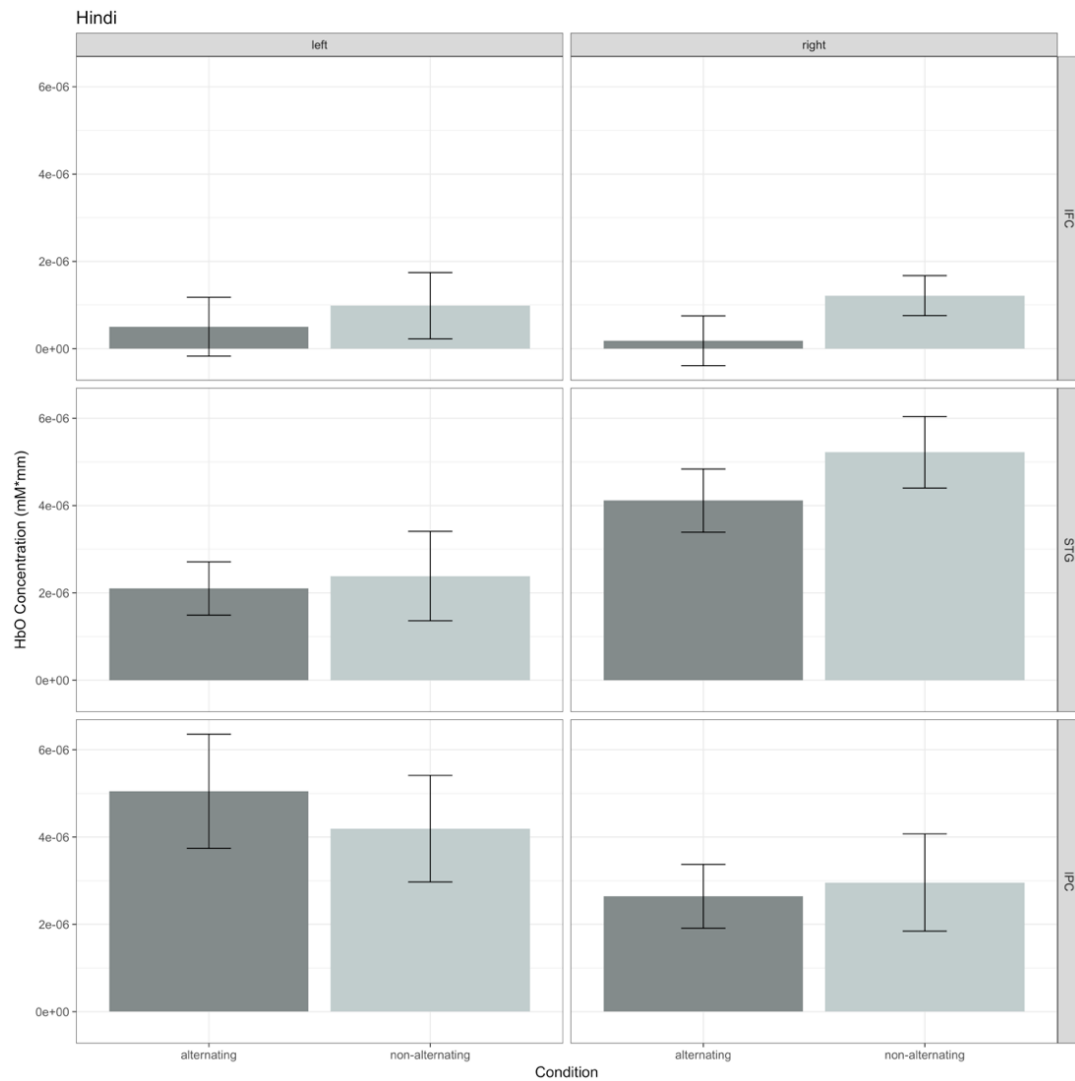


Figure 4.7. Mean HbO values for the Hindi alternating (dark grey) and non-alternating (grey) conditions. Columns indicate hemispheres and rows indicate ROIs.

In summary, our confirmatory findings showed that 5-7-month-old infants responded similarly to English, Mandarin, and Hindi stimuli by eliciting the same pattern of activation that was significantly observed in the right STG. Exploratory analyses further showed that for each language, infants exhibited a uniform pattern of activation observed in the left IPC and right STG. However, the effect of Condition was shown to be insignificant, thus conclusions could not be drawn on infants'

discrimination of the language contrasts. Therefore, from this point forward, all three languages will be referred to as *conditions* and not contrasts.

All infants underwent the same experimental conditions, and there were no major issues regarding the alignment and placement of the headband. Additionally, our pre-processing parameters required that each participant had to have valid data in at least three blocks in each condition to be included in the final analysis. Upon visual inspection, raw data from each participant appeared stable and uncontaminated with considerable noise (e.g. bad channel connectivity). It is also important to note that our sample had a low attrition rate (17%). We therefore conclude that our data are robust, clearly supported by the consistent and precise activation patterns across all languages, which will be discussed in the next section.

## Discussion

The current study examined activation patterns during native and non-native language processing in English-learning infants aged 5 to 7 months old. We demonstrated that it was possible to use fNIRS to localize hemodynamic activation in young infants during phonological perception. Additionally, we were able to use this neuroimaging technique to elucidate brain regions involved in processing familiar (English consonant) and unfamiliar (Mandarin lexical tone, Hindi dental/retroflex) language conditions.

In short, our exploratory findings revealed similar activation patterns localized in the left IPC and right STG for all three languages, whereas our confirmatory findings only revealed right-lateralized STG activation. However, differences in cortical activation between alternating and non-alternating conditions for any language were not significant. Two possible explanations may have contributed to this unexpected

finding. First, the presentation of experimental stimuli might have affected brain activation. Previous studies have used the block alternating paradigm in which silence periods were not used. Instead, blocks containing non-alternating stimuli (A-A-A-A-) were used as the baseline and blocks with alternating stimuli (A-B-A-B-) were used as the test condition. Then, the baseline and test blocks alternated for a predetermined amount of repetitions (e.g. Arimitsu et al., 2011; Furuya & Mori, 2003; Minagawa-Kawai et al., 2013; Sato, Sogabe, & Mazuka, 2010; Sato, Utsugi, Yamane, Koizumi, & Mazuka, 2013). The present study used a silence period between alternating and non-alternating blocks to avoid signal contamination from adjacent stimulation periods. However, the silence period of 25 seconds between 20-second blocks of auditory stimulation might have been too long for infants to retain the information learned from the previous stimulation period and compare with the following block. This might explain why similar levels of activation were observed for the alternating and non-alternating conditions across all languages and brain regions. A closer look into verbal working memory in infants under 12 months of age might help confirm whether or not silence periods are useful in phonemic discrimination in fNIRS experimental paradigms. To the best of our knowledge, there is no published research on phonological working memory in infants under 12 months of age.

Another explanation for the absence of discrimination concerns the effect of task demands. A major difference should be considered when comparing outcomes from active listening in behavioral paradigms and passive listening during neuroimaging sessions. Behavioral paradigms require an overt response from the infant (e.g. head turn, fixation time), whereas infant neurophysiological measures do not require such a response. Early imaging studies on adults found that temporal regions activated bilaterally during passive listening tasks (Binder et al., 1994), whereas active

listening tasks that require attention such as phonemic discrimination or identification induced greater left-hemispheric activity in the STG and to a lesser extent, in the left IFC (Démonet et al., 1992). As the focus of attention affects how language input is processed, future neuroimaging studies that include an attentional component and/or active behavioral responses from the infant would help clear up equivocal results on brain activation during passive listening procedures.

Our findings with regard to 5-7-month-old English monolingual infants suggest that they processed native and non-native language conditions equally, which support previous literature on the universal, acoustic processing of speech before the consolidation of native phonemic categories. These findings provide new contributions to the current literature through identifying specific brain regions involved in universal phonetic processing. We discuss next the roles of the left IPC and right STG and how they contribute to the perception of native and non-native speech information.

### ***LEFT IPC AS A SENSORIMOTOR INTERFACE IN SPEECH PERCEPTION***

The exploratory findings from the current study revealed that all infants exhibited robust responses in the left IPC in response to all language stimuli. This region is part of the dorsal pathway of speech perception, which comprises cortical structures from inferior prefrontal cortex, premotor cortex, superior temporal gyrus, and inferior parietal cortex (see Figure 4.8). The main function of this dorsal stream circuitry is to translate acoustic speech input received in the temporoparietal regions into articulatory representations in the prefrontal cortex, or Broca's Area (Hickok & Poeppel, 2007). Thus, the involvement of sensorimotor brain regions in speech perception supports the view that speech perception entails motor processes (Bruderer, Danielson, Kandhadai, & Werker, 2015).



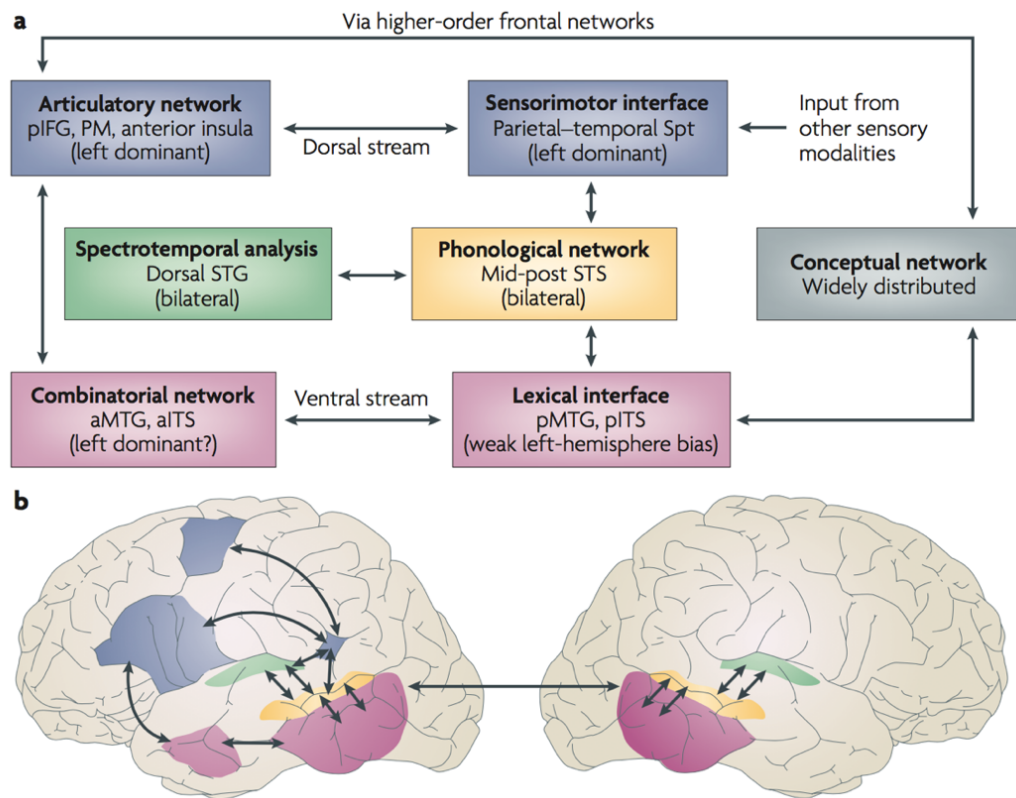


Figure 4.8. The dual-stream model of the functional anatomy of language. Reproduced from Hickok & Poeppel (2007).

Motor theories of speech perception and computational modeling assume a link between sensory input and motor speech systems (Liberman & Mattingly, 1985; Westermann & Miranda, 2004), suggesting a neural mechanism that codes and stores sensory speech input must be present to regulate the fine motor control of articulatory movements for accurate speech reproduction. This area is known as Sylvian parietal temporal (Spt), and it is located in the Sylvian fissure at the boundary between the posterior superior temporal gyrus and the inferior parietal cortex (Hickok & Poeppel, 2007). The current findings show that for Hindi, a significant main effect of ROI was reported with the bilateral IPC eliciting the greatest activation. In contrast, for English and Mandarin, significant main effects of ROI were seen in the STG. These findings

suggest that increased sensorimotor involvement is required for processing unfamiliar speech sounds that are produced with more complex articulatory movements.

In language acquisition, infants must also learn to accurately produce sound sequences in their native language in addition to acquiring a language-specific repertoire of phonemes. Interestingly, at 5-7 months of age, infants engage in vocal play and start to show signs of canonical babbling where they are gaining control over their oral articulators and vocal productions. During vocal play, infants experiment with their control over pitch and loudness through producing vocalizations that resemble squeals, growls, and yells. Additionally, adult-like vowels are being produced as well as marginal babbling, which contains CV or VC features (e.g. “daaaa”, “uuuum”) but lack the mature regular-syllable timing properties of canonical babbling (Vihman, 2014).

### ***MATURATIONAL DIFFERENCES IN THE RIGHT TEMPORAL CORTEX AND ITS ROLE IN THE ANALYSIS OF SPEECH***

In accordance with previous research, we expected to observe bilateral and possibly left-dominant activation in the STG in response to all stimuli, as both hemispheres first process all incoming linguistic and non-linguistic auditory information. Then, further linguistic analyses would be carried out in the left hemisphere (Binder et al., 1994; Wilson & Iacoboni, 2006). Contrary to our predictions, the current confirmatory results revealed a robust right-lateralized response in the STG to each language in both alternating and non-alternating conditions. Interestingly, our results were in accordance with a previous MEG study of phonemic discrimination that found a right hemispheric bias in English monolingual 11-month-olds in a double-oddball paradigm (Ferjan Ramírez, Ramírez, Clarke, Taulu, & Kuhl, 2016). The standard stimulus in that study was an ambiguous and mutual phonetic unit between Spanish and English, and the two deviant stimuli were only exclusive to either

language. The monolingual infants' neural responses to the Spanish and English oddballs were widespread over the right inferior frontal cortex, superior temporal areas, and sensorimotor regions. Areas in the left hemisphere were also involved but to a lesser extent. Similarly, in a previous fNIRS study (Petitto et al., 2012), an overall greater right hemispheric activation in the STG was seen in all monolingual, bilingual, younger (4-6 months), and older (10-12 months) infants during phonemic discrimination in a classic oddball paradigm. Moreover, the right-lateralized pattern of activity was more prevalent in monolinguals. Although an explanation of the finding was not given in Petitto et al. (2012), we postulate three possible explanations that might account for our own finding of a rightward dominance in the STG.

First, in our study the duration between each word presentation (jittered between 1000-1500 ms) did not resemble the fast-paced nature of natural speech. Previous studies on hemispheric lateralization have suggested that the left hemisphere analyzes fast, temporal aspects of speech and the right hemisphere for slower, more spectral properties (Arimitsu et al., 2012). The slower frequency with which speech stimuli were presented in the present study may therefore have favored the characteristics that tend to be processed in the right hemisphere. Second, our stimuli were presented in strings of auditory sounds, without context. Young, preverbal infants have not yet acquired an adequate lexicon and thus cannot contextualize sounds from a stream of syllables (Ferjan Ramírez et al., 2016). Previous EEG research has suggested that left-hemispheric specialization of speech emerges slowly as the size of the lexicon increases (Mills, Plunkett, Prat, & Schafer, 2005), which might explain the absence of a robust left-hemispheric response in our 5-7-month-olds. Lastly, infants' vascular systems might not have been mature enough to give reliable hemodynamic responses due to insufficient cerebral blood flow (Issard & Gervain, 2018; Meek et al., 1998). Rapid

maturation changes of the auditory system take place during the first year of life. In an fMRI experiment, Leroy and colleagues (2011) assessed the maturation of the linguistic dorsal pathway in fourteen sleeping infants 2.6 – 16.3 weeks of age. The authors discovered that the right superior temporal sulcus matured faster than its left homologue. Coupled with the overall larger right hemispheric cerebral blood flow at rest (Chiron et al., 1997; Roche-Labarbe et al., 2012) and during phonetic discrimination tasks (Mahmoudzadeh et al., 2013), maturational factors may have contributed to our unilateral findings in the STG.

### ***NATIVE AND NON-NATIVE PHONOLOGICAL PROCESSING IN THE LANGUAGE-UNIVERSAL BRAIN***

The present study sheds new light on the neural substrates of phonetic processing at 5-7 months of age, where infants' perceptual systems have not yet fully attuned to their native language. Previous studies have demonstrated that at this age, infants are sensitive to phonological information on a universal basis, and as a result, they are able to discriminate any phonemic contrast to which they are tested (Eimas, Siqueland, Jusczyk, & Vigorito, 1971; Werker & Tees, 1984). Our findings complement existing evidence in behavioral and electrophysiological domains by demonstrating that our sample of infants showed no overall hemodynamic differences in processing native (English) and non-native (Mandarin, Hindi) language conditions. This was suggested by a uniform pattern of activation in the left IPC and right STG. We expected that there would be some activation in Broca's Area within the left IFC during native speech processing. Although no overall activation of the IFC was significant in the current analyses, this result is supported by an earlier finding that suggested the region did not significantly activate under passive listening to meaningless speech, which in the current case was the presentation of spontaneous

syllables (Wilson, Saygin, Sereno, & Iacoboni, 2004). Furthermore, activation in Broca's Area signifies phonemic processing (Myers, Blumstein, Walsh, & Eliassen, 2009; Zatorre, Evans, Meyer, & Gjedde, 1992). It is likely that the absence of cortical activation in the left IFC indicated that this region was not yet specialized for phonemic processing in 5-7-month-old monolingual infants, and that younger infants' perception of native and non-native speech was solely based on acoustic (STG) and not phonemic analyses (IFC) (Zatorre, Meyer, Gjedde, & Evans, 1996). The finding of left IPC (sensorimotor cortex) activation indicates that although younger infants at 5-7 months of age have not started producing words yet, listening to speech activated this area to possibly allow for the mapping of motor articulatory representations in executive regions (e.g. left IFC) in preparation for the production of more complex sound sequences (i.e. words) at a later stage in development.

In summary, the present study was successful in using fNIRS to confirm that monolingual infants at 5-7 months of age did not show differential activation to speech sounds from native and non-native languages. This finding suggests that the infants processed all languages in the same way, and that they did not show differential processing for any language. Therefore, the present findings might possibly indicate a universal pattern of neural activation to all language stimuli during the early stages of speech perception. Further, our exploratory results found activation in areas suggesting the involvement of sensorimotor processes in speech perception. This is in accordance with the onset of canonical babbling, around 6 to 9 months of age, where non-verbal infants are gaining increased motor control of their articulators for speech reproduction. We also provide evidence in support of the notion that the sensorimotor integration in speech perception is closely linked to speech production (Hickok, Houde, & Rong, 2011).

## CHAPTER 5

**Bilingual 10-12-month-old infants show greater sensitivity to non-native  
phonemic information in Broca's Area**

## Introduction

In the previous chapter we demonstrated that English monolingual infants at 5-7 months of age did not show differential activation to either English, Mandarin, or Hindi language conditions, which was indicative of a possible universal pattern of brain activity in response to native and non-native phonemic information. This finding was in accordance with previous evidence for universal speech perception before 6-8 months of age (Bertoncini, Bijeljac-Babic, Blumstein, & Mehler, 1987; Eimas, 1974, 1975; Eimas, Siqueland, Jusczyk, & Vigorito, 1971; Streeter, 1976; Trehub, 1976; Werker, Gilbert, Humphrey, & Tees, 1981; Werker & Tees, 1984, 1999). In many of these studies, however, groups of older infants between 9 and 12 months of age were also tested on the same contrasts but in most cases were not able to discriminate the non-native contrasts. These results have indicated that with age, a loss of perceptual sensitivity to non-native speech information occurs as a result of native-language experience and commitment. While this loss of non-native phonemic perception has been extensively studied in monolingual infants, similar processes in bilingual phonemic perception are still under-explored.

### ***NATIVE PHONEMIC PERCEPTION IN BILINGUAL INFANTS***

Behavioral evidence has shown that bilinguals successfully form native phonemic representations of both their languages by the end of the first year of life. Using a habituation paradigm, Burns, Yoshida, Hill, and Werker (2007) studied native phonemic boundaries of English monolinguals and French-English bilinguals aged 6-8, 10-12, and 14-20 months. They habituated the infants to an ambiguous, medial stimulus [pa], which would be perceived as /p/ by adult French listeners and /b/ by adult English listeners and then tested them on [ba] and [p<sup>h</sup>a], perceived as /b/ and /p/ by both

groups, respectively. From 10-12 months of age, English monolingual infants only dishabituated to the test trial corresponding to the English boundary (i.e. [p<sup>h</sup>a]), whereas bilinguals demonstrated significant recovery in looking times to both test trials. The results suggested that bilingual infants at 10-12 months and 14-20 months of age were sensitive to the phonemic boundaries of each of their two languages, indicated by a recovery in looking time. Although behavioral measures have suggested that bilinguals are equally sensitive to each of their native phonemic categories, neurophysiological evidence might be able to uncover differential processing of those representations.

For example, 11-month-old bilingual infants showed distinct differences from monolinguals in the perception of native phonemic contrasts. Ferjan Ramírez, Ramírez, Clarke, Taulu, and Kuhl (2016) employed whole-head magnetoencephalography (MEG) in a double oddball paradigm that used a medial token [ta] against the Spanish /da/ and English /t<sup>h</sup>a/ to assess native phonemic perception in the form of MMRs in English monolingual and Spanish-English bilingual 11-month-olds. MMR, or mismatch response, consists of an early component typically occurring between 100 and ~260 ms post stimulus, and a late component with a typical latency of ~260-460 ms. Their findings revealed that both groups were equally sensitive to English contrasts, but that bilinguals showed a stronger response to Spanish contrasts compared to monolinguals. However, when the infants' responses were split into early and late MMR time windows, bilinguals showed a significantly larger English MMR in the early time window, whereas monolinguals showed a significantly larger English MMR in the late time window. As early MMR components signify a less mature and universal encoding of information and late MMR for specialized phonetic analyses, the authors suggested that bilinguals neurally discriminated the contrasts within each of their two native languages only at the acoustic level, whereas monolinguals discriminated the



English contrast at the phonemic level. These results further support that bilingual infants' sensitivity to the phonemic boundaries of both their languages suggests a slower transition from the acoustic (universal) to phonemic (language-specific) analysis of native speech (Kuhl et al., 2008). This can result from dealing with a complex workload, in which the increased amount of phonological information that bilingual infants need to learn facilitates a higher-functioning and flexible perceptual system.

According to the NLNC hypothesis that was introduced in Chapter 1, it is plausible that bilinguals undergo a protracted development in acquiring native phonology. Although it appears that bilingual infants are knowledgeable with the phonemic boundaries in both of their native languages by 12 months of age (Burns, Yoshida, Hill, & Werker, 2007; Ferjan Ramírez, Ramírez, Clarke, Taulu, & Kuhl, 2016; Sundara, Polka, & Molnar, 2008), studies that examined neural activation patterns have suggested that the development of native phonological systems in bilingual infants differ from that of monolingual infants, which was characterized by a slower transition from acoustic to phonemic analysis (Ferjan Ramírez, Ramírez, Clarke, Taulu, & Kuhl, 2016; Garcia-Sierra et al., 2011). Testing bilinguals' perception of a third contrast belonging outside of their native phonemic repertoire would help elucidate whether bilinguals retain perceptual plasticity in universal phonetic discrimination (i.e. a slower commitment to native phonology) for a longer period of time than monolinguals. This protocol, however, has only been done in a few studies.

### ***NON-NATIVE PHONEMIC PERCEPTION IN BILINGUAL INFANTS***

In a visual infant-controlled habituation paradigm, 11-month-old English monolingual and English-Mandarin bilingual infants were assessed on the

discrimination of a non-native Hindi dental-retroflex contrast and a native English consonant contrast, as well as own-race and other-race faces (Singh, Loh, & Xiao, 2017). It was found that bilinguals showed perceptual plasticity in the language domain by discriminating the non-native contrast, whereas monolinguals did not. The results, however, did not extend across to the perceptual domain of face recognition.

To date, there has been only one study that used neuroimaging to examine brain activation patterns in the bilateral STG and IFC during non-native phonemic perception in English monolingual infants and a heterogeneous sample of bilingual infants who had been receiving simultaneous exposure to French, Spanish, or Chinese in addition to English. Petitto and colleagues (2012) used fNIRS to test 4-6- and 10-12-month-olds on a native English /ba - pa/ contrast and a non-native Hindi dental-retroflex /ɖa - ʈa/ contrast in an event-related oddball paradigm. Upon initial whole-brain analyses of the activation patterns, their findings showed that there was no significant difference between the left and right hemispheres across all experimental conditions. When analyzed by region of interest (STG, IFC), their findings revealed a greater right-hemispheric activation in the STG by all infants (i.e. 4-6- and 10-12-month-old monolinguals and bilinguals). In the left STG, infants of all ages and language groups showed similar brain activation patterns. The authors suggested that activation in the left STG to all auditory stimuli is observed early and remains stable across the first year of life. In the inferior frontal regions, the right IFC showed a decrease in activation from the younger to older age groups. The authors interpreted the finding as a developmental shift in lateral dominance, characteristic of perceptual narrowing, where activity in the language network in the left hemisphere increases with linguistic experience (Sato, Sogabe, & Mazuka, 2010). Regarding the left IFC in isolation, all babies showed differences in activation between the native and non-native contrasts.

English monolingual infants showed a much greater pattern of activation to their native language, whereas bilingual babies had similar activity levels for both native and non-native contrasts.

These findings allowed Petitto et al. (2012) to find support for the Perceptual Wedge Hypothesis, in which exposure to more than one language “wedges” open the closing doors of native language commitment and perceptual attunement. As a result, the Perceptual Wedge allows language sensitivity to remain open for longer due to the increased neural and computational demands of bilingual language processing (Petitto et al., 2012). Simply put, Petitto and colleagues have shown through fNIRS brain imaging that bilinguals’ perceptual systems remain open at the time where monolinguals’ have already attenuated to the native language.

#### The present study

In line with the third objective of the present thesis, the goal of the current study was to assess the neural flexibility of non-native phonemic perception between monolingual and bilingual infants around 10-12 months of age when the onset of perceptual narrowing is taking place. Previous research assessing bilingual phonemic perception have only tested bilingual infants on phonemic contrasts from both of their native languages (e.g. Burns, Yoshida, Hill, & Werker, 2007; Ferjan Ramírez, Ramírez, Clarke, Taulu, & Kuhl, 2016; Garcia-Sierra et al., 2011; Sundara, Polka, & Molnar, 2008) but seldom on a third contrast belonging outside of the bilinguals’ phonemic repertoire. A direct way to test the effect bilingualism has on perceptual flexibility is to present a non-native phonemic contrast to monolingual and bilingual infants at the age when monolinguals’ phonological perception abilities start to attune to the native

language (i.e. 10-12 months). If bilinguals remain sensitive to non-native contrasts while monolinguals do not, it can be inferred that the onset perceptual narrowing is protracted in bilinguals, thus providing supporting evidence for the Perceptual Wedge Hypothesis. The current work was unique in its exploration to replicate and extend the findings of young monolingual and bilingual infants from Petitto et al. (2012) through assessing bilinguals' discrimination of a non-native Hindi dental-retroflex contrast on a homogeneous sample of Mandarin-English bilingual infants and aimed to see which cortical areas were activated following the presentation of the contrast. The current work also measured additional cortical regions in the inferior parietal cortex and implemented a more traditional block stimulation paradigm rather than an event-related oddball paradigm.

Guided by previous research, the present study hypothesized that (1) monolingual and bilingual infants would show similar patterns of activation to the native English phonemic contrast, particularly in the left IFC as it is known for phonological processing, (2) monolinguals and bilinguals would show differential processing of the Mandarin lexical tone contrast, in which tonal listeners (bilinguals) would exhibit left-hemispheric activity while non-tonal listeners (monolinguals) would show right-hemispheric activity, and (3) a clear distinction of left inferior frontal activity would be observed only in bilingual infants in response to the non-native Hindi contrast, providing further evidence for the Perceptual Wedge Hypothesis.

## Method

### *Participants*

A total of 38 infants were included in the final analysis. Twenty-one were English monolinguals (mean age = 11.34 months, SD = 17.66 days) and 17 were

Mandarin-English bilinguals (mean age = 11.82 months, SD = 23.66 days). An additional 35 infants were tested but excluded due to fussiness and lack of cooperation (n = 16), extensive motion artifacts (n = 11), failing to meet the criteria for each condition to contain valid data in at least three blocks (n = 6), and unsuccessful channel calibration (n = 2). All infants were healthy and born full-term, with no auditory or cognitive disabilities. The monolingual infants were tested in the UK, and their parents reported that their child was not being regularly exposed to a second language and has had no previous experience listening to Mandarin or Hindi. The bilingual infants were tested in Singapore. Prior to the experiment, their parents completed the Language and Social Background Questionnaire (LSBQ) that inquired about some general demographic information and their child's language background and use as well as their own (Anderson, Mak, Chahi, & Bialystok, 2018). The criteria for bilingualism were for infants to be receiving at least 20% exposure from each language. The participants received, on average, 47.5% Mandarin and 52.5% English exposure. They also had no previous exposure to Hindi.

### *Stimuli*

Three minimal pair CVC contrasts in English, Mandarin, and Hindi were used. Native female speakers from each language naturally produced five tokens for each word in adult-directed speech. The stimuli consisted of an English consonant contrast /dæb - tæb/, a Mandarin lexical tone contrast /taw<sup>2</sup> - taw<sup>3</sup>/, and a Hindi dental-retroflex contrast /ɖa:l - ɖa:l/. All tokens were unfiltered and equalized to an intensity of 75 dB and adjusted to a duration of 650 ms using Praat (Boersma & Weenink, 2012). Detailed information about the stimuli can be referred to in Chapter 3.

*fNIRS recording*

Monolingual infants were tested using a multichannel NIRx NIRScout system emitting two continuous wavelengths at 760 and 850 nm and sampled at a rate of 7.81 Hz. Bilingual infants were tested using a NIRx NIRScout-extended system emitting continuous wavelengths at 785 and 830 nm with a slower sampling rate of 3.91 Hz. All other experimental factors remained as identical as possible, including the headband and testing procedure. Eight infrared emitters and eight detectors were arranged in a flexible headband in two staggered 2 x 4 arrays positioned bilaterally on each side of the head, with 9 channels per hemisphere (Figure 5.1). The separation between sources and detectors was 20 mm, and placement of the headband was guided by the international 10-20 system for EEG.

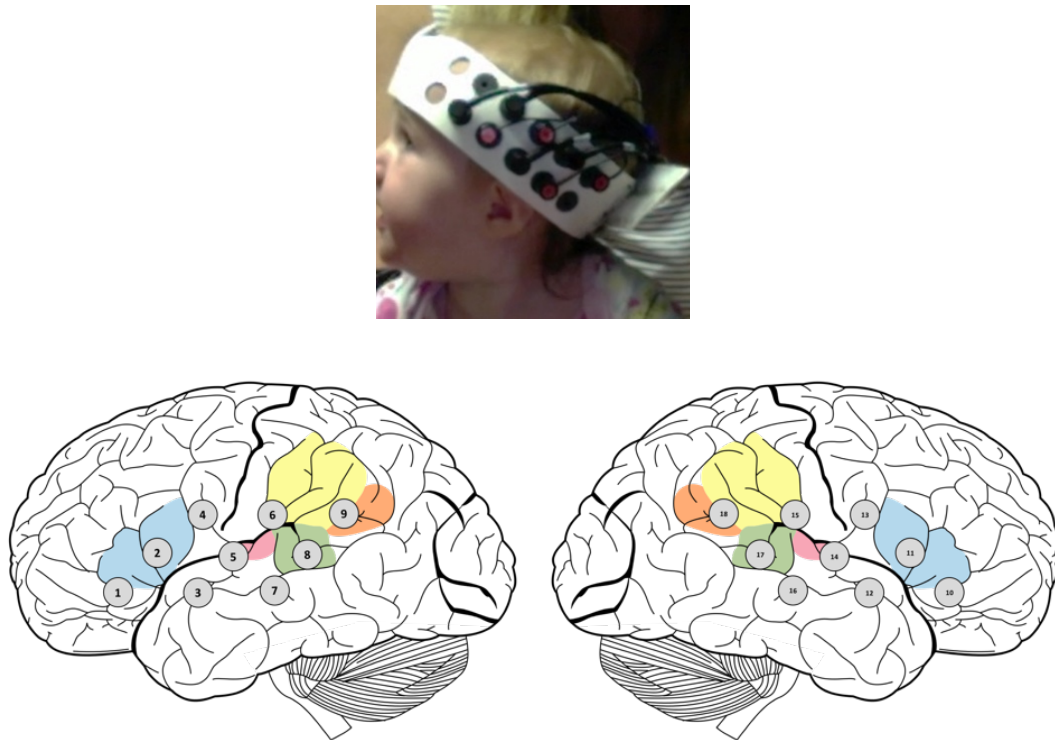


Figure 5.1. (Top) Image of an infant wearing the fNIRS headband showing the optode configuration during a testing session. The red optodes signify light source emitters and the black optodes represent infrared light detectors. (Bottom) Anatomical diagram of channel placement over the approximate underlying cortical structures.

### *Procedure*

The experiment was conducted in a noise-isolated laboratory room. There, the infant sat on their caregiver's lap facing 1 meter away from a computer monitor. The experimenter carefully aligned the fNIRS headband on the infant. Sound stimuli were then administered through concealed front-facing loudspeakers at approximately 70 dB. The timing and presentation of the stimuli was controlled via MATLAB, and NIRStar 15.0 was used for data acquisition.

The three language conditions (English, Mandarin, Hindi) were presented in a fixed order that was counterbalanced across the participants (see Figure 5.2). Each

language condition had one non-alternating and one alternating condition. Twenty-five second silence periods appeared after each auditory block to allow for the hemodynamic response to return to baseline before the start of the next auditory block. A silent, black-and-white video of moving shapes played on the monitor as a visual filler while the experimenter waved silent toys to keep infants entertained throughout the study. The video did not synchronize with the auditory stimuli and was not associated with the study. The duration of the experiment was approximately 20 minutes. Parents and caregivers were debriefed at the end of the study. A more thorough description of the experimental procedure can be found in Chapter 4.

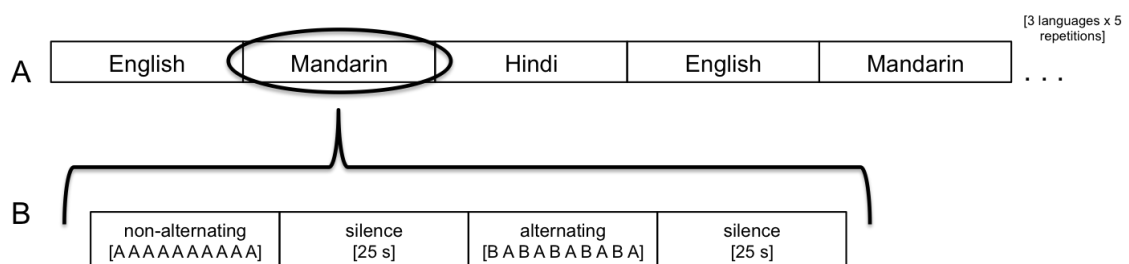


Figure 5.2. Illustration of the procedure. (A) The three language conditions were repeated in a fixed order for a total 15 repetitions. (B) Each language condition contained a non-alternating and an alternating auditory block each followed by a silence period. The order of presentation for the auditory blocks was counterbalanced across participants.

### *fNIRS pre-processing*

Optical data collected from two wavelengths were transformed into HbO, HbR, and HbT. HomER2 was used to carry out motion artifact correction and filtering procedures (Huppert, Diamond, Franceschini & Boas, 2009), and a wavelet-based



motion artifact removal was applied to reconstruct the signal (Brigadoi et al., 2014; Molavi & Dumont, 2012). The data were bandpass-filtered between 0.01 and 0.7 Hz. Baseline correction was taken from a five-second interval preceding the onset of auditory stimulation, and the HbO responses for each condition were averaged across all blocks. The grand averaged time course of each condition (auditory stimulation block + silence block) was plotted for further inspection. Figure 5.3 depicts the HbO responses of Channel 14 to each condition for monolinguals and bilinguals, separately. Channel 14 was chosen as a visual example as it showed clear, canonical responses across all auditory conditions.

In order to be included in the final analysis, it was required that each channel had to contain valid data in all three language conditions, as well as a minimum of three valid trials per condition. If any auditory block was invalid, the entire language condition was excluded from the final analysis.

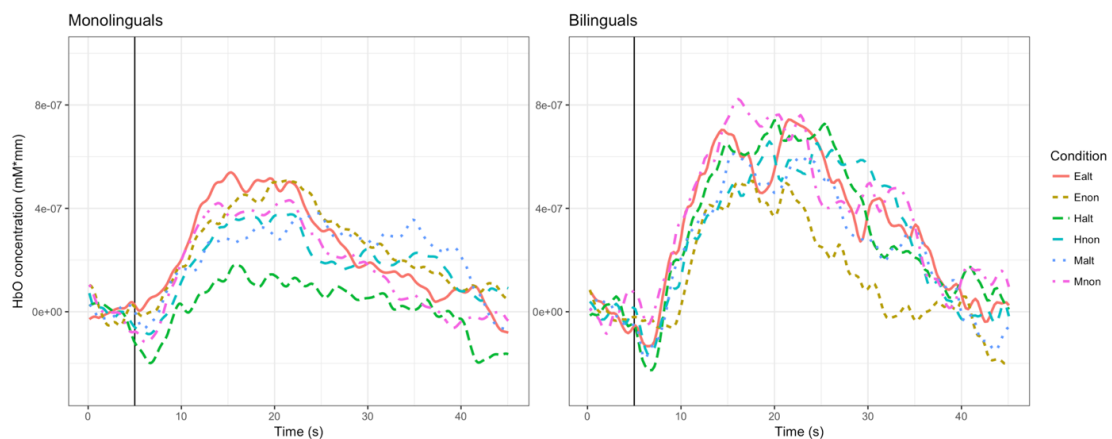


Figure 5.3. Monolingual (left) and Bilingual (right) HbO time courses for each auditory condition (E = English; H = Hindi; M = Mandarin; alt = Alternating; non = non-Alternating) from Channel 14 in the right STG. The vertical line indicates the onset of auditory stimulation.

### *Defining cortical regions of interest (ROIs)*

Three main ROIs were selected based on the co-registration of the underlying cortical structures using the international 10-20 referencing system, where all channels within each ROI were averaged. The channels laid bilaterally over the IFC, STG, and IPC (Figure 5.4).

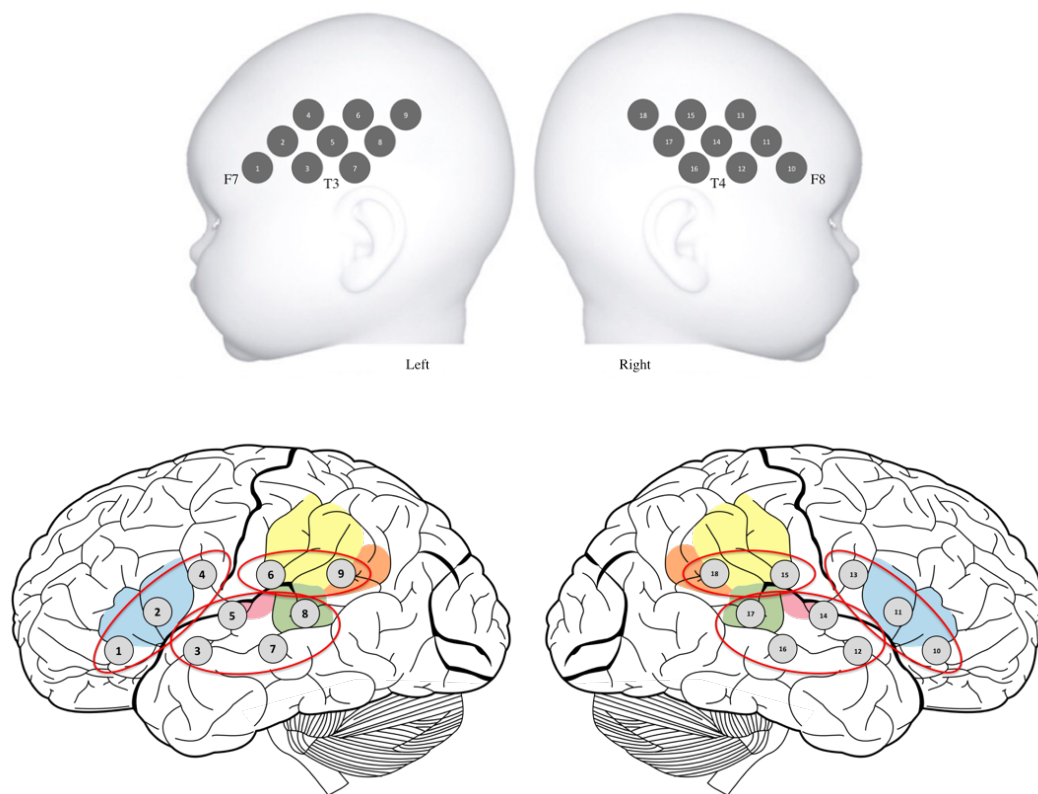


Figure 5.4. (Top) Channel placement on the left and right hemispheres. (Bottom) Approximate projections onto the cortical regions of interest (circled in red).

### *Statistical analyses*

A 20-second time window, between 5 and 25 seconds after stimulus onset that incorporated the peak values for each condition, was applied to all channels. Then, the data points in each window were averaged. Prior to statistical analysis, boxplots were

generated from the averaged values to observe the overall distributions of peak activity and for the identification of outliers in each channel and condition. Four participants were removed from the final analysis because their values were outliers for over 10% of their data (condition x channel;  $6 \times 18 = 108$  data points), or for more than eleven instances. The outlying values of participants who were not excluded from the final analysis were not removed.

## Results

Cortical hemodynamic responses to English, Mandarin, and Hindi alternating and non-alternating conditions were measured across three distinct brain regions that make up the neural network of speech perception. These responses were compared between 10-12-month-old monolingual and bilingual infants to assess whether the window of universal phonetic perception remains open for a longer period time in bilinguals. Since English was native to both groups, we expected no differences in cortical activation to the English contrast between monolingual and bilingual infants. However, group differences were anticipated for Mandarin, as the change in pitch may be perceived lexically in the left hemisphere by tone learners (bilinguals) or acoustically in the right hemisphere by non-tone learners (monolinguals). It was hypothesized that tonal and non-tonal language learners would perceive the Mandarin lexical tone/pitch distinction, but exhibit lateralization effects in different hemispheres. Lastly, to test the Perceptual Wedge Hypothesis described in Petitto et al. (2012) in which bilingualism prolongs the ability to remain sensitive to phonological information belonging outside of the infants' phonemic repertoire, we expected only bilinguals to remain sensitive to Hindi by showing discrimination of the contrast.

### ***Confirmatory analyses***

In an overall analysis, a mixed 3 x 3 x 2 x 2 x 2 repeated-measures ANOVA was conducted on HbO with Language (English, Mandarin, Hindi), ROI (IFC, STG, IPC), Condition (non-Alt, Alt), and Hemisphere (left, right) as within-subjects factors, and Group (Monolingual, Bilingual) as the between-subjects factor. The results revealed a main effect of Language that was nearly significant, in which Hindi exerted the highest activation from all infants,  $F(2, 72) = 2.92$ ,  $MS_e = 5926$ ,  $p = .06$ ,  $\eta_p^2 = .08$ .

### ***Exploratory analyses***

In the following, each language was analyzed separately in a series of mixed repeated-measures ANOVAs with HbO values as the dependent variable, ROI, Condition, and Hemisphere as the within-subjects factors, and Group as the between-subjects factor.

#### ***English***

A three-way interaction was found between ROI, Condition, and Hemisphere,  $F(2, 72) = 5.22$ ,  $MS_e = 2621$ ,  $p = .008$ ,  $\eta_p^2 = .13$ . Analyses were conducted separately for each ROI and revealed an interaction between Condition and Hemisphere in the IFC,  $F(1, 37) = 9.80$ ,  $MS_e = 6382$ ,  $p = .003$ ,  $\eta_p^2 = .21$ . From this, it was found that during the alternating condition, brain activation was lateralized to the left hemisphere in both language groups,  $F(1, 37) = 5.48$ ,  $MS_e = 2835$ ,  $p = .025$ ,  $\eta_p^2 = .13$ . Moreover, the left IFC was found to have a greater activation to the alternating condition,  $F(1, 37) = 5.98$ ,  $MS_e = 8242$ ,  $p = .019$ ,  $\eta_p^2 = .14$ . The left IFC represents Broca's Area, which has been found to activate during native phonemic processing. Figure 5.6 reports the three-way interaction.

Another significant two-way interaction was found between Hemisphere and Group,  $F(1, 36) = 4.29$ ,  $MS_e = 4655$ ,  $p = .046$ ,  $\eta_p^2 = .11$ , showing that bilinguals had an overall higher activation in the right hemisphere,  $F(1, 16) = 9.17$ ,  $MS_e = 2614$ ,  $p = .008$ ,  $\eta_p^2 = .36$ , whereas monolinguals had higher left-hemispheric activation than bilinguals,  $F(1, 36) = 4.13$ ,  $MS_e = 1340$ ,  $p = .049$ ,  $\eta_p^2 = .10$ . No further interactions were found in the STG or IPC. Despite hemispheric differences between groups, monolinguals and bilinguals, overall, performed similarly in response to the native English contrast (alternating condition), evidenced by an increase in activation in the left IFC.

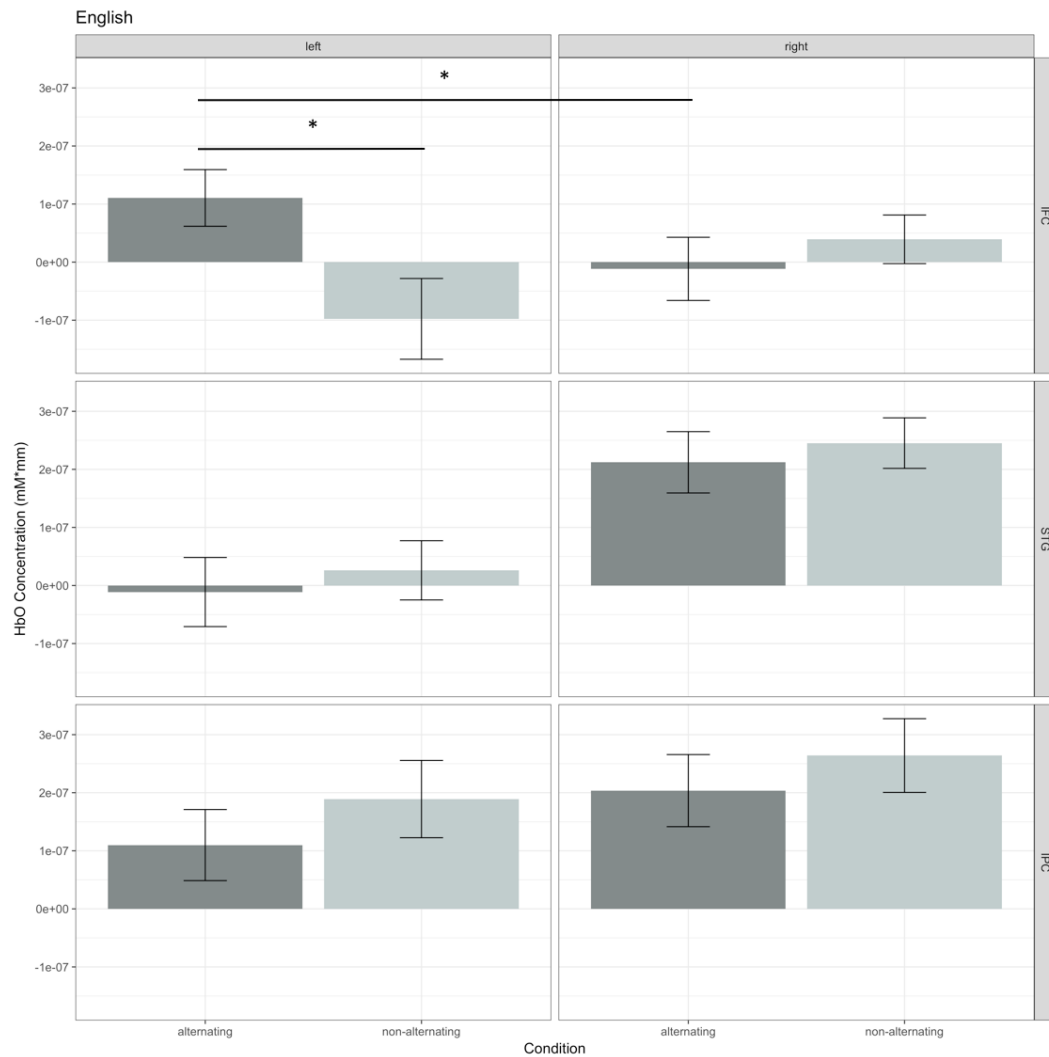


Figure 5.5. Combined monolingual and bilingual mean HbO values for the English non-alternating and alternating conditions, split across hemispheres and ROIs,  $*p < .05$ ,  $**p < .01$ .

### *Mandarin*

There was a significant four-way interaction between ROI, Condition, Hemisphere, and Group,  $F(2, 72) = 3.51$ ,  $MS_e = 1884$ ,  $p = .035$ ,  $\eta_p^2 = .09$ . The data were split by ROI to examine whether any interactions were present between Condition, Hemisphere, and Group. A significant three-way interaction was found for IPC,  $F(1, 36) = 9.56$ ,  $MS_e = 4573$ ,  $p = .004$ ,  $\eta_p^2 = .21$ ; thus, the interaction was broken down

further by Group to examine any differences between Condition and Hemisphere. The interaction between Condition and Hemisphere was present only in bilinguals,  $F(1, 16) = 4.88$ ,  $MS_e = 3088$ ,  $p = .042$ ,  $\eta_p^2 = .23$ , in which bilinguals exhibited a lateralized response to the contrast (alternating condition) in the right IPC,  $F(1, 16) = 14.16$ ,  $MS_e = 14882$ ,  $p = .002$ ,  $\eta_p^2 = .47$ . No other interactions or simple main effects were found.

Figure 5.7 reports the four-way interaction.

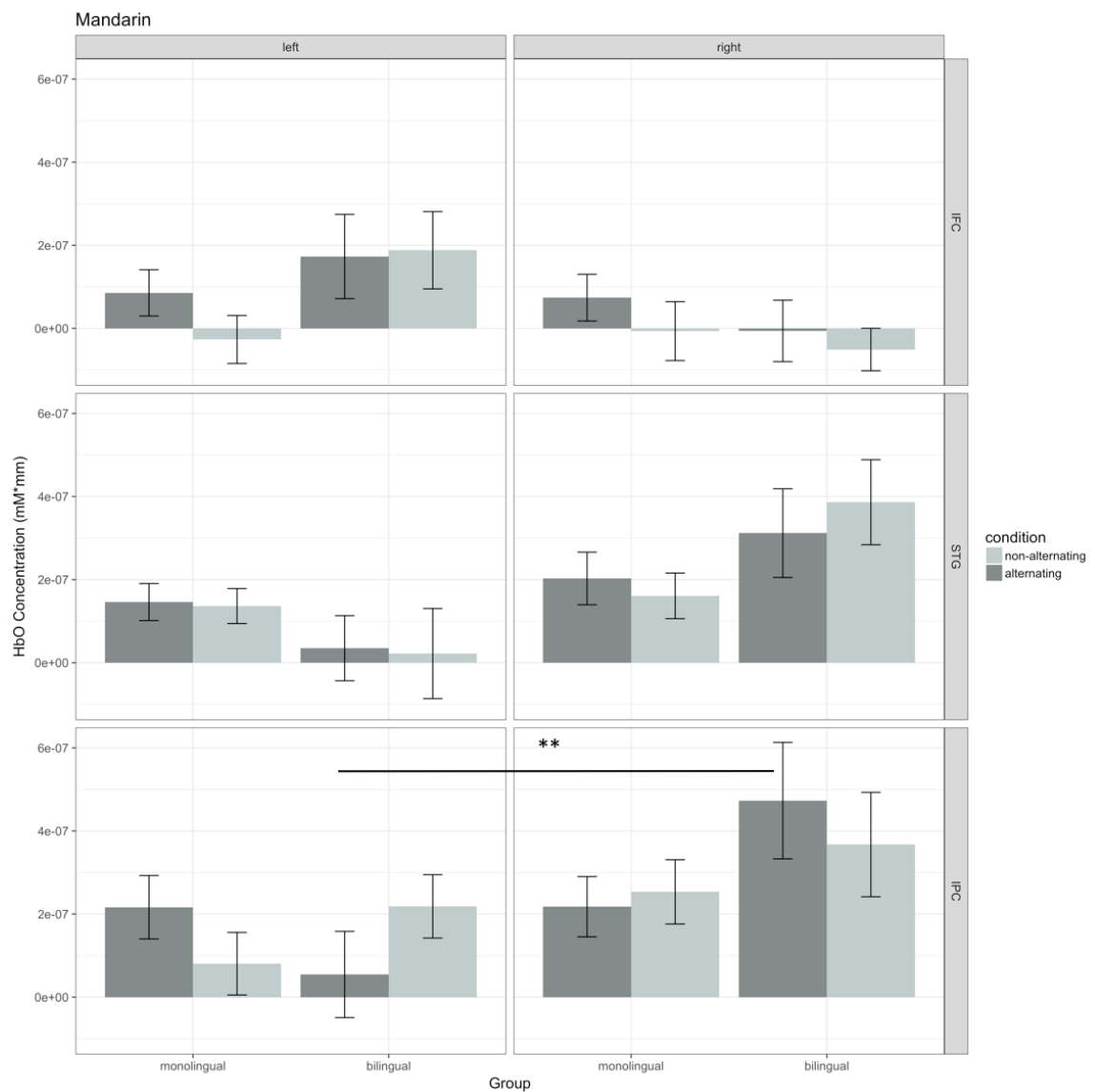


Figure 5.6. Mean HbO values for the Mandarin non-alternating and alternating conditions in monolinguals and bilinguals, split across hemispheres and ROIs,  $*p < .05$ ,  $**p < .01$ .

Although significance was not found for more anterior regions IFC or STG, the means indicate, however, that in the IFC, bilinguals elicited a greater activation in the left hemisphere than monolinguals, who appeared to process the contrast equally in both hemispheres (see Table 5.3). This suggests that bilinguals had a left-lateralized processing of the native lexical tone contrast in the region of the brain responsible for native phonemic processing, that is, the IFC; but the sample size might have not been sufficient for the observation to reach significance. In the STG, bilinguals showed greater right-hemispheric activation to the contrast than monolinguals. Taken together, monolinguals and bilinguals were shown to process Mandarin stimuli mainly in the right hemisphere, with activation centralized in the right IPC. However, significance was found in the rightward lateralization of the IPC in bilinguals during the alternating condition. Results from the IFC, although not reaching significance, showed that bilinguals had a greater left-hemispheric processing difference than monolinguals.

Table 5.1. Mean HbO values for the Mandarin ROIs in the left and right hemispheres, by condition and group.

		IFC				STG				IPC			
		LH	(se)	RH	(se)	LH	(se)	RH	(se)	LH	(se)	RH	(se)
Mono-lingual	non-Alt	-2.62	(7.04)	-.67	(7.36)	13.64	(7.22)	16.10	(7.37)	8.05	(7.53)	25.34	(7.74)
	Alt	8.54	(6.11)	7.40	(6.10)	14.61	(5.76)	20.29	(7.96)	21.64	(7.63)	21.76	(7.25)
Bilingual	non-Alt	18.80	(7.83)	-5.09	(8.18)	2.21	(8.02)	38.63	(8.19)	21.84	(7.63)	36.73	(12.56)
	Alt	17.29	(6.79)	-.62	(6.78)	3.51	(6.40)	31.20	(8.84)	5.45	(10.38)	47.29	(14.01)



### *Hindi*

There was a four-way interaction was found between ROI, Condition, Hemisphere, and Group,  $F(2, 72) = 3.96$ ,  $MS_e = 1835$ ,  $p = .023$ ,  $\eta_p^2 = .10$ . The interaction is reported in Figure 5.8. Subsequent analyses were administered by ROI to investigate whether there was an interaction between Condition, Hemisphere, and Group. The three-way interaction was found only in the IFC,  $F(1, 36) = 7.61$ ,  $MS_e = 5076$ ,  $p = .009$ ,  $\eta_p^2 = .18$ , where the data were further split by Group to look for an interaction between Condition and Hemisphere. The interaction was found in bilinguals, in which simple main effects analyses found that bilinguals showed a higher left-lateralized response to the non-alternating condition in the IFC,  $F(1, 16) = 8.31$ ,  $MS_e = 11249$ ,  $p = .011$ ,  $\eta_p^2 = .34$ .

Through observing means, however, bilinguals exhibited greater HbO activation to the alternating condition in contrast to the non-alternating condition in the right IPC (mean difference monolinguals: 2.48; mean difference bilinguals: 22.44) (see Table 5.5). In summary, bilinguals indeed showed an unusual response to the non-native stimuli in the left IFC, albeit with higher activation to the non-alternating stimulus. For the alternating condition, although not significant, the means indicate that bilinguals had a greater difference in activation to the contrast in the right IPC compared to monolinguals.

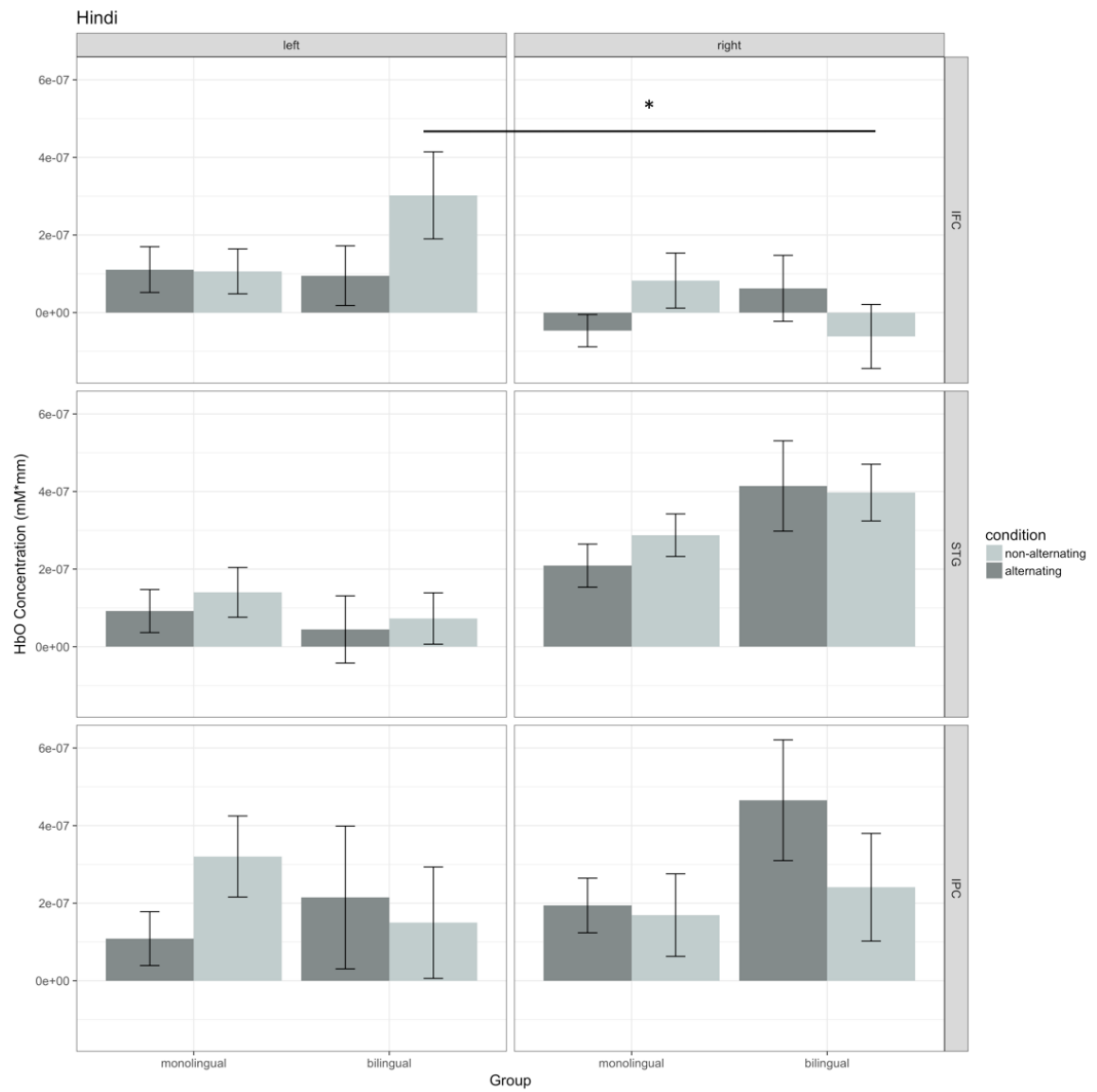


Figure 5.7. Mean HbO values for the Hindi non-alternating and alternating conditions in monolinguals and bilinguals, split across hemispheres and ROIs,  $*p < .05$ ,  $**p < .01$ .

Table 5.2. Mean HbO values for the Hindi ROIs in the left and right hemispheres, by condition and group.

		IFC				STG				IPC			
		LH	(se)	RH	(se)	LH	(se)	RH	(se)	LH	(se)	RH	(se)
Mono-lingual	non-Alt	10.62	(5.78)	8.25	(7.09)	14.00	(6.41)	28.74	(5.48)	32.04	(10.45)	16.94	(10.64)
	Alt	11.09	(5.89)	-4.66	(4.15)	9.19	(5.54)	20.89	(5.56)	10.87	(6.94)	19.42	(7.04)
Bilingual	non-Alt	30.22	(11.21)	-6.16	(8.25)	7.27	(6.62)	39.73	(7.30)	14.98	(14.36)	24.11	(13.88)
	Alt	9.52	(7.69)	6.26	(8.51)	4.45	(8.66)	41.43	(11.64)	21.48	(18.40)	46.55	(15.56)

## Discussion

The goal of the present study was to test the Perceptual Wedge Hypothesis by examining the brain regions associated with phonological processing in 10-12-month-old English monolingual and Mandarin-English bilingual infants. English, Mandarin, and Hindi phonemic contrasts were presented to the participants while their hemodynamic activity was recorded using fNIRS. Motivated by previous research (e.g. Petitto et al., 2012), this study addressed three specific hypotheses. First, both language groups were expected to show similar activation patterns to the native English contrast. Under exploratory analyses, the current results confirmed our hypothesis by finding that monolinguals and bilinguals exhibited a higher response in HbO to the English alternating condition, which was localized in the cortical region responsible for native-phonemic processing (i.e. left IFC). Second, the Mandarin lexical tone contrast was unique in a way that it could be processed lexically in the left hemisphere or

acoustically/prosodically in the right hemisphere depending on the listener's language experience (Klein, Zatorre, Milner, & Zhao, 2001; Wang, Behne, Jongman, & Sereno, 2004; Wang, Jongman, & Sereno, 2001). The current findings suggest that bilinguals successfully detected the change in pitch during the Mandarin alternating condition, which was indicated by a significant increase in HbO in the right IPC. Additionally, although not significant, the means indicated that bilinguals showed a left-lateralized response to the alternating condition in the IFC while monolinguals showed equal and bilateral responses in the same region. In fact, monolinguals, overall, showed fairly equal and bilateral HbO activation to the Mandarin alternating condition across all ROIs (see Figure 5.7). Finally, we expected monolinguals and bilinguals to process the non-native Hindi contrast differently, and specifically, only bilinguals to discriminate this contrast. Our findings indicated that, indeed, both groups showed differential responses to the non-native contrast. Interestingly, bilinguals' responses were significantly left-lateralized to the non-alternating condition in the IFC. The following discusses the exploratory findings for each language in more detail, with the aim of shedding more light on the neural underpinnings of infant phonemic perception.

***ENGLISH: Monolingual and bilingual infants exhibited native language processing patterns***

The current study found no differences in performance between monolinguals and bilinguals for the discrimination of the native English phonemic contrast. Both groups showed significantly higher activation in the left IFC to the alternating condition than to the non-alternating condition. These findings were in accordance with previous research that has found evidence for increasing left inferior frontal activation with age to native English phonemic contrasts between 6- and 12-month-old infants using MEG (Imada et al., 2006). This effect has also been observed in 7-12-year-olds as well as

adults using fMRI (Conant, Liebenthal, Desai, & Binder, 2014; Burton, 2001; Zatorre, Evans, Meyer, & Gjedde, 1992; Zatorre, Meyer, Gjedde, & Evans, 1996). Our findings are also supported by a previous fNIRS study (i.e. Pettito et al., 2012) that found a leftward shift of activation with age from the right to the left IFC in 10-12-month-old infants. These results suggest that the hemispheric shift in activation indicates a developmental change in lateral dominance at the precise age where infants begin to produce first words at 12 months, and they confirm that activation seen in the left IFC reflects processes associated specifically with native phonemic perception.

Previous studies have shown that left superior temporal regions were also involved with speech and phonemic perception in 3-month-old infants (Dehaene-Lambertz, Dehaene, & Hertz-Pannier, 2002). The absence of differential activation to the native English contrast in the left STG in the current study might be a consequence from the lack of spatial accuracy fNIRS yields in comparison to fMRI. For example, one fMRI study of adults showed differential processing in adjacent areas of the temporal lobe, where the middle and anterior regions of the left superior temporal sulcus (STS), underneath the STG, were activated during phonemic perception, whereas the dorsal STG containing the primary auditory cortex was bilaterally activated by phonemic and non-phonemic stimuli (Liebenthal, Binder, Spitzer, Possing, & Medler, 2005). The placement of the fNIRS channels in the present study did not allow us to clearly separate the STS from the STG, thus it may have contributed to the absence of activation to the native contrast in the alternating condition. However, the current results indicated a main effect of Hemisphere, in which all participants exhibited prominent right-lateralized, and not bilateral, responses to the native contrast in the STG.

Our results were not in line with previous findings on the role of the left STG in phonemic processing (Imada et al., 2006). A possibility for this difference in lateralization could be caused by hemispheric specialization for auditory linguistic information. Generally, the left hemisphere is known to process the fast, temporal characteristics of the speech signal, whereas the right hemisphere is specialized on the slower, more spectral properties of speech processing such as intonation and prosody (Arimitsu et al., 2011). The stimuli presented were indeed phonemic in nature, but the interval at which the tokens were spaced apart (1000-1500 ms) did not resemble the fast-paced nature of spoken language. Telkemeyer et al. (2009) conducted simultaneous EEG and fNIRS recordings on varying temporal structures of acoustic signals and found that newborns processed slower modulations in the right temporal areas, whereas faster signals, resembling phonetic information, elicited bilateral activation. Thus, in the present study, the responses seen in the right STG could be explained by the slower speed in which the stimuli were presented. Taken together, phonemic processing was observed in the left IFC, but the differential lateralization of the contrast in the STG could be related to the temporal factors of the stimuli.

***MANDARIN: Contextual cues are important for bilinguals to derive lexical representations from auditory information***

The current study assessed discrimination of a Mandarin lexical tone contrast on tonal (bilingual) and non-tonal (monolingual) language learners. The tones used were a relatively subtle contrast made up of rising tone T2 and low-level tone T3. It was hypothesized that tonal and non-tonal language learners would be able to discriminate the pitch contrast by showing differential activation to the alternating condition but exhibit differences in lateralization. In accordance with previous evidence on the hemispheric specialization of speech processing (Zatorre & Belin, 2001), it was

hypothesized that bilinguals would discriminate the phonemic contrast in the left hemisphere and monolinguals would detect the change in pitch in the right hemisphere. Our results showed that monolingual infants did not show any lexical processing of the Mandarin in the left hemisphere. Interestingly, bilingual infants also did not show left-lateralized processing to the same contrast.

A similar null effect in bilinguals has recently been demonstrated in a behavioral modified stimulus alternating paradigm (Singh et al., 2018) that revealed a unique developmental profile for lexical tone discrimination in Mandarin-English bilinguals. In Singh et al.'s study, infants at 6, 9, 12, and 13 months of age did not show evidence of discriminating either subtle (Tones 2 and 3) or salient (Tones 1 and 3) Mandarin lexical tone contrasts. On the other hand, their study indicated that Mandarin monolingual infants demonstrated successful tone discrimination at all ages and that English monolingual infants showed discrimination of tone from 9 months of age. The authors postulated that the pattern of responses seen in bilinguals could be explained by the nature of dual language learning, in which the bilingual brain requires time to organize phonological input from two language systems and thus may result to a later transition period from acoustic to phonemic analysis of speech (Ferjan Ramírez, Ramírez, Clarke, Taulu, & Kuhl, 2016). Moreover, lexical tone discrimination in bilinguals who are simultaneously learning tonal (Mandarin) and non-tonal (English) languages may find it difficult to ascribe meaning to conflicting linguistic and/or pragmatic functions of pitch. For example, T2 is a rising tone, which in English is pragmatically attributed to the inflection of interrogative statements, or questions. Such an interpretation of suprasegmental features of speech would be expected to be processed in the right hemisphere (Zatorre, Evans, Meyer, & Gjedde, 1992). This evidence suggests that referential and contextual information is critical for bilingual

infants to ascribe linguistic meaning of speech. More importantly, it raises the question of whether traditional phonemic discrimination tasks are suitable for measuring cortical regions in response to native and non-native speech processing in bilingual populations, as they are predominately syllable discrimination paradigms that are lacking in context and may lead to ambiguous interpretations. As the testing procedure for the current study was conducted in English, bilinguals may have been biased to process the lexical tone contrast in an English context, leading to an absence of significant left-hemispheric activation to the contrast.

Although left-lateralized activation to the alternating Mandarin condition was not found in bilinguals, the group showed a significant right-lateralized pattern of activation in the inferior parietal cortex. The right IPC has been shown to activate under processes related to working memory during pitch discrimination in adults (Royal et al., 2016). In their study, Royal et al. applied fMRI to compare pitch discrimination in melodic and non-melodic contexts. In the melodic task, participants listened to a musical melody and had to judge whether the last note was in tune, out of key, or out of tune. In the non-melodic task, participants listened to a sequence of acoustic tones (AAAA) or (AAAB) and had to judge whether the fourth tone was higher or lower in pitch than the previous ones. The right IPC was found to be significantly more active in the melodic task, where the region preferentially processed out-of-tune pitch violations. Moreover, increased activity in the region reflects the processing of pitch information in the context of tonal structure. The results could be extended to melodic patterns found in speech prosody. The notion that native tone language learners have heightened sensitivity to pitch variations (Bidelman, Hutka, & Moreno, 2013; Giuliano, Pfordresher, Stanley, Narayana, & Wicha, 2011) support the current findings of responses to the tone contrast in the right IPC.



***HINDI: Monolinguals and bilinguals show differential processing to the non-native contrast in the left IFC***

The main objective of the current study was to use fNIRS brain imaging to elucidate whether bilingualism facilitates perceptual plasticity in the discrimination of a non-native Hindi dental/retroflex contrast. The non-native contrast was tested at the age where monolinguals' language-universal sensitivity to all phonetic contrasts has considerably diminished (Burns, Yoshida, Hill, & Werker, 2007; Kuhl et al., 2006; Rivera-Gaxiola, Silva-Pereyra, & Kuhl, 2005; Singh, Loh, & Xiao, 2017; Werker & Tees, 1984). Given that dental/retroflex contrasts are not phonemic in English or Mandarin; this contrast was non-native to all participants. We hypothesized that only 10-12-month-old bilingual infants would show differential activation to the non-native contrast, in support of the Perceptual Wedge Hypothesis. We expected our results to replicate a previous finding from which activation during the discrimination of the Hindi contrast would be observed in the left IFC (Petitto et al., 2012). Interestingly, our results found that only bilinguals showed a significant increase in HbO in left IFC to the non-alternating condition but not the alternating condition. Moreover, no hemispheric differences were found for the alternating condition in either monolinguals or bilinguals. These results suggest tentative evidence for differential processing of the non-native contrast in bilinguals, in support of the Perceptual Wedge Hypothesis. However, it should be noted that the direction of the discrimination, greater activation in the non-alternating condition, was opposite to our predictions.

Novelty preferences to non-alternating stimulus sequences have been found in previous studies that utilized a familiarization paradigm. For example, 6- and 8-month-old infants from English-speaking families were tested on a unimodal or bimodal frequency distribution along an eight-step continuum of [da] and [ta] (Maye, Werker, & Gerken, 2002). Following a 2.3-minute familiarization phase of varied syllables from

either frequency distribution, infants were tested on Tokens 3 and 6 (medial tokens) for the non-alternating, baseline trials and then Tokens 1 and 8 (endpoint stimuli) for the alternating trials (see Figure 1.1, pg. 5). The authors found that the infants who had been familiarized to the unimodal distribution showed no preference to the alternating trials, whereas infants from the bimodal distribution showed a novelty preference to the non-alternating, medial trials. A later study found the same pattern of results, where 10-month-old English monolingual infants were exposed to flat and bimodal frequency distributions to learn the non-native Hindi dental-retroflex distinction (Yoshida, Pons, Maye, & Werker, 2010). The authors found that 10-month-old infants familiarized on the bimodal distribution were able to discriminate the non-native contrast, but only following a longer familiarization period (i.e. 224 s vs. 112 s). More importantly, the infants fixated longer to the non-alternating than alternating test trials. In contrast, studies that do not implement a familiarization phase (e.g. habituation studies) have found a novelty preference to the alternating stimuli (Weikum, Oberlander, Hensch, & Werker, 2012). The crucial finding in these studies was the infants' ability to detect a change between the two stimulus trials that were being presented, not the direction of the preference: any difference in response to differing stimuli suggests that a contrast between the stimuli was detected (de Groot, 2011; Houston-Price & Nakai, 2004).

Although the present study did not use a familiarization paradigm typical of behavioral studies, increased brain activation to the non-alternating condition might therefore indicate differential processing to stable and repeated presentations of stimuli. For the current study, stimuli from non-alternating conditions were more familiar than the alternating conditions, as standard tokens (e.g. A-A-A-A-) were presented more frequently than target tokens (e.g. B-A-B-A-) (see Figure 5.2). An explanation for the pattern of results could possibly be that bilinguals prefer stable and consistent stimuli

to extract information from their usually mixed (i.e. bilingual) environments. Future research investigating bilingual auditory familiarity preferences would help increase our understanding of how bilinguals take greater advantage of the environmental cues in order to learn and understand new language systems.

In this study, we expected to reproduce the finding from Petitto et al. (2012) that demonstrated how a heterogeneous sample of 10-12-month-old bilinguals showed greater HbO activation in the left IFC in response to the presentation of a non-native Hindi dental/retroflex contrast, while monolinguals exhibited a greater activation only to the native language in the same region. The left IFC plays an important role in phonological processing and verbal working memory (Dehaene-Lambertz et al., 2006; Imada et al., 2006). It is also known as Broca's Area, which integrates information from different domains such as motor articulatory information to aid in speech production (Kuhl, Ramírez, Bosseler, Lin, & Imada, 2014). Petitto et al.'s study implemented an oddball paradigm in which cortical hemodynamic responses were measured using fNIRS. In their paradigm, each run contained 40 events of a mixture of standard, deviant (oddball), and catch (silent) trials. Each run was constrained under the following rules: a deviant event was presented only after a minimum of three standard events, and no more than two catch events were to be presented in succession. Each event was 500 ms in duration, and they were presented 1000 ms apart. Analyses were conducted on the "standard" tokens, which were the stimuli that followed the "target" or oddball token. The minimum length of presentation for the standard events was 4.5 s, which may have been difficult to capture a complete canonical hemodynamic response to the contrast following the oddball. As hemodynamic responses unfold much slower than ERPs, it is difficult to be certain whether the oddballs were driving the changes in hemodynamic activation in this paradigm. It is possible that the

responses measured were from the standard, non-changing events. Under this interpretation, our finding of left IFC activation in bilinguals to the non-alternating Hindi stimuli converges with the finding from Petitto et al. (2012).

As two fNIRS systems were used to acquire monolingual and bilingual infant data from the UK and Singapore, respectively, the difference in the frequencies emitted and sampling rate from each system might have impacted our results. To account for this potential confounding factor, the monolingual data were down-sampled to match the data on bilingual infants.

In summary, the current findings suggest evidence for the differential processing of the non-native contrast in bilinguals, in support of the Perceptual Wedge Hypothesis. Our findings are in accordance with previous behavioral and neurophysiological studies that have shown how learning at least two languages can prolong the onset of perceptual narrowing and facilitate flexibility in non-native phonemic discrimination (Graf Estes & Hay, 2015; Petitto et al., 2012; Singh, Loh, & Xiao, 2017). Indeed, for bilinguals to acquire native phonology is a complex undertaking, in which they must learn to understand and separate two possibly overlapping phonological systems (Singh et al., 2018). As a result, the trajectory of native phonological acquisition in bilingual infants might take a longer time than those who are learning only one language.

## CHAPTER 6

**Changes in the organization of cortical activity between younger and older monolingual infants reveal a possible neural signature of universal language processing in the left inferior parietal cortex**

## Introduction

In Chapters 4 and 5 we examined brain activation to native and non-native phonemic contrasts in monolingual 5-7- and 10-12-month-old infants; however, the chapters investigated each age group separately. The first objective of the present thesis was to examine developmental changes in brain activation during phonological processing. Therefore, we explicitly compare the results of younger and older monolingual infants in the current chapter. As we did not assess a sample of younger, 5-7-month-old bilingual infants, comparisons between younger and older bilinguals were not made.

Petitto et al. (2012) examined the developmental changes between younger and older monolingual infants and showed a decrease in activation in the right IFC from 4-6 months to 10-12 months of age. The apparent leftward migration of brain activity in the IFC was explained as a shift in lateral dominance across development, specifically around the time when infants begin to produce their first words by 12 months of age. Moreover, the left IFC, or Broca's Area, has been shown to be associated with articulatory representations in speech production in adults (Myers, Blumstein, Walsh, & Eliassen, 2009). Regarding the STG, Petitto et al. (2012) found that all monolingual infants (younger and older) exhibited a greater overall activation in the right hemisphere, and that there was a significant difference in activation between the non-native (Hindi) and pure tone conditions. Additionally, by examining the left STG alone, the authors found that younger and older infants showed similar activations, suggesting that the activity in the left STG comes in early and remains stable across the first year of life. To date, Petitto et al. (2012) remains to be the only published developmental fNIRS study that examined cortical activation in native and non-native phonemic processing at two time points within the first year of life.

A theory accounting for the changing brain activation with age, Interactive Specialization (IS), postulates that the changing patterns are a result of cortical regions competing with each other to acquire their final role in processing information (Johnson, 2011). From this perspective, some regions begin with broadly defined functionality and may activate in response to a wide range of stimuli and tasks. With age and experience, these regions become more refined or *specialized* such that their activity becomes localized to a specific area. Additionally, IS assumes that activation from a specific cortical region is partly determined by its own activity as well as its connectivity to other regions and their patterns of activity. Therefore, the onset of new processing competencies during development is associated with changes in activation patterns over a network made up of neighboring regions (e.g. the neural network of speech processing) and not just by the activation of one cortical region.

In support of the IS theory, an ERP study conducted by Mills, Coffey-Corina, and Neville (1993) investigated patterns of neural activity relevant to language processing in two groups of 20-month-old infants with relatively higher and lower productive vocabulary scores. The authors found that the level of competency was positively linked to the specialization of the brain areas related to language processing. The participants passively listened to known, unknown, and backward words while their ERPs were being measured. Infants with lower productive vocabulary showed larger ERPs between 200-400 ms to known than unknown words. Further, these ERPs were broadly distributed over anterior and posterior cortical regions of both hemispheres. In contrast, infants with higher productive vocabulary showed more localized neural activity observed in the temporal and parietal regions of the left hemisphere in the same time window. A subsequent analysis that compared the infants' vocabulary sizes with age held constant revealed that the differences in brain activity

was related more to vocabulary than age. These results were in accordance with IS in that more focal patterns of activation are positively correlated with experience. Motivated by IS (Johnson, 2011) and the supporting findings from Mills, Coffey-Corina, and Neville (1993), we compared cortical responses from younger and older monolingual infants from Chapters 4 and 5 to examine differences in brain activation with age, if any, to native and non-native phonemic contrasts, which in turn would allow us to better capture the changes in the organization of brain activity as a function of language development.

## Method

### *Participants*

A total of twenty-four 5-7-month-old (Younger) infants from Chapter 4 and twenty-one 10-12-month-old (Older) infants from Chapter 5 were included in the analysis. Both age groups were being raised in monolingual English language environments and had not had any linguistic exposure to Mandarin or Hindi.

### *Statistical analyses 1*

Datasets from each age group were compiled into one spreadsheet for analysis. Time windows for analyses were set between 5 and 25 seconds after stimulus onset for each channel within each ROI (IFC, STG, IPC) in each language (English, Mandarin, Hindi) and condition (Alt, non-Alt). All of the values within each time window and ROI were averaged. Each language was analyzed separately in a series of mixed 3 x 2 x 2 x 2 repeated-measures ANOVAs using HbO values as the dependent variable, region of interest (IFC, STG, IPC), condition (non-Alt, Alt), and hemisphere (left, right) as the within-subject factors, and age (younger, older) as the between-subject factor.



## Results

### *English*

A three-way interaction was found between ROI, Hemisphere, and Age,  $F(2, 86) = 4.49$ ,  $MS_e = 3154$ ,  $p = .014$ ,  $\eta_p^2 = .10$  (Figure 6.1). Further, there was a significant interaction between ROI and Age,  $F(2, 86) = 7.62$ ,  $MS_e = 6463$ ,  $p = .001$ ,  $\eta_p^2 = .15$ , in which simple main effect analyses revealed that the younger infants showed the greatest activation in the bilateral STG,  $F(2, 46) = 9.98$ ,  $MS_e = 3935$ ,  $p < .001$ ,  $\eta_p^2 = .30$ , whereas the older infants significantly showed the greatest activation more superior in the bilateral IPC,  $F(2, 40) = 8.02$ ,  $MS_e = 21.01$ ,  $p = .001$ ,  $\eta_p^2 = .29$ .

There was also a significant interaction between ROI and Hemisphere  $F(2, 86) = 5.12$ ,  $MS_e = 3593$ ,  $p = .008$ ,  $\eta_p^2 = .11$ , where simple main effects were found in both directions. First, in the right hemisphere, the STG had the highest level of activation in both age groups,  $F(2, 88) = 12.22$ ,  $MS_e = 5181$ ,  $p < .001$ ,  $\eta_p^2 = .22$ . Pairwise comparisons using a Bonferroni correction found significant differences between all ROIs: STG-IFC (mean difference 21.46,  $p < .001$ ), IPC-IFC (mean difference 10.42,  $p = .038$ , and STG-IPC (mean difference 11.04,  $p = .044$ ). Second, and in the opposite direction, there was a simple main effect of hemisphere in the STG,  $F(1, 44) = 4.84$ ,  $MS_e = 2282$ ,  $p = .033$ ,  $\eta_p^2 = .10$ , with brain activation significantly lateralized to the right hemisphere in both age groups.

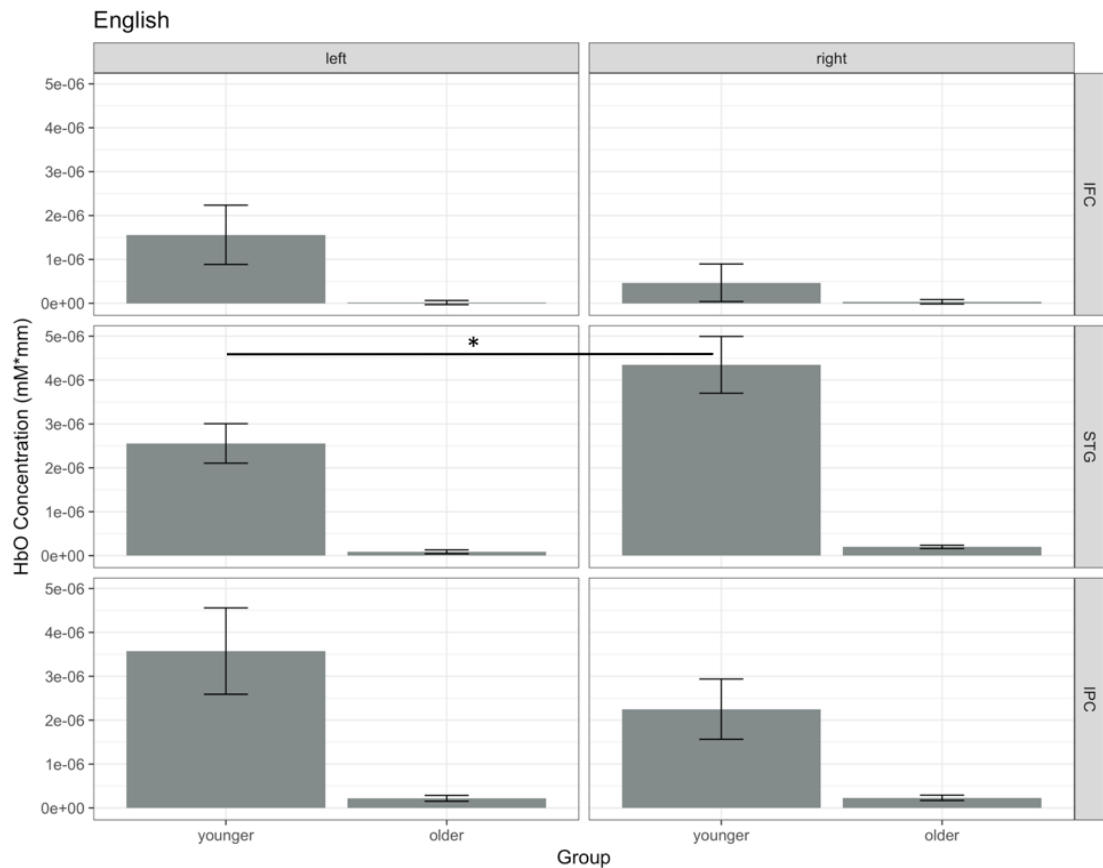


Figure 6.1. Mean HbO values for English in younger and older English monolingual infants, split across hemispheres and ROIs. As there was no effect of condition, the alternating and non-alternating values were averaged together. \* $p < .05$ , \*\* $p < .01$ .

### *Mandarin*

A three-way interaction was found between ROI, Hemisphere, and Age,  $F(2, 86) = 5.25$ ,  $MS_e = 3496$ ,  $p = .007$ ,  $\eta_p^2 = .11$  (Figure 6.2). There was a significant interaction between ROI and Age,  $F(1.55, 66.53) = 12.40$ ,  $MS_e = 10644$ ,  $p < .001$ ,  $\eta_p^2 = .22$ , in which simple main effect analyses revealed that the younger infants significantly showed the greatest activation to the tonal stimuli in the bilateral STG,  $F(2, 46) = 16.03$ ,  $MS_e = 4937$ ,  $p < .001$ ,  $\eta_p^2 = .41$ , whereas the older infants significantly showed the greatest activation more superior in the bilateral IPC  $F(2, 40) = 5.16$ ,  $MS_e = 15.27$ ,  $p = .01$ ,  $\eta_p^2 = .21$ .

There was also a significant interaction between ROI and Hemisphere  $F(2, 86) = 5.12$ ,  $MS_e = 3593$ ,  $p = .008$ ,  $\eta_p^2 = .11$ , where simple main effects were found in both directions. First, there was a simple main effect of ROI in the left hemisphere,  $F(2, 88) = 4.79$ ,  $MS_e = 1932$ ,  $p = .011$ ,  $\eta_p^2 = .10$ , showing that the greatest activation was found in the left IPC in both age groups. Pairwise comparisons (Bonferroni corrected) found a significant difference between the IFC and IPC (mean difference 13.05,  $p = .013$ ). Second, there was another simple main effect of ROI in the right hemisphere,  $F(2, 88) = 15.32$ ,  $MS_e = 5819$ ,  $p < .001$ ,  $\eta_p^2 = .26$ , with STG showing the greatest activity in both age groups. Pairwise comparisons revealed that all ROIs were significantly different from each other: IFC-STG (mean difference 22.74,  $p < .001$ ), IFC-IPC (mean difference 11.15,  $p = .037$ ), and STG-IPC (mean difference 11.59,  $p = .012$ ). Third, and in the opposite direction, there was a simple main effect of hemisphere in the STG,  $F(1, 44) = 8.41$ ,  $MS_e = 6531$ ,  $p = .006$ ,  $\eta_p^2 = .16$ , showing that brain activation was significantly lateralized to the right hemisphere in both age groups. No other significant main effects or interactions were found.

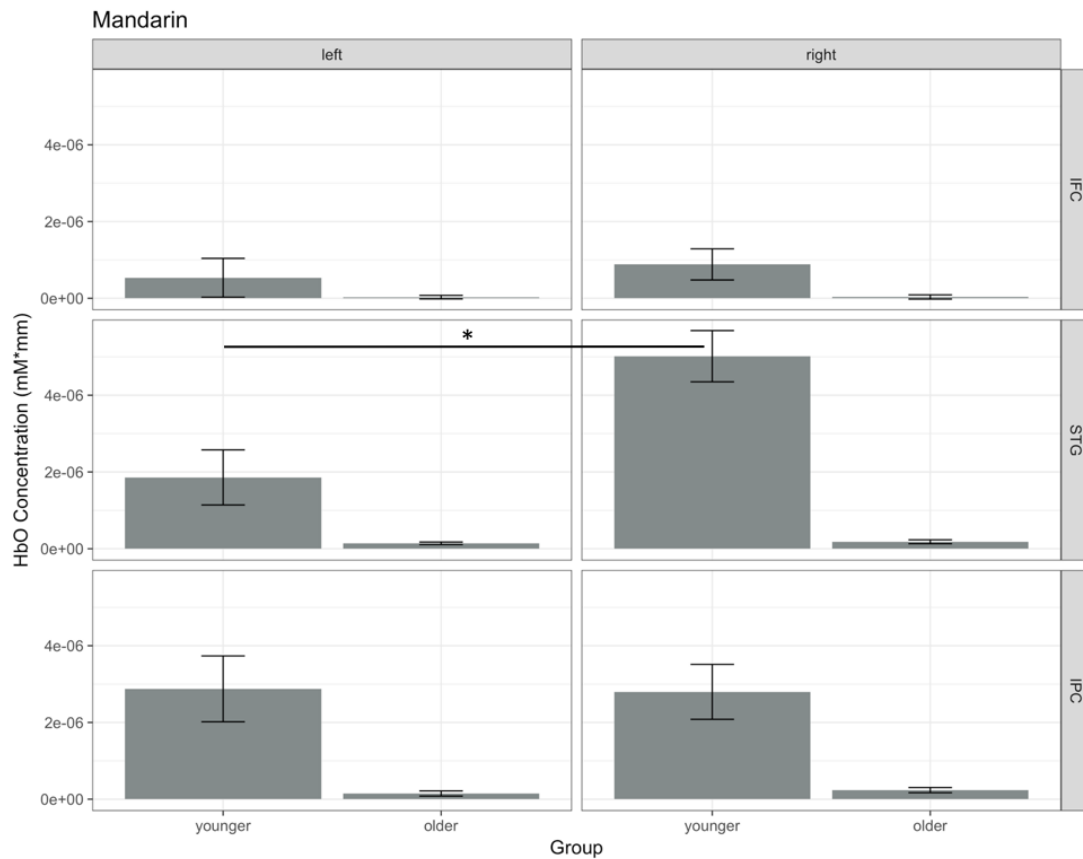


Figure 6.2. Mean HbO values for Mandarin in younger and older English monolingual infants, split across hemispheres and ROIs. As there was no effect of condition, the alternating and non-alternating values were averaged together. \* $p < .05$ , \*\* $p < .01$ .

### *Hindi*

A three-way interaction was found between ROI, Hemisphere, and Age,  $F(2, 86) = 5.57$ ,  $MS_e = 4678$ ,  $p = .005$ ,  $\eta_p^2 = .12$  (Figure 6.3). Subsequent analyses were conducted by Age to look for any further main effects or interactions between ROI and Hemisphere. There was a significant interaction between ROI and Hemisphere in the younger infants with simple main effects in both directions,  $F(2, 46) = 6.99$ ,  $MS_e = 5456$ ,  $p = .002$ ,  $\eta_p^2 = .23$ . First, there was a simple main effect of ROI in the left hemisphere, in which the IPC had the greatest activation,  $F(2, 46) = 7.46$ ,  $MS_e = 9174$ ,  $p = .002$ ,  $\eta_p^2 = .25$ . Second, there was another simple main effect of ROI in the right

hemisphere, where the STG had the greatest activation,  $F(2, 46) = 13.95$ ,  $MS_e = 9467$ ,  $p < .001$ ,  $\eta_p^2 = .38$ . Third, and in the opposite direction, there was a simple main effect of hemisphere in the STG where brain activation was lateralized to the right hemisphere,  $F(1, 23) = 10.88$ ,  $MS_e = 14111$ ,  $p = .003$ ,  $\eta_p^2 = .32$ . As for the older infants, there was only a main effect of ROI, in which brain activation was greatest in the bilateral IPC, followed by the STG then IFC,  $F(2, 40) = 3.31$ ,  $MS_e = 22.81$ ,  $p = .047$ ,  $\eta_p^2 = .14$ .

There was also a significant interaction between ROI and Hemisphere  $F(2, 86) = 6.59$ ,  $MS_e = 5536$ ,  $p = .002$ ,  $\eta_p^2 = .13$ , where simple main effects were found in both directions. First, there was a simple main effect of ROI in the left hemisphere,  $F(2, 88) = 6.97$ ,  $MS_e = 5141$ ,  $p = .002$ ,  $\eta_p^2 = .14$ , showing that the IPC had the greatest activation in both age groups. Pairwise comparisons found a significant difference between the IFC and IPC (mean difference 21.17,  $p = .015$ , Bonferroni corrected). Second, there was another simple main effect of ROI in the right hemisphere,  $F(2, 88) = 12.49$ ,  $MS_e = 5579$ ,  $p < .001$ ,  $\eta_p^2 = .22$ , with STG showing the greatest amount of activity in both age groups. Pairwise comparisons revealed a significant difference between IFC and STG (mean difference 22.25,  $p < .001$ ). Third, and in the opposite direction, there was a simple main effect of hemisphere in the STG,  $F(1, 44) = 10.13$ ,  $MS_e = 4131$ ,  $p = .003$ ,  $\eta_p^2 = .19$ , which showed that activation was significantly lateralized to the right hemisphere in both age groups. No other main effects or interactions were found.

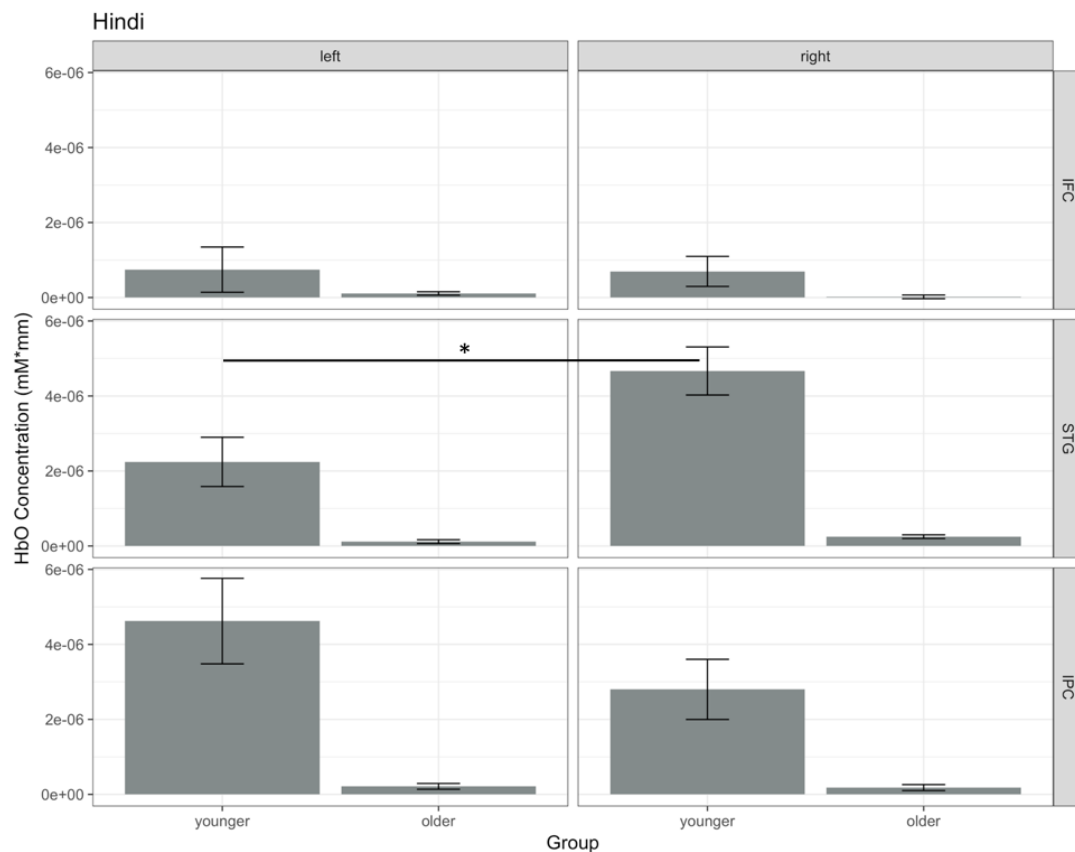


Figure 6.3. Mean HbO values for Hindi in younger and older English monolingual infants, split across hemispheres and ROIs. As there was no effect of condition, the alternating and non-alternating values were averaged together.  $*p < .05$ ,  $**p < .01$ .

### *Statistical analyses 2*

As there was no main effect or interaction that involved the factor Condition, an additional set of analyses were conducted on the dataset in which the non-Alt values were subtracted from the Alt values to retrieve a single difference measure for each channel in all three languages per participant (similar to the method used for the adult study in Chapter 3). Individual channels were examined for this analysis, rather than ROIs, to see if there were any significant channel-specific changes in brain activity with age in native and non-native phonemic discrimination. A mixed  $3 \times 2 \times 9 \times 2$  repeated-measures ANOVA was conducted with Language (English, Mandarin, Hindi),

Hemisphere (left, right), and Channel (LH: 1-9, RH: 10-18) as the within-subject factors, and Age (younger, older) as the between-subjects factor.

## Results

No significant main effects or interactions were found for HbO values. However, upon analyzing HbR, a mixed, repeated-measures ANOVA revealed a significant interaction between Language and Hemisphere in which the effect was significant in both directions  $F(1.76, 73.78) = 3.35$ ,  $MS_e = 1051$ ,  $p = .043$ ,  $\eta_p^2 = .07$ . Simple main effect analyses found that Hindi elicited the greatest increase in HbR in the left hemisphere,  $F(1, 44) = 7.91$ ,  $MS_e = 168$ ,  $p = .007$ ,  $\eta_p^2 = .15$  (see Figure 6.4). It was also found that within the left hemisphere, Hindi showed the greatest significant increase in HbR (or cortical deactivation) out of all the languages,  $F(2, 88) = 4.01$ ,  $MS_e = 239$ ,  $p = .022$ ,  $\eta_p^2 = .08$ . Pairwise comparisons (Bonferroni corrected) indicated a significant difference between Hindi and Mandarin (mean difference = 4.12,  $p = .046$ ). As an effect of Age was not found, the data from Chapters 4 and 5 suggest that Younger and Older infants did not show significant differences in brain activation patterns in the processing of each of the languages from 5-7 months to 10-12 months of age. Figure 6.4 depicts hemodynamic responses from the left and right hemispheres of all infants to Hindi.

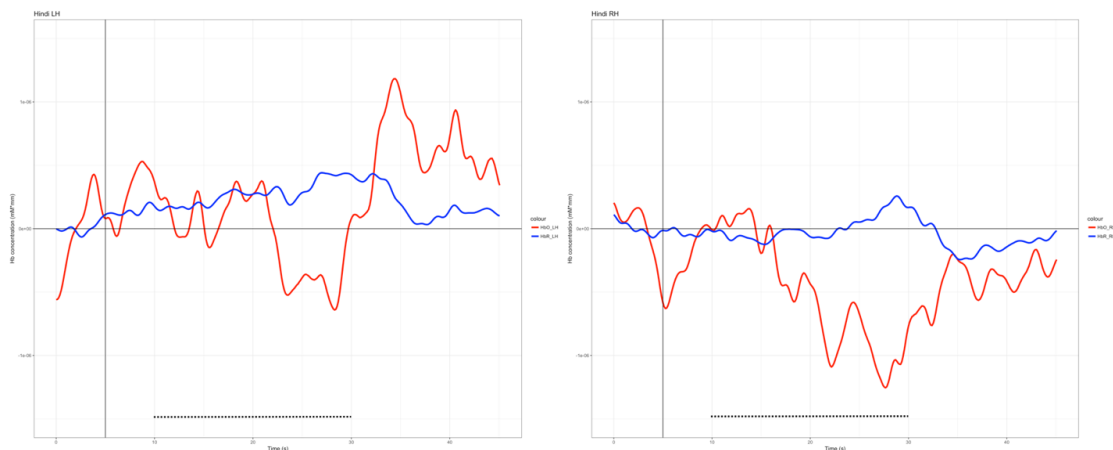


Figure 6.4. The grand averaged time courses of HbO (red line) and HbR (blue line) for Hindi (Alt – non-Alt) in younger and older infants and across all channels in the left and right hemispheres. The vertical line indicates the onset of auditory stimulation, and the horizontal dotted line indicates the time window (between 5 and 25 seconds after stimulus onset) for statistical analysis. All infants showed a significant increase in HbR to Hindi in the left hemisphere.

## Discussion

Taken together, younger infants exhibited an increase in HbO to all language conditions in the right STG, whereas the left sensorimotor cortex (left IPC) was recruited in the processing of the two non-native conditions (Mandarin and Hindi). Older infants, on the other hand, showed an increase in HbO in the bilateral IPC to all conditions. In the second analysis of HbR, there was no significant main effect of Age in the processing of all languages. However, both younger and older monolinguals were found to have a significant increase in HbR in the left hemisphere to Hindi (see Figure 6.4), suggesting that the non-native contrast elicited cortical deactivation in the dorsal pathway of speech processing that mainly resides in the left hemisphere (Hickok & Poeppel, 2007, Chapter 1). This finding was partially in line with our prediction that older monolingual infants would not exhibit a response indicating the detection of the



non-native contrast (i.e. an increase in HbR). Instead, we found that *both* 5-7- and 10-12-month-old English monolinguals showed cortical deactivation. Moreover, all infants showed differential HbR activity between the two non-native language conditions (Mandarin and Hindi), suggesting that the infants had distinct activation patterns between the two non-native conditions in which Hindi elicited greater cortical deactivation than Mandarin.

A possible explanation for these results is that most fNIRS systems only measure relative changes of HbO, HbR, and HbT concentrations, therefore, non-canonical or inverted responses are not straightforward to interpret, because it can be difficult to understand the physiological and functional meaning of a relative decrease in oxygenation (Issard & Gervain, 2018). However, it is generally accepted that a decrease in deoxygenation signifies an increase in oxygenation (or cortical activation), whereas an increase in deoxygenation signifies a decrease in oxygenation (cortical deactivation). Non-canonical responses are commonly observed in young infants, as their brains are still maturing with age, and the significant difference in cortical deactivation seen between Hindi and Mandarin suggests that both age groups differentiated the two non-native languages, with higher deactivation to Hindi compared to Mandarin. This result is supported by previous evidence for lexical tone perception, which has shown that non-tonal adult listeners were still proficient in acoustically differentiating changes in pitch (Hallé, Change, & Best, 2004), whereas non-native adult listeners have failed to discriminate the Hindi dental-retroflex contrast (Werker & Tees, 1984). Thus, the results from the present study suggest that the Mandarin language condition was more salient than Hindi to all infants, which was indicated by a relatively greater activation to the change in pitch.

Whereas older infants recruited the bilateral IPC in the processing of native and non-native language conditions, the increase in HbO in younger infants to only the non-native languages in the left IPC (i.e. sensorimotor cortex) suggests that this region could be a neural signature of universal, non-native speech perception. As discussed in Chapter 1, the left IPC functions as a sensorimotor interface in the dorsal pathway of speech perception. The interface, also referred to as Spt, codes sensory speech input and regulates the fine motor control of articulatory movements for accurate speech reproduction (Hickok, Buchsbaum, Humphries, & Muftuler, 2003). The leftward lateralization of cortical activation towards the Spt in younger infants was indicative of the universal perceptual system encoding non-native phonemic information for later speech reproduction. On the other hand, the absence of lateralization in the older infants shows that non-native information was no longer necessary for the encoding of articulatory processes. Interestingly, English did not exhibit a lateralized response as would be expected. It is possible that /t/ and /d/ have already been encoded into the infants' native phonemic repertoire and are already being produced in canonical babbling or as first words. This finding is also supported by the IS theory, in that brain activation patterns to a certain stimulus may change with age and experience. As previously mentioned, cortical regions may become more refined or specialized such that their activity becomes localized to a specific area (Johnson, 2011). However, the opposite can also occur, where a specific region activated at an earlier age prior to specialization may not activate in the same way at a later age after more experience has been acquired. Our finding of unilateral activation in the left IPC in younger infants and bilateral activation in older infants is consistent with the IS account, in which the mapping of a cognitive function to brain regions is fluid and dynamic over development (Joseph, Gathers, & Bhatt, 2011).

It was also interesting to see that the HbO means between younger and older infants were largely dissimilar, with the means for older infants much lower than those for younger infants (see Figures 6.1, 6.2, and 6.3). Three possible explanations may have accounted for the difference in these values. First, the older infants had more hair and slightly longer hair than the younger infants. The obstruction of hair might have contributed to the attenuation of the signal as it could have absorbed some of the infrared light. Second, an improved prototype of the infant fNIRS headband was developed between the testing of the older and younger infants. The newer version of the headband (tested on younger infants) was designed to have better contact between the optodes and the scalp, as well as an improved tolerance to motion artifacts. Third, under the IS framework, it is supposed that brain activation would become more focal with age. As our ROIs incorporated multiple channels and the number of channels varied for each ROI, it is possible that the weaker responses seen in the older infants were explained by the grouping of the channels. Therefore, these factors may have refined the fNIRS signal in younger infants, but they also might have made it more complicated to compare hemodynamic measurements across the two age groups.

In summary, the results from the current chapter indicate that older English monolingual infants recruited the bilateral IPC in the processing of all native and non-native contrasts, whereas younger infants recruited the right STG. More importantly, only the younger infants showed left IPC activation in the processing of the two non-native contrasts, both of which had significantly different levels of activation between each other. These results suggest that brain activation patterns are flexible during phonological development, consistent with IS, and that it is a possibility that the left IPC, or sensorimotor interface, could potentially be a neural signature for non-native phonological perception in the universal stages of speech processing.

## CHAPTER 7

### **The influence of bilingualism on phonological perception: A discussion**

## Objectives

The current thesis explored three main objectives: the first objective was to examine the development of brain activation patterns in phonological perception across the first year of life and in adulthood. The second was to study the neural correlates of lexical tone perception in Mandarin-English bilingual infants and adults who were learning and using two languages that contrast in the use of pitch. The third and final objective was to try and replicate Petitto et al. (2012) using a different experimental paradigm and a homogenous sample of bilingual infants to extend the evidence for the Perceptual Wedge Hypothesis. English monolingual and Mandarin-English bilingual infants and adults were tested on their native and non-native phonemic discrimination abilities. The following addresses each objective in greater detail.

### ***THE INFLUENCE OF LANGUAGE EXPERIENCE ON CORTICAL ACTIVATION DURING PHONEMIC PROCESSING ACROSS INFANCY AND IN ADULTHOOD***

The empirical research presented in the current thesis compared cortical activation patterns in the developing brain during phonological processing at 5-7 and 10-12 months of age. These findings, along with those from adults in Chapter 3, were examined to assess any similarities or differences in native and non-native speech perception between the developing infant and adult brain. The main observation taken from examining brain activation of all of our participants across the four empirical chapters revealed that younger (5-7 months) monolingual infants showed no differences in cortical activation across all native (English) and non-native (Mandarin, Hindi) language conditions. Their brain activity was indicated by a uniform pattern of activation localized in the sensorimotor interface in the left IPC and the auditory cortex in the right STG. The absence of differential activation in the left IFC (or Broca's Area)

suggests that younger infants' perception of native and non-native speech information was solely based on acoustic (STG) and not phonemic analyses (IFC) (Zatorre, Meyer, Gjedde, & Evans, 1996). As the main role of the left IFC is for the computation of phonetic categories (Myers, Blumstein, Walsh, & Eliassen, 2009; Zatorre, Evans, Meyer, & Gjedde, 1992), the absence of any differential processing or lateralization suggests that this region might not have been specialized for phonemic perception in 5-7-month-old monolingual infants.

Our additional finding of left IPC (sensorimotor) activation indicates that although younger infants at 5-7 months of age have not yet started producing words, listening to speech activated the left sensorimotor cortex to possibly allow for the subsequent mapping of motor articulatory representations in executive regions (e.g. left IFC) in preparation for the production of more complex sound sequences (i.e. words) at a later stage in development. Further, around 5-7 months of age, infants are beginning to engage in vocal play and start to show signs of canonical babbling. In other words, the activity observed in the left IPC in younger infants might suggest that increased sensorimotor involvement is required for the processing and regulating of fine motor-articulatory control for later speech reproduction.

Between all younger and older infants, our results showed an emergence of left IFC activation to the English contrast with age, suggesting that by the end of the first year of life, infants' perceptual networks have become specialized to process the native language in the region where native category representations are formed and distinguished (Myers, Blumstein, Walsh, & Eliassen, 2009; Zatorre, Evans, Meyer, & Gjedde, 1992). This finding is closely aligned with the Interactive Specialization (IS) theory introduced in Chapter 6, which states that cortical regions may become more refined and specialized with age and experience such that their activity becomes

localized to a specific area (Johnson, 2011), and in this case, the left IFC. Moreover, new processing competencies are likely to be associated with changes in activation patterns over a network made up of neighboring regions (Johnson, 2011). As the bilateral STG, left IFC and left IPC make up the dorsal pathway of speech perception (Hickok & Poeppel, 2007), the emergence of left IFC activation at 10-12 months of age comes as expected.

However, unlike the older infants, monolingual and bilingual adults did not show differential activation in any specific region to the same English contrast. We postulated that the absence of any differential processing was due to the reduced complexity of the stimulus. It is possible that the cognitive demand for the processing of a familiar and repeated /dæb – tæb/ contrast was reduced for experienced native listeners (e.g. Dehaene-Lambertz et al., 2006). Lastly, monolingual and bilingual adults exhibited activation in different regions of the brain while actively listening to the non-native Hindi contrast. Monolinguals recruited posterior areas in the bilateral IPC, whereas bilinguals exhibited bilateral activation in the anterior areas in the IFC. The IFC is commonly known to activate under phonemic discrimination tasks, as well as speech production and articulation. The IPC, however, is known for its role in the sensorimotor integration of speech (Hickok, Buchsbaum, Humphries, & Muftuler, 2003; Hickok & Poeppel, 2007), as well as the allocation of attention (Behrmann, Geng, & Shomstein, 2004; Liu, Slotnick, Serences, & Yantis, 2003). Thus, it is possible that during the presentation of Hindi, monolingual adults substantially shifted their attention towards the contrast that was relatively difficult, and sometimes impossible, to perceive by recruiting attentional mechanisms located in the bilateral IPC. This finding has also been observed in monolingual Japanese listeners, where MEG brain imaging revealed prolonged activity in the bilateral IPC during the processing of the non-native English

/r – l/ contrast (Zhang, Kuhl, Imada, Kotani, & Tohkura, 2005). On the other hand, bilingual adults recruited regions that activate when an acoustic change crossed a categorical boundary (Myers, Blumstein, Walsh, & Eliassen, 2009), suggesting that they utilized phonological and articulatory mechanisms to analyze subtle acoustic changes in linguistic auditory stimuli. The differential activation seen in the processing of the non-native Hindi contrast suggests that language experience influences the recruitment of executive brain regions to manipulate perceptual information.

### ***LEXICAL TONE PROCESSING IN MONOLINGUALS AND BILINGUALS WHOSE TWO LANGUAGES CONFLICT IN THE USE OF PITCH***

Through implementing fNIRS brain imaging, we found that brain activation to lexical tone was observed predominantly in the right hemisphere in monolingual non-tone-learning adults, as well as monolingual 5-7- and 10-12-month-old infants. This finding was in accordance with the literature on the hemispheric lateralization of speech, in which the right hemisphere is functionally specialized to process suprasegmental properties of spoken language such as prosody, or changes in pitch contour (Arimitsu et al., 2011; Homae, Watanabe, Nakano, Asakawa, & Taga, 2006; Homae, Watanabe, Nakano, & Taga, 2012; Telkemeyer et al., 2009; Zatorre, Evans, Meyer, & Gjedde, 1992). On the other hand, our results for bilingual infants and adults tell a more complex story. First, we found that bilingual adults showed classic brain activation patterns to native phonemic perception in the left IFC (Broca's Area) and bilateral STG to the Mandarin lexical tone contrast (Hickok & Poeppel, 2000; Zatorre, Evans, Meyer, & Gjedde, 1992). Interestingly, bilingual 5-7- and 10-12-month-old infants did not show left IFC activation to the native lexical tone contrast but exhibited robust, right-lateralized responses in the STG (younger infants) and IPC (older infants). This pattern of activation was in contrast to our findings of left IFC activation to the



native English contrasts in older monolingual and bilingual infants. Previous research supports our findings for older bilingual infants by showing that the right IPC indeed has activated under melodic, pitch judgment tasks (Royal et al., 2016), which compared to the current work can be extended to the melodic patterns found in lexical tone.

The absence of left IFC activation suggests that older bilingual infants did not perceive the lexical tone contrast phonemically, but rather acoustically in the right hemisphere. Our finding can be explained by a recent behavioral study conducted by Singh et al. (2018), in which the authors found that 6-, 9-, 12-, *and* 13-month-old Mandarin-English bilingual infants failed to discriminate both salient (T1-T3) and subtle (T2-T3) Mandarin tonal contrasts in a modified stimulus alternating paradigm, whereas monolingual infants at 9 months were able to discriminate the salient contrast and then both salient and subtle contrasts from 12 months onwards. The authors explained that their findings were due to the varied role that pitch plays in spoken language (i.e. phonemic vs. pragmatic), where the ambiguity resulting from a lack of language context may have been exacerbated in bilingual infants who were acquiring two language systems that use pitch in contrasting ways. Thus, the acquisition of the native phonology of two language systems that contrast in their use of pitch may be a more complex undertaking for bilingual infants. This may result in their perceptual systems requiring more time to consolidate and organize dual language phonological input. However, fNIRS brain imaging allowed us to observe what behavioral paradigms could not. Our neurophysiological findings complement well with those from Singh et al. (2018) by showing that bilingual infants at 10-12 months of age were more sensitive to pitch changes than monolinguals by detecting the change in pitch acoustically in the right IPC. In bilingual adults, however, the left IFC was activated in the discrimination of the lexical tone contrast. This finding is in support of previous MEG and EEG

research that have shown how bilingual infants indeed experience a slower transition from acoustic to phonemic analysis compared to monolinguals, whereby bilinguals might remain in the earlier stages of language perception for a longer period of time until a sufficient amount of experience has been accumulated (Ferjan Ramírez, Ramírez, Clarke, Taulu, & Kuhl, 2016; Garcia-Sierra et al., 2011).

To date, the research described in Chapter 5 remains to be the only study that used fNIRS brain imaging to assess Mandarin-English bilingual infants on their perception of Mandarin lexical tone. While Singh et al. (2018) did not find evidence for lexical tone discrimination in 6-13-month-old Mandarin-English bilingual infants in a behavioral phonemic discrimination paradigm, our neurophysiological results replicate and support these findings by showing no indication of significant brain activation between 5-12-month-olds in the region responsible for phonemic category representations. Additionally, we were able to extend Singh et al.'s findings by revealing that older bilingual infants remained more sensitive than monolinguals to acoustic pitch changes, which was a finding that could not have been detected with a behavioral phonemic discrimination paradigm. The current work has important implications for future research, whereby implementing contextual cues might be helpful to speech perception performance in bilingual subjects. Furthermore, brain imaging measures may unveil more complex inferences and more fine-grained development to those only derived from behavioral tests.

### ***COMPARING THE CURRENT WORK TO THE FINDINGS OF PETITTO ET AL. (2012)***

The third objective of the current thesis sought to replicate and extend the findings from a previous study that had also used fNIRS to assess native and non-native phonemic discrimination in 4-6- and 10-12-month-old monolingual and bilingual

infants. Petitto et al. (2012) employed fNIRS brain imaging in an event-related oddball paradigm on native English /ba - pa/ and non-native Hindi dental-retroflex /ʈa - ʈa/ contrasts, along with a non-linguistic pure tone condition consisting of one 250 Hz tone. Our experimental paradigm in Chapters 4 and 5 differed from Petitto et al. in three ways: first, we used a traditional block design that assessed differences in cortical activation to alternating (B-A-B-A-) and non-alternating (A-A-A-A-) blocks. This experimental design allowed us to test the processing of these two types of stimulus sequences separately, whereas Petitto et al. more explicitly measured the responses to a change in stimulus (oddball). Second, Petitto et al. studied English monolingual and bilingual infants exposed to French, Spanish, or Chinese in addition to English, whereas here (Chapter 5) we assessed English monolingual and Mandarin-English bilingual infants. Testing a homogeneous group of bilinguals can eliminate potential confounds resulting from language distance. For example, the language distance between Mandarin and English is greater than the distance between Spanish and English. Third, and finally, we measured brain regions that encompassed the entire neural network of speech processing, that is, the IFC, STG, and IPC (Hickok & Poeppel, 2007). On the other hand, Petitto et al. only measured brain activation in the IFC and STG. Measuring an additional region in the pre-motor cortices (IPC) allowed us to have the advantage in examining all functional aspects related to phonological processing in the speech perception network, especially since previous research has suggested that speech perception is a sensorimotor task (Bruderer, Danielson, Kandhadai, & Werker, 2015; Liberman & Mattingly, 1985; Westermann & Miranda, 2004; Wilson, Saygin, Sereno, & Iacoboni, 2004). In the following, similarities and differences of the findings between the current research (Chapters 4 and 5) and that of Petitto et al. are reviewed and discussed.

*All infants*

Petitto et al. (2012) found that all of their infant participants (4-6 and 10-12 months) exhibited a greater right-hemispheric activation in the STG to all auditory conditions, which was equivalent to our findings of right-lateralized activation in the STG in younger (Chapter 4) monolingual infants and older (Chapter 5) monolingual and bilingual infants. Although an explanation of this finding was not given in Petitto et al., we postulated that the rightward dominance in the STG was due to the slower speed at which the language stimuli were presented (Arimitsu et al., 2012), a smaller lexicon in infants aged under 12 months (Ferjan Ramírez et al., 2016; Mills, Plunkett, Prat, & Schafer, 2005), and the faster maturation and stronger cerebral blood flow of the right hemisphere in young infants (Chiron et al., 1997; Leroy et al., 2011; Roche-Labarbe et al., 2012).

Next, Petitto et al.'s (2012) findings showed that in the bilateral STG of all infants, the pure tone (250 Hz) condition significantly differed from the non-native (Hindi) condition, but not from the native (English) condition. Interestingly, this result aligned with our finding reported in Chapter 6, where all monolingual infants showed a significant difference in the left hemisphere between the two non-native language conditions, Hindi and Mandarin (which would be comparable to Petitto et al.'s pure tone condition), but not English. Although our findings were lateralized to the left STG and Petitto et al.'s findings in the bilateral STG, our results indicated that younger and older monolingual infants showed differential activation between the two types of non-native languages.

### *Younger vs. Older monolingual infants*

Additionally, in an interaction that approached significance ( $p < .08$ ), Petitto et al. (2012) found that the *left* IFC exhibited similar levels of activation between all younger (4-6 months) and older (10-12 months) infants, whereas the *right* IFC showed a decline in activation with age. The authors concluded that there was a shift in lateral dominance from bilateral to left-hemispheric in the IFC across development, approximately around the same time when infants begin producing their first words at 12 months. Our findings did not show any significant decrease in right IFC activity with age, however, our results did show an *increase* in left IFC activation with age. Our younger, 5-7-month-old monolingual infants did not show differential activation in the IFC to the English contrast (alternating condition), however, our older 10-12-month-old monolinguals and bilinguals exhibited robust left-lateralized activity in the same region. Our finding further strengthens and extends the observation made in Petitto et al., by showing that infants' neural language networks are increasingly becoming specialized towards the native language at the age when they begin to produce their first words, which was shown by the emergence of activation in Broca's Area, known to be associated with language production, at 10-12 months of age.

### *Monolingual vs. Bilingual infants*

Lastly, and of great importance to the main question of the current thesis, we explored whether there were any similarities or differences in cortical activation between older 10-12-month-old monolingual and bilingual infants to the non-native Hindi phonemic contrast. It was predicted (in Petitto et al. and in this thesis) that bilinguals would show perceptual flexibility in discriminating the non-native contrast. This was explained by the Perceptual Wedge Hypothesis in which bilinguals' universal

phonemic discrimination capacities would remain open for a longer period of time than monolinguals. Petitto et al. found that bilingual 10-12-month-old infants exhibited greater left IFC activation than their monolingual peers following the presentation of the deviant (i.e. oddball) stimulus during the non-native Hindi condition. Further, they also found that 10-12-month-old monolinguals showed greater left IFC activation only to the native language (English), whereas bilinguals did not show a significant difference in brain activation between native (English) and non-native (Hindi) languages. The authors thus concluded that 10-12-month-old bilingual infants remained sensitive to non-native phonemic contrasts for a longer period of time than monolingual infants of the same age, therefore providing evidence for the Perceptual Wedge Hypothesis. Our finding in Chapter 5 extends Petitto et al.'s results by showing that older 10-12-month-old bilingual infants exhibited a significantly greater increase in cortical activation in the left IFC to the Hindi *non-alternating* condition relative to older monolingual infants, whereas the only significant left IFC activation found in older monolingual infants was in response to the native language, English, and not Hindi.

It was unexpected to find that older bilingual infants had a greater response to the non-alternating block and not the alternating block that presented the contrast. However, it has been shown before that infants preferred non-alternating trials over alternating trials in previous behavioral studies (e.g. Maye, Werker, & Gerken, 2002; Yoshida, Pons, Maye, & Werker 2010). It is possible that because bilingual infants were able to discriminate the Hindi contrast, they might have exhibited differential processing to stable and repeated presentations of unfamiliar, non-native stimuli as they might prefer consistent streams of stimulation in order to extract information from their usually mixed, auditory environments. The main finding, however, is that bilingual infants were able to exhibit differential activation for a stimulus sequence that was

being presented within a language, and any difference in response to differing stimuli suggests that a contrast between the stimuli was detected (de Groot, 2011; Houston-Price & Nakai, 2004). Future research investigating the evidence for a bilingual auditory familiarity preference would increase our understanding of how bilinguals take greater advantage of the environmental cues in order to learn and understand new language systems.

### ***EXTENDING THE WORK OF PETITTO ET AL. (2012)***

Taken together, the work in Chapter 5 enriches previous findings on early speech processing from Petitto et al. (2012) in two ways. First, we were able to provide evidence for neural specialization in monolinguals to the native language between 5-7 and 10-12 months of age, as well as an absence in differential activation to the non-native contrast at 10-12 months (Chapters 4 and 5). This finding also provides support for the Native Language Neural Commitment (NLNC) Hypothesis, which states that universal phonetic perception must give way to language-specific information in order for native language acquisition to take place (Kuhl, 2004; Kuhl & Rivera-Gaxiola, 2008). This in turn increases learning for more complex language patterns, such as words, that are compatible with the learned phonemic structure of the native language. At the same time, there would be a reduction in attention to non-native language patterns so that the learning of them is no longer facilitated. This process thus shapes the brain through native linguistic experience during the critical period in language development in the first year of life. Similar to the results reported in Petitto et al., our sample of bilingual 10-12-month-olds appeared to remain in the universal stages of speech perception while monolinguals have already transitioned to language-specific analysis.

Second, our findings indicated that bilingual phonological acquisition develops differently from monolinguals, whereby bilingualism facilitates perceptual plasticity during the time when monolinguals' perceptual sensitivities have attuned to the native phonology. Previous neurophysiological studies on bilingual phonemic perception were also in support of these results. For example, Ferjan Ramírez, Ramírez, Clarke, Taulu, and Kuhl (2016) used MEG to assess Spanish and English phonemic contrasts in monolingual and bilingual 11-month-olds in a double oddball paradigm. Through examining MMR components elicited from the infants, the authors found that bilingual infants neurally discriminated both native contrasts at the acoustic level (early time window), whereas monolinguals discriminated the English contrast at the phonemic level (late time window). These results suggest that bilingual infants undergo a slower transition from acoustic (universal) to phonemic (language-specific) analysis of native speech (Kuhl et al., 2008). This was a result of dealing with a complex phonological space, in which the increased amount of phonetic information that bilingual infants need to learn facilitates a higher-functioning and flexible perceptual system. In turn, it will take bilinguals more time than monolinguals to acquire enough experience in each language to transition to the phonemic analysis of native speech, thus facilitating perceptual plasticity in phonological processing and protracting the perceptual narrowing time window.

### ***LIMITATIONS***

The work in the present thesis is limited by the lack of a sample of younger, 5-7-month-old bilingual infants, individual variability regarding age of acquisition in bilingual adults, as well as the testing procedure in Chapters 4 and 5 which might have been the result of contextual and referential issues for bilingual infants and the



unexpected pattern of activation between the alternating and non-alternating stimulation blocks.

We would have had a more complete and robust understanding of bilingual phonological development if an additional sample of 5-7-month-olds was tested in Singapore. However, there were not enough resources (i.e. time) for it to be possible. Rather, we chose to test 10-12-month-olds as it was of greater importance to assess perceptual flexibility of bilinguals at that age to see whether the findings would be in support of the Perceptual Wedge Hypothesis. Fortunately, Petitto et al. (2012) did test a younger and older age group of bilinguals, however, one of the main differences between the current research and that of Petitto et al. was that we assessed a homogenous sample of bilingual infants. Future work examining younger and older Mandarin-English bilingual infants on non-native phonemic perception would deepen our understanding of the changing activation patterns with age in the neural network of speech processing and how these developmental patterns are similar or different from monolinguals.

A second limitation is that we did not assess varying levels of language proficiency and age of acquisition in bilingual adults. As previous research has shown that either factor plays a role in the language processing and neural organization of bilinguals (Archila-Suerte, Zevin, Bunta, & Hernandez, 2013; Perani et al., 1998), more detailed investigations of these factors are needed to expand our understanding of how the availability of cognitive processes at the age of second language exposure and attained L2 proficiency would have an effect on the recruitment of brain regions to encode non-native phonemic information. Further, additional factors (e.g. quality and quantity of first (L1) and second (L2) language input, parental education, socioeconomic status (SES), and cultural biases) might have enriched our

understanding of the current findings. For example, it has been shown that vocabulary development and language processing skills were significantly affected by SES, whereby 24-month-old infants from lower and higher SES families experience a 6-month gap in processing skills critical to language development between the two groups (Fernald, Marchman, Weisleder, 2013). In preschoolers, it has been shown that L1 and L2 input from home was significantly related to vocabulary skills in L1 and L2 (Cheung, Kan, Winicour, Yang, 2018). If these additional factors were taken into consideration at the time of testing, then they might have helped provide further explanations of some of the surprising results found in the current work.

Another limitation of the current work was the way in which the stimuli were presented. Chapters 4 and 5 used silence periods between alternating and non-alternating stimulation blocks to avoid signal contamination from adjacent stimulation periods. However, the silence period of 25 seconds between 20-second blocks of auditory stimulation might have been too long for young infants to retain the information learned from the previous stimulation period and compare with the following stimulation block. This might explain why similar levels of activation were observed for the alternating and non-alternating conditions across all languages and brain regions, as well as the greater rightward dominance of activation. Although there is currently no published research on verbal working memory in infants under 12 months of age, future research might help confirm whether or not silence periods are useful in phonemic discrimination in fNIRS experimental paradigms. But in any case, it can be argued that the paradigms in Chapters 4 and 5 assessed the processing of two different types of stimulus sequences (i.e. alternating vs. non-alternating), whereas alternating block paradigms (e.g. e.g. Arimitsu et al., 2011; Minagawa-Kawai, Mori, & Sato, 2005; Sato, Sogabe, & Mazuka, 2010, discussed in Chapter 2) used non-

alternating stimulus blocks as a baseline. To prevent the possibility of having no significant differences in activation between alternating and non-alternating stimulus blocks, measuring the response to a change in stimulus in an alternating block paradigm might be better suited.

Lastly, previous research has emphasized how it is not best practice to assess bilinguals under traditional monolingual testing procedures. Phoneme discrimination under challenging conditions (e.g. absence of language context, lack of referential cues) might be difficult for bilinguals, as they may be more reliant on contextual and referential support to navigate a richer and more complex phonological space (Singh et al., 2018). As the testing sessions in Chapter 5 were conducted in English, bilingual infants might have been biased to process the lexical tone contrast in the English context. Therefore, the lack of contextual cues in this case is one of the possible explanations for the absence of left hemispheric activation to the native Mandarin lexical tone contrast.

### **Conclusion**

Through using fNIRS brain imaging, we successfully demonstrated language-specific perceptual narrowing in the brains of monolingual infants, as well as the neuroplasticity of the bilingual perceptual system in processing non-native speech by 12 months of age. Further, we provided neurophysiological evidence to extend the findings of a behavioral lexical tone discrimination study that showed how phonological acquisition in bilinguals is a more complex undertaking than monolinguals (Singh et al., 2018). As more cognitive resources are required to learn two separate and possibly overlapping phonological systems, bilingual infants might therefore take advantage of their environmental resources and use contextual and

referential cues to extract linguistic information from general auditory input. In Singh et al. (2018) and in Chapter 5, the research suggests that the lack of referential cues for bilingual infants had a negative impact on their phonemic discrimination abilities of Mandarin lexical tone. Although Singh et al. (2018) and the current work did not explicitly test for the effect of referential cues during phonological perception, this finding warrants future research to implement these types of cues.

The current work also demonstrated that language experience undoubtedly modifies the cortical architecture of speech perception. Akin to the NLNC Hypothesis (Kuhl, 2004; Kuhl & Rivera-Gaxiola, 2008), brain activation to the English contrast was shown to emerge in the left inferior frontal cortex in monolingual infants by 10-12 months of age. Additionally, monolingual infants no longer exhibited phonemic sensitivity to non-native Mandarin and Hindi speech sounds by the end of the first year of life. On the other hand, bilingual infants of the same age showed flexibility in non-native phonological perception by showing differential activation in each language condition, in support of the Perceptual Wedge Hypothesis. We also demonstrated how language experience modifies the cortical architecture in adults through showing that monolingual adults recruited attention-related mechanisms in the posterior regions to look for and analyze subtle acoustic changes in non-native linguistic auditory stimuli while bilingual adults used anterior, articulatory motor mechanisms.

In summary, the exploratory nature of the work presented in the current thesis shows us how acquiring two phonological systems from birth affects phonemic perception across infancy and in adulthood. Whereas monolingual and bilingual infants use the same cognitive resources to acquire language, bilinguals need to allocate their limited resources across two language systems. Due to the greater amount of phonological information that bilinguals are required to learn, native acquisition of both

languages will naturally take a longer time than acquiring only one. As a result, bilinguals may take longer to neurally commit to the native language by remaining sensitive to non-native phonemic contrasts for a longer time than monolinguals. Therefore, the present work should stand as a demonstration for the complexities of dual language acquisition that bilingual infants may face, and how it can affect the recruitment of cortical regions during the perception of phonological units.

## References

- Anderson, J. A., Mak, L., Chahi, A. K., & Bialystok, E. (2018). The language and social background questionnaire: Assessing degree of bilingualism in a diverse population. *Behavior Research Methods*, 50(1), 250-263.
- Anderson, J. L., Morgan, J. L., & White, K. S. (2003). A statistical basis for speech sound discrimination. *Language and Speech*, 46(2-3), 155-182.
- Archila, P., Ramos, A. I., Zevin, J., & Hernandez, A. E. (2010). *How age of acquisition and proficiency predicts sound detection in bilinguals*. Poster presented at Armadillo Conference, College Station, TX.
- Archila-Suerte, P., Zevin, J., Bunta, F., & Hernandez, A. E. (2012). Age of acquisition and proficiency in a second language independently influence the perception of non-native speech. *Bilingualism: Language and Cognition*, 15(1), 190-201.
- Archila-Suerte, P., Zevin, J., Ramos, A. I., & Hernandez, A. E. (2013). The neural basis of non-native speech perception in bilingual children. *NeuroImage*, 67, 51-63.
- Arimitsu, T., Uchida-Ota, M., Yagihashi, T., Kojima, S., Watanabe, S., Hokuto, I., ... & Minagawa-Kawai, Y. (2011). Functional hemispheric specialization in processing phonemic and prosodic auditory changes in neonates. *Frontiers in Psychology*, 2, 202.
- Aslin, R. N., Shukla, M., & Emberson, L. L. (2015). Hemodynamic correlates of cognition in human infants. *Annual Review of Psychology*, 66, 349-379.
- Ball, R., Shu, C., Xi, P., Rioux, M., Luximon, Y., & Molenbroek, J. (2010). A comparison between Chinese and Caucasian head shapes. *Applied Ergonomics*, 41(6), 832-839.

- Behrmann, M., Geng, J. J., & Shomstein, S. (2004). Parietal cortex and attention. *Current Opinion in Neurobiology*, 14(2), 212-217.
- Benavides-Varela, S., Hochmann, J. R., Macagno, F., Nespor, M., & Mehler, J. (2012). Newborn's brain activity signals the origin of word memories. *Proceedings of the National Academy of Sciences*, 201205413.
- Bertoncini, J., Bijeljac-Babic, R., Blumstein, S. E., & Mehler, J. (1987). Discrimination in neonates of very short CVs. *The Journal of the Acoustical Society of America*, 82(1), 31-37.
- Best, C. C., & McRoberts, G. W. (2003). Infant perception of non-native consonant contrasts that adults assimilate in different ways. *Language and Speech*, 46(2-3), 183-216.
- Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. C. Goodman & H. C. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 167-224). Cambridge, MA: MIT Press.
- Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, 14(3), 345-360.
- Bidelman, G. M., Hutka, S., & Moreno, S. (2013). Tone language speakers and musicians share enhanced perceptual and cognitive abilities for musical pitch: Evidence for bidirectionality between the domains of language and music. *PloS One*, 8(4), e60676.

- Binder, J. R., Rao, S. M., Hammeke, T. A., Yetkin, F. Z., Jesmanowicz, A., Bandettini, P. A., ... & Hyde, J. S. (1994). Functional magnetic resonance imaging of human auditory cortex. *Annals of Neurology*, 35(6), 662-672.
- Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language*, 44(2), 235-249.
- Boersma, P., & Weenink, D. (2012). Praat, v. 5.3.
- Born, P., Leth, H., Miranda, M. J., Rostrup, E., Stensgaard, A., Peitersen, B., ... & Lou, H. C. (1998). Visual activation in infants and young children studied by functional magnetic resonance imaging. *Pediatric Research*, 44(4), 578-583.
- Born, P., Rostrup, E., Leth, H., Peitersen, B., & Lou, H. (1996). Change of visually induced cortical activation patterns during development. *The Lancet*, 347(9000), 543.
- Bosseler, A., Taulu, S., Pihko, E., Mäkelä, J., Imada, T., Ahonen, A., & Kuhl, P. (2013). Theta brain rhythms index perceptual narrowing in infant speech perception. *Frontiers in Psychology*, 4, 690.
- Bouchon, C., Nazzi, T., & Gervain, J. (2015). Hemispheric asymmetries in repetition enhancement and suppression effects in the newborn brain. *PloS One*, 10(10), e0140160.
- Brigadoi, S., Ceccherini, L., Cutini, S., Scarpa, F., Scatturin, P., Selb, J., ... & Cooper, R. J. (2014). Motion artifacts in functional near-infrared spectroscopy: A comparison of motion correction techniques applied to real cognitive data. *NeuroImage*, 85, 181-191.



- Bruderer, A. G., Danielson, D. K., Kandhadai, P., & Werker, J. F. (2015). Sensorimotor influences on speech perception in infancy. *Proceedings of the National Academy of Sciences*, 112(44), 13531-13536.
- Burnham, D. K. (1986). Developmental loss of speech perception: Exposure to and experience with a first language. *Applied Psycholinguistics*, 7(3), 207-239.
- Burnham, D., Francis, E., Webster, D., Luksaneeyanawin, S., Lacerda, F., & Attapaiboon, C. (1996). Facilitation or attenuation in the development of speech mode processing? Tone perception over linguistic contexts. In *Sixth Australian International Conference on Speech Science and Technology* (pp. 587-592).
- 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(3), 455-474.
- Byers-Heinlein, K., Burns, T. C., & Werker, J. F. (2010). The roots of bilingualism in newborns. *Psychological Science*, 21(3), 343-348.
- Chao, Y. R. (1965). *Grammar of spoken Chinese*. Berkeley, CA: University of California Press.
- Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., & Näätänen, R. (1998). Development of language-specific phoneme representations in the infant brain. *Nature Neuroscience*, 1, 351-353.
- Cheung, S., Kan, P. F., Winicour, E., & Yang, J. (2019). Effects of home language input on the vocabulary knowledge of sequential bilingual children. *Bilingualism: Language and Cognition*, 22(5), 986-1004.
- Chiron, C., Jambaque, I., Nabbout, R., Lounes, R., Syrota, A., & Dulac, O. (1997). The right brain hemisphere is dominant in human infants. *Brain*, 120(6), 1057-1065.

- Conant, L. L., Liebenthal, E., Desai, A., & Binder, J. R. (2014). fMRI of phonemic perception and its relationship to reading development in elementary- to middle-school-age children. *NeuroImage*, 89, 192-202.
- Cooper, R., Selb, J., Gagnon, L., Phillip, D., Schytz, H. W., Iversen, H. K., ... & Boas, D. A. (2012). A systematic comparison of motion artifact correction techniques for functional near-infrared spectroscopy. *Frontiers in Neuroscience*, 6, 147.
- Dehaene, S., Dupoux, E., Mehler, J., Cohen, L., Paulesu, E., Perani, D., ... & Le Bihan, D. (1997). Anatomical variability in the cortical representation of first and second language. *NeuroReport*, 8(17), 3809-3815.
- Dehaene-Lambertz, G. (1997). Electrophysiological correlates of categorical phoneme perception in adults. *NeuroReport*, 8(4), 919-924.
- Dehaene-Lambertz, G., Dehaene, S., Anton, J. L., Campagne, A., Ciuciu, P., Dehaene, G. P., ... & Pallier, C. (2006). Functional segregation of cortical language areas by sentence repetition. *Human Brain Mapping*, 27(5), 360-371.
- Dehaene-Lambertz, G., Dehaene, S., & Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science*, 298(5600), 2013-2015.
- Dehaene-Lambertz, G., Montavont, A., Jobert, A., Alliol, L., Dubois, J., Hertz-Pannier, L., & Dehaene, S. (2010). Language or music, mother or Mozart? Structural and environmental influences on infants' language networks. *Brain and Language*, 114(2), 53-65.
- Démonet, J. F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J. L., Wise, R., ... & Frackowiak, R. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain*, 115(6), 1753-1768.

- DeWitt, I., & Rauschecker, J. P. (2012). Phoneme and word recognition in the auditory ventral stream. *Proceedings of the National Academy of Sciences*, 109(8), E505-E514.
- Eilers, R. E., & Minifie, F. D. (1975). Fricative discrimination in early infancy. *Journal of Speech, Language, and Hearing Research*, 18(1), 158-167.
- Eilers, R. E., Wilson, W. R., & Moore, J. M. (1977). Developmental changes in speech discrimination in infants. *Journal of Speech, Language, and Hearing Research*, 20(4), 766-780.
- Eimas, P. D. (1974). Auditory and linguistic processing of cues for place of articulation by infants. *Perception & Psychophysics*, 16(3), 513-521.
- Eimas, P. D. (1975). Auditory and phonetic coding of the cues for speech: Discrimination of the [r-l] distinction by young infants. *Perception & Psychophysics*, 18(5), 341-347.
- Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, 171(3968), 303-306.
- Fenson, L. (2007). *MacArthur-Bates communicative development inventories*. Baltimore, MD: Brookes Publishing Company.
- Ferjan Ramírez, N., Ramírez, R. R., Clarke, M., Taulu, S., & Kuhl, P. K. (2016). Speech discrimination in 11-month-old bilingual and monolingual infants: A magnetoencephalography study. *Developmental Science*, 20(1), e12427.
- Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science*, 16(2), 234-248.

- Flom, R. (2014). Perceptual narrowing: Retrospect and prospect. *Developmental Psychobiology*, 56(7), 1442-1453.
- Furuya, I., & Mori, K. (2003). Cerebral lateralization in spoken language processing measured by multi-channel near-infrared spectroscopy (NIRS). *No to Shinkei = Brain and Nerve*, 55(3), 226-231.
- Gandour, J., Wong, D., Hsieh, L., Weinzapfel, B., Lancker, D. V., & Hutchins, G. D. (2000). A crosslinguistic PET study of tone perception. *Journal of Cognitive Neuroscience*, 12(1), 207-222.
- Garcia-Sierra, A., Rivera-Gaxiola, M., Percaccio, C. R., Conboy, B. T., Romo, H., Klarman, L., ... & Kuhl, P. K. (2011). Bilingual language learning: An ERP study relating early brain responses to speech, language input, and later word production. *Journal of Phonetics*, 39(4), 546-557.
- Gervain, J., Berent, I., & Werker, J. F. (2012). Binding at birth: The newborn brain detects identity relations and sequential position in speech. *Journal of Cognitive Neuroscience*, 24(3), 564-574.
- Gervain, J., Macagno, F., Cogoi, S., Peña, M., & Mehler, J. (2008). The neonate brain detects speech structure. *Proceedings of the National Academy of Sciences*, 105(37), 14222-14227.
- Gervain, J., Mehler, J., Werker, J. F., Nelson, C. A., Csibra, G., Lloyd-Fox, S., ... & Aslin, R. N. (2011). Near-infrared spectroscopy: A report from the McDonnell infant methodology consortium. *Developmental Cognitive Neuroscience*, 1(1), 22-46.

- Giuliano, R. J., Pfordresher, P. Q., Stanley, E. M., Narayana, S., & Wicha, N. Y. (2011). Native experience with a tone language enhances pitch discrimination and the timing of neural responses to pitch change. *Frontiers in Psychology*, 2, 146.
- Glover, G. H. (2011). Overview of functional magnetic resonance imaging. *Neurosurgery Clinics*, 22(2), 133-139.
- Goldberg, E., Podell, K., & Lovell, M. (1994). Lateralization of frontal lobe functions and cognitive novelty. *The Journal of Neuropsychiatry and Clinical Neurosciences*, 6(4), 371.
- Golestani, N., & Zatorre, R. J. (2004). Learning new sounds of speech: Reallocation of neural substrates. *NeuroImage*, 21(2), 494-506.
- Graf Estes, K., & Hay, J. F. (2015). Flexibility in bilingual infants' word learning. *Child Development*, 86(5), 1371-1385.
- de Groot, A. (2011). Language and cognition in bilinguals and multilinguals. New York: Psychology Press.
- Grosjean, F. (2010). *Bilingual: Life and reality*. Harvard University Press.
- Hallé, P. A., Chang, Y. C., & Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics*, 32(3), 395-421.
- Hampson, M., Peterson, B. S., Skudlarski, P., Gatenby, J. C., & Gore, J. C. (2002). Detection of functional connectivity using temporal correlations in MR images. *Human Brain Mapping*, 15(4), 247-262.
- Hannon, E. E., & Trehub, S. E. (2005a). Metrical categories in infancy and adulthood. *Psychological Science*, 16(1), 48-55.

- Hannon, E. E., & Trehub, S. E. (2005b). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences*, *102*(35), 12639-12643.
- Hay, J. F., Graf Estes, K., Wang, T., & Saffran, J. R. (2015). From flexibility to constraint: The contrastive use of lexical tone in early word learning. *Child Development*, *86*(1), 10-22.
- Herwig, U., Satrapi, P., & Schönfeldt-Lecuona, C. (2003). Using the international 10-20 EEG system for positioning of transcranial magnetic stimulation. *Brain Topography*, *16*(2), 95-99.
- Hickok, G., Buchsbaum, B., Humphries, C., & Muftuler, T. (2003). Auditory-motor interaction revealed by fMRI: Speech, music, and working memory in area Spt. *Journal of Cognitive Neuroscience*, *15*(5), 673-682.
- Hickok, G., Houde, J., & Rong, F. (2011). Sensorimotor integration in speech processing: Computational basis and neural organization. *Neuron*, *69*(3), 407-422.
- Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Sciences*, *4*(4), 131-138.
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*, 393.
- Homae, F., Watanabe, H., Nakano, T., Asakawa, K., & Taga, G. (2006). The right hemisphere of sleeping infant perceives sentential prosody. *Neuroscience Research*, *54*(4), 276-280.
- Homae, F., Watanabe, H., Nakano, T., & Taga, G. (2007). Prosodic processing in the developing brain. *Neuroscience Research*, *59*(1), 29-39.

- Homae, F., Watanabe, H., Nakano, T., & Taga, G. (2012). Functional development in the infant brain for auditory pitch processing. *Human Brain Mapping, 33*(3), 596-608.
- Houston-Price, C. & Nakai, S. (2004). Distinguishing novelty and familiarity effects in infant preference procedures. *Infant and Child Development, 13*, 341-348.
- Hunter, M. A., & Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. *Advances in Infancy Research, 5*, 69-95.
- Huppert, T. J., Diamond, S. G., Franceschini, M. A., & Boas, D. A. (2009). HomER: A review of time-series analysis methods for near-infrared spectroscopy of the brain. *Applied Optics, 48*(10), D280-D298.
- Imada, T., Zhang, Y., Cheour, M., Taulu, S., Ahonen, A., & Kuhl, P. K. (2006). Infant speech perception activates Broca's Area: A developmental magnetoencephalography study. *NeuroReport, 17*(10), 957-962.
- Issard, C., & Gervain, J. (2017). Adult-like processing of time-compressed speech by newborns: A NIRS study. *Developmental Cognitive Neuroscience, 25*, 176-184.
- Issard, C., & Gervain, J. (2018). Variability of the hemodynamic response in infants: Influence of experimental design and stimulus complexity. *Developmental Cognitive Neuroscience, 33*, 182-193.
- Jackson-Maldonado, D., Thal, D., Fenson, L., Marchman, V. A., Newton, T., & Conboy, B. T. (2003). *MacArthur inventarios del desarrollo de habilidades comunicativas: User's guide and technical manual*. Baltimore, MD: Brookes Publishing Company.

- Jöbsis-vanderVliet, F. F. (1999). Discovery of the near-infrared window into the body and the early development of near-infrared spectroscopy. *Journal of Biomedical Optics*, 4(4), 392-397.
- Johnson, M. H. (2011). Interactive specialization: A domain-general framework for human functional brain development? *Developmental Cognitive Neuroscience*, 1(1), 7-21.
- Jongman, A., Wang, Y., Moore, C. B., & Sereno, J. A. (2006). Perception and production of Mandarin Chinese tones. Submitted to: *Handbook of Chinese psycholinguistics*. E. Bates, L.H. Tan, and O.J.L Tzeng (eds.). Cambridge University Press.
- Joseph, J. E., Gathers, A. D., & Bhatt, R. S. (2011). Progressive and regressive developmental changes in neural substrates for face processing: Testing specific predictions of the interactive specialization account. *Developmental Science*, 14(2), 227-241.
- Kelly, D. J., Liu, S., Lee, K., Quinn, P. C., Pascalis, O., Slater, A. M., & Ge, L. (2009). Development of the other-race effect during infancy: Evidence toward universality? *Journal of Experimental Child Psychology*, 104(1), 105-114.
- Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L., & Pascalis, O. (2007). The other-race effect develops during infancy: Evidence of perceptual narrowing. *Psychological Science*, 18(12), 1084-1089.
- Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PloS One*, 7(5), e36399.



- Klein, D., Zatorre, R. J., Milner, B., & Zhao, V. (2001). A cross-linguistic PET study of tone perception in Mandarin Chinese and English speakers. *NeuroImage*, 13, 646-653.
- Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Research Reviews*, 29(2-3), 169-195.
- Knudsen, E. I. (2004). Sensitive periods in the development of the brain and behavior. *Journal of Cognitive Neuroscience*, 16(8), 1412-1425.
- Kobayashi, M., Otsuka, Y., Nakato, E., Kanazawa, S., Yamaguchi, M. K., & Kakigi, R. (2011). Do infants represent the face in a viewpoint-invariant manner? Neural adaptation study as measured by near-infrared spectroscopy. *Frontiers in Human Neuroscience*, 5, 153.
- Kotilahti, K., Nissilä, I., Näsi, T., Lipiäinen, L., Nojonen, T., Meriläinen, P., ... & Fellman, V. (2010). Hemodynamic responses to speech and music in newborn infants. *Human Brain Mapping*, 31(4), 595-603.
- Kovelman, I., Shalinsky, M. H., Berens, M. S., & Petitto, L. A. (2008). Shining new light on the brain's "bilingual signature": A functional near infrared spectroscopy investigation of semantic processing. *NeuroImage*, 39(3), 1457-1471.
- Kozel, F. A., Tian, F., Dhamne, S., Croarkin, P. E., McClintock, S. M., Elliott, A., ... & Liu, H. (2009). Using simultaneous repetitive transcranial magnetic stimulation/functional near infrared spectroscopy (rTMS/fNIRS) to measure brain activation and connectivity. *NeuroImage*, 47(4), 1177-1184.

- Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, 5, 831-843.
- Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1493), 979-1000.
- Kuhl, P. K., Conboy, B. T., Padden, D., Nelson, T., & Pruitt, J. (2005). Early speech perception and later language development: Implications for the “critical period”. *Language Learning and Development*, 1(3-4), 237-264.
- Kuhl, P. K., & Miller, J. D. (1975). Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants. *Science*, 190(4209), 69-72.
- Kuhl, P. K., & Miller, J. D. (1978). Speech perception by the chinchilla: Identification functions for synthetic VOT stimuli. *The Journal of the Acoustical Society of America*, 63(3), 905-917.
- Kuhl, P. K., Ramírez, R. R., Bosseler, A., Lin, J. F. L., & Imada, T. (2014). Infants’ brain responses to speech suggest analysis by synthesis. *Proceedings of the National Academy of Sciences*, 111(31), 11238-11245.
- Kuhl, P. K., & Rivera-Gaxiola, M. (2008). Neural substrates of language acquisition. *Annual Review of Neuroscience*, 31, 511-534.
- 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(2), F13-F21.

- 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(5044), 606-608.
- Lee, Y. S., Turkeltaub, P., Granger, R., & Raizada, R. D. (2012). Categorical speech processing in Broca's Area: An fMRI study using multivariate pattern-based analysis. *Journal of Neuroscience*, 32(11), 3942-3948.
- Leroy, F., Glasel, H., Dubois, J., Hertz-Pannier, L., Thirion, B., Mangin, F., & Dehaene-Lambertz, G. (2011). Early maturation of the linguistic dorsal pathway in human infants. *The Journal of Neuroscience*, 31(4), 1500-1506.
- Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, 54(5), 358-368.
- Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21(1), 1-36.
- Liebenthal, E., Binder, J. R., Spitzer, S. M., Possing, E. T., & Medler, D. A. (2005). Neural substrates of phonemic perception. *Cerebral Cortex*, 15(10), 1621-1631.
- Liu, L., & Kager, R. (2014). Perception of tones by infants learning a non-tone language. *Cognition*, 133(2), 385-394.
- Liu, T., Slotnick, S. D., Serences, J. T., & Yantis, S. (2003). Cortical mechanisms of feature-based attentional control. *Cerebral Cortex*, 13(12), 1334-1343.
- Lloyd-Fox, S., Blasi, A., & Elwell, C. E. (2010). Illuminating the developing brain: The past, present and future of functional near infrared spectroscopy. *Neuroscience & Biobehavioral Reviews*, 34(3), 269-284.

- Lloyd-Fox, S., Blasi, A., Elwell, C. E., Charman, T., Murphy, D., & Johnson, M. H. (2013). Reduced neural sensitivity to social stimuli in infants at risk for autism. *Proceedings of the Royal Society B: Biological Sciences*, 280(1758), 20123026.
- Lloyd-Fox, S., Papademetriou, M., Darboe, M. K., Everdell, N. L., Wegmuller, R., Prentice, A. M., ... & Elwell, C. E. (2014). Functional near infrared spectroscopy (fNIRS) to assess cognitive function in infants in rural Africa. *Scientific Reports*, 4, 4740.
- Mahmoudzadeh, M., Dehaene-Lambertz, G., Fournier, M., Kongolo, G., Goudjil, S., Dubois, J., ... & Wallois, F. (2013). Syllabic discrimination in premature human infants prior to complete formation of cortical layers. *Proceedings of the National Academy of Sciences*, 201212220.
- Martin, E., Joeri, P., Loenneker, T., Ekatodramis, D., Vitacco, D., Hennig, J., & Marcar, V. L. (1999). Visual processing in infants and children studied using functional MRI. *Pediatric Research*, 46(2), 135-140.
- Mattock, K., & Burnham, D. (2006). Chinese and English infants' tone perception: Evidence for perceptual reorganization. *Infancy*, 10(3), 241-265.
- Mattock, K., Molnar, M., Polka, L., & Burnham, D. (2008). The developmental course of lexical tone perception in the first year of life. *Cognition*, 106(3), 1367-1381.
- Maurer, D., & Werker, J. F. (2014). Perceptual narrowing during infancy: A comparison of language and faces. *Developmental Psychobiology*, 56(2), 154-178.
- May, L., Gervain, J., Carreiras, M., & Werker, J. F. (2017). The specificity of the neural response to speech at birth. *Developmental Science*, 21(3), e12564.

- Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, 82(3), B101-B111.
- Mechelli, A., Crinion, J. T., Noppeney, U., O'doherty, J., Ashburner, J., Frackowiak, R. S., & Price, C. J. (2004). Structural plasticity in the bilingual brain. *Nature*, 431, 757.
- Meek, J. H., Firbank, M., Elwell, C. E., Atkinson, J., Braddick, O., & Wyatt, J. S. (1998). Regional hemodynamic responses to visual stimulation in awake infants. *Pediatric Research*, 43, 840-843.
- Mercure, E., Kushnerenko, E., Goldberg, L., Bowden-Howl, H., Coulson, K., Johnson, M. H., & MacSweeney, M. (2019). Language experience influences audiovisual speech integration in unimodal and bimodal bilingual infants. *Developmental Science*, 22(1), e12701.
- Mills, D. L., Coffey-Corina, S. A., & Neville, H. J. (1993). Language acquisition and cerebral specialization in 20-month-old infants. *Journal of Cognitive Neuroscience*, 5(3), 317-334.
- Mills, D. L., Plunkett, K., Prat, C., & Schafer, G. (2005). Watching the infant brain learn words: Effects of vocabulary size and experience. *Cognitive Development*, 20(1), 19-31.
- Minagawa-Kawai, Y., Cristià, A., & Dupoux, E. (2011). Cerebral lateralization and early speech acquisition: A developmental scenario. *Developmental Cognitive Neuroscience*, 1(3), 217-232.
- Minagawa-Kawai, Y., Cristià, A., Long, B., Vendelin, I., Hakuno, Y., Dutat, M., ... & Dupoux, E. (2013). Insights on NIRS sensitivity from a cross-linguistic study on the emergence of phonological grammar. *Frontiers in Psychology*, 4, 170.

- Minagawa-Kawai, Y., van der Lely, H., Ramus, F., Sato, Y., Mazuka, R., & Dupoux, E. (2010). Optical brain imaging reveals general auditory and language-specific processing in early infant development. *Cerebral Cortex*, 21(2), 254-261.
- Minagawa-Kawai, Y., Mori, K., Hebden, J. C., & Dupoux, E. (2008). Optical imaging of infants' neurocognitive development: Recent advances and perspectives. *Developmental Neurobiology*, 68(6), 712-728.
- Minagawa-Kawai, Y., Mori, K., Naoi, N., & Kojima, S. (2007). Neural attunement processes in infants during the acquisition of a language-specific phonemic contrast. *Journal of Neuroscience*, 27(2), 315-321.
- Minagawa-Kawai, Y., Mori, K., & Sato, Y. (2005). Different brain strategies underlie the categorical perception of foreign and native phonemes. *Journal of Cognitive Neuroscience*, 17(9), 1376-1385.
- Minagawa-Kawai, Y., Naoi, N., Nishijima, N., Kojima, S., & Dupoux, E. (2007). Developmental changes in cerebral responses to native and non-native vowels: A NIRS study. In *Proceedings of the International Conference of Phonetic Sciences XVI* (pp. 1877-1880).
- Molavi, B., & Dumont, G. A. (2012). Wavelet-based motion artifact removal for functional near-infrared spectroscopy. *Physiological Measurement*, 33(2), 259-270.
- Moore, C. B., & Jongman, A. (1997). Speaker normalization in the perception of Mandarin Chinese tones. *The Journal of the Acoustical Society of America*, 102(3), 1864-1877.
- Myers, E. B. (2014). Emergence of category-level sensitivities in non-native speech sound learning. *Frontiers in Neuroscience*, 8, 238.

- Myers, E. B., Blumstein, S. E., Walsh, E., & Eliassen, J. (2009). Inferior frontal regions underlie the perception of phonetic category invariance. *Psychological Science*, 20(7), 895-903.
- Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., ... & Allik, J. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, 385(6615), 432.
- Narayan, C. R., Werker, J. F., & Beddor, P. S. (2010). The interaction between acoustic salience and language experience in developmental speech perception: Evidence from nasal place discrimination. *Developmental Science*, 13(3), 407-420.
- Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, 24(3), 756-766.
- Nazzi, T., Nelson, D. G. K., Jusczyk, P. W., & Jusczyk, A. M. (2000). Six-month-olds' detection of clauses embedded in continuous speech: Effects of prosodic well-formedness. *Infancy*, 1(1), 123-147.
- Nespor, M., Peña, M., & Mehler, J. (2003). On the different roles of vowels and consonants in speech processing and language acquisition. *Lingue e Linguaggio*, 2(2), 203-230.
- Nordenhake, M., & Svantesson, J.-O. (1983). Duration of standard Chinese word tones in different sentence environments. *Working Papers*, 25, 105-111.
- Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2010). Broca's Area and language processing: Evidence for the cognitive control connection. *Language and Linguistics Compass*, 4(10), 906-924.

- Obrig, H., Mock, J., Stephan, F., Richter, M., Vignotto, M., & Rossi, S. (2017). Impact of associative word learning on phonotactic processing in 6-month-old infants: A combined EEG and fNIRS study. *Developmental Cognitive Neuroscience*, 25, 185-197.
- Pascalis, O., Loevenbruck, H., Quinn, P. C., Kandel, S., Tanaka, J. W., & Lee, K. (2014). On the links among face processing, language processing, and narrowing during development. *Child Development Perspectives*, 8(2), 65-70.
- Peña, M., Maki, A., Kovačić, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., & Mehler, J. (2003). Sounds and silence: An optical topography study of language recognition at birth. *Proceedings of the National Academy of Sciences*, 100(20), 11702-11705.
- Peña, M., Werker, J. F., & Dehaene-Lambertz, G. (2012). Earlier speech exposure does not accelerate speech acquisition. *Journal of Neuroscience*, 32(33), 11159-11163.
- Perani, D., Dehaene, S., Grassi, F., Cohen, L., Cappa, S. F., Dupoux, E., ... & Mehler, J. (1996). Brain processing of native and foreign languages. *NeuroReport*, 7(15-17), 2439-2444.
- Perani, D., Paulesu, E., Galles, N. S., Dupoux, E., Dehaene, S., Bettinardi, V., ... & Mehler, J. (1998). The bilingual brain. Proficiency and age of acquisition of the second language. *Brain*, 121(10), 1841-1852.
- Petitto, L. A., Berens, M. S., Kovelman, I., Dubins, M. H., Jasinska, K., & Shalinsky, M. (2012). The “Perceptual Wedge Hypothesis” as the basis for bilingual babies’ phonetic processing advantage: New insights from fNIRS brain imaging. *Brain and Language*, 121(2), 130-143.



- Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, 20(2), 421-435.
- Rivera-Gaxiola, M., Klarman, L., Garcia-Sierra, A., & Kuhl, P. K. (2005). Neural patterns to speech and vocabulary growth in American infants. *NeuroReport*, 16(5), 495-498.
- Rivera-Gaxiola, M., Silva-Pereyra, J., & Kuhl, P. K. (2005). Brain potentials to native and non-native speech contrasts in 7- and 11-month-old American infants. *Developmental Science*, 8(2), 162-172.
- Roche-Labarbe, N., Fenoglio, A., Aggarwal, A., Dehaes, M., Carp, S. A., Franceschini, M. A., & Grant, P. E. (2012). Near-infrared spectroscopy assessment of cerebral oxygen metabolism in the developing premature brain. *Journal of Cerebral Blood Flow & Metabolism*, 32(3), 481-488.
- Royal, I., Vuvan, D. T., Zendel, B. R., Robitaille, N., Schönwiesner, M., & Peretz, I. (2016). Activation in the right inferior parietal lobule reflects the representation of musical structure beyond simple pitch discrimination. *PloS One*, 11(5), e0155291.
- Sato, Y., Sogabe, Y., & Mazuka, R. (2010). Development of hemispheric specialization for lexical pitch–accent in Japanese infants. *Journal of Cognitive Neuroscience*, 22(11), 2503-2513.
- Sato, Y., Utsugi, A., Yamane, N., Koizumi, M., & Mazuka, R. (2013). Dialectal differences in hemispheric specialization for Japanese lexical pitch accent. *Brain and Language*, 127(3), 475-483.

- Seidl, A., & Cristià, A. (2008). Developmental changes in the weighting of prosodic cues. *Developmental Science*, 11(4), 596-606.
- Shmuel, A., Yacoub, E., Chaimow, D., Logothetis, N. K., & Ugurbil, K. (2007). Spatio-temporal point-spread function of fMRI signal in human gray matter at 7 Tesla. *NeuroImage*, 35(2), 539-552.
- Singh, L., Fu, C. S., Seet, X. H., Tong, A. P., Wang, J. L., & Best, C. T. (2018). Developmental change in tone perception in Mandarin monolingual, English monolingual, and Mandarin–English bilingual infants: Divergences between monolingual and bilingual learners. *Journal of Experimental Child Psychology*, 173, 59-77.
- Singh, L., Loh, D., & Xiao, N. G. (2017). Bilingual infants demonstrate perceptual flexibility in phoneme discrimination but perceptual constraint in face discrimination. *Frontiers in Psychology*, 8, 1563.
- Streeter, L. A. (1976). Language perception of 2-month-old infants shows effects of both innate mechanisms and experience. *Nature*, 259, 39-41.
- Stroganova, T. A., Orekhova, E. V., & Posikera, I. N. (1998). Externally and internally controlled attention in infants: An EEG study. *International Journal of Psychophysiology*, 30(3), 339-351.
- Sundara, M., Ngon, C., Skoruppa, K., Feldman, N. H., Onario, G. M., Morgan, J. L., & Peperkamp, S. (2018). Young infants' discrimination of subtle phonetic contrasts. *Cognition*, 178, 57-66.
- Sundara, M., Polka, L., & Molnar, M. (2008). Development of coronal stop perception: Bilingual infants keep pace with their monolingual peers. *Cognition*, 108(1), 232-242.

- Taga, G., Homae, F., & Watanabe, H. (2007). Effects of source-detector distance of near infrared spectroscopy on the measurement of the cortical hemodynamic response in infants. *NeuroImage*, 38(3), 452-460.
- Telkemeyer, S., Rossi, S., Koch, S. P., Nierhaus, T., Steinbrink, J., Poeppel, D., ... & Wartenburger, I. (2009). Sensitivity of newborn auditory cortex to the temporal structure of sounds. *Journal of Neuroscience*, 29(47), 14726-14733.
- Thierry G., Boulanouar, K., Kherif, F., Ranjeva, J., & Démonet, J. (1999). Temporal sorting of neural components underlying phonological processing. *NeuroReport*, 10, 2599-2603.
- Trehub, S. E. (1973). Infants' sensitivity to vowel and tonal contrasts. *Developmental Psychology*, 9(1), 91-96.
- Trehub, S. E. (1976). The discrimination of foreign speech contrasts by infants and adults. *Child Development*, 47(2), 466-472.
- Vihman, M. M. (2014). *Phonological development: The first two years*. Boston, MA: Wiley-Blackwell.
- Wang, Y., Behne, D. M., Jongman, A., & Sereno, J. A. (2004). The role of linguistic experience in the hemispheric processing of lexical tone. *Applied Psycholinguistics*, 25(3), 449-466.
- Wang, Y., Jongman, A., & Sereno, J. A. (2001). Dichotic perception of Mandarin tones by Chinese and American listeners. *Brain and Language*, 78, 332-348.
- Watanabe, E., Maki, A., Kawaguchi, F., Takashiro, K., Yamashita, Y., Koizumi, H., & Mayanagi, Y. (1998). Non-invasive assessment of language dominance with near-infrared spectroscopic mapping. *Neuroscience Letters*, 256(1), 49-52.

- Wei, M., Joshi, A. A., Zhang, M., Mei, L., Manis, F. R., He, Q., ... & Xue, F. (2015). How age of acquisition influences brain architecture in bilinguals. *Journal of Neurolinguistics*, 36, 35-55.
- Weikum, W. M., Oberlander, T. F., Hensch, T., K., & Werker, J. F. (2012). Prenatal exposure to antidepressants and depressed maternal mood alter trajectory of infant speech perception. *Proceedings of the National Academy of Sciences*, 109(2), 17221-7.
- Werker, J. F., Gilbert, J. H., Humphrey, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child Development*, 53(1), 349-355.
- Werker, J. F., & Hensch, T. K. (2015). Critical periods in speech perception: New directions. *Annual Review of Psychology*, 66, 173-196.
- 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(1), 49-63.
- Werker, J. F., & Tees, R. C. (1999). Influences on infant speech processing: Toward a new synthesis. *Annual Review of Psychology*, 50(1), 509-535.
- Westermann, G., & Miranda, E. R. (2004). A new model of sensorimotor coupling in the development of speech. *Brain and Language*, 89(2), 393-400.
- Wilcox, T., Bortfeld, H., Woods, R., Wruck, E., & Boas, D. A. (2005). Using near-infrared spectroscopy to assess neural activation during object processing in infants. *Journal of Biomedical Optics*, 10(1), 011010.

- Wilson, S. M., & Iacoboni, M. (2006). Neural responses to non-native phonemes varying in producibility: Evidence for the sensorimotor nature of speech perception. *NeuroImage*, 33(1), 316-325.
- Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience*, 7, 701-702.
- Xi, J., Zhang, L., Shu, H., Zhang, Y., & Li, P. (2010). Categorical perception of lexical tones in Chinese revealed by mismatch negativity. *Neuroscience*, 170(1), 223-231.
- Yeung, H. H., Chen, K. H., & Werker, J. F. (2013). When does native language input affect phonetic perception? The precocious case of lexical tone. *Journal of Memory and Language*, 68(2), 123-139.
- Yip, M. (2002). *Tone*. Cambridge University Press.
- Yoshida, K. A., Pons, F., Maye, J., & Werker, J. F. (2010). Distributional phonetic learning at 10 months of age. *Infancy*, 15(4), 420-433.
- Zatorre, R. J., & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex*, 11(10), 946-953.
- Zatorre, R. J., Evans, A. C., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science*, 256(5058), 846-849.
- Zatorre, R. J., Meyer, E., Gjedde, A., & Evans, A. C. (1996). PET studies of phonetic processing of speech: Review, replication, and reanalysis. *Cerebral Cortex*, 6(1), 21-30.

- Zhang, Y., Kuhl, P. K., Imada, T., Kotani, M., & Tohkura, Y. I. (2005). Effects of language experience: Neural commitment to language-specific auditory patterns. *NeuroImage*, 26(3), 703-720.
- Zinszer, B. D., Chen, P., Wu, H., Shu, H., & Li, P. (2015). Second language experience modulates neural specialization for first language lexical tones. *Journal of Neurolinguistics*, 33, 50-66.