EXPLORING THE PERCEPTION OF PHONEMIC VOWEL LENGTH CONTRASTS: EVIDENCE FROM INFANTS AND ADULTS

Hui Chen

(BA (Chinese Language & Literature), MA (Chinese Linguistics), MPhil (Linguistics))

A thesis submitted in fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Department of Linguistics
ARC Centre of Excellence in Cognition and its Disorders
Faculty of Human Sciences
Macquarie University
Sydney, Australia

Submitted: 6th April 2016
Awarded: 4th August 2016
Conferred: September 2016

Table of Contents

Thesis Abstract

Infants tune into native sound categories as early as in their first year, but in order to understand the language, they must become aware of the *phonemic* function of the sounds as well. A number of studies have investigated infants' phonemic awareness of consonant and vowel quality contrasts at the early word learning stage, however, there has been no research directly examining infants' early understanding of *phonemic vowel length* contrasts in a language where vowel duration can signal word meaning alone, such as Japanese, Finnish, Arabic, and even Australian English. Since vowels can also vary in duration as a function of prosodic context, an investigation of how *phonemic vowel length* is acquired is essential for understanding early phonological development more generally.

This thesis therefore focuses on the perception of phonemic vowel length contrasts. It is comprised of three studies, targeting three populations respectively: Japanese infants, Australian English infants, and bi-dialectal Australian English adults. The first study revealed that Japanese infants have developed awareness of phonemic vowel length contrasts by 18 months, which is probably related to the systematicity and robustness of the contrasts manifesting in the language. The second study showed that Australian English-learning infants have become sensitive to mispronunciations of phonemic vowel length by at least 24 months, possibly earlier than often thought. The third study indicated that native Australian English adults, who have had early exposure to another English dialect that does not have contrastive vowel length, might have established more flexible phonological categories of phonemic vowel length, compared to those without this early exposure.

Taken together, the findings of this thesis suggest that the development of phonemic vowel length contrasts is tied to the systematicity and stability of these contrasts in the language being learned.

Declaration

The research presented in this thesis is my original work and it has not been submitted for a higher degree in any other institution. In addition, I certify that all information sources and literature used are indicated in the thesis. The research projects presented in this thesis have gain approval from Macquarie University Human Research Ethics Committee (Ref 5201200814).

Some of the material in this thesis is in preparation for submitting for publication. Chapter Two is based on the article (1), Chapter Three is based on the article (2), and Chapter Four is based on the article (3):

(1) Chen, H., Yamane, N., Xu Rattanasone, N., Demuth, K., & Mazuka, R. (under review). Japanese infants are aware of phonemic vowel length in novel words at 18 months. *Infancy*.

(2) Chen, H., Xu Rattanasone, N., Cox, F., & Demuth, K. (under review). Understanding the acquisition of phonemic vowel length contrasts in Australian English-learning 18- and 24-month-olds. *Language Learning and Development*.

(3) Chen, H., Xu Rattanasone, N., Cox, F., & Demuth, K. (in submission). Effect of early dialectal exposure on adult perception of phonemic vowel length. *Journal of the Acoustical Society of America*.

I additionally certify that I was the first author of all the chapters of this thesis, and that all the experimental design, data collection, coding, analysis and writing of the articles have been performed and completed by me, in consultation with my principle supervisor Katherine Demuth, co-supervisors Felicity Cox and Nan Xu Rattanasone, and with my collaborators Reiko Mazuka and Naoto Yamane.

Signed:

Hui Chen (Student Number: 42663431)

Acknowledgments

I would like to express my gratitude to many people who have offered me invaluable help and encouragement during my PhD journey.

First and foremost, I owe my profoundest gratitude to my principal supervisor, Distinguished Professor Katherine Demuth. She has spent enormous amounts of time and energy guiding me through the process of my PhD study. She has stimulated my thinking, broadened my thoughts and helped me to discover the beauty of research in phonological development. She has consistently pushed me to sharpen my ideas as reflected in both my thinking and writing, and reminded me to be always mindful of the bigger picture. I thank her for providing unfailing support, encouragement, and care throughout my academic career.

I would like to extend my sincere gratitude to my co-supervisors Associate Professor Felicity Cox and Dr. Nan Xu Rattanasone. Felicity's comprehensive lectures on acoustics of Australian English inspired me to study Australian English vowels and phonemic vowel length. She has provided careful and invaluable feedback on my thesis. Nan has guided me through the process of conducting research and analyzing results with infant studies. She has also spent large amounts of time reading my drafts and providing detailed comments.

My heartfelt appreciation also goes to my collaborators Professor Reiko Mazuka and Dr. Naoto Yamane. I am greatly indebted to Reiko for her interest on my topic and the opportunity she offered me to conduct experiments on Japanese infants at her lab in Tokyo. She has spent lots of time and energy on me and the project. I have benefited from her expertise in infant studies and Japanese phonology in every discussion with her. Thank you to Naoto for providing training on using E-Prime, conducting experiments with Japanese infants, Perl and R scripting, and comments to the design and write up of my Japanese study.

I would like to thank all the Japanese and Australian infants, their parents, and the

Australian adult participants for their participation. I am also deeply grateful to Kaori Myogan and Kelly Miles, for their incredible patience and fantastic performance in the stimulus recordings. It was a real blessing for me to have them being my "voice ladies".

For my Japanese colleagues, I would like to express my gratitude to the fellows in the Laboratory for Language Development at RIKEN Brain Science Institute, including Ako Ohori, Kyoko Teraoka, Mari Kanamura, Yuri Hatano, Ruka Hayano, Naoko Kijima, Sho Tsuji, Natsumi Shibata, Ying Deng, and Miki Takahashi. I thank them for providing assistance with participant recruitment, data collection, processing of demographic information, and training on data coding. I also thank them for hosting me as a fellow colleague with enthusiasm.

For my colleagues in the Child Language Lab at Macquarie University, including Katherine Revius, Benjamin Davies, Kiri Mealings, Gretel Macdonald, and Fabia Andronos. I am deeply grateful for their efficient help with participant recruitment, data collection, processing of demographic information, administrative and emotional support. Special thanks go to my "eye-tracking buddies" Katherine Revius and Benjamin Davies. You are the most joyful people to work with.

I have also benefited from engaging in professional discussions with the following post-doctoral fellows, Ivan Yuen, Susan Lin, Carmen Kung, and Elaine Schmidt. I thank Ivan Yuen, Carmen Kung, and Elaine Schmidt for their warm support and friendship at the final stage of my thesis writing. I also thank Titia Benders and Michael Proctor for holding various technical and writing workshops which have been very helpful for my research. For my fellow PhD colleagues, Ekaterina Tomas, Kiri Mealings, Sithembinkosi Dube, Qandeel Hussain, Benjamin Davies, Tamara Schembri, Paragati Mandikal Vasuki, and Ping Tang, I thank them for their friendship, support, and encouragement throughout my PhD journey. For other colleagues, including but not exclusive to Mili Mathew, Samantha Bentink, Rebecca Holt, Nay San, Alyssa Dyball, Amy German, Phillip Cheng, David Rosson, Vasfiye Geckin, Vishnu K. K. Nair, Neha Khetrapal, Shasha An, Lisa

Ceccherini, Rebecca Gelding, Wei He, Aijun Huang, Weiyi Ma, David Meng, Gao Na, Valerie Pu, Alex Woolgar, and Peng Zhou, who had interacted with me however briefly at the Child Language Lab and the ARC Centre of Excellence for Cognition and its Disorders, thank you for making the journey joyful and colourful. I also thank Collette Ryan for her administrative assistance in the department.

For other national and international colleagues, Christina Quattropani, and Isabella Dautriche, thank you for providing assistance with participant recruitment and information on eye-tracking data analysis. Also, I have discussed my research at many conferences and workshops with eminent researchers in my field including professors Catherine Best, Paola Escudero, Daniel Swingley, Thierry Nazzi, and Ota Mitsuhiko, and I thank them for their time and insightful discussions.

Thanks also go to Macquarie University for the iMQRES scholarship as a financial support, the Department of Linguistics for funding my research, the ARC Centre of Excellence for Cognition and its Disorders for all the facilities and lovely working environment, and RIKEN Brain Science Institute for the Short-term International Program to fund my research in Japan.

I am indebted to my friends Huong Duong, Sany Dwita, Stella Ye, Soumaya Abbassi, Minna Hsu, Xuejing Lu, Yanan Sun, Ami Sambai, Joann Tang, Ricqle Wei, and especially Sophie Zhu, for their warm friendship and spiritual support, which helped me overcome the frustrating moments in my PhD journey. Special thanks also go to the little angel Bella, who has been touching, inspiring, and encouraging me, with all her incredible endeavors, wisdoms, and enthusiasms for life. Working and Playing with her was always a highlight of my weeks.

Finally, I would like to express my deep gratitude to my beloved parents. Their continued support and appreciation on my work over these years helped me go through every difficulty in the lengthy process of my study. It is their endless love that gives me the strength to continue my academic pursuit ahead.

Chapter One: Introduction

Introduction

One of the central questions of language acquisition is how learners establish phonological categories. To do this, they must determine what kinds of phonetic variation are linguistically meaningful. More specifically, in the lexical domain, language learners need to associate sounds with word meanings, and discover which sound changes are phonemic, altering word meaning. This is essential for the establishment of phonological representations. For some languages, phonemic vowel length is an important contrast that must be acquired. This thesis aims to explore the development of phonemic vowel length contrasts and the effect of native language input on this process.

Development of Phonological Representations

During the past decades, it has been widely established that infants' perceptual sensitivity becomes attuned to language-specific phonetic contrasts in the first year of life. Several developmental patterns have been reported in this process. Infants have been most commonly reported to maintain their sensitivity to native phonetic contrasts, but attenuate the awareness of nonnative ones (Burnham, 1986; Burnham, Earnshaw, & Quinn, 1987; Kuhl, Williams, Lacerda, Steven, & Linblom, 1992; Polka & Werker, 1994; Werker & Lalonde, 1988; Werker & Polka, 1993; Werker & Tees, 1983, 1984a, 1984b; among others). This happens first with lexical tones by four months (Yeung, Chen, & Werker, 2013), then with vowels at around six months (Kuhl et al., 1992; Polka & Werker, 1994), and later with consonants between ten and twelve months (e.g., Aslin, Pisoni, Hennesy, & Perey, 1981; Best & McRoberts, 2003; Burnham, 1986; Werker, Gilbert, Humphrey, & Tees, 1981; Werker & Lalonde, 1988; Werker & Tees, 1983, 1984a). Also, infants' sensitivity to some of the native phonetic contrasts is enhanced with more native language input (Burnham et al., 1987; Kuhl, 2004; Kuhl et al., 2006; Polka, Colantonio, & Sundara, 2001; Sundara, Polka, & Genesee, 2006; Tsao, Liu, & Kuhl, 2006). In addition, some phonetic contrasts regarded as difficult can be perceived later when infants have more native language experience (e.g., Filipino initial consonants /n-ŋ/ by 10-12 months, Narayan, Werker, & Beddor, 2010; and Japanese short-long vowel duration by 9.5 months, Sato, Sogabe, & Mazuka, 2010). Taken together, these findings suggest that infants' phonetic awareness is crucially shaped by the native language input and undergoes re-organization to conform to the native sound system in their first year of age in life.

However, in addition to the above mentioned language specific factors that shape early phonological development, some acoustically salient cues can bias speech perception across languages. Some phonetic contrasts have been reported to remain highly discriminable throughout early development. They are either difficult to assimilate into the native phonemic inventory, such as Zulu clicks (Best & McRoberts, 2003; Best, McRoberts, & Sithole, 1988), or constrained by language-universal perceptual bias, such as changes towards vowels with extreme articulatory-acoustic properties (Polka & Bohn, 1988, 2003, 2011). Further, even for native phonetic contrasts, acoustically less salient contrasts are perceived later than more salient ones (Narayan et al., 2010; Pons, Albareda-Castellot, & Sebastián-Gallés, 2012). These findings suggest that acoustic factors constrain (i.e., either facilitate or delay) the developmental process of language-specific attunement in phonetic sensitivity in early infancy (see Cutler & Mehler (1993) and Yeung et al. (2013) for further discussion on the relationship between acoustic saliency and discrimination time course of tones > vowels > consonants during the first year of life).

To develop phonological representations in the lexicon, infants must further specify the contrastive phonetic detail in words and understand their phonemic status in the language. The interaction between universal acoustic biases and early language-specific experience becomes more complex in phonemic perception at the word learning stage in the second year of life.

Infants demonstrate difficulty in using their newly developed phonetic sensitivity to learn novel words (see Werker & Fennell (2004) for a review). Using a habituation-switch task, Stager and Werker (1997) first reported that infants at 14 months are not able to perceive the fine phonetic detail in *consonants* with this test paradigm, and this result was replicated by a number of subsequent studies (e.g., Pater, Stager, & Werker, 1998, 2004). Only by 20 months, can infants succeed in this task, while 17-month-olds demonstrate an intermediate level (Werker, Fennell, Corcoran, & Stager, 2002). These results suggest a second re-organization in phonological perception in the second year of life, when language-specific phonological tuning is still taking place (Stager & Werker, 1997; Werker & Fennell, 2004).

However, later studies have revealed that infants succeed in perceiving mispronunciations of known words in an Intermodal Preferential Looking (IPL, Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987) task by 14-15 months of age (Fennell & Waxman, 2010; Fennell & Werker, 2003; Mani & Plunkett, 2007; Swingley & Aslin, 2002; among others). More recent studies that use the IPL testing paradigm have also shown that young infants do notice the minimal phonetic difference in novel words by 14-15 months (Mani & Plunkett, 2008; Yoshida, Fennel, Swingley, & Werker, 2009). Together these studies suggest infants do tune into the contrastive phonemic function of vowels and consonants earlier than previously thought. The failure in previous studies testing novel words in a habituation-switch task might be then possibly due to the attention demands in the learning and testing phase, lexical competition and the likelihood of the context for encoding a novel object (cf. Swingley & Aslin, 2007). The findings further suggest that language-specific phonological tuning and the development of phonetic categories over the course of the first year is relevant for the phonological representations in infants' first words (Swingley & Aslin, 2002).

Language-specific phonological tuning may be affecting phonological development in another way at the word learning stage. Recent studies have shown that infants recognize words in a nonnative regional/artificial accent by 19 months of age but not earlier at 15 months (Best, Tyler, Gooding, Orlando, & Quann, 2009; Mulak, Best, Tyler, Kitamura, & Irwin, 2013; White & Aslin, 2011). These findings suggest that, besides mastering the contrastive function of cross-category variation, infants also learn to attenuate within-category variation and understand *phonological constancy* (Best, 1994; Best et al., 2009). This is regarded as another milestone in the development of abstract phonological categories. Interestingly, it has also been reported that when infants were exposed to a familiar story read in the unfamiliar accent prior to test, the early inability to recognize accented words in the 15-month-olds can be overcome (van Heugten & Johnson, 2014). Taken these together, the top-down abstraction ability infants develop in the second year, along with an enlarging vocabulary, appears to play an important role in the process of developing phonological representations.

On the other hand, acoustic saliency affects the development of phonemic representations at the early word learning stage in a prominent way. Infants have been shown to have a linear perceptual sensitivity to 1-feature, 2-feature, and 3-feature mispronunciations in initial consonants at 19 months (White & Morgan, 2008). As for vowels, British English infants have been shown that they can detect a 1-feature change in vowel quality (i.e., height and backness) in known words by 18 months, but not earlier at 15 months (Mani et al., 2008; Mani & Plunkett, 2007), and their sensitivity to different sizes of mispronunciations is not linear (Mani & Plunkett, 2011). For testing novel words in a habituation-switch task, some studies have showed that infants can perceive vowel contrasts with bigger phonetic differences even when they are nonnative (e.g., Canadian English /i/ vs. /ɪ/) at 15 months, but not those with smaller native contrasts (e.g., Canadian English /i/ vs. /u/ and Australian English /i/ vs. /ɪ/) (Curtin, Fennell, & Escudero, 2007, 2009; Escudero, Best, Kitamura, & Mulak, 2014). These findings suggest that infants' sensitivity to phonological contrasts is closely related to the magnitude of acoustic difference (Escudero et al., 2014).

Phonemic Vowel Length

Vowel length is understood as the duration of vowel. It can vary depending on intrinsic and extrinsic factors, and serves many functions in language. For intrinsic vowel length, which is determined by the nature of the vowel itself, its variations can signal changes in word meaning. For example, the words *to* ('door') and *too* ('tower') in Japanese can only be differentiated by duration of the vowels (Ota, 1999). Intrinsic vowel length therefore can constitute phonemic contrasts in language, i.e., *phonemic vowel length*. For extrinsic vowel length, which is affected by contextual factors, its variations can inform voicing of coda consonants, prosodic context, syntactic structure, pragmatic meanings, etc. In these cases, variations of extrinsic vowel length do not affect the abstract phonemic status of vowels nor the identity of individual words, but can influence the phonetic manifestation of intrinsic vowel length.

Phonemic vowel length contrasts differ from vowel quality contrasts in the way the contrasts manifest. Phonemic vowel length exploits relative quantitative (durational) oppositions along the time dimension, rather than qualitative (spectral) oppositions at a certain point/period. To figure out phonemic length categories, learners need to listen to more words and abstract the length category from the phonological context. In this sense, phonemic vowel length is not as straightforward as vowel quality (e.g., vowel height/backness) to perceive.

However, acquisition of phonemic vowel length plays a crucial role in phonological development. This is not only because it can signal word meanings in some languages, but also because vowel length is closely related to syllable weight which determines lexical stress/pitch accent in many languages (Hayes, 1987, 1989, 1995).

Languages with phonemic vowel length contrasts differ in whether vowel duration alone can constitute a vowel contrast in the language. Some languages have minimal spectral quality difference involved in the phonemic vowel length contrasts, such as Japanese, Arabic, Finnish, Danish, Estonian and Kamba (Ladefoged, 1971; Lehiste, 1970), whereas other languages have prominent spectral quality difference in the vowel length pairs, such as Dutch, German, Swedish and Cantonese. Moreover, languages with phonemic vowel length contrasts differ in how systematic the contrasts manifest in the vowel system. In languages like Japanese, Arabic and Swedish, all vowels have short and long counterparts, while in languages like Australian English and Cantonese, the distinctions are restricted to a subset of vowels.

For those languages employing vowel length as a prominent contrastive feature in a vowel system, a sharp mean durational difference between short and long vowels can normally be observed, with approximately 1:2 short-to-long ratios (e.g., Japanese and Dutch). However, the distribution of short and long vowels may be uneven. In Japanese, it has been reported that long vowels constitute only 6% of the vowels in all syllable counts in an infant-directed speech (IDS) corpus (Bion, Miyazawa, Kikuchi, & Mazuka, 2013), and occur in 8-9% of all utterances in another Japanese IDS database (Mugitani, Pons, Fais, Dietrich, Werker, & Amano, 2009). Also, the duration of short vowels has been reported to be variable (Bion et al., 2013), which might possibly be due to interaction with contextual extrinsic vowel length. The overlaps between short and long vowels in distribution can also be observed in Dutch (Swingley, 2006).

Acquisition of Phonemic Vowel Length

Phonemic vowel length is often thought to be later acquired than vowel quality. The evidence is primarily from some of production studies in children, showing that the development of vowel length contrasts was protracted until children produced coda consonants and developed branching syllable nucleus (e.g., Demuth & Fee, 2005; Fikkert, 1994). Nevertheless, later acoustic analyses on Swedish (Buder & Stoel-Gammon, 2002), German (Kehoe & Lleó, 2003) and Japanese

(Ota, 1999, 2001) children's production indicate that children have already produced significant durational differences between target long and short vowels by the age of two or even earlier.

Another set of evidence suggesting late mastery of phonemic vowel length lies on phonetic discrimination studies in infants (Mugitani et al., 2009; Sato et al., 2010). It is reported that Japanese infants cannot discriminate vowel durational differences until 9.5 months, while they succeed with vowel spectral quality differences by 4 months (Sato et al., 2010). Further, Japanese 18-month-olds have been reported to have a transient asymmetric pattern in vowel length perception, discriminating only the long to short vowel change (Mugitani et al., 2009). This has been taken as a sign of phonemic awareness of the vowel length categories.

Studies directly testing infants' sensitivity to *phonemic* vowel length contrasts are mostly on languages that employs both vowel duration and spectral quality in the contrasts (Dietrich, Swingley, & Werker, 2007; van der Feest & Swingley, 2009). It has been shown that Dutch infants can detect changes of vowel duration in both directions in a habituation-switch task at 18 months (Dietrich et al., 2007). However, when using known words and the IPL paradigm, which is supposed to be an easier task, Dutch 21-month-olds can only identify the shortening mispronunciations of long vowels, revealing asymmetric perception (van der Feest & Swingley, 2009).

Motivation for Thesis

While the acquisition of vowel length in languages also involving prominent spectral quality difference has been studied extensively (Fikkert, 1994; Buder & Stoel-Gammon, 2002; Kehoe & Lleó, 2003; Dietrich et al., 2007; van der Feest & Swingley, 2009; among others), studies on the acquisition of phonemic vowel length in languages where vowel length is the only contrastive feature are limited – some have only looked at phonetic discrimination (Mugitani et al., 2009; Sato et al., 2010), but few have been done on phonemic perception, and even fewer on non-systematic vowel length contrasts. Therefore, this thesis aims to investigate perception of phonemic vowel length contrasts in languages exploiting vowel length alone in the contrasts.

Previous studies have shown that infants by 18 months can detect a 1-feature change in vowel quality mispronunciations (Mani et al., 2008). Phonemic vowel length contrasts also involve a 1-feature change, so we wondered whether infants can perceive a mispronunciation at the same age. We expected that Japanese infants, who are learning a language employing a systematic phonemic vowel length contrast, would be able to perceive a phonemic vowel length difference at 18 months. However, for a language like Australian English, which has a non-systematic phonemic vowel length contrast, we expected that the development of phonemic vowel length representations might be delayed, due to non-systematic input. The issue of native input also raised interesting questions for adults who have had exposure to various dialectal varieties with and without a phonemic vowel length contrast. We therefore explored this issue to determine if mono-dialectal and bi-dialectal adults would provide any evidence for different representations of vowel length.

Organization of Thesis

Chapters Two to Four of this thesis present three studies (in journal article format) that aim at providing an in-depth view of how infants and adults represent phonemic vowel length contrasts in their lexicon. Each article reviews relevant literature, outlines the employed methodology, presents and discusses the results, and draws conclusions from the research. Below is an outline of each article.

Chapter Two: Japanese infants are aware of phonemic vowel length in novel words at 18 months. Previous studies have shown that infants at 18 months become sensitive to a single feature vowel quality change in familiar words (e.g., Mani et al., 2008). This article therefore examines whether monolingual Japanese infants are sensitive to phonemic vowel length contrasts at the same age. Two groups of Japanese 18-month-olds were taught two novel words with either a short or long vowel, and then were tested with a mispronunciation detection task using the Looking-While-Listening paradigm (Fernald, Zangl, Portillo, & Marchman, 2008; Swingley, 2011). Infants who were taught with novel words containing a long vowel succeeded in learning the novel words and detecting shortening mispronunciations, whilst infants who were taught with novel words containing short vowels failed to learn the novel words. This may be due to the acoustic salience of long vowels, and the variability of short vowels in Japanese speech input. The findings provide evidence for a well-specified phonemic length category in long vowels in infants who are leaning a language where vowel length categories are relevant to word learning. The study hence provides a significant contribution to the current literature of phonological development, as it reveals an important developmental achievement in representations of phonemic vowel length by 18 months.

Chapter Three: Understanding the acquisition of phonemic vowel length contrasts in Australian English-learning 18- and 24-month-olds. This article explores two issues: First, whether the systematicity of a phonemic contrast will influence the age of emergence of its phonological representation in the language; and second, whether phonemic vowel length is specified in the lexicon later than other 1-feature vowel quality contrasts in phonological development. The article therefore targets Australian English, which has only a subset of vowels employing phonemic vowel length alone in vowel contrasts, and that allows comparison between phonemic vowel length and single feature vowel quality contrasts. In two studies, monolingual Australian English 18- and 24-month-olds were tested on their sensitivity to mispronunciations of familiar words using the IPL paradigm. The results show that, while the younger group did not seem to perform the task, 24-month-olds can successfully perform the task and display sensitivity to mispronunciations of vowel height, vowel backness, and more importantly, vowel length. The findings reveal that Australian English infants are aware of phonemic vowel length contrasts at least by 24 months. The results on infants' perception of different 1-feature vowel contrasts also suggest that phonemic vowel length is acquired earlier than it is often thought. The study provides an important complement to the first article (Chapter Two) for a more complete picture of early development of phonemic vowel length across different vowel systems with systematic vs. non-systematic phonemic vowel length contrasts. The study also demonstrates a well-controlled comparison of infants' sensitivity to phonemic vowel length and other 1-feature vowel quality contrasts for the first time, and therefore contributes to our understanding of the relationship between vowel length and vowel quality acquisition.

Chapter Four: Effect of early dialectal exposure on adult perception of phonemic vowel length. This article aims to understand whether and how early exposure to another dialect will influence the development of phonemic categories in the canonical dialect system, especially when the canonical dialect has a phonemic contrast that is non-contrastive in the other. The study therefore compares the perceptual sensitivity to mispronunciations of phonemic vowel length in Australian English adults with and without early exposure to another English dialect that does not have this contrast, using a similar IPL task as the studies in the second article (Chapter Three). The results revealed that bi-dialectal adults are more tolerant to mispronunciations of vowel length in their lexicon compared to mono-dialectal population, suggesting they have developed more flexible phonological categories of vowel length and may not specify phonemic vowel length feature in their lexicon. This study successfully uses the IPL paradigm to examine the perceptual sensitivity to mispronunciations in adult listeners, and demonstrates the flexibility of phonological categories in bi-dialectal adults. The results also suggest that a canonical phonemic feature is not always specified in the lexicon, especially when another phonological feature is exploited in the non-canonical dialect. Specification of phonological features is suggested to be highly determined by the acoustic saliency and specific characteristics of the features in the system. This study thus deepens our understanding of the complex effect of native input on the development of phonological representations and the generalization ability required in phonological acquisition, and provides implications for phonological development in bilingual populations.

Finally, **Chapter Five** summarizes the general findings of the three articles, discusses what the results reveal in terms of the development of phonemic vowel length contrasts, highlights the implications and significance of these findings, outlines the limitations of the studies, and suggests future directions for research.

References

Aslin, R. N., Pisoni, D. B., Hennesy, B. L., & Perey, A. J. (1981). Discrimination of voice onset time by human infants: New findings and implications for early experience. *Child Development*, *52*, 1135-1145.

Best, C. T. (1994). Learning to perceive the sound pattern of English. In C. Rovee-Collier & L. Lipsitt (Eds.), *Advances in Infancy Research* (Vol. 9, pp. 217-304). Hillsdale, NJ: Ablex Publishers.

Best, C. T., & 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. doi:10.1177/00238309030460020701

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. doi:10.1037/0096-1523.14.3.345

Best, C. T., Tyler, M. D., Gooding, T. N., Orlando, C. B., & Quann, C. A. (2009). Development of phonological constancy: Toddlers' perception of native- and Jamaican-accented words. *Psychological Science*, *20*(5), 539-542. doi:10.1111/j.1467-9280.2009.02327.x

Bion, R. A. H., Miyazawa, K., Kikuchi, H., & Mazuka, R. (2013). Learning phonemic vowel length from naturalistic recordings of Japanese infant-directed speech. *PLoS ONE*, *8*(2), e51594. doi:10.1371/journal.pone.0051594

Buder, E. H., & Stoel-Gammon, C. (2002). American and Swedish children's acquisition of vowel duration: Effects of vowel identity and final stop voicing. *Journal of the Acoustical Society of America*, *111*(4), 1854-1864.

Burnham, D. K. (1986). Developmental loss of speech perception: Exposure to an experience with a first language. *Applied Psycholinguistics*, *7*(3), 207-240. doi:10.1017/S0142716400007542

Burnham, D. K., Earnshaw, L. J., & Quinn, M. C. (1987). The development of categorical identification of speech. In B. E. McKenzie & R. H. Day (Eds.), *Perception in infancy: Problems and issues* (pp. 237-275). Hillsdale, NJ: Lawrence Erlbaum.

Curtin, S., Fennel, C., & Escudero, P. (2007). Infants' recognition of vowel contrasts in a word learning task. In H. Caunt-Nulton, S. Kulatilake, & I. Woo (Eds.), *BUCLD 31: Proceedings of the 31st annual Boston University Conference on Language Development* (Vol. 1, pp. 141-152). Somerville, MA: Cascadilla Press.

Curtin, S., Fennel, C., & Escudero, P. (2009). Weighting of vowel cues explains patterns of word-object associative learning. *Developmental Science*, *12*, 725-731.

Cutler, A., & Mehler, J. (1993). The periodicity bias. *Journal of Phonetics*, *21*, 103-108.

Demuth, K., & Fee, E. J. (1995). *Minimal prosodic words in early phonological development*.

Unpublished manuscript, Brown University and Dalhousie University.

Dietrich, C., Swingley, D., & Werker, J. (2007). Native language governs interpretation of salient speech sound differences at 18 months. *Proceedings of the National Academy of Sciences*, *104*(41), 16027-16031. doi:10.1073/pnas.0705270104

Escudero, P., Best, C. T., Kitamura, C., & Mulak, K. E. (2014). Magnitude of phonetic distinction predicts success at early word learning in native and non-native accents. *Frontiers in Psychology*, *5*, 1-11. doi:10.3389/fpsyg.2014.01059

Fennell, C. T., & Waxman, S. R. (2010). What paradox? Referential cues allow for infant use of phonetic detail in word learning. *Child Development*, *81*(5), 1376-1383. doi:10.1111/j.1467-8624.2010.01479.x

Fennell, C. T., & Werker, J. F. (2003). Early word learners' ability to access phonetic detail in well-known words. *Language and Speech*, *46*(2-3), 245-264. doi:10.1177/00238309030460020901

Fernald, A., Zangl, R., Portillo, A.L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In Sekerina, I. A., Fernandez, E.M., & Clahsen, H. (Eds.), *Developmental psycholinguistics: On-line methods in children's language processing* (pp. 97-135). Amsterdam: John Benjamins.

Fikkert, P. (1994). *On the acquisition of prosodic structure* (Doctoral dissertation, HIL dissertations 6, Leiden University). The Hague: Holland Academic Graphics.

Golinkoff, R., Hirsh-Pasek, K., Cauley, K., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, *14*, 23- 45.

Hayes, B. (1987). A revised parametric metrical theory. In J. McDonough and B. Plunkett (Eds.), *Proceedings of the North East Linguistic Society (Vol. 17)*. Ameherst, MA: Graduate

Linguistic Student Association.

Hayes, B. (1989). Compensatory lengthening in moraic phonology. *Linguistic Inquiry*, 20, 253-306.

Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. Chicago: University of Chicago Press.

Kehoe, M. M., & Lleó, C. (2003). The acquisition of nuclei: A longitudinal analysis of phonological vowel length in three German-speaking children. *Journal of Child Language*, *30*, 527-556.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience, 5,* 831–843.

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science, 9,* F13-F21.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, 255, 606-608.

Ladefoged, P. (1971). *Preliminaries to linguistic phonetics.* Chicago: University of Chicago Press.

Lehiste, I. (1970). *Suprasegmentals.* Cambridge, Mass.: MIT Press.

Mani, N., Coleman, J., & Plunkett, K. (2008). Phonological specificity of vowel contrasts at 18-months. *Language and Speech*, *51*(1&2), 3-21.

Mani, N., & Plunkett, K. (2007). Phonological specificity of vowels and consonants in early lexical representations. *Journal of Memory and Language*, *57*(2), 252-72.

Mani, N., & Plunkett, K. (2008). 14-month-olds pay attention to vowels in novel words.

Developmental Science, *11*(1), 53–59.

Mani, N., & Plunkett, K. (2011). Does size matter? Subsegmental cues to vowel mispronunciation detection. *Journal of Child Language*, *38*, 606-627.

Mugitani, R., Pons, F., Fais, L., Dietrich, C., Werker, J. F., & Amano, S. (2009). Perception of vowel length by Japanese- and English-learning infants. *Developmental Psychology*, *45*(1), 236-247.

Mulak, K. E., Best, C. T., Tyler, M. D., Kitamura, C., & Irwin, J. R. (2013). Development of phonological constancy: 19-month-olds, but not 15-month-olds, identify words in a non-native regional accent. *Child Development*, *84*(6), 2064-2078. doi:10.1111/cdev.12087

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. doi:10.1111/j.1467-7687.2009.00898.x

Ota, M. (1999). *Phonological theory and the acquisition of prosodic structure: Evidence from child Japanese* (Unpublished doctoral dissertation). Georgetown University, Washington D. C.

Ota, M. (2001). Phonological theory and the development of prosodic structure: Evidence from child Japanese. *Annual Review of Language Acquisition*, *1*, 65-118.

Pater, J., Stager, C., & Werker, J. F. (1998). Additive effects of phonetic distinctions in word learning. *Proceedings of the 16th International Congress on Acoustics and 135th Meeting of the Acoustical Society of America*, 2049-2050.

Pater, J., Stager, C., & Werker, J. F. (2004). The lexical acquisition of phonological contrasts. *Language*, *80*(3), 361-379.

Polka, L., & Bohn, O.-S. (1996). A cross-language comparison of vowel perception in English-learning and German-learning infants. *Journal of the Acoustical Society of America*, *100*(1), 577-592. doi:10.1121/1.415884

Polka, L., & Bohn, O.-S. (2003). Asymmetries in vowel perception. *Speech Communication*, *41*(1), 221-231. doi:10.1016/S0167-6393(02)00105-X

Polka, L., & Bohn, O.-S. (2011). Natural Referent Vowel (NRV) framework: An emerging view of early phonetic development. *Journal of Phonetics*, *39*(4), 467-478. doi: 10.1016/j.wocn.2010.08.007

Polka, L., Colantonio, C., & Sundara, M. (2001). A cross-language comparison of /d/–/ð/ perception: Evidence for a new developmental pattern. *Journal of the Acoustical Society of America, 109,* 2190–2201.

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.

Pons, F., Albareda-Castellot, B., & Sebastián-Gallés, N. (2012). The interplay between input and initial biases: Asymmetries in vowel perception during the first year of life. *Child Development*, *83*(3), 965-976. doi:10.1111/j.1467-8624.2012.01740.x

Sato, Y., Sogabe, Y., & Mazuka, R. (2010). Discrimination of phonemic vowel length by Japanese infants. *Developmental Psychology*, *46*(1), 106-199. doi:10.1037/a0016718

Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, *388*(6640), 381-382.

Sundara, M., Polka, L., & Genesee, F. (2006). Language-experience facilitates discrimination of /d–ð/ in monolingual and bilingual acquisition of English. *Cognition, 100,* 369–388.

Swingley, D. (2003). Phonetic detail in the developing lexicon. *Language and Speech*, *46*(2-3), 265-294. doi:10.1177/00238309030460021001

Swingley, D. (2006). *Data-driven phonological distinction without phonetic opposition*. Paper presented at the International Conference on Infant Studies (XV ICIS), June 2006, Kyoto, Japan.

Swingley, D. (2009). Onsets and codas in 1.5-year-olds' word recognition. *Journal of Memory and Language*, *60*(2), 252-269. doi:10.1016/j.jml.2008.11.003

Swingley, D. (2011). The looking-while-listening procedure. In E. Hoff (Ed.), *Research methods in child language: A practical guide* (pp. 29-42). Malden, MA: Wiley-Blackwell.

Swingley, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, *76*, 147-166.

Swingley, D., & Aslin, R. N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, *13*(5), 480-484. doi: 10.1111/1467-9280.00485

Swingley, D., & Aslin, R. N. (2007). Lexical competition in young children's word learning. *Cognitive Psychology*, *54*, 99-132. doi:10.1016/j.cogpsych.2006.05.001

Tsao, F. M., Liu, H. M., & Kuhl, P. K. (2006). Perception of native and non-native affricate-fricative contrasts: Cross-language tests on adults and infants. *Journal of the Acoustical Society of America, 120,* 2285-2294.

van der Feest, S. V. H., & Swingley, D. (2009). Language-specific interpretation of vowel duration in 21-month-olds' word recognition: A cross-linguistic study of phonetic attribution. Paper presented at the 34th Boston University Conference on Language Development, November, Boston, MA.

van Heugten, M., & Johnson, E. K. (2014). Learning to contend with accents in infancy: Benefits of brief speaker exposure. *Journal of Experimental Psychology: General*, *143*(1), 340-350. doi:10.1037/a0032192

Watson, C. I., & Harrington, J. (1999). Acoustic evidence for dynamic formant trajectories in

Australian English vowels. *Journal of the Acoustical Society of America*, *106*(1), 458-468. doi:10.1121/1.427069

Werker, J. F., & Fennell, C. T. (2004). Listening to sounds versus listening to words: Early steps in word learning. In D. G. Hall & S. R. Waxman (Eds.), *Weaving a Lexicon* (pp. 79-109). Cambridge, M.A.: MIT Press.

Werker, J. F., Fennell, C. T., Corcoran, K. M., & Stager, C. L. (2002). Infants' ability to learn phonetically similar words: Effects of age and vocabulary size. *Infancy*, *3*, 1-30.

Werker, J. F, Gilbert, J. H., Humphrey, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child Development*, *52*, 349-355.

Werker, J. F., & Lalonde, C. E. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology*, *24*, 672–683.

Werker, J. F., & Polka, L. (1993). Developmental changes in speech perception: New challenges and new directions. *Journal of Phonetics*, *21*(1-2), 83-101.

Werker, J. F., & Tees, R. C. (1983). Developmental changes across childhood in the perception of non-native speech sounds. *Canadian Journal of Psychology*, *37*(2), 178-286.

Werker, J. F., & Tees, R. C. (1984a). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *7*, 49-63.

Werker, J. F., & Tees, R. C. (1984b). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America*, *75*, 1866-1878.

White, K. S., & Aslin, R. N. (2011). Adaptation to novel accents by toddlers. *Developmental Science*, *14*(2), 372-384. doi:10.1111/j.1467-7687.2010.00986.x

White, K. S., & Morgan, J. L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language*, *59*, 114-132.

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. doi:10.1016/j.jml.2012.09.004

Yoshida, K. A., Fennell, C. T., Swingley, D., & Werker, J. F. (2009). Fourteen-month-old infants learn similar-sounding words. *Developmental Science*, *12*(3), 412-418. doi:10.1111/j.1467-7687.2008.00789.x

Chapter Two: Japanese Infants are Aware of Phonemic Vowel Length in Novel Words at 18 Months

This chapter is based on the following paper under review for publication:

Chen, H., Yamane, N., Xu Rattanasone, N., Demuth, K., & Mazuka, R. (under review). Japanese infants are aware of phonemic vowel length in novel words at 18 months. *Infancy*.

All components of this paper, both experimental and written, have been completed by me, with advice from the co-authors (my supervisors and collaborators) when needed.

Abstract

Although much research has examined the acquisition of vowel quality contrasts, and vowels that combine quality and length distinctions, there has been little research directly examining the early awareness of phonemic vowel length during word learning. In this study, two groups of Japanese 18-month-olds were taught novel words that contained either long or short vowels, and were then tested on their sensitivity to mispronunciations in vowel length during a Looking-While-Listening task. Infants who were taught novel words with a long vowel successfully learnt the novel words and identified vowel length mispronunciation of the words, whereas infants who were taught novel words with a short vowel failed to learn the novel words. This is possibly due to the fact that Japanese short vowels have a wider range of durations than long vowels. The results suggest that Japanese infants have developed well-specified long vowels in their lexicon, showing an emerging awareness of phonemic vowel length at 18 months. The implications for learning phonemic vowel length contrasts in other languages are discussed.

Key words: phonemic vowel length, perception, Japanese 18-month-olds, novel words, long vowels

Japanese Infants are Aware of Phonemic Vowel Length in Novel Words at 18 Months

Introduction

Phonemic vowel length contrasts employ quantitative (durational) oppositions along the time dimension, differing from vowel quality contrasts which use qualitative (spectral) oppositions. Many languages employ vowel length as a contrastive feature, such as Japanese, Finnish, Danish, Arabic, Estonian, Czech, Slovak, Kamba, amongst others, where a contrast in vowel duration alone can be used to distinguish word meanings (Ladefoged, 1971; Lehiste, 1970). For instance, in Japanese, the words /seki/ 'seat' and /seːki/ 'century' are differentiated only by the duration of the vowel /e/ in the first syllable. These languages therefore exhibit typical *phonemic vowel length* contrasts without prominent spectral difference, and are often referred to as *quantity* languages. In languages such as Dutch, Swedish, German, Cantonese, vowel length combines with vowel *quality* to signal different word meanings. Although a number of studies have investigated infants' phonological awareness of vowel *quality* contrasts during early word learning (e.g., Curtin, Fennel, & Escudero, 2007, 2009; Mani, Coleman, & Plunkett, 2008; Mani & Plunkett, 2007, 2008, 2011; Nazzi, 2005; Swingley & Aslin, 2000, 2002), and several have examined vowel length contrasts in the languages where vowel length *co-occurs* with prominent vowel quality differences (Benders, 2013; Benders & Mandell, 2011; Dietrich, Swingley, & Werker, 2007; van der Feest & Swingley, 2009), there has been little research directly examining children's early sensitivity only to *phonemic vowel length* contrasts in a quantity language. Whether the acquisition of *phonemic vowel length* reveals a different developmental process compared to that of vowel *quality* contrasts is therefore of great theoretical interest, especially regarding how and when these different phonological representations develop.

A number of early word recognition studies suggest that English infants tune into vowel *quality* contrasts in their lexical representations at as early as 14 months. Swingley and Aslin (2002) tested American English-learning 14-month-olds with mispronunciations in either consonants or vowels in familiar words using the Intermodal Preferential Looking (IPL) paradigm (Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987). They found infants took significantly longer to switch from the distracter to the target object when the target label was mispronounced than when it was correctly pronounced. This effect was significant by item, suggesting that infants showed no difference between their sensitivity to consonant and vowel *quality* mispronunciations. Mani and Plunkett (2008) further showed that both 14-month-old and 18-month-old British English infants can detect the mispronunciation of vowel *quality* in newly learnt novel words in an IPL task (i.e., *padge* [pæʤ] mispronounced as *poudge* [puːʤ], and *mot* [mɔt] as *mit* [mɪt]). However, these studies contained multiple feature changes that combined both quality and quantity.

More recent studies have found that British English-learning infants can recognize a single-feature change in vowel *quality* when a familiar word is mispronounced in vowel height or vowel backness (but not in vowel roundedness) at 18 months (Mani et al., 2008; Mani & Plunkett, 2007, 2011), whereas at 15 months infants failed to do so (Mani & Plunkett, 2007). We also know that, at 19 months, American English-learning infants can identify a single-feature change in *consonants*, i.e., voicing, place and manner of articulation, in familiar words (White & Morgan, 2008). This leads to the question of whether infants around this age can perceive a change in vowel *length*, also a single feature change, in a word recognition task.

Previous studies have revealed that Dutch-acquiring 21-month-olds can detect the shortening of long vowels when familiar words are mispronounced in an IPL task, but not the lengthening of short vowels (van der Feest & Swingley, 2009), and the same pattern has been reported in Dutch adults (Nooteboom & Doodeman, 1980; van der Feest & Swingley, 2011). However, long vs. short vowels in Dutch differ not only in duration, but also in spectral quality

(Adank, van Hout, & Smits, 2004; Escudero, Benders, & Lipski, 2009). Thus, listeners might misperceive a shortened long vowel since the spectral information is reduced, but still be able to correctly identify a lengthened short vowel, since the spectral information is maintained. Furthermore, younger Dutch infants do not show this asymmetric pattern at 18 months – they can detect a switch in vowel duration in both directions after habituating with both the long and short versions of a native short vowel /ɑ/ or of a nonnative vowel /æ/ in a word learning task (Dietrich et al., 2007). However, it is unclear whether these conflicting results reveal a developing sensitivity to vowel duration contrasts in languages such as Dutch, or if these different results are due to the different types of tasks employed. The time is therefore right to explore these issues more thoroughly in a *quantity* language where vowel length alone can signal word meanings.

Japanese is one of the quantity languages which has contrasts in *phonemic* vowel *length* with minimal spectral quality differences (Vance, 1987). Previous acoustic studies have demonstrated that the formant differences between Japanese long and short vowels are very small, especially when at a slow speech rate (e.g., Hirata & Tsukada, 2009), and this is also evidenced in infant-directed speech (IDS) in both read and spontaneous speech modes (Werker, Pons, Dietrich, Kajikawa, Fais, & Amano, 2007). Vowel duration alone can therefore be used to distinguish word meanings, and the long vs. short contrast is systematically represented across the entire vowel (and consonant) system. The long to short durational ratio of vowels, on average, approximates or exceeds 2 to 1 (Han, 1962; Hirata, 2004), and this is reflected in IDS as well, for all five vowels (Bion, Miyazawa, Kikuchi, & Mazuka, 2013; Werker et al., 2007). We would therefore expect Japanese infants to exhibit an early mastery of *phonemic vowel length* contrasts given the systematic and significant durational differences between long and short vowel pairs in the input they hear. However, it is also important to note that the Japanese long and short vowels are distributed unevenly in the input infants hear. In a study of IDS from the Japanese Mother-Infant

Conversation Corpus, long vowels constituted only 6% of the vowels in all syllable counts (Bion et al., 2013). A similar distributional pattern has been found in utterance counts in two mother-infant interaction speech samples from another Japanese IDS database (Amano, Nakatani, & Kondo, 2006), where only 8-9% of the utterances contained long vowels (Mugitani, Pons, Fais, Dietrich, Werker, & Amano, 2009). The consistent results from these two IDS analyses suggest that Japanese infants are exposed to a limited number of long vowels compared to short vowels in their native language. This might then present more of a challenge for Japanese infants' learning this *phonemic vowel length* contrast than otherwise thought.

Previous behavioural research has shown that Japanese 10-month-old infants can discriminate durational difference in word-medial long/short vowels (as in /taːku/ vs. /taku/ and in /maːna/ vs. /mana/) (Mugitani et al., 2009; Sato, Sogabe, & Mazuka, 2010), whereas the 4-month-olds cannot, and the 7.5- to 8.5-month-olds might have an emerging phonetic awareness of the difference (Sato et al., 2010). On the other hand, neuropsychological evidence from an NIRS study has revealed that Japanese infants display auditory sensitivity to both across- and within-category in word-final vowel durational differences (as in /manaː/ vs. /mana/) as early as 3-4 months, but show greater cerebral response to across-category difference at 6-7 months at 13-14 months (Minagawa-Kawai, Mori, Naoi, & Kojima, 2007). Critically, though, it is only at 13-14 months that the specificity to the across-category contrast becomes consistently left hemisphere lateralized (as found in Japanese adults (Minagawa-Kawai, Mori, Furuya, Hayashi, & Sato, 2002)), suggesting that it is only in this latter stage that the long/short vowel contrast is processed as linguistically-relevant information in the brain. These findings provide important insight into infants' ability in the perception of phonemic vowel length contrasts in the second year of life. However, there is a lack of behavioural data indicating when Japanese infants understand the *phonemic function* of vowel duration, and be able to use vowel length contrasts to acquire or access lexical form. There has been some production evidence showing that Japanese infants can produce significant durational difference between target long and short vowels at 17-19 months, though the long-to-short durational ratios were not yet adult-like (Ota, 1999). Furthermore, the naturalistic production data comes from a limited number of participants (n = 3) and only a few tokens of familiar words (n = 8-12). It is therefore necessary to directly test whether Japanese infants can use the vowel duration differences contrastively to identify a change in word meaning.

Interestingly, an asymmetric directionality effect has been documented in vowel duration discrimination in Japanese infants at 18 months (Mugitani et al., 2009). These infants only discriminated a change from long to short vowels, but not *vice versa*. This directionality effect cannot be found in younger infants in similar tasks (Mugitani et al., 2009; Sato et al., 2010), which has been interpreted as an evidence of the influence of the native phonology. However, in the NIRS study, infants habituated to only short vowels, and infants at both 13-14 months and 25-28 months demonstrated neural responses to across-category contrasts, and more crucially the responses were left lateralized (Minagawa-Kawai et al., 2007). This suggests that Japanese infants do detect changes from short to long vowels when processing the differences linguistically. This raises the question of whether the directionality effect would be found in a word learning/word recognition task.

The present study therefore investigated whether Japanese infants show sensitivity to changes in *phonemic vowel length* in a word recognition task at 18 months, and whether there might be any directionality effects. To avoid possible lexical effects which might influence performance, we taught infants two novel words via recorded videos (adapted from Dautriche, Swingley, & Christophe, 2015), and tested their perception of vowel length mispronunciations in the newly taught words using the Looking-while-Listening paradigm (Fernald, Zangl, Portillo, & Marchman, 2008; Swingley, 2011). This simulated a natural word learning environment. In contrast to the task designed by Dautriche et al. (2015), our task not only challenged the infants to learn the novel words, but also to detect the mispronunciations of these newly learnt words. We predicted that, despite the small number of long vowels in the input, Japanese infants would be aware of *phonemic vowel length* contrasts by the age of 18 months, given the mora-timed prosody of Japanese as well as the systematic and prominent long vs. short contrasts found throughout the entire sound system for both vowels and consonants. Ota's (1999, 2001) longitudinal study of infants' spontaneous productions has shown that Japanese infants have largely understood the moraic structure of the language by 16-18 months. This suggests that the emergence of *phonemic vowel length* contrasts may also occur around this age.

Method

Participants

Two groups of full-term monolingual Japanese-acquiring 18-month-olds were tested with their caretakers in Tokyo, Japan. The Long Vowel group consisted of 17 infants (10 boys and 7 girls) with a mean age of 18.17 (months.days, range: 18.00 – 18.29), and the Short Vowel group included 19 infants (11 boys and 8 girls) with a mean age of 18.13 (range: 18.01-18.29). Data from an additional 20 infants were excluded due to fussiness or disinterest resulting in more than 50% of trials missing (n = 9), experimental error (n = 3), sibling influence (n = 1), and not meeting the analysis criteria (n = 6, see **Measurement and Analysis**). Caretakers were asked to complete the Japanese Adaptation of the MacArthur Communicative Development Inventory (JCDI, Ogura & Watamaki, 2004). There was no significant difference in age or expressive or receptive vocabulary size between the two groups (see Table 1).

Table 1. *Mean, Standard Deviation, and Range of Age, Receptive Vocabulary and Expressive Vocabulary for Each Group of Participants*

		Receptive Vocabulary		Expressive Vocabulary	
Group	Age in day	Score	Percentile	Score	Percentile
Long Vowel	564.4	205.4	49.4	69.2	58.2
	SD: 10.7	*SD*: 109.4	*SD*: 25.9	*SD*: 61.6	*SD*: 29.8
	(547-576)	(26-371)	(5-85)	(5-234)	(5-95)
Short Vowel	560.1	205.1	49.2	49.3	45.5
	SD: 9.5	*SD*: 99.8	*SD*: 23.4	*SD*: 52.1	*SD*: 31.4
	(548-576)	(48-352)	(10-80)	(0-198)	(5-95)

Note. No significant difference of Age (p = .214), Receptive Vocabulary Score (p = .994) or Percentile (p = .981), or Expressive Vocabulary Score (p = .301) or Percentile (p = .222) between two groups.

Stimuli and Design

In the experiment we implemented a between-subject design where the Long Vowel Group were taught the long vowel words and tested on mispronunciations with short vowels, and *vice versa* for the Short Vowel Group. Both groups were first taught the two novel words with training videos, and then tested on their sensitivity to mispronunciations in the Looking-While-Listening paradigm.

To test infants' sensitivity to vowel length contrasts, two sets of disyllabic novel words were used: the long vowel pair *baato* /baːto/ and *giite* /giːte/, as well as the contrastive short vowel pair *bato* /bato/ and *gite* /gite/ (target vowels underlined). Each was paired with a novel toy (see

Figure 1). Each group of infants was trained with one pair of the novel words as the correct pronunciations (CP) and tested with the contrastive pair as mispronunciations in vowel length (ML). Three points need to be made about the target novel words. First, they were disyllabic, as the 2-syllable word has been found to be the most frequent word type in infant-directed vocabulary (IDV) in Japanese (around 63%) (Mazuka, Kondo, & Hayashi, 2008). Second, target vowels were always on the first syllable of the words (i.e., word-medial position). This allowed us to test infants' sensitivity to phonemic vowel length contrasts independent of phrase-final lengthening. Third, the target words were always lexically pitch-accented on the first syllable and unaccented on the second. Japanese pitch accent (H*L) is realized primarily by fundamental frequency (f0). The f0 contour of a H*L consists of a high f0 on the accented syllable followed by a steep fall to low f0. This fall takes place during the first syllable when the vowel is long, but between the first and second syllable when the vowel is short. It has been reported that the effect of pitch accent on vowel duration appears to be small and not significant (Beckman, 1986; Nazzi, Floccia, & Bertoncini, 1998).

Two sets of control items were used in the test as well: the Novel Word (NW) and the Known Word (KW) pairs. The novel word was used to provide a looking control for words that are mispronounced in both consonants and vowels from the two taught words, *zake* /dzake/ for the Long Vowel Group and *zaake* /dzaːke/ for the Short Vowel Group. Three pairs of known words were used to maintain infants' interest in the test and provide a looking control for learnt real words, with matched animacy in each pair: *inu* (dog) - *neko* (cat), *booru* (ball) - *pan* (bread), and *ringo* (apple) - *banana* (banana). (See Table 2 for a full set of test items in each group).

Table 2. *Experimental Items Used in the Long Vowel Group and the Short Vowel Group, Including Test Items in Correct Pronunciation (CP) and Mispronunciation in Vowel Length (ML) Conditions, and Control Items in Novel Word (NW) and Known Word (KW) Conditions*

Group	Test Items		Control Items	
	Correct Pronunciation (CP)	Mispronunciation (ML)	Novel Word (NW)	Known Word (KW)
Long Vowel	*baato*	*bato*	*zake*	*inu, neko*
	giite	*gite*		*booru, pan*
				ringo, banana
Short Vowel	*bato*	*baato*	*zaake*	*inu, neko*
	gite	*giite*		*booru, pan*
				ringo, banana

Training videos. The training videos were recorded by a 26-year-old female native speaker of Tokyo Japanese in infant-directed speech in an acoustically shielded recording booth. The videos were recorded with a digital camera as 1920 x 1080 AVCHD videos and transferred into WMV format with software Movie Maker.

The training videos were made to closely replicate the method used in Dautriche et al. (2015), which successfully demonstrated the teaching of novel words to French 18-month-olds. Two novel objects were used to teach the two novel word labels and each novel word was demonstrated in two different training videos, approximately 30 seconds each. In each training video, the speaker described and played with one of the novel objects, labeling it with a novel word five times to help infants make the object and label associations. The associations of the novel objects with the novel labels were counterbalanced across infants. An additional introductory video was also made and shown to infants before other videos, where a toy car was labeled and described to help infants understand the purpose of the training videos. (See Figure 1 for the two novel objects and Appendix I for sample training scripts).

Figure 1. Two novel objects used in the training videos and test trials.

Auditory stimuli. The auditory test stimuli were recorded by the same Japanese speaker for test trials, sampled at 44.1 kHz. Each test word (illustrated with *baato* below) was recorded three times in the following carrier sentences[1]:

	Part (1)	Part (2)	Part (3)
(a)	*"mite mite!*	*[baato]!*	*[baato] dane!"*
	("*Look, Look!*	*[baato]!*	*It's [baato]!*)
(b)	*"hora mite!*	*[baato]!*	*[baato] dayo!"*
	("*Look again!*	*[baato]!*	*It's [baato]!*)

(c) *"mite mite!* *[baato]!* *[baato] wa dotchi?"*

("*Look, Look!* *[baato]!* *Which is [baato]?*)

The best token of each sentence was chosen and spliced into three parts: (1) the pre-target utterance "*mite mite!"* or "*hora mite!*"; (2) the test word in isolation "*[baato]!*"; and (3) the repetition of the test word in an offset utterance "*[baato] dane!*", "*[baato] dayo!*" or "*[baato] wa dotchi?*". Then the best token for each test word in (2) was selected and used consistently throughout the test. The best carriers were also selected for (a), (b), and (c), which were alternated throughout the test to maintain infants' interest and attention. The three parts were then scripted into the E-Prime 2.0 program (Schneider, Eschman, & Zuccolotto, 2012) in order to form a test sentence and aligned the onsets of each part across all trials.

The details of the target vowel durations produced in the training videos and used in the test trials are shown in Table 3. The mean duration of the long vowels was 138 ms (*SD* = 23, range = 93-180) in the training videos for the Long Vowel Group, while the mean duration of the short vowels was 64 ms (*SD* = 19, range = 32-91) in the training for the Short Vowel Group. The long vowels were 183-197 ms and the short vowels were 53-78 ms for the first mention in the test trials for both groups, which maintained a constancy of durational cues for infants. Consistent with the literature, formant values between the long and short vowels in the test trials were very similar, with less range for the short vowel: long vowel /aː/ - F1 (*M* = 884 Hz, range = 873-919 Hz), F2 (*M* = 1716 Hz, range = 1689-1739 Hz) vs. short vowel /a/ -F1 (*M* = 873 Hz, range = 853-895 Hz), F2 (*M* = 1692 Hz, range = 1666-1720 Hz), and long vowel /iː/ - F1 (*M* = 360 Hz, range = 342-383 Hz), F2 (*M* = 2723 Hz, range = 2510-3086 Hz) vs. short vowel /i/ - F1 (*M* = 325 Hz, range = 322-329 Hz), F2 (*M* = 2834 Hz, range = 2657-2915 Hz).

Table 3. *Mean Duration, Standard Deviation, and Range of the Target Vowels in Training and Test Phases (in ms)*

	Training Phase		Test Phase			
			CP		ML	
	Target Word		1st mention	2nd mention	1st mention	2nd mention
Group		(*SD*, range)		(range)		(range)
Long	*baato*	156	197	167	78	71
Vowel		(15, 128-180)		(156-176)		(66-80)
	giite	120	183	132	53	42
		(14, 93-142)		(123-139)		(37-51)
Short	*bato*	80	78	71	197	167
Vowel		(8, 66-91)		(66-80)		(156-176)
	gite	48	53	42	183	132
		(10, 32-71)		(37-51)		(123-139)

Note. CP = Correct Pronunciation; ML = Mispronunciation in Vowel Length. All acoustic measurements were performed using Praat (Boersma & Weenink, 2012).

Visual stimuli. In the test phrase, photographs of the two novel objects were shown on a light-gray background as a yoked-pair, aligned horizontally and of similar size. Known objects were also presented as photographs and were also in yoked-pairs (e.g., a dog always appeared with a cat).

Procedure

The experiment was conducted in a sound-attenuated laboratory room. The infant sat on the caretaker's lap while looking at a large polycarbonate rear-projection screen located one meter in front of him/her. The two stimulus pictures were projected at a size of 50 x 35 cm on the screen, at a distance of 60 cm from each other. A black-and-white video camera with a zoomable lens was positioned below the screen and focused on the face of the infant. The speech stimuli were delivered at a conversation level (mean ≈ 65 dBA) from two loudspeakers located behind the screen on both sides. During the entire session, the caretaker wore a SENNHEISER NoiseGard HMEC 322 headset playing masking music to ensure that they could not hear the stimuli and influence the infant.

The experiment was conducted using the E-Prime 2.0 software (Schneider et al., 2012) which controlled the presentation of the stimuli and test trials. A mixer integrated the video signal from the video recorder with the graphic information about the test events from E-Prime onto a VCR for later coding. The experimenter (the first author) controlled the process of the experiment and started each video and experimental trial once the infant looked at the center of the screen.

The experiment was composed of four phases: training phase 1 – test phase 1 – training phase 2 – test phase 2. In the first training phase, infants first saw the introductory video, then they were taught the two novel words by viewing four training videos (two for each novel word). The order of the presentation of the two novel words alternated in the training videos (i.e., *baato-giite-baato-giite* or *giite-baato-giite-baato* for the Long Vowel Group, and the same for the short vowel words for the Short Vowel Group) and was counterbalanced across infants (Dautriche et al., 2015). After the first test phase, the second training phase began, where the infants saw the second training video of each novel word again to ensure they did not forget the taught labels mid-way through the experiment.

Infants attended to 20 trials in four blocks in the two test phases, including two orientation trials, two CP and two ML for each taught word, four NW, and six KW. In the two orientation trials, infants saw the two taught novel objects shown on the screen and heard each of the novel labels once in each trial. This helped infants to orient their looks to the left and the right side of the screen, and become familiar with the test procedure. In the following test trials, the ML and NW trials were blocked, so infants would hear either mispronunciations in vowel length or a control novel pronunciation but not both in a block. The presentation order of the blocks alternated between the ML and NW conditions and was counterbalanced across infants. The six KW trials were dispersed into the two test phases in random order, with two in the first and four in the second. They were used to separate the test trials and maintain infant's interest on the task. The order of the test trials was randomized in each block. Target and distractor pictures appeared on the left and the right side of the screen equally, and the target side did not repeat more than twice on consecutive trials.

In each test trial, infants saw either the two taught novel objects or two known objects presented side-by-side on the screen. After 2 seconds, the auditory stimulus started: *"mite mite! [XX]! [XX] dane!"* ("*Look, Look! [XX]! It's [XX]!*"). The trial ended 4 seconds after the onset of the first mention of the test word, with the entire trial lasting for 7.5 seconds. The entire experiment lasted about 6.5 minutes (see Figure 2 for details of the trial procedure).

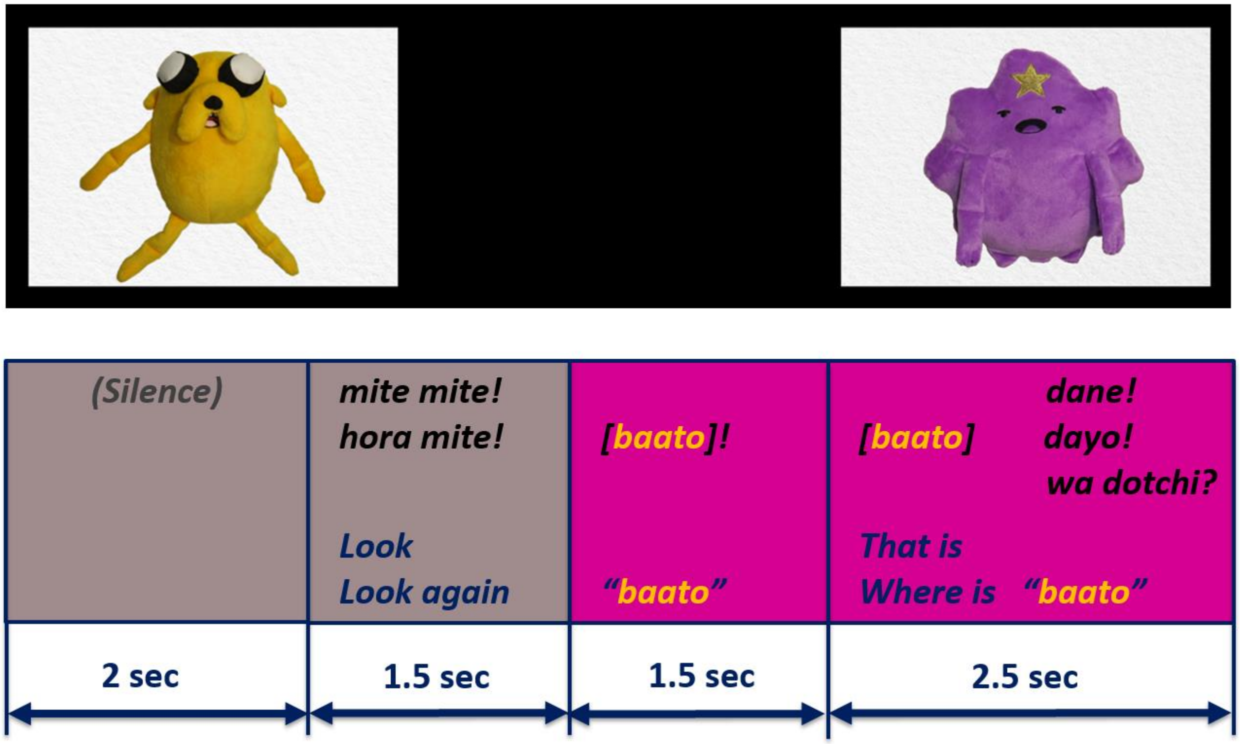


Figure 2. Test trial procedure (illustrated with *baato*).

Coding

Videotapes of the experiments were digitized into MOV files with software FinalCut Pro on a Macintosh computer and then coded using ELAN (Brugman & Russell, 2004). Each video had a frame rate of 30 frames per second. The principle coder (the first author) blind-coded (without audio input) all the MOV films frame by frame, by coding whether the infant was looking at the left picture, the right picture, the center of the screening, shifted between the left and right pictures, or away. An experienced second coder coded the data from four randomly selected infants (11% of the 36 infants). Data from the onset of the first mention of the test word until the end of trial (4 s) were extracted from these infants and compared with the coding of the five types of looking behaviours. Pearson's r between the two coders averaged .984, $p < .001$, indicating a high level of consistency between the two coders. The codes of right and left were then transformed to target or distractor according to the correct answers in the actual test for each participant[2].

Measurement and Analysis

As reported in the **Participants** section, data from an additional 6 participants were not included in the final analysis. Inclusion into the final analysis was determined by measuring infants' reaction time in looking shifts from distractor to the target picture in the KW condition. All infants who were more than 2 standard deviations slower than the mean (n = 2), or who could not provide such shifting data on reaction time in the KW condition (n = 4), were then excluded from further analysis. This ensured that we included only those infants who demonstrated good understanding of the task.

We then conducted a cluster-based permutation analysis on infants' looking data from the onset of the first mention of the test word until the end of the trial to find a time-window where we observed a significant increase in looks towards the target object. This analysis was originally developed for EEG data (Maris & Oostenveld, 2007), but has also been used for eye-tracking studies with infants (e.g., Dautriche et al., 2015; de Carvalho, Dautriche, & Christophe, 2016), providing an on-line assessment of looking patterns as they evolve over time (see Dautriche et al. (2015) for further justification).

The fixation proportion of looks towards the target object for each frame in each participant was calculated by dividing the looks at the target picture by the total looks at both the target and the distractor pictures. All fixation proportions of looks towards the target object were then transformed using arcsine-square-root transformation to better fit the assumptions of the *t*-tests for test against chance. At each time point we conducted a one-tailed *t*-test on fixations to the target compared to 0.5 chance level for each test condition, i.e., CP, ML, NW, and KW. To test whether there was a significant difference between conditions of CP and ML, we conducted an additional cluster-based permutation analysis in which clusters were formed on the basis of paired two-tailed *t*-tests comparing the looking proportion between the two conditions at each time point. The same cluster-based permutation analyses were conducted to compare the looking behaviours between ML and NW as well.

Results

Eye movement results from the two groups of infants are shown in Figure 3 and Figure 4, which demonstrates the average proportion of looks towards the target object for CP, ML, NW, and KW conditions from the onset of the first mention of the test word (0 s) until the end of the trial (4 s).

Infants in the Long Vowel Group did not fixate at the target object until they heard the second mention of the test word for CP (blue line in Figure 3), whereas they looked more towards the target object after the first mention of the test word for ML, but this was not maintained very long before shifting away quickly (red line in Figure 3). For the control conditions, as we expected, infants did not show recognition for NW where their proportional looks remained around 0.5 chance level all the time (yellow line in Figure 3), whereas they did show recognition of KW (green line in Figure 3). The cluster-based permutation analyses indicated a significant time-window when the proportion of looks towards the target object was significantly above chance level for CP condition (2.508-3.036 s time-window, blue-shaded area in Figure 3; $p < .05$), however, no significant time-window could be found for ML condition ($p = .27$), nor for NW condition ($p = .48$). As a result, the difference between the proportion of target looks for the CP and ML conditions was approaching significance in the time-window from 2.673 s to 2.970 s (gray-shaded area in Figure 3; $p = .059$), but no significant difference was found immediately after the first mention of the test word. Finally, no time-window of clusters was found between the ML and NW conditions, suggesting that infants were not able to recognize the reference of the label better in ML condition than in NW condition.

In the Short Vowel Group, infants' looks remained at 0.5 chance level for the first mention of the test word for CP, and tended to shift their looks towards the distractor after the second mention (blue line in Figure 4). Looks towards the target object remained at chance level as well after the first mention of ML, and showed minimal shift to the target object after the second mention (red line in Figure 4). For the control conditions, infants showed recognition of KW (green line in Figure 4), but not NW (yellow line in Figure 4). However, the cluster-based permutation analyses did not reveal any significant time-windows when the proportion of looks towards the target object was significantly above chance level in either CP condition (no time-window) or ML condition ($p = .39$). A difference between the identification of the target object in the two conditions was found to be significant in time-window from 2.376 s to 2.838 s (gray-shaded area in Figure 4; $p < .05$), but in an unexpected direction as they looked toward the distractor in CP but the target in ML condition.

The findings showed that the infants in the Long Vowel Group could successfully recognize a correct pronunciation and a mispronunciation of vowel length in the taught new words, whereas the Short Vowel Group failed to recognize even the correct pronunciations of the newly taught words. We also observed that the Long Vowel Group and the Short Vowel Group had different looking patterns for the correct pronunciations of the target object. We will return to this finding later in the discussion.

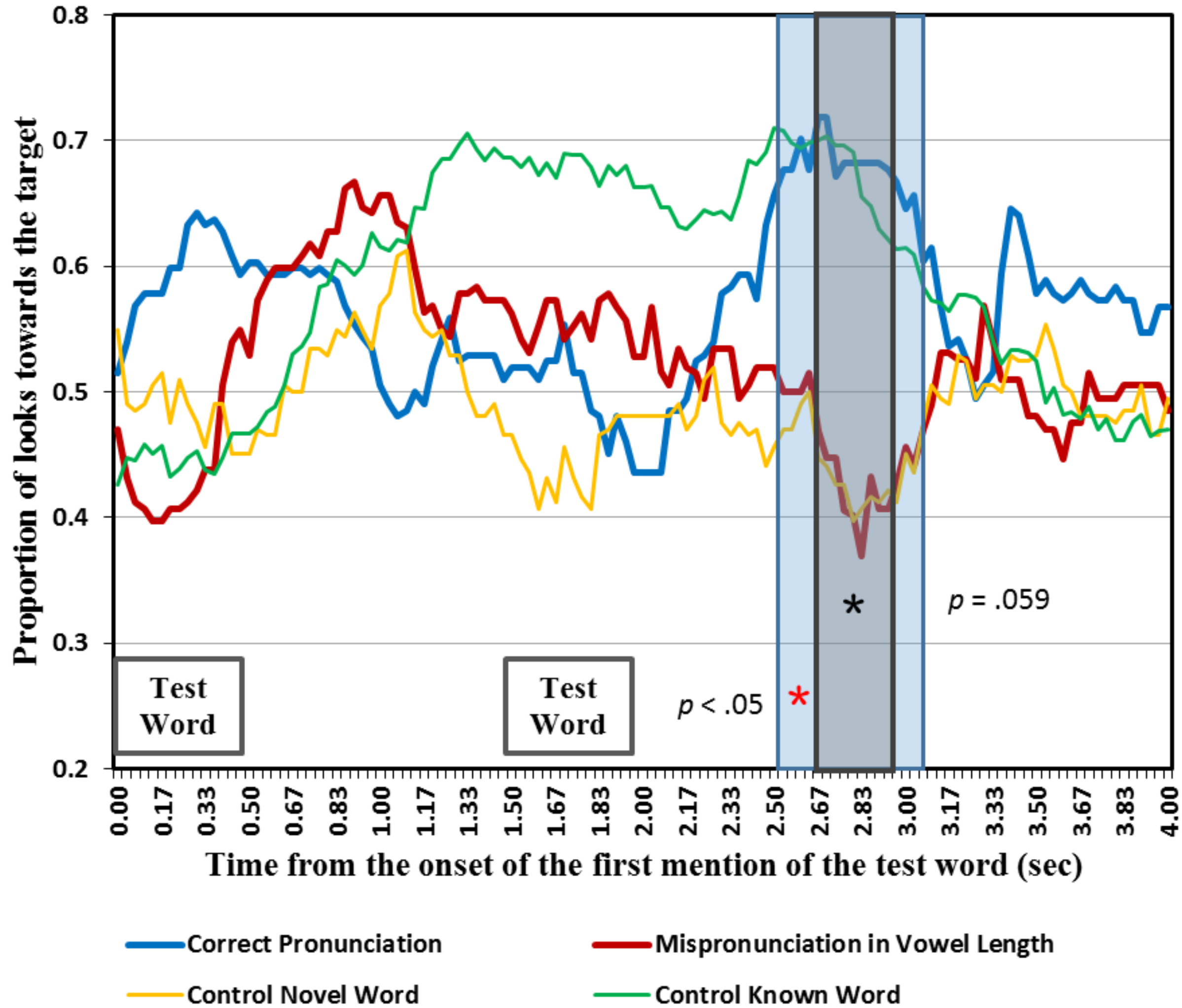


Figure 3. Mean proportion of looks towards the target object from the onset of the first mention of the test word for the correct pronunciations (CP, blue), vowel length mispronunciations (ML, red), the control novel words (NW, yellow), and the control known words (KW, green) in the Long Vowel Group. The infants successfully learnt the correct pronunciations (blue-shaded time-window) as shown by an increase of looks towards the target object, but failed to recognize the mispronunciations, staying at chance level. The gray-shaded time-window corresponds to the region where the infants were more likely look at the target object when asked on the correct pronunciations than on the mispronunciations in vowel length. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

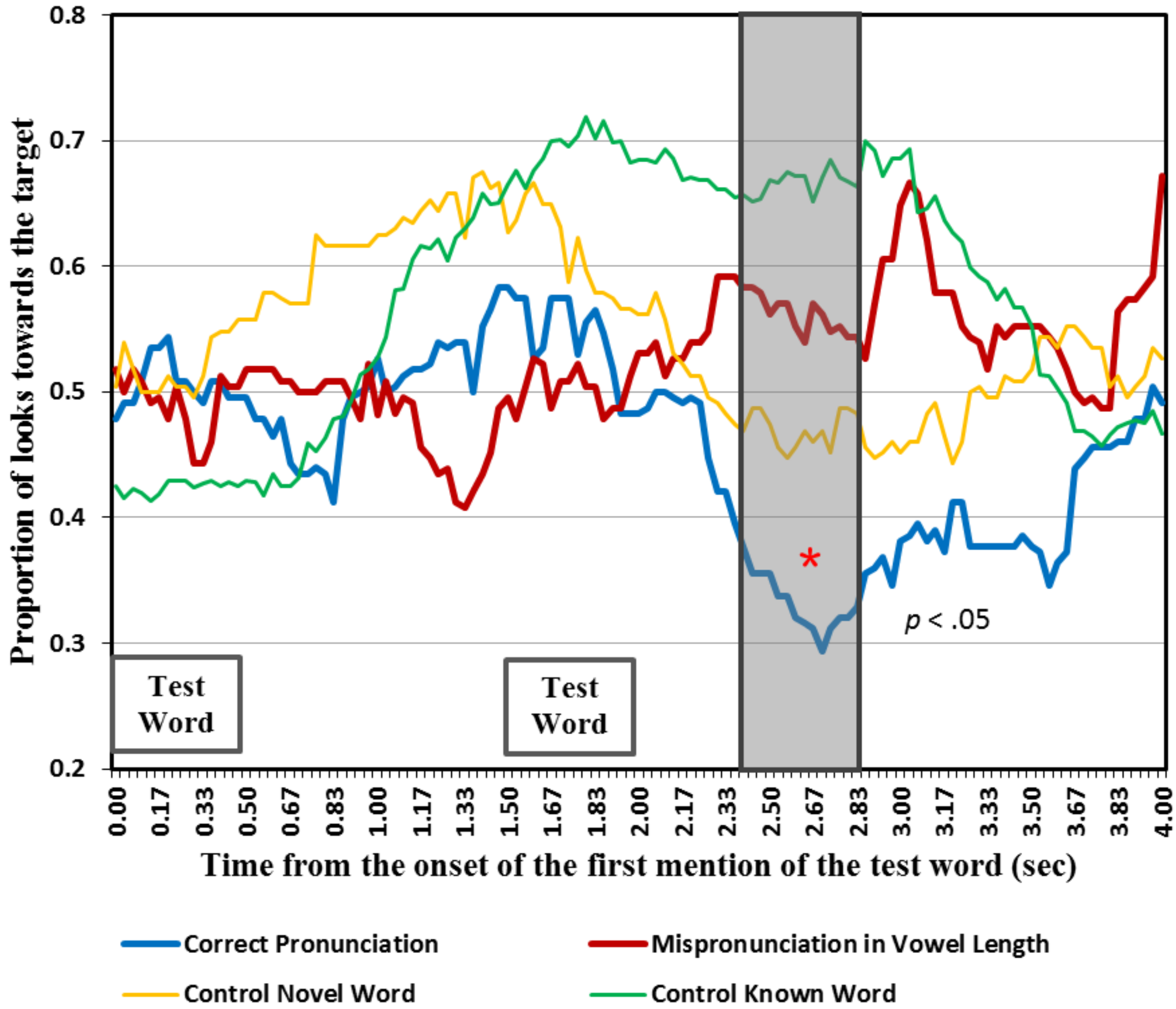


Figure 4. Mean proportion of looks towards the target object from the onset of the first mention of the test word for the correct pronunciations (CP, blue), vowel length mispronunciations (ML, red), the control novel words (NW, yellow), and the control known words (KW, green) in the Short Vowel Group. The infants failed to learn the correct pronunciations as shown by a decrease of looks towards the target object, and failed to recognize the mispronunciations, staying at chance level. The gray-shaded time-window corresponds to the region where the infants were more likely look at the target object when asked on the mispronunciations than on the correct pronunciations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Discussion

The present study set out to investigate whether, at 18 months, Japanese infants could identify a vowel length change in newly learnt novel words, and whether they could detect a change in both long-to-short and short-to-long mispronunciations. We taught two groups of infants two novel words, either both with long vowels or both with short vowels, and tested their sensitivity to the mispronunciations of vowel length in a Looking-While-Listening task. The results showed that only the infants in the Long Vowel Group could successfully recognize the newly taught long vowel words as well as identify a mispronunciation in vowel length. In contrast, the infants in the Short Vowel Group failed to recognize the newly taught short vowel words, showing a less robust representation of the short vowels after training.

Our results therefore indicate that, by 18 months, Japanese infants have built up a robust representation of the long vowels in the lexicon and were not tolerant to a change in the length of the vowel. This challenges the conclusion of the previous syllable discrimination study, which suggested that Japanese 18-month-olds treat the long vowel as an atypical short vowel, and are therefore sensitive to the phonetic change from long to short (atypical to typical) but not *vice versa* (Mugitani *et al*., 2009). By using a word learning and recognition task rather than phonetic discrimination, our results demonstrate that Japanese infants are not only sensitive to the shortening of a long vowel, but also resist accepting a shortened mispronunciation as the same referent at 18 months. This result is therefore consistent with the finding in the word recognition study with Dutch 21-month-olds (van der Feest & Swingley, 2009), where infants show sensitivity to the shortening of a long vowel in a word learning/recognition task. Importantly, in our study this sensitivity cannot be driven by the reduction of vowel spectral information since Japanese vowel length contrasts rarely involve spectral differences. Our findings therefore suggest that, by 18 months, Japanese-learning infants have a well-specified length feature for long vowels in their lexical representation, despite its limited input in the language. One possible reason might be lie on the acoustic/perceptual salience of the long vowels in Japanese, which are, on average, twice the length of the short vowels. Also, as all vowels have short and long counterparts in Japanese, this systematicity might possibly help infants to generalize abstract short/long length categories with exemplars from various vowels, despite their uneven distributions in the input they hear.

It has been reported that Japanese infants demonstrate an understanding of the moraic structure of their language very earlier in production – they lengthen the vowel in word-medial position only when a moraic coda was dropped (about 60-80% before 18 months), but never do so for the dropping of non-moraic onsets (Ota, 1999, 2001). The sense of the timing unit in the language might thus facilitate the establishment of the contrast between short vs. long vowels, which contain one vs. two moras of structure.

Our results on the asymmetry of responses to long vs. short vowels appears to pattern with the directionality effect reported in the Japanese syllable discrimination study (Mugitani et al., 2009), as well as the word recognition studies with Dutch 21-month-olds (van der Feest & Swingley, 2009) and adults (Nooteboom & Doodeman, 1980; van der Feest & Swingley, 2011). However, given the different tasks used between our study and these previous studies, the asymmetric pattern revealed in our results must be interpreted differently. To successfully identify a vowel length mispronunciation of the newly taught novel words in our task, the Japanese infants had to first correctly learn the new words, indicated by looking significantly towards the target objects after hearing the correctly pronounced target word. Unfortunately, the infants in the Short Vowel Group did not show this looking pattern in the correct pronunciation condition. We therefore cannot interpret their performance in the mispronunciation condition as being able or unable to identify a mispronunciation in vowel length. As a result, it would be inappropriate to simply take these results as an asymmetric directionality effect in the perception of vowel duration changes in Japanese 18-month-old infants.

However, these different results from our two test groups raise the question of why infants did not learn the new words in the Short Vowel condition, whereas those in the Long Vowel condition did. This difference was unlikely due to differences in the language ability or experience of the infants, as no significant differences were found in the age, receptive or expressive vocabulary size between the two groups of infants. It is also inconsistent with influence from the input, as infants hear many more short vowels than long vowels (Bion et al., 2010; Mugitani et al., 2009). The reason for the difference between the two groups could be quite complex. First, perhaps the long vowel words were easier to learn than the short vowel words due to the greater acoustic/perceptual salience of the former, giving infants more information in terms of the identity of the vowel. Song, Demuth and Morgan (2010) showed that infants had an easier time recognizing a word in slow speech that was half the rate of fast (adult-directed) speech. In our case, it could be easier for infants to remember a word when a syllable is longer.

Another possible explanation for why the long vowel words might have been easier to learn than the short vowel words is that the long vowel might have a prosodic form that better fits the canonical form of words that infants typically hear. It has been reported that Japanese infant-directed vocabulary often takes a disyllabic Heavy-Light form, and it appears to be the form that Japanese infants prefer (Mazuka, 2015; Mazuka et al., 2008). If this is true, *baato* and *giite,* with two moras in the first syllable, fit this template, but *bato* and *gite* do not. Perhaps infants attended to and learned more effectively in the Long Vowel condition, giving the asymmetrical result.

Last but not the least is the possibility that the inconsistent pitch alignments in disyllabic CVCV words in actual speech might have hindered the successful learning of the target short vowel words (i.e., *bato* and *gite*). As mentioned in the Method section, in disyllabic words with the first syllable pitch-accented, the pitch reaches its target (high f0) on the first syllable and the steep fall of pitch takes place between the first and second syllable when the vowel is short. However, this typically happens when the word is produced in isolation (phrase-finally). In actual speech, the pitch target is rarely realized on the first syllable of words containing a short vowel, as it takes time to reach the peak. Instead, the pitch target is more typically implemented on the following vowel, so the second syllable is more likely to contain a high pitch (except when phrase-final). In the training videos, the Short Vowel stimuli *bato* and *gite* were both presented in two different phrasal positions, one phrase-initial and the other phrase-final. This resulted in different tokens of the novel words in the training video having different pitch alignment patterns. Given the limited exposure to the new words in a short period of time, this (natural) inconsistency of the phonetic form of the new words might prove distracting, leading to incomplete learning of the short vowel words. In contrast, in the words carrying long vowels, *baato* and *giite*, the pitch target can be implemented consistently on the long vowel, then falling on the second syllable regardless of the phrasal position. Thus, infants in the Long Vowel Group received more consistent phonetic realizations of the new target words, which might then facilitate the reliable learning of novel words carrying long vowels.

Our results also indicate the importance of investigating infants' vowel length perception abilities in a time course analysis rather than region of interest in looking time differences. As revealed in our data, the infants in the Long Vowel Group took more than 2.5 seconds after the onset of the first mention of the test word to settle their looks at the target object in the correct pronunciation condition, which was much longer than their identification for known word conditioin (which took about 1 second). As a result, the time-window for differentiating their looking behaviours between correct pronunication and mispronunciations conditions was much later than the usual time-window used in other studies (i.e., within 2 seconds or 2.5 seconds after the onset of test word) (e.g., Mani et al., 2008; Mani & Plunkett, 2007, 2008, 2011; Swingley &

Aslin, 2000, 2002). However, this is not surprising given the nature of vowel length contrasts, which encode relative changes (duration) rather than absolute changes (spectral properties). To detect a vowel length change, it is necessary for infants to listen to more words in an utterance before they can possibly compute relative length. Therefore, it is crucial to look at their looking behaviour as it unfolds overtime rather than in a set time-window.

For future studies it will be important to better control for the variations in pitch pattern for the short vowel words, and test whether Japanese infants have also specified the length feature for short vowels in their lexicon by 18 months. It would also be interesting to better understand the difference between phonemic vowel length and vowel quality development, and whether Japanese infants have equal sensitivity to mispronunciations in vowel length contrasts and one-feature vowel quality contrasts at the same age. Conducting a word recognition task on Japanese adults would also indicate how both long and short vowels are encoded in the lexicon, and compare this to Dutch, where adults only specify the length feature in long vowel words (Chládková, Escudero, & Lipski, 2015; Escudero et al., 2009). Finally, it would be very interesting to test infants from other languages with a systematic phonemic vowel length contrast which are not mora-timed, such as in Finnish or Arabic, as well as languages where vowel duration is contrastive for only a few vowel pairs, such as in Australian English.

In summary, this study showed that, by 18 months, Japanese infants can successfully learn novel words with long vowels and detect shortening mispronunciations of these newly learnt words. This result, together with previous studies on word learning and word recognition in Dutch infants (Dietrich et al., 2007; van der Feest & Swingley, 2009), provides evidence for a well-specified phonemic length category in long vowels in infants learning a language where vowel length categories are relevant to word meanings. This reveals an important developmental achievement in sensitivity to phonemic vowel length by 18 months. The findings presented in this paper therefore help us to better understand the nature of phonemic vowel length development, and how infants' begin to specify phonemic features in word learning.

Footnotes

[1] Note that Japanese is an SOV (subject-object-verb) language.

[2] The target objects for the control pronunciations were randomly assigned and counter-balanced between the two novel toys.

Acknowledgments

We thank all the infants and parents involved in the study for their participation. We also thank Kaori Myogan, Ako Ohori, Kyoko Teraoka, Mari Kanamura, Yuri Hatano, Ruka Hayano, Naoko Kijima, Ying Deng, Natsumi Shibata, Sho Tsuji, Ivan Yuen, Carmen Kung, Isabella Dautriche, the Laboratory for Language Development at RIKEN Brain Science Institute, and the Child Language Lab at Macquarie University for their helpful assistance and feedback. This research was supported, in part, by funding from the first author's iMQRES at Macquarie University, the Short-term IPA Program at RIKEN BSI, and the following grants: ARC Centre of Excellence in Cognition and its Disorders CE110001021 and ARC FL130100014. Portions of this study were presented at the Workshop on Infant Speech Perception: Phonological and Lexical Development (WISP) – Macquarie University, Sydney, 1-2 September 2015; and the 40th Boston University Conference on Child Language Development (BUCLD40) – Boston University, Boston, 13-15 December 2015.

References

Adank, P., van Hout, R., & Smits, R. (2004). An acoustic description of the vowels of Northern and Southern standard Dutch. *Journal of the Acoustical Society of America*, *116*, 1729-1738. doi:10.1121/1.1779271

Amano, S., Nakatani, T., & Kondo, T. (2006). Fundamental frequency of infants' and parents' utterances in longitudinal recordings. *Journal of the Acoustical Society of America*, *119*, 1636-1647.

Beckman, M. E. (1986). *Stress and Non-Stress Accent*. Dordrecht, Holland; Riverton, N.J.: Foris.

Benders, T. (2013). *Nature's distributional-learning experiment: Infant' input, infants' perception, and computational modeling* (Unpublished doctoral dissertation). University of Amsterdam, Netherlands.

Benders, T., & Mandell, D. (2011). *The integration of acoustic dimension in infants' phoneme perception: Introducing and testing a categorization paradigm.* Poster session presented at the Society for Research in Child Development (SRCD) Biennial Meeting, 31 March-2 April, Montréal, Canada.

Bion, R. A. H., Miyazawa, K., Kikuchi, H., & Mazuka, R. (2013). Learning phonemic vowel length from naturalistic recordings of Japanese infant-directed speech. *PLoS ONE*, *8*(2), e51594. doi:10.1371/journal.pone.0051594

Boersma, P., & Weenink, D. (2012). *Praat: doing phonetics by computer* [Computer program]. Version 5.3.29, retrieved 4 October 2012 from http://www.praat.org/

Brugman, H., & Russel, A. (2004). Annotating multimedia/multi-modal resources with ELAN. In *Proceedings of LREC 2004*, Fourth International Conference on Language Resources and Evaluation.

Chládková, K., Escudero, P., & Lipski, S. C. (2015). When "AA" is long but "A" is not short: Speakers who distinguish short and long vowels in production do not necessarily encode a short-long contrast in their phonological lexicon. *Frontiers in Psychology*, *6*, *438*. doi:10.3389/fpsyg.2015.00438

Curtin, S., Fennel, C., & Escudero, P. (2007). Infants' recognition of vowel contrasts in a word learning task. In H. Caunt-Nulton, S. Kulatilake, & I. Woo (Eds.), *BUCLD 31: Proceedings of the 31st annual Boston University Conference on Language Development* (Vol. 1, pp. 141-152). Somerville, MA: Cascadilla Press.

Curtin, S., Fennel, C., & Escudero, P. (2009). Weighting of vowel cues explains patterns of word-object associative learning. *Developmental Science*, *12*, 725-731.

Dautriche, I., Swingley, D., & Christophe, A. (2015). Learning novel phonological neighbors: Syntactic category matters. *Cognition*, *143*, 77-86. doi:10.1016/j.bcognition.2015.06.003

de Carvalho, A., Dautriche, I. & Christophe, A. (2016). Three-year-olds use prosody on-line to constrain syntactic analysis. *Developmental Science*, *19*, 235-250.

Dietrich, C., Swingley, D., & Werker, J. (2007). Native language governs interpretation of salient speech sound differences at 18 months. *Proceedings of the National Academy of Sciences*, *104*(41), 16027-16031. doi:10.1073/pnas.0705270104

Escudero, P., Benders, T., & Lipski, S. (2009). Native, non-native and L2 perceptual cue weighting for Dutch vowels: The case of Dutch, German, and Spanish listeners. *Journal of Phonetics*, *37*, 452–465. doi:10.1016/j.wocn.2009.07.006

Fernald, A., Zangl, R., Portillo, A.L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In Sekerina, I. A., Fernandez, E.M., & Clahsen, H. (Eds.), *Developmental psycholinguistics: On-line methods in children's language processing* (pp. 97-135).

Amsterdam: John Benjamins.

Golinkoff, R., Hirsh-Pasek, K., Cauley, K., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, *14*, 23-45.

Han, M. S. (1962). The feature of duration in Japanese. *Onsei no kenyuu*, *10*, 65-80.

Hirata, Y. (2004). Effects of speaking rate on the vowel length distinction in Japanese. *Journal of Phonetics*, *32*(4), 565-589. doi:10.1016/j.wocn.2004.02.004

Hirata, Y., & Tsukada, K. (2009). Effects of speaking rate and vowel length on formant frequency displacement in Japanese. *Phonetica*, *66*(3), 129-149. doi:10.1159/000235657

Ladefoged, P. (1971). *Preliminaries to linguistic phonetics.* Chicago: University of Chicago Press.

Lehiste, I. (1970). *Suprasegmentals.* Cambridge, Mass.: MIT Press.

Mani, N., Coleman, J., & Plunkett, K. (2008). Phonological specificity of vowel contrasts at 18-months. *Language and Speech*, *51*(1&2), 3-21.

Mani, N., & Plunkett, K. (2007). Phonological specificity of vowels and consonants in early lexical representations. *Journal of Memory and Language*, *57*(2), 252–72.

Mani, N., & Plunkett, K. (2008). 14-month-olds pay attention to vowels in novel words. *Developmental Science*, *11*(1), 53–59.

Mani, N., & Plunkett, K. (2011). Does size matter? Subsegmental cues to vowel mispronunciation detection. *Journal of Child Language*, *38*, 606-627.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177-190.

Mazuka, R. (2015). *Prosodic forms of infant-directed vocabulary can facilitate word learning.* Talk presented in the Workshop on Infant Speech Perception (WISP), September 2015, Macquarie University, Sydney, Australia.

Mazuka, R., Kondo, T., & Hayashi, A. (2008). Japanese mothers' use of specialized vocabulary in infant-directed speech: Infant-directed vocabulary in Japanese. In N. Masataka (Ed.), *The Origins of Language: Unraveling Evolutionary Forces* (pp. 39-58). Tokyo: Springer Japan.

Minagawa-Kawai, Y., Mori, K., Furuya, I., Hayashi, R., & Sato, Y. (2002). Assessing cerebral representations of short and long vowel categories by NIRS. *Neuroreport*, *13*(5), 581-584.

Minagawa-Kawai, Y., Mori, K., Naoi, N., & Kojima, S. (2007). Neural attunement processes in infants during the acquisition of a language-specific phonemic contrast. *The Journal of Neuroscience*, *27*(2), 315-321. doi:10.1523/jneurosci.1984-06.2007

Mugitani, R., Pons, F., Fais, L., Dietrich, C., Werker, J. F., & Amano, S. (2009). Perception of vowel length by Japanese- and English-learning infants. *Developmental Psychology*, *45*(1), 236-247.

Nazzi, T. (2005). Use of phonetic specificity during the acquisition of new words: Differences between consonants and vowels. *Cognition*, *98*, 13-30.

Nazzi, T., Floccia, C., & Bertoncini, J. (1998). Discrimination of pitch contours by neonates. *Infant Behavior and Development*, *21*(4), 779-784. doi:10.1016/S0163-6383(98)90044-3

Nooteboom, S. G., & Doodeman, G. J. N. (1980). Production and perception of vowel length in spoken sentences. *Journal of the Acoustical Society of America*, *67*, 276–287.

Ogura, T., & Watamaki, T. (2004). *The Japanese MacArthur Communicative Development Inventory: Words and Gestures.* Kyoto: Kyoto International Social Welfare Exchange Center.

Ota, M. (1999). *Phonological theory and the acquisition of prosodic structure: Evidence from child Japanese* (Unpublished doctoral dissertation). Georgetown University, Washington D. C.

Ota, M. (2001). Phonological theory and the development of prosodic structure: Evidence from child Japanese. *Annual Review of Language Acquisition*, *1*, 65-118.

Sato, Y., Sogabe, Y., & Mazuka, R. (2010). Discrimination of phonemic vowel length by Japanese infants. *Developmental Psychology*, *46*(1), 106-199. doi:10.1037/a0016718

Schneider, W., Eschman, A., & Zuccolotto, A. (2012). *E-Prime User's Guide*. Pittsburgh: Psychology Software Tools, Inc. [Computer program]. Version 2.0, retrieved from http://www.pstnet.com/

Song, J. Y., Demuth, K., & Morgan, J. (2010). Effects of the acoustic properties of infant-directed speech on infant word recognition. *Journal of the Acoustical Society of America*, *128*(1), 389-400. doi:10.1121/1.3419786

Swingley, D. (2011). The looking-while-listening procedure. In E. Hoff (Ed.), *Research methods in child language: A practical guide* (pp. 29-42). Malden, MA: Wiley-Blackwell.

Swingley, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, *76*, 147-166.

Swingley, D., & Aslin, R. N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, *13*(5), 480-484. doi:10.1111/1467-9280.00485

van der Feest, S. V. H., & Swingley, D. (2009). Language-specific interpretation of vowel duration in 21-month-olds' word recognition: A cross-linguistic study of phonetic attribution. Paper presented at the 34th Boston University Conference on Language Development, November, Boston, MA.

van der Feest, S. V. H., & Swingley, D. (2011). Dutch and English listeners' interpretation of vowel duration. *Journal of the Acoustical Society of America*, *129*(3), EL57-EL63. doi:10.1121/1.3532050.

Vance, T. J. (1987). *An Introduction to Japanese Phonology*. Albany: State University of New

York Press.

Werker, J. F., Pons, F., Dietrich, C., Kajikawa, S., Fais, L., & Amano, S. (2007). Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition*, *103*(1), 147-162. doi:10.1016/j.cognition.2006.03.006

White, K. S., & Morgan, J. L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language*, *59*, 114-132.

Appendix: Examples of Video Scripts Used in Training

A. *baato* associated with the yellow toy – No. 1

“mite mite! kore wa *baato* dayo. *baato* wa ookina me o shi teru ne! *baato* no kao ni wa omoshiroi hana ga tsuiteru ne! *baato* ni wa te to ashi tsuiteru yo. mite, kiiroi *baato*!”

English Translation:

“Look! Look! This is a *baato*! A *baato* has big eyes! There is a funny nose on a *baato*’s face! A *baato* has arms and legs too! Look at the yellow *baato*!”

Japanese script:

“見てみて！これは *baato* だよ。*baato* はおっきな目をしてるね〜。*baato* の顔には面白い鼻がついてるね！*baato* には手と足ついてるよ。見て。黄色い *baato*！”

B. *baato* associated with the yellow toy – No. 2

“mite mite! *baato* ga iru yo. *baato*, odori ga sukina nda. *baato,* tanoshi-soo ni waratteru ne! *baato* no ha mieru? kawaii ne *baato*!”

English Translation:

“Look! A *baato* is here. *baatos* like dancing! The *baato* is happy and laugh! Can you see the *baato*’s teeth! Good job, *baato*!”

Japanese script:

“見てみて！*baato* がいるよ。*baato* 踊りが好きなんだ。*baato* 楽しそうに笑ってるね〜。*baato* の歯見える？かわいいね *baato*！”

C. *giite* associated with the purple toy – No. 1

“mite mite! kore wa *giite* dayo. *giite* wa murasakiirona nda. *giite* no atama ni wa ohoshisama ga tsuiteru ne! *giite* te ga futatsu tsuiteru ne. *giite* tte hen'na karada shi teru ne!”

English Translation:

“Look! Look! This is a *giite*! A *giite* is purple! There is a star on a *giite*’s head! A *giite* has two arms too! What a funny body a *giite* has!”

Japanese script:

“見てみて！これは *giite* だよ。*giite* は紫色なんだ。*giite* の頭にはお星様がついてるね！*giite* 手が 2 つついてるね。*giite* って変な体してるね！”

D. *giite* associated with the purple toy – No. 2

“mite mite! *giite* ga iru yo. *giite* ohoshisama ga, sukina nda ne! *giite* wa tokidoki kono ohoshisama kakusu nda yo. *giite* tottemo kawaii ne! mite! *giite* un un tte, hakushu shi teru yo!”

English Translation:

“Look! A *giite* is here. The *giite* loves its star! The *giite* hides the star sometimes. The *giite* is so cute! See! The *giite* agrees and claps its hands!”

Japanese script:

“見てみて！*giite* がいるよ。*giite* お星様が、好きなんだね〜。*giite* は時々このお星様隠すんだよ。*giite* とってもかわいいね。見て！*giite* うんうんって、拍手してるよ！”

Chapter Three: Understanding the Acquisition of Phonemic Vowel length Contrasts in Australian English-Learning 18- and 24-Month-Olds

This chapter is based on the following paper under review for publication:

Chen, H., Xu Rattanasone, N., Cox, F., & Demuth, K. (under review). Understanding the acquisition of phonemic vowel length contrasts in Australian English-learning 18- and 24-month-olds. *Language Learning and Development*.

All components of this paper, both experimental and written, have been completed by me, with advice from the co-authors (my supervisors) when needed.

Abstract

Phonemic vowel length is often thought to be mastered later than vowel quality. However, there are few studies directly examining infants' sensitivity to vowel length in comparison to vowel quality at the early word learning stage. This study therefore targeted Australian English, which has non-systematic phonemic vowel length contrasts. Two groups of infants were tested on their sensitivity to mispronunciations involving a 1-feature change in vowel height, backness or length in familiar words. The results showed that, although the 18-month-olds were unable to perform the task, the 24-month-olds displayed similar sensitivity to all the single vowel feature changes. The findings reveal that, at least by 24 months, these infants are aware of phonemic vowel length contrasts in a word recognition task. This suggests that for Australian English-learning infants, the phonemic vowel length contrast is acquired around the same time as the vowel height and backness contrasts.

Key words: phonemic vowel length, perception, Australian English, non-systematic, 1-feature

Understanding the Acquisition of Phonemic Vowel Length Contrasts in Australian English-Learning 18- and 24-Month-Olds

Introduction

Vowel length serves many functions in language. Variations in the duration of a vowel can signal phonological contrasts in the environment (such as voicing of the coda consonants), prosodic context (such as word and phrasal boundary), syntactic structure, emphatic stress, speech rate, and so on. In these situations, the duration of a vowel varies due to contextual factors and does not affect the phonemic status of the vowel. This is called *extrinsic* vowel length. Previous acquisition literature suggests that infants start developing perceptual sensitivity to *extrinsic* vowel length that contributes to identification of voicing contrasts between 8 and 14 months (Ko, Soderstrom, & Morgan, 2009), and even earlier to vowel duration cues that signal a clausal boundary at 4 months (Seidl & Cristià, 2008).

On the other hand, the duration of the vowel can also vary depending on its phonemic status in a language. This is called *intrinsic* vowel length and its variation can signal a difference in word meanings. Phonemic vowel length contrast refers to the distinctive *intrinsic* length of vowels independent of vowel quality (spectral differences). This is one of the vowel contrasts employed in some languages, which functions alongside other vowel features, such as vowel height, vowel backness, and vowel roundness. For instance, in Australian English, *hut* /hɐt/ and *heart* /hɐːt/[1, 2] can be distinguished only by the duration of the vowels, just as *hut* /hɐt/ and *hat* /hæt/ can be distinguished primarily by the degree of backness of the vowels. In the acquisition literature, infants' phonemic perception of vowel quality features has been investigated extensively (Curtin, Fennell, & Escudero, 2007, 2009; Mani, Coleman, & Plunkett, 2008; Mani & Plunkett, 2007, 2008, 2011; Nazzi, 2005; Swingley & Aslin, 2000, 2002; among others). However, only a few studies have looked at the perception of vowel length, and this has been mainly in languages where vowel length is combined with other vowel quality cues, as in Dutch (Benders, 2013; Benders & Mandell, 2011; Dietrich, Swingley, & Werker, 2007; van der Feest & Swingley, 2009). There have been few studies on infants' perception of *phonemic* vowel length in languages where a vowel contrast can be signaled by duration alone, such as Japanese (Chen, Yamane, Xu Rattanasone, Demuth, & Mazuka, 2015). We know of no other studies that have compared infants' perception of phonemic vowel length to other vowel features in a language where each feature cue alone can signal a change in word meaning such as Australian English. Therefore, this study aims to investigate infants' sensitivity to phonemic vowel length contrasts in comparison to vowel height and vowel backness contrasts in Australian English.

Previous studies on phonological and lexical development have generally looked at when and how infants developed sensitivity to vowel *quality* contrasts at the early word learning stage. Swingley and Aslin (2000, 2002) tested 14-, 18- and 23-month-old American English infants' sensitivity to mispronunciations of either consonants or vowels of familiar words, using the Intermodal Preferential Looking (IPL) paradigm (Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987). They found that all age groups looked significantly less at the target when its label was mispronounced than correctly pronounced (though both were above chance), and infants also took significantly longer to shift their looks from the distractor to the target in the mispronunciation condition. These results held for all test items, so infants were not only sensitive to consonants but also to vowel quality changes in familiar words *apple* (*'opple'*) and *car* (*'cur'*). Similar results were found in a later study with British English 15-, 18-, and 24-month-olds, using a set of vowel-medial monosyllabic familiar words (Mani & Plunkett, 2007). The British infants showed the same latency effect between correct pronunciations and vowel mispronunciations (e.g., *book – bik*, *bread – brod*, *dog – dig*), and showed a difference in looking preference between the two conditions.

Note that the vowel mispronunciations tested in the above studies involved combinations of two or more vowel quality features (i.e., height, backness, or roundness). One of the questions, then, is when infants become sensitive to a single vowel feature change, and whether they show equal sensitivity to all vowel features. Mani et al. (2008) found that British English 18-month-olds were sensitive to mispronunciations of familiar words in both vowel height and backness, but not in roundness. Mani and Plunkett (2007) also reported that British English 15-month-olds performed at the same level as older 18- and 24-month-olds in detecting 1-feature mispronunciations of vowel quality and consonant contrasts. However, these infants' performance did not differ between correct pronunciations and mispronunciations of the vowel. Thus, it remains unclear if 15-month-olds can detect a 1-feature change in vowel quality in a word recognition task.

Given that contrastive vowel length is also one of the important vowel features employed in some languages, such as Japanese, our question was whether infants can understand the phonemic function of vowel length at 18 months. Previous studies on vowel length perception at the early word learning stage have mainly focused on languages where the vowel length contrasts also involve prominent quality spectral differences, such as Dutch. Van der Feest and Swingley (2009) reported that Dutch 21-month-olds could recognize only shortening but not lengthening mispronunciations of vowels in familiar words, similar to findings with Dutch adults (Nooteboom & Doodeman, 1980; van der Feest & Swingley, 2011). Nevertheless, given that long/short vowels in Dutch occur in combination with certain vowel quality features, infants might rely on either vowel quality or the combination of quality and duration in word recognition, and use duration as a cue only when quality information is reduced. Recent perception and EEG studies (Chládková, Escudero, & Lipski, 2013, 2015; Escudero, Benders, & Lipski, 2009) have revealed that Dutch adults only specify the length feature in words with long vowels but not with short, suggesting that vowel length may have a variable status in Dutch vowels. Therefore, infants acquiring languages such as Dutch might develop vowel length representations differently from those who are learning a language where vowel length contrasts can take a phonemic function alone, such as Japanese and Australian English.

A recent study found that Japanese 18-month-olds could correctly identify newly learned disyllabic words with a long vowel and detect the shortening mispronunciations (Chen et al., 2015). This suggests that Japanese-learning infants as young as 18 months are aware of phonemic vowel length contrasts, and that vowel length contrasts may not necessarily be acquired later than vowel quality contrasts. But Japanese has systematic vowel length contrasts across the entire vowel inventory, so given the consistency and robustness of phonemic vowel length in the vowel system, Japanese infants might tune into the contrasts early. Therefore, our question was *when* infants acquiring a language with non-systematic vowel length contrasts in the vowel inventory become aware of the phonemic status of vowel length, and *whether* infants would show similar sensitivity to phonemic vowel length as with other vowel quality features.

Australian English, as a non-rhotic variety, employs phonemic vowel length (in the absence of spectral features) in only two vowel pairs - /ɐ/ vs. /ɐː/ (e.g. *hut* vs. *heart*) and /e/ vs. /eː/[3] (e.g. *bed* vs. *bared*). Critically, it has been shown that these vowel pairs exhibit minimal spectral quality differences (Bernard, 1967; Cox, 2006; Watson & Harrington, 1999). Australian English therefore provides a well-controlled comparison across vowel length, vowel height, and vowel backness. As shown in Figure 1, /ɐ/ vs. /ɐː/ constitutes a vowel length contrast with minimal change in vowel height/backness. In contrast, /e/ vs. /æ/ and /ɜː/ vs. /ɐː/ differ similarly in vowel height contrast, with minimal change in vowel backness within each pair. For /æ/ vs. /ɐ/, there is minimal change in vowel height but the degree of vowel backness contrast is similar in extent to the vowel height contrast in the previous set (i.e., 2 bark). Note, however, that the vowel height and backness contrasts in these vowels are smaller than those tested in the previous British English study (e.g., Mani et al., 2008) – instead of *high* vs. *low* and *front* vs. *back* 2-degree differences, the above vowels contrasting in *mid-high* vs. *low* in height, and *front* vs. *central* in backness, constituting 1-degree differences. Thus, one might predict that detecting such a small change in a mispronunciation task would be more challenging for young learners.

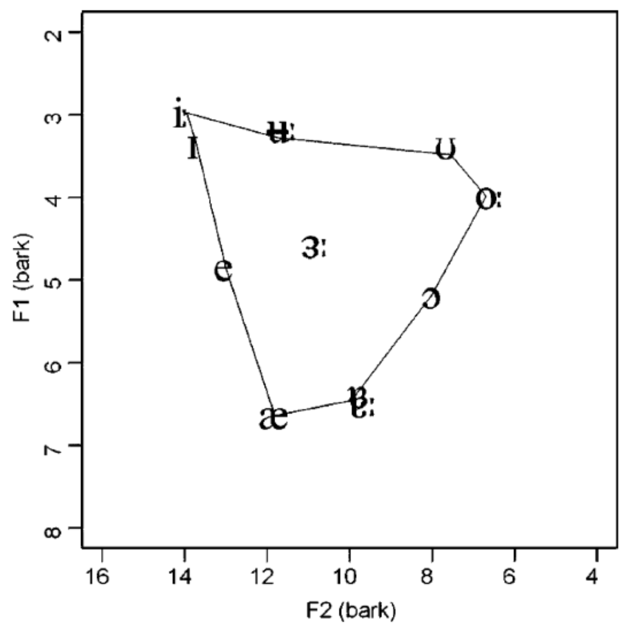


Figure 1. Australian English: Monophthong vowel formant plot based on citation form /hVd/ words from five male speakers (reproduced from Cox & Palethorpe, 2007:346).

Therefore, the present study investigated when Australian English infants become sensitive to phonemic vowel length compared to a 1-degree vowel height and vowel backness change in a mispronunciation detection task using the IPL paradigm. We first tested infants at 18 months who might possibly be able to identify the mispronunciations, as previous studies suggest that 18-month-old infants display sensitivity to single feature vowel quality contrasts and vowel length contrasts for British English (Mani et al., 2008) and Japanese (Chen et al., 2015). However, it is also possible that Australian English infants develop this fine-grained sensitivity later than 18 months, given that vowel length contrasts found in Australian English are non-systematic compared to Japanese. Also, the smaller 1-degree vowel quality contrasts tested here may be more challenging for 18-month-olds compared to the task used in Mani et al.'s (2008) study of British English.

Experiment 1: 18-Month-Olds

Method

Participants. Twenty monolingual Australian English-learning 18-month-olds (6 boys and 14 girls) with a mean age of 18.17 (months.days, range: 17.29-19.14) were tested with their parents in Sydney, Australia. An additional 35 infants were tested but excluded from the final analysis, due to fussiness and inability to complete the task (n = 4), parent interference (n = 1), equipment failure (n = 1), calibration failure (n = 3), low sampling rate[4] (< 45%, n = 12), and inability to complete at least one trial per condition (n = 14).

All infants had Australian English as the only language spoken by the parents at home, with minimal exposure to other English dialects or languages (≤ 7 hours per week). They had no reported language or hearing problems. Parents completed an adapted toddler short form of the MacArthur Communicative Development Inventories (MCDIs; Fenson, Pethick, Renda, Cox, Dale, & Reznick, 2000) with both receptive and productive checklists, an extra vocabulary questionnaire (see Appendix I for a report), and a language background questionnaire (see Appendix II). The receptive MCDI scores of the short form for the 20 infants ranged from 30 to 81 with a mean of 50, while the expressive scores for the same 100 words ranged from 1 to 77 with a mean of 26.85.

Stimuli and design. Ten monosyllabic CVC English inanimate nouns with a voiceless coda consonant, e.g., *bus*, were used as familiar stimuli in the experiment (see Table 1). They included seven words selected from the American CDI norms (Dale & Fenson, 1996) and comprehended by more than 55% of the American infants by 18 months according to the reported norms. An additional three words (*sock, tap*, *heart*) were chosen from a pilot study with the highest familiarity to the infants (90%, 70%, 40%, respectively for the 3 words). These were added to the original seven words to make a test set of 10 words.

For six out of the 10 familiar stimuli, vowel substitutions were made to create the mispronounced words in the test. Each stimulus was mispronounced in two different ways, both with only one single vowel feature change (e.g., the vowel /æ/ in familiar stimulus *tap* was changed in vowel height or in vowel backness). To ensure that the infants did not hear the same word mispronounced more than once during the test, two lists were created. Each word was mispronounced in two different vowel features, one for each list. For instance, the familiar stimulus *tap* /tæp/ was mispronounced as *tep* [tep] in vowel height (VH_{MP}) in List1 and as *tup* [tɐp] in vowel backness (VB_{MP}) in List 2; Likewise, the familiar stimulus *heart* /hɐːt/ was mispronounced as *hurt* [hɜːt] in vowel height (VH_{MP}) and as *hut* [hɐt] in vowel length (VL_{MP}) in the two lists, respectively (see Table 1 for details).

It would be ideal to have each familiar word mispronounced in three ways, however, given the constraints of the vowel contrasts to be tested in Australian English, and the requirement of finding high frequency familiar words, we could not provide a complete three-way contrast for each word.

With the exception of the test item *hurt*, all the mispronunciations of the familiar stimuli were either nonce words or low frequency words in English not familiar to 18-month-olds, e.g., *bass* [bæs]. The low frequency words were not included in the American CDI norms (Dale & Fenson, 1996) and 0% of the infants understood the items based on the parent reports from the pilot study. The word *hurt* had 20% comprehension according to the parent reports from the pilot, but it was interpreted as a verb rather than a noun. All test items were presented as the labels of objects in the experiment, where they were required to be nouns. Previous studies have shown that a familiar verb can be easily learnt as the label of a novel object by infants (Ferguson, Graf, & Waxman, 2014; Dautriche, Swingley, & Christophe, 2015).

For the remaining four familiar stimuli, two items in each list were correctly pronounced (CP) while the other two were produced as novel pronunciations (NP). For the novel pronunciations, none of the segments matched the originals in the corresponding familiar words (e.g., *gup* [gɐp] vs. *foot* [fʊt]). The CPs and NPs were reversed in the two lists, which were randomly presented across participants.

In addition, we included five familiar animate nouns, three (*duck, shark, cat*) for the training trials and two (*dog, bird*) for the filler trials during the test. Two novel pronunciations were then designed, one (*nep* for *cat*) used in training and the other (*dirk* for *dog*) as a filler item during the test. Each infant was presented with two filler items, one correct (*bird*) and one novel (*dirk*). Performance on these filler items was not analyzed.

Table 1. *Trial conditions, familiar stimuli and test items in the two test lists*

Pronunciation Condition	List 1		List 2	
	Familiar	Test	Familiar	Test
Vowel Height Misp. (VH$_{MP}$)				
/æ/ → /e/	(tap)	tep [tep]	(hat)	het [het]
/ɐː/ → /ɜː/	(heart)	hurt [hɜːt]	(bath)	birth [bɜːθ]
Vowel Backness Misp. (VB$_{MP}$)				
/ɐ/ → /æ/	(bus)	bass [bæs]	(cup)	cap [kæp]
/æ/ → /ɐ/	(hat)	hut [hɐt]	(tap)	tup [tɐp]
Vowel Length Misp. (VL$_{MP}$)				
/ɐ/ → /ɐː/	(cup)	carp [kɐːp]	(bus)	barss [bɐːs]
/ɐː/ → /ɐ/	(bath)	buth [bɐθ]	(heart)	hut [hɐt]
Correct Pronunciation (CP)	(sock)	sock [sɔk]	(book)	book [bʊk]
	(fork)	fork [foːk]	(foot)	foot [fʊt]
Novel Pronunciation (NP)	(book)	geap [giːp]	(sock)	nart [nɐːt]
	(foot)	gup [gɐp]	(fork)	noss [nɔs]

Note. VH$_{MP}$ = Vowel Height Mispronunciation; VB$_{MP}$ = Vowel Backness Mispronunciation; VL$_{MP}$ = Vowel Length Mispronunciation. Australian English is non-rhotic and therefore none of these words is produced with an /ɹ/ or rhoticised vowel.

Auditory stimuli. The auditory stimuli were recorded by a native Australian English female speaker in an infant-friendly manner. Recordings were made in an acoustically shielded recording booth, sampled at 48 kHz. The target test item was recorded three times in a carrier sentence: "*Look at the [Target]!*" We then chose the best token for each target test item according to the judgments of all four authors and used them as the final auditory stimuli in the study.

The acoustics of the vowels used in the mispronunciation conditions are shown in Table 2. The mean duration of the short vowels was 256 ms (*SD* = 46 ms, range = 221-349 ms), while the mean duration of the long vowels was 562 ms (*SD* = 45 ms, range = 524-624 ms). The durations were also consistent with those reported for child-directed speech in a previous production study on Australian English (Yuen, Cox, & Demuth, 2014), where the short vowels in a CVC word in utterance-final position with focus were 251 ms on average, while the long vowels in a CVːC word were 364 ms on average in the same environment in adults. The F1 and F2 values in the tested vowels were also consistent with previous studies (e.g., Cox, 2006). Acoustic measurements were completed using Praat (Boersma & Weenink, 2012). Due to the very high pitch used in the child-friendly manner (max. 350-460 Hz), formant tracking in Praat was largely influenced by the first harmonic (F0) and unreliable. Therefore, the formant measurements for each token were based on the spectral slice obtained from the point where the formants were resolved once the pitch returned to a value below 300 Hz. The spectral slice was taken no later than 70% of the vowel.

Table 2. *Mean durations (ms) and F1 and F2 values (Hz) for the vowel stimuli used in the mispronunciation conditions (with range)*

Short			Long		
Vowel	Duration	Formants	Vowel	Duration	Formants
/e/	225 (221-228)	*F1*: 773 (765-781) *F2*: 2259 (2204-2313)	/ɜː/	580 (535-624)	*F1*: 718 (703-732) *F2*: 1608 (1596-1619)
/æ/	304 (258-349)	*F1*: 1090 (1074-1105) *F2*: 1818 (1808-1828)			
/ɐ/	246 (226-281)	*F1*: 989 (928-1095) *F2*: 1537 (1518-1547)	/ɐː/	545 (524-565)	*F1*: 928 (909-947) *F2*: 1438 (1430-1446)
Grand Mean	256 (221-349)		Grand Mean	562 (524-624)	

Note. The vowel /ɐ/ had three tokens; all others had two.

Visual stimuli. Child-friendly cartoon pictures were created for each familiar stimulus and judged by the authors and their colleagues as typical exemplifiers of the corresponding objects. A novel object was created to match the style, colour, shading, size and visual complexity for each familiar stimulus. These were yoked as familiar-novel pairs. All the novel objects were assessed for their novelty to young children. During the experiment, the yoked familiar-novel cartoon pictures were depicted against two off-white background frames, and displayed side-by-side horizontally on a black screen (see example in Figure 2). A full set of 10 picture pairs for the test items and 4 pairs for training and fillers is shown in Appendix III.

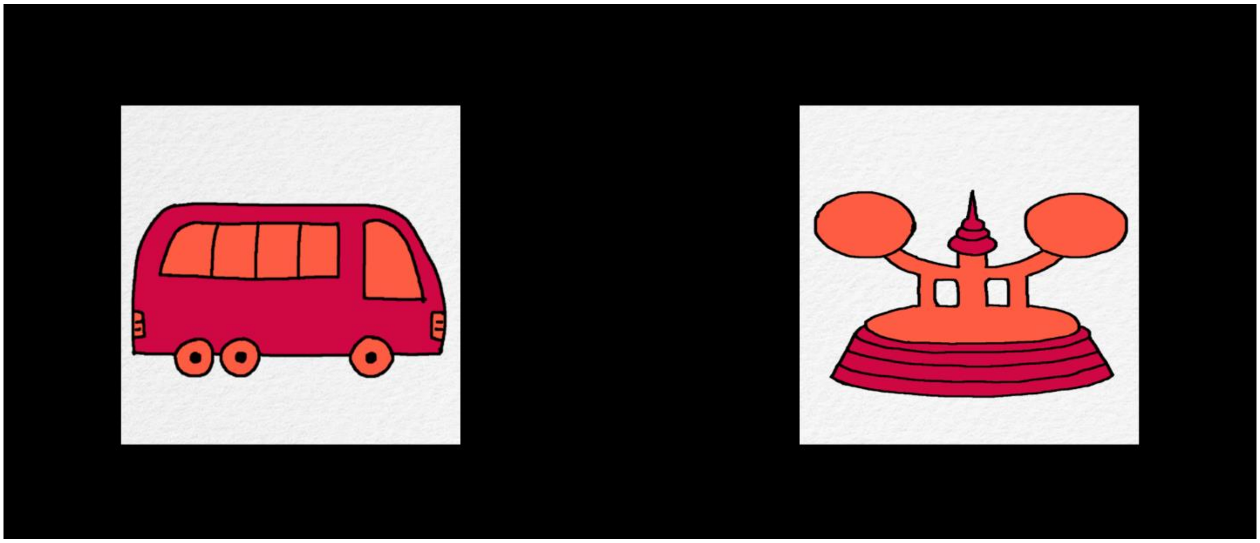


Figure 2. An example of the yoked cartoon pictures: *bus – barss [ɐː] / bass [æ]*

Procedure. The study comprised of a familiarization session and an IPL test session. The mispronunciation detection task and IPL paradigm used in the test session was based on the mutual exclusivity constraint that children possibly apply to restrict word meanings (Markman, 1984, 1987, 1990, 1991; Markman & Wachtel, 1988; see also Merriman & Bowman, 1989). Mutual exclusivity is a word learning constraint whereby category terms are mutually exclusive, so each object can have only one category label and each label can refer to only one category of objects. It is suggested that children are biased by this constraint, so that they resist a second label for a familiar object, and prefer to map a novel label to a novel rather than a familiar object. Although evidence for the existence of the constraint is mostly based on 2- and 3-year-olds, there have been studies showing that infants reject second labels for known objects as young as 15-16 months (Markman, Wasow, & Hansen, 2003), and they prefer a novel referent for a novel label at 17 months (Halberda, 2003). The paradigm should therefore be appropriate in our study (18-month-olds).

Familiarization. To ensure infants could easily and correctly identify the familiar cartoon pictures, they were first familiarized with real photos of all the tested familiar stimuli, collated into a picture book containing four coloured photos of the familiar stimuli on each page. Each familiar stimulus was shown to infants three times with different photos to help infants generalize the images of the familiar stimulus. These real photos were used rather than the cartoon pictures in order to maintain the visual novelty of the cartoon pictures in the later test. During the familiarization session, infants sat with their parent in front of the experimenter (the first author) and a research assistant in a playroom. Either the research assistant, who was a native female speaker of Australian English, or the infant's parent named the photos in the album book one-by-one to the infant, and made sure that the infant looked at each photo.

Test. During the test session, infants sat on their parent's lap in a sound attenuated test room while looking at a widescreen 27" monitor displayed at 1920 x 1080 pixels, tilted at 3°, and located approximately 80 cm in front of the infant. The two off-white background frames containing the yoked stimulus pictures were displayed at a size of 16.2 x 16.2 cm each on the monitor, at a distance of 16.2 cm from each other, providing a minimal 11.5° and a maximal 33.7° horizontal gaze angle from the infant. Infants' looking behaviour was recorded by a Tobii X120 Eye Tracker, tilted at 30°, and positioned 65 cm in front of the infant and 30 cm under the monitor. The eye-tracker collected gaze data from both eyes, sampling at 120 Hz with a 100 ms recovery time for lost tracking. The auditory stimuli were delivered at a conversation level (≈ 65 dBA) from two computer speakers located on both sides of the monitor. During the entire session, the parent wore a pair of masking glasses to ensure that they could not see the stimuli and potentially influence the infant. A SONY digital video camera with a zoom lens was positioned beside the eye-tracker to record infants' looking behaviour as a secondary recording of the sessions.

The entire test session was conducted using the Tobii Studio software which controlled the presentation of stimuli and test trials. Each trial was delivered in a video (.AVI) file encoded in JPEG codec 3.2.4 at 24 frames-per-second. The video files were created using FinalCut Pro software on a Macintosh computer by integrating the corresponding auditory and visual stimuli.

The test session contained 15 trials, including three training trials, two test trials for each pronunciation condition (i.e., 2 VH_{MP}, 2 VB_{MP}, 2 VL_{MP}, 2 CP, 2 NP), and two filler trials (1 CP, 1 NP). In the first two training trials, infants saw the same pair of familiar animals (*duck* - *shark*) on the screen, but heard *duck* in one trial and *shark* in the other. At the end of training trials, the correct picture danced to cheerful music while the incorrect one stayed still. This helped infants orient their looks to the both sides of the screen. In the third training trial, infants saw a familiar animal *cat* paired with a novel animate object on the screen, and heard "*Look at the nep!*". The novel animate object then danced to cheerful music, which helped infants orient their looks to a novel object when they heard something other than the label of a familiar object.

In the test trials, infants saw pictures of a familiar and novel object presented side-by-side on the screen. After 4 seconds, the two pictures were replaced with a looming red ball in the middle of the black screen. The gaze-centering period lasted for 1 second before the looming ball disappeared and the same set of pictures were displayed on the screen again. The auditory stimulus then started: "*Look at the [Target]!*", with the vowel onset of the target test item aligned with the onset of the 7th second of the trial. The trial ended 4 seconds after the vowel onset of the target test item, with the entire trial lasting for 10 seconds. The entire test session lasted about 3.5 minutes for each infant (see Figure 3 for an example of a test trial).

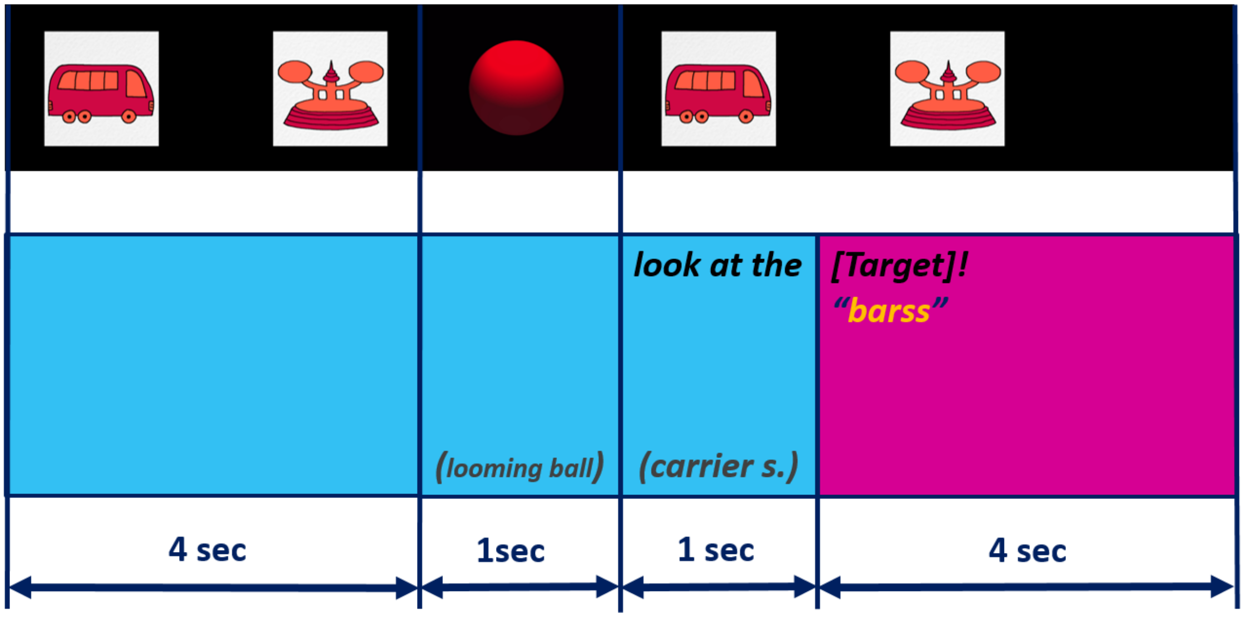


Figure 3. An example of a test trial: "*Look at the barss[ɐː]!*"

Infants were randomly assigned to four test versions, including two different randomizations of test items for List 1 and List 2. The order of the test trials in each version was pseudo-randomized so that no two consecutive trials had the same condition. Familiar and novel pictures appeared equally often on the left and right side of the screen in each list, and were counterbalanced across all four versions. The familiar pictures did not appear more than twice on the same side in consecutive trials.

Data Analysis

Gaze position on each trial was recorded by the Tobii X120 Eye Tracker approximately every 8.33 ms. Raw looking data were converted into fixations using the I-VT fixation filter in Tobii Studio (version 3.2.3). Missing data points were interpolated for sections having a duration of less than 60 ms, and fixations that were less than 75 ms were discarded. Areas of Interest (AOIs) were defined as two 19 x 19 cm squares covering the off-white background frames in each trial, given the typical 1-0.5° accuracy for remote eye-trackers. The AOIs were marked as *Familiar* or *Novel* in each trial according to the corresponding cartoon pictures.

In order to determine whether infants' looking behaviour would differ as a function of mispronunciation condition, the difference score was used as the dependent measure. This has been widely used in previous IPL studies to investigate infants' looking preference shift after hearing the auditory stimulus (Mani & Plunkett, 2007, 2008, 2011; Mani et al., 2008; White & Morgan, 2008; among others). Infants' total fixation duration on each AOI was calculated for two different time-windows in each test trial - the *Pre-naming* phase and the *Post-naming* phase. The pre-naming phase was defined as the first 4 seconds of the trial, and the post-naming phase was defined as the period from 367 ms after the vowel onset of the target test item to the end of the trial, as it has been widely assumed that it takes 367 ms for infants to initiate an eye movement (Swingley & Aslin, 2000, 2002; among others).

For each phase, we calculated the Proportion of Fixation duration on the Familiar object (PFF) by dividing the fixation duration on the *Familiar* AOI by the total fixation duration on both the *Familiar* and the *Novel* AOIs. For each infant, a difference score was then calculated for each condition, by averaging differences of PFF between the two phases across trials in the condition (i.e., PFF (*Post-naming*) – PFF (*Pre-naming*)). A positive score indicated an increment of proportional fixation on the familiar object in the post-naming phase compared to the pre-naming phase.

To establish a baseline preference for each trial, only the trials in which the infant fixated both the *Familiar* and the *Novel* AOIs at some point during the pre-naming phase were included in the analysis. This criterion excluded 14 trials across the infants. Also, trials in which the sum duration of the AOI fixations equaled less than 50% of the test phase duration were excluded. This was to ensure that we only analyze those trials where infants were largely attending and their looking behaviours were reasonably relevant to a response to the auditory stimulus. Across all the participants, another 13 trials were excluded for this reason. Altogether, a total 13.5 % of the overall 200 trials was excluded.

Results and Discussion

The parental vocabulary questionnaire indicated that, on average, these familiar stimuli were understood by 80% (range 25-100%) of the infants, and that 40% (range 20-60%) of the infants could produce them (see detailed report in Appendix IA). The word *heart* was understood only by 25% of the infants and produced by 20%, suggesting that *heart* was the least familiar to the infants. On the other hand, the low frequency items used as mispronunciations were unfamiliar to the infants, with very low reported comprehension rates (0-25%). The word *hurt* was the only one understood by 25% of the infants, but again as a verb rather than a noun.

Average difference scores for each of the five pronunciation conditions are shown in Figure 4. We expected to see a significant increment in the PFF from the pre-naming to the post-naming phase in a trial if infants regarded the label they heard as the correct label for the familiar object. In Figure 4, an increment of PFF was found in all five conditions, regardless whether the labels were correctly pronounced, mispronounced, or novel to the infants.

Two primary analyses were conducted on the difference scores. First, to establish whether infants would recognize test items as the legitimate names of the familiar objects (i.e., an effect of naming), we compared their difference scores to 0 (meaning no change of PFF between the two phases) for each condition. Two-tailed *t*-tests showed that the increment of PFF from the pre- to post-naming phase was significant only for VH_{MP}, $t(19) = 2.705$, $p = .014$, but not significant for either the CP condition, $t(19) = 1.856$, $p = .079$, or the conditions VB_{MP}, $t(19) = .956$, $p = .351$, VL_{MP}, $t(19) = .652$, $p = .522$, and NP, $t(19) = .645$, $p = .527$. The fact that fixations on the familiar object did NOT significantly increased in the CP condition, suggesting that these infants might not understand the familiar words or the task.

Second, to understand whether there was an effect of mispronunciation on responses, with alpha set at .05, a repeated measures ANOVA analysis was conducted. The analysis revealed that the main effects of *pronunciation condition* (CP, VH_{MP}, VB_{MP}, VL_{MP}, NP) was not significant, $F(4, 76) = 1.023$, $p = .401$, $\eta_p^2 = .051$, indicating that the infants did not perform differently across the five conditions.

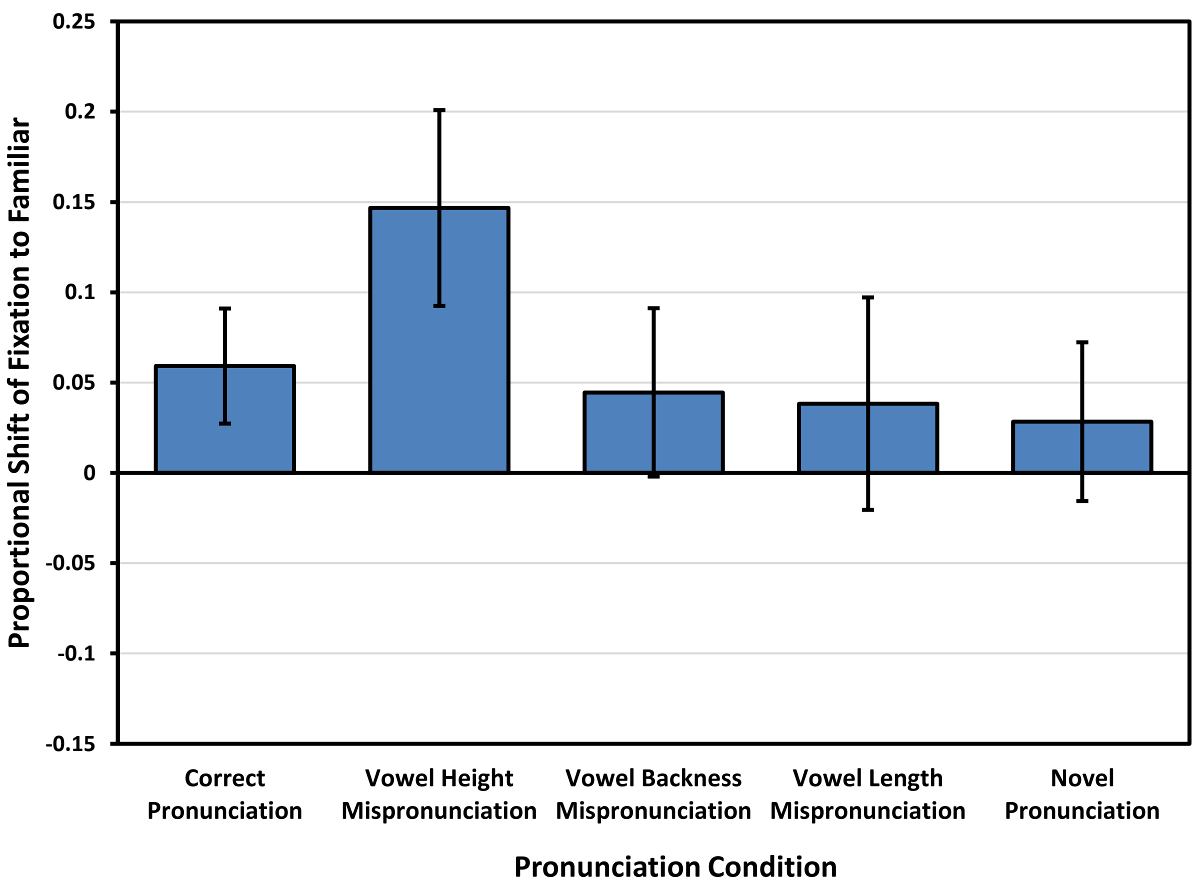


Figure 4. The 18-month-olds' average proportional shift of fixation duration to the *Familiar* AOI from the *Pre-naming* phase to the *Post-naming* phase in the five pronunciation conditions. Error bars indicate +/- 1.00 standard error of mean.

These results for Experiment 1 suggest that 18-month-olds might not yet have developed familiarity to the tested familiar items, as indicated by chance level performance on the CP condition. Alternatively, these infants might not fully understand the task given the chance level performance across the different conditions with the exception of the VH_{MP} condition. In the VH_{MP} condition, instead of increasing fixations to the novel object, the infants showed increased fixations on the *familiar* object in the post-naming phase. This might be related to the low comprehension and production rates reported in the parental reports, especially in VH_{MP} condition, e.g., *heart* (25% & 20%) and *tap* (50% and 20%).

Note that the vowel contrasts tested here were less acoustically distinct than those tested in previous studies (e.g., Mani et al., 2008); the current study employed a 1-degree difference /e/ vs. /æ/ on VH whereas Mani et al. (2008) used some larger 2-degree difference (e.g., /ɪ/ vs. /a/). Our task should therefore be more difficult for 18-month-olds than that used in Mani et al. Furthermore, the vowel length contrast is implemented in only a subset of two vowels in Australian English. This could mean that the perception task is harder for Australian infants compared to Japanese 18-month-olds, where the phonemic vowel length contrast is systematic throughout the vowel system. Perhaps, then, the phonemic vowel length contrast in Australian English, as well as 1-degree high and backness feature differences in English more generally, are acquired later. In Experiment 2 we therefore tested an older group of 24-month-old infants using the same procedure.

Experiment 2: 24-Month-Olds

Method

Participants. Twenty monolingual Australian English-learning 24-month-olds (5 boys and 15 girls) were tested with their parents in Sydney, with a mean age of 24.18 (range: 23.29-25.11). An additional 17 infants were tested but excluded in the final analysis, due to inability to be calibrated (n = 1), low sampling rate (n = 3), and returning insufficient trials per condition (n = 13). Data from a further infant who was identified as an outlier was excluded in the final analysis (see **Data Analysis** below).

The infants all came from Australian English-speaking families and had no reported language or hearing problems. Parents completed the same set of vocabulary checklist and questionnaires as was used for the 18-month-olds. The receptive scores of the short form MCDIs for the 20 24-month-olds ranged from 14 to 97 with a mean of about 72.32, while the expressive scores ranged from 21 to 97 with a mean of 57.6.

Stimuli, design and procedure. The stimuli, the design and the procedure were the same as in Experiment 1.

Data Analysis

Data from the 24-month-olds were extracted and analyzed in the same way as in Experiment 1. Across all the participants, 11 trials were excluded due to no AOI fixation recorded during the pre-naming phase, and an additional 17 trials were excluded as the total AOI fixation duration was less than 50% of the post-naming phase. In total, approximately 14% of the data were discarded from the final results based on these criteria. One participant with a z-score of larger than 2 was excluded as an outlier.

Results and Discussion

Again, results from the parental vocabulary questionnaire indicated that, on average, the 24-month-olds understood these familiar stimuli (94% (range 70-100%)), and were producing of them (83% (ranged 55-100%)) (see Appendix IB). For the low frequency items which were used as mispronunciations, most had lower than 30% reported comprehension rates and even lower reported production rates. The only exception was *hurt*, which was understood by 85% of the 24-month-olds and produced by 50%, again, as a verb.

Average difference scores for each pronunciation condition across the 24-month-old Australian English-learning infants are shown in Figure 5. Again, a positive score indicates an increment of PFF from pre- to post-naming phase.

As in Experiment 1, two major analyses were conducted. First, to establish whether there was an effect of naming on infants' responses in each condition, two-tailed *t*-tests were used to compare the difference scores to 0 for each condition. The analyses revealed that the increment of PFF from the pre-naming to the post-naming phase in the CP condition was significant, $t(19) = 2.539$, $p = .020$, showing the infants understood the familiar words and was performing the task. However, the shifts of PFF were not significantly different from chance in the mispronunciation conditions (VH_{MP}: $t(19) = 1.927$, $p = .069$; VB_{MP}: $t(19) = -.519$, $p = .610$; VL_{MP}: $t(19) = -.476$, $p = .640$), or in the NP condition, $t(19) = -1.791$, $p = .089$. As no significant increments of fixations on the familiar object were found when the labels were mispronounced or novel words were pronounced, the results suggest the infants did not accept the mispronunciations or novel words as the names of the familiar objects.

Second, to explore whether there was an effect of mispronunciation on infants' behaviour, a repeated measures ANOVA analysis across the five pronunciation conditions was conducted. The analysis revealed significant main effect of *pronunciation condition*, $F(4, 76) = 3.065$, $p = .021$, $\eta_p^2 = .139$, indicating that the infants performed differently across the five pronunciation conditions. Moreover, planned comparisons between CP and all other conditions were conducted (with Bonferroni adjustment): There was a significant difference between CP and $VL_{MP,}$ $F(1, 19) = 6.312$, $p = .021$, $\eta_p^2 = .249$, showing that the proportional shifts towards the familiar object were larger in CP ($M = .129$) than in VL_{MP} ($M = -.022$). There was also a significant difference between CP and NP, $F(1, 19) = 6.922$, $p = .016$, $\eta_p^2 = .267$, showing that the proportional shifts towards the familiar object were larger in CP ($M = .129$) than in NP ($M = -.0848$) as well. These suggests that the infants looked significantly less at the familiar object when provided mispronunciations in vowel length and novel pronunciations, indicating a sensitivity to these mispronunciations. Other comparisons were not significant (CP and VH_{MP}, $F(1, 19) = .547$, $p = .469$, $\eta_p^2 = .028$; CP and VB_{MP}, $F(1, 19) = 3.20$, $p = .090$, $\eta_p^2 = .144$). The mean proportional shifts towards the familiar object in VB_{MP} (M = -.027) was even smaller than that in VL_{MP} (M = -.022), but the variation in VB_{MP} (SD = .233, CI: -.136–.082) was bigger than that in VL_{MP} (SD = .207, CI: -.119–.075), which might explain the lack of significant effect in the comparison between CP and VB_{MP}.

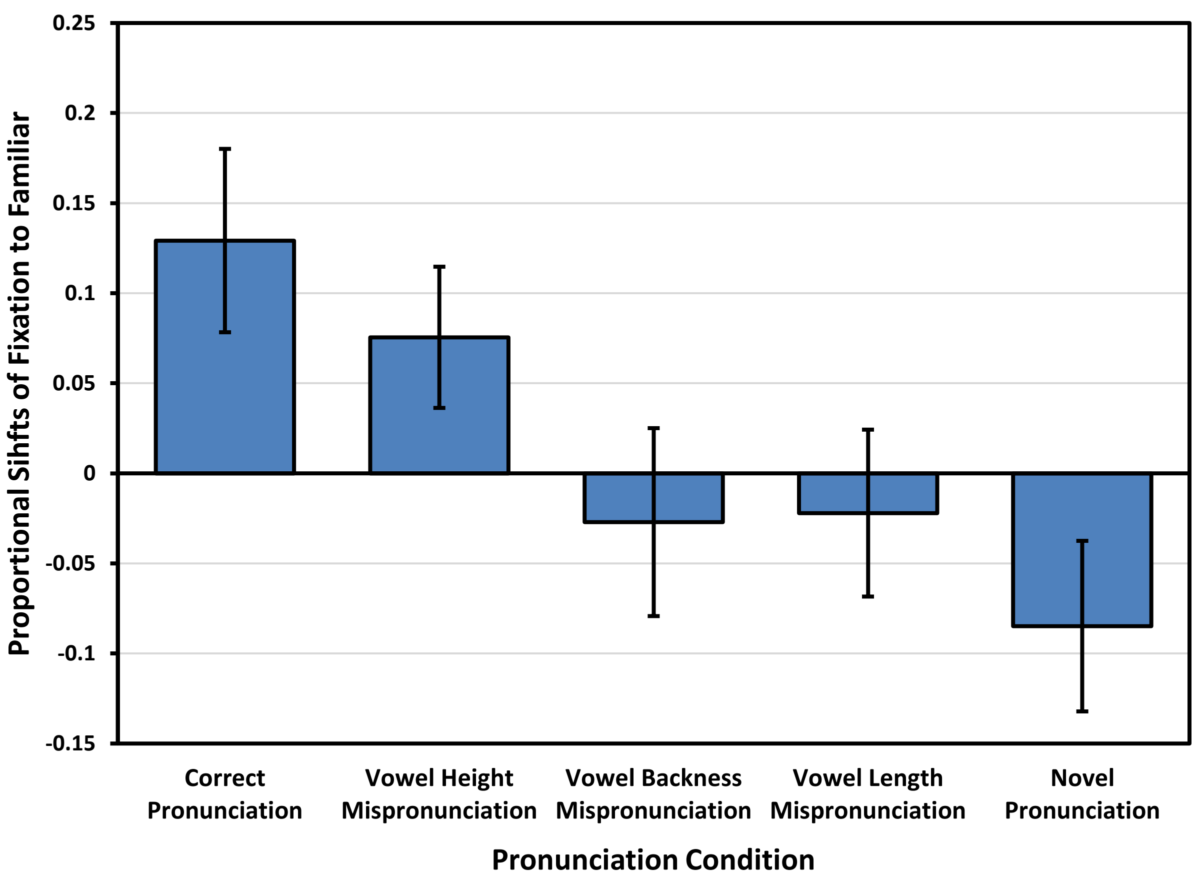


Figure 5. The 24-month-olds' average proportional shift of fixation duration to the *Familiar* AOI from the *Pre-naming* phase to the *Post-naming* phase in the five pronunciation conditions. Error bars indicate +/- 1.00 standard error of mean.

Together the results indicated that the 24-month-olds increased fixations on the familiar object in the CP condition. Performance in the NP condition was different from CP indicating that performance is consistent with the mutual exclusivity constraint. Additionally, performance in VL_{MP} differed from CP, indicating that 24-month-olds are not tolerant to mispronunciations in VL. Furthermore, performance on all mispronunciation conditions were not significantly different from chance, whilst CP was significantly above chance, indicating infants did show some sensitivity to mispronunciations in the other vowel features as well.

General Discussion

In this study we investigated when Australian English-learning infants become sensitive to phonemic vowel length contrasts in comparison to other vowel features. We tested two groups of infants aged 18 and 24 months. In **Experiment 1** we found that 18-month-olds showed no difference in performance across the five pronunciation conditions. They did not consistently look more at the familiar objects when the labels were mispronounced in vowel backness and vowel length, but neither did they do so when the labels were correctly pronounced. This suggests that the 18-month-olds in this study might not have performed the task, or that they did not show different sensitivities to the mispronunciations in various vowel features. Parental reports suggested that some of the items were not very familiar to this group of infants, e.g., *heart* and *tap*. Also, the target test words were embedded in utterance-final position and underwent final lengthening, which made the recognition of phonemic vowel length contrasts harder, despite the fact that they were different in duration. Furthermore, to maintain a single feature change, the vowel contrasts tested in the study were smaller than those used in previous studies of British English 18-month-olds (Mani et al., 2008). In **Experiment 2**, we therefore tested 24-month-olds and found that they consistently looked more at the familiar objects after hearing a correct pronunciation, but looked randomly between the familiar and novel objects when the label was mispronounced. This indicates that 24-month-olds were sensitive to the mispronunciations in single vowel feature change. Moreover, their performance in the vowel length mispronunciation condition was significantly different from the correct pronunciation condition, showing they are not tolerant to situations where familiar words are changed in vowel length.

The results of **Experiment 2** suggest that, at least by 24 months, Australian English infants have already tuned into the phonemic function of vowel length contrasts in familiar words, and understood that changes in vowel length may lead to changes in word reference. Furthermore, 24-month-olds displayed the ability of identifying phonemic vowel length contrasts even when the intrinsic length of vowel is interacting with higher level prosody as a result of phrase-final lengthening. As all vowels will be lengthened at phrase-final position, even for short vowels, infants have to factor in the prosody when interpreting the vowel identity. This suggest that Australian English infants might be sensitive to vowel length contrasts in an easier task without final lengthening at an age earlier than 24 months.

Australian English infants therefore appear to develop sensitivity to phonemic vowel length contrasts later than Japanese infants who showed awareness of phonemic vowel length in newly learned words as early as 18 months (Chen et al., 2015). One possible explanation for this difference is that the Australian English 18-month-olds might simply have failed the task. The perception task we used in the present study differed in several ways from the Japanese study. In the Japanese study, infants were taught two novel label-object associations in training and then presented with the same pair of objects again and again in the test, whereas in the current study infants saw each pair of objects only once and one of the objects was always novel to them. It could be difficult for infants to process a new object in every single trial and to apply the mutual exclusivity constraint even if they could hear the difference. Although the mutual exclusivity constraint has been evidenced in the 15-16-month-olds who reject second labels to familiar objects (Markman et al., 2003), previous word meaning acquisition literature has shown that infants younger than 17 months have difficulties in mapping a novel label to a novel object even when they see the yoked familiar-novel object pair again and again (Halberda, 2003). In our study, infants never saw the yoked pair more than once, so it might be difficult for the younger group to map the mispronunciations/novel words to the novel object in each trial. Moreover, the Australian English infants listened to each test item only once, whereas the test words were repeated twice in a trial in the Japanese study, and they were placed in utterance-initial position to adhere to the syntactic properties of Japanese. Thus, unlike in the current study, there were no phrasal lengthening effects in the Japanese study.

Alternatively, Australian English infants might not be able to tune into phonemic vowel length contrasts as early as Japanese infants. Given that phonemic vowel length is restricted to a subset of vowels and non-systematic in Australian English, infants acquiring this language might not receive consistent input with phonemic vowel contrasts to the extent that Japanese infants do. Conversely, Australian English-learning infants will have encountered many more vowels not contrasting in vowel length. They therefore have to hear more input and take longer to build up the representations of long vs. short, and also learn which vowels and words are relevant to these contrasts. We could expect that infants learning other quantity languages, such as Arabic, Finnish, Estonian, etc., would become sensitive to vowel length contrasts as early as Japanese infants given the systematicity and robustness of the contrasts in these phonological systems.

Our results also showed that Australian English-learning infants had similar, if not better, sensitivity to phonemic vowel length contrasts compared to other 1-feature vowel quality contrasts at 24 months. This finding is contrary to the suggestion that vowel length contrasts are later acquired than vowel quality, evidenced in some production studies (Demuth & Fee, 1995; Fikkert,

1994) and a phonetic discrimination study in very young infants (Sato, Sogabe, & Mazuka, 2010). In contrast, the present study directly tested infants' sensitivity to *phonemic* vowel length contrasts at the early word learning stage, providing a well-controlled comparison between vowel length and vowel quality contrasts.

One remaining question is why Australian English infants did not display robust sensitivity to vowel height and backness contrasts at 24 months or even earlier at 18 months, given the prior finding from British English infants, which shows that the 18-month-olds can successfully identify a change in vowel height and backness in a similar word recognition task (Mani et al., 2008). We think that there might be several reasons for these different results: Firstly, Australian English allows us to test contrastive vowel height and backness pairs with minimal difference in other dimensions. Infants therefore listened to vowel height (or backness) cues with little additional spectral cues on vowel backness (or height). Secondly, we tested smaller 1-degree changes in vowel height and backness compared to the previous British English study. Whilst the British English infants were tested with vowel height pairs such as /ɪ/ vs. /a/ and vowel backness pairs such as /ɪ/ vs. /ʌ/ with a larger contrasts in the dimension, our stimuli employed smaller 1-degree contrasts in the dimension (i.e., mid-height vs. low, and front vs. central). Previous studies have shown that magnitude of phonetic difference does influence young infants' word learning and recognition (Escudero, Best, Kitamura, & Mulak, 2014). Thirdly, the task in our study was different from the British English one in that we used familiar-novel object pairings instead of two familiar objects as visual stimuli, which possibly made the perception process more difficult for infants. A later study on British English infants by Mani and Plunkett (2011) adopting a more comparable design to the present study revealed results similar to our Australian English one, where the 18-month-olds (when exposed to 4 second pre-naming phase in their Experiment 2) did NOT perform differently between conditions of correct pronunciations and 1-feature vowel mispronunciations.

In our results, the 24-month-olds did show a certain degree of awareness of 1-degree changes in vowel height and backness, given the chance level looking behaviours in the two mispronunciation conditions. We therefore would expect that Australian English-learning 24-month-olds have similar sensitivity to larger vowel quality contrasts and vowel length contrasts in general.

For future studies it would therefore be interesting to test whether the younger 18-month-old Australian English infants can identify a change in these 1-feature vowel contrasts in an easier task – for instance, embedding test items in non-final position, and using two familiar objects as visual stimuli. Moreover, considering the importance of consistent and systematic input in early phonemic acquisition, it would be interesting to see what happens to phonemic perception when one received *variable* input in his/her early life, for example, as in bilingual or bi-dialectal experience at home. We might see competition between vowel quality and length with variable input in populations learning two different languages or dialects with and without phonemic vowel length.

To summarize, our results showed that, at least by 24 months, Australian English infants can successfully identify changes in phonemic vowel length in familiar words even with phrase-final lengthening, and they display similar sensitivity to vowel length and vowel height/backness contrasts. Together with previous research on vowel length acquisition in Japanese infants (Chen et al., 2015), this study provides a more complete picture of infants' development of phonemic vowel length contrasts across very different vowel systems, e.g., systematic vs. variable phonemic vowel length contrasts. It also provides a well-controlled comparison of infants' sensitivity to phonemic vowel length and other 1-feature vowel quality contrasts for the first time. The findings therefore contribute to our understanding of the relationship between vowel length and vowel quality acquisition, providing a better understanding of how infants learn various dimensions of vowel features during early word learning.

Footnotes

[1] As Australian English is the target language in the current study, the International Phonetic Alphabet (IPA) transcriptions here reflect Australian English vowels (see Harrington, Cox, & Evans, 1997).

[2] Australian English is a non-rhotic variety of English.

[3] The vowel /eː/ presents quite variably in Australian English, from monophthongal to fully diphthongal but often disyllabic in open syllables (Cox, 2006; Harrington, Cox, & Evans, 1997).

[4] The sampling rate is automatically computed by Tobii Studio and indicates how often participants' gaze was tracked throughout the recording session. It is calculated by dividing the number of eye tracking samples that were correctly identified, by the number of attempts. In infant studies, a low sampling rate suggests fussiness or inattentiveness, and the 45% criterion has been used in similar studies employing a Tobii eye-tracker (e.g., Mulak, Best, Tyler, Kitamura, & Irwin, 2013).

Acknowledgments

We thank all the infants and parents involved in the study for their participation. We also thank Katherine Revius, Benjamin Davies, Kelly Miles, Kiri Mealings, Ivan Yuen, Gretel Macdonald, Fabia Andronos, Christina Quattropani, and the Child Language Lab at Macquarie University for their helpful assistance and feedback. This research was partially supported by funding from the first author's iMQRES and REF at Macquarie University and the following grants: ARC Centre for Cognition and its Disorders CE110001021 and ARC FL130100014. Portions of this study were presented at the 14th Conference on Laboratory Phonology (LabPhon14) – NINJAL, Tokyo, 25-27 July 2014; and the Australian Linguistic Society Annual Conference 2015 (ALS2015) - Western Sydney University, Sydney, 9-11 December 2015.

References

Benders, T. (2013). *Nature's distributional-learning experiment: Infant' input, infants' perception, and computational modeling* (Unpublished doctoral dissertation). University of Amsterdam, Netherlands.

Benders, T., & Mandell, D. (2011). *The integration of acoustic dimension in infants' phoneme perception: Introducing and testing a categorization paradigm.* Poster session presented at the Society for Research in Child Development (SRCD) Biennial Meeting, 31 March-2 April, Montréal, Canada.

Bernard, J. (1967). Length and identification of Australian English vowels. *Journal of the Australasian Universities Modern Language Association*, *27*, 37-58.

Boersma, P., & Weenink, D. (2012). *Praat: doing phonetics by computer* [Computer program]. Version 5.3.29, retrieved 4 October 2012 from http://www.praat.org/

Buder, E. H., & Stoel-Gammon, C. (2002). American and Swedish children's acquisition of vowel duration: Effects of vowel identity and final stop voicing. *Journal of the Acoustical Society of America*, *111*(4), 1854-1864.

Chládková, K., Escudero, P., & Lipski, S. C. (2013). Pre-attentive sensitivity to vowel duration reveals native phonology and predicts learning of second-language sounds. *Brain & Language*, *126*, 243-252.

Chládková, K., Escudero, P., & Lipski, S. C. (2015). When "AA" is long but "A" is not short: Speakers who distinguish short and long vowels in production do not necessarily encode a short-long contrast in their phonological lexicon. *Frontiers in Psychology*, *6*, 438. doi:10.3389/fpsyg.2015.00438

Chen, H., Yamane, N., Xu Rattanasone, N., Demuth, K., & Mazuka, R. (2015). *Japanese infants* *are aware of phonemic vowel length in novel words at 18 months*. Paper presented at the 40th Boston University Conference of Language Development (BUCLD40), Boston University, Boston.

Cox, F. (2006). The acoustic characteristics of /hVd/ vowels in the speech of some Australian Teenagers. *Australian Journal of Linguistics*, *26*(2), 147-179.

Cox, F., & Palethorpe, S. (2007). Australian English. *Journal of the International Phonetic Association*, *37*, 341-350. doi:10.1017/S0025100307003192

Curtin, S., Fennel, C., & Escudero, P. (2007). Infants' recognition of vowel contrasts in a word learning task. In H. Caunt-Nulton, S. Kulatilake, & I. Woo (Eds.), *BUCLD 31: Proceedings of the 31st annual Boston University Conference on Language Development* (Vol. 1, pp. 141-152). Somerville, MA: Cascadilla Press.

Curtin, S., Fennel, C., & Escudero, P. (2009). Weighting of vowel cues explains patterns of word-object associative learning. *Developmental Science*, *12*, 725-731.

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavioral Research Methods, Instruments, & Computers*, *28*, 125-127.

Dautriche, I., Swingley, D., & Christophe, A. (2015). Learning novel phonological neighbors: Syntactic category matters. *Cognition*, *143*, 77-86. doi:10.1016/j.bcognition.2015.06.003

Demuth, K., & Fee, E. J. (1995). *Minimal prosodic words in early phonological development*. Unpublished manuscript, Brown University and Dalhousie University.

Dietrich, C., Swingley, D., & Werker, J. (2007). Native language governs interpretation of salient speech sound differences at 18 months. *Proceedings of the National Academy of Sciences*, *104*(41), 16027-16031. doi:10.1073/pnas.0705270104

Escudero, P., Benders, T., & Lipski, S. (2009). Native, non-native and L2 perceptual cue weighting for Dutch vowels: The case of Dutch, German, and Spanish listeners. *Journal* *of Phonetics*, *37*, 452–465. doi:10.1016/j.wocn.2009.07.006

Escudero, P., Best, C. T., Kitamura, C., & Mulak, K. E. (2014). Magnitude of phonetic distinction predicts success at early word learning in native and non-native accents. *Frontiers in Psychology*, *5*, 1-11. doi:10.3389/fpsyg.2014.01059

Fenson, L., Pethick, S. J., Renda, C., Cox, J. L., Dale, P. S., & Reznick, J. S. (2000). Short-form versions of the MacArthur Communicative Development Inventories. *Applied Psycholinguistics*, *21*(01), 95–115. doi:10.1017/S0142716400001053

Ferguson, B., Graf, E., Waxman, S. R. (2014). Infants use known verbs to learn novel nouns: evidence from 15- and 19-month-olds. *Cognition*, *131*, 139-146.

Fikkert, P. (1994). *On the acquisition of prosodic structure* (Doctoral dissertation, HIL dissertations 6, Leiden University). The Hague: Holland Academic Graphics.

Golinkoff, R., Hirsh-Pasek, K., Cauley, K., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, *14*, 23-45.

Harrington, J., Cox, F., & Evans, Z. (1997). An acoustic phonetic study of broad, general, and cultivated Australian English vowels. *Australian Journal of Linguistics, 17*, 155-184. doi:10.1080/07268609708599550

Halberda, J. (2003). The development of a word-learning strategy. *Cognition*, *87*(1), B23-B34. doi:10.1016/S0010-0277(02)00186-5

Kehoe, M. M., & Lleó, C. (2003). The acquisition of nuclei: A longitudinal analysis of phonological vowel length in three German-speaking children. *Journal of Child Language*, *30*, 527-556.

Ko, E.-S., Soderstrom, M., & Morgan, J. (2009). Development of perceptual sensitivity to extrinsic vowel duration in infants learning American English. *Journal of the Acoustical Society of America*, *126*(5), EL134-EL139. doi: 0.1121/1.3239465

Mani, N., Coleman, J., & Plunkett, K. (2008). Phonological specificity of vowel contrasts at 18-months. *Language and Speech*, *51*(1&2), 3-21.

Mani, N., & Plunkett, K. (2007). Phonological specificity of vowels and consonants in early lexical representations. *Journal of Memory and Language*, *57*(2), 252–72.

Mani, N., & Plunkett, K. (2011). Does size matter? Subsegmental cues to vowel mispronunciation detection. *Journal of Child Language*, *38*, 606-627.

Markman, E. M. (1984). The acquisition and hierarchical organization of categories by children. In C. Sophian (Ed.), *Origins of cognitive skills* (pp. 371-406). Hillsdale, NJ: Erlbaum.

Markman, E. M. (1987). How children constrain the possible meanings of words. In U. Neiser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization* (pp. 255-287). Cambridge: Cambridge University Press.

Markman, E. M. (1990). Constraints children place on word meanings. *Cognitive Science*, *14*(1), 57-77. doi:10.1016/0364-0213(90)90026-S

Markman, E. M. (1991). The whole-object, taxonomic, and mutual exclusivity assumptions as initial constraints on word meanings. In S. A. Gelman & J. P. Byrnes (Eds.), *Perspectives on language and thought: Interrelations in development* (pp. 72-106). Cambridge: Cambridge University Press.

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, *20*(2), 121-157. doi:10.1016/0010-0285(88)90017-5

Markman, E. M., Wasow, J. L., & Hansen, M. B. (2003). Use of the mutual exclusivity assumption by young word learners. *Cognitive Psychology*, *47*(3), 241-275. doi: 10.1016/S0010-0285(03)00034-3

Merriman, W. E., & Bowman, L. L. (1989). The mutual exclusivity bias in children's word learning (with Commentary by Brian MacWhinney) [Monograph]. *Monographs of the Society for Research in Child Development*, *54*(3-4, Serial No. 220).

Mulak, K. E., Best, C. T., Tyler, M. D., Kitamura, C., & Irwin, J. R. (2013). Development of phonological constancy: 19-month-olds, but not 15-month-olds, identify words in a non-native regional accent. *Child Development*, *84*(6), 2064-2078. doi:10.1111/cdev.12087

Nazzi, T. (2005). Use of phonetic specificity during the acquisition of new words: Differences between consonants and vowels. *Cognition*, *98*, 13-30.

Nooteboom, S. G., & Doodeman, G. J. N. (1980). Production and perception of vowel length in spoken sentences. *Journal of the Acoustical Society of America*, *67*, 276–287.

Sato, Y., Sogabe, Y., & Mazuka, R. (2010). Discrimination of phonemic vowel length by Japanese infants. *Developmental Psychology*, *46*(1), 106-199. doi:10.1037/a0016718

Seidl, A., & Cristià, A. (2008). Developmental changes in the weighting of prosodic cues. *Developmental Science*, *11*(4), 596-606. doi:10.1111/j.1467-7687.2008.00704.x

Swingley, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, *76*, 147-166.

Swingley, D., & Aslin, R. N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, *13*(5), 480-484. doi:10.1111/1467-9280.00485

van der Feest, S. V. H., & Swingley, D. (2009). *Language-specific interpretation of vowel duration in 21-month-olds' word recognition: A cross-linguistic study of phonetic attribution*. Paper presented at the 34th Boston University Conference on Language Development, November, Boston, MA.

van der Feest, S. V. H., & Swingley, D. (2011). Dutch and English listeners' interpretation of vowel duration. *Journal of the Acoustical Society of America*, *129*(3), EL57-EL63. doi:10.1121/1.3532050.

Watson, C. I., & Harrington, J. (1999). Acoustic evidence for dynamic formant trajectories in Australian English vowels. *Journal of the Acoustical Society of America*, *106*(1), 458-468. doi:10.1121/1.427069

White, K. S., & Morgan, J. L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language*, *59*, 114-132.

Yuen, I., Cox, F., & Demuth, K. (2014). Three-year-olds' production of Australian English phonemic vowel length as a function of prosodic context. *Journal of the Acoustical Society of America*, *135*(3), 1469-1479.

Appendix I: Parental Report on Infants' Understanding and Using of the Test Words

A. Experiment 1 - 18 months

	Children Had Understood		Children Had Used	
Word	Count*	Percentage	Count*	Percentage
1. Words used as familiar stimulus in Mispronunciation conditions				
cup	20	100%	4	20%
bus	16	80%	11	55%
bath	19	95%	9	45%
heart	5	25%	4	20%
hat	18	90%	10	50%
tap	10	50%	5	25%
Mean	*14.7*	*73.3%*	*7.3*	*35.8%*
2. Words used as familiar stimulus in Correct/Novel Pronunciation conditions				
sock	20	100%	11	55%
foot	20	100%	8	40%
book	19	95%	12	60%
fork	14	70%	7	35%
Mean	*18.3*	*91.2%*	*9.5*	*47.5%*
Grand Mean	***16.1***	***80.5***	***8.1***	***40.5%***
3. Some actual words used as Mispronunciations				
carp	0	0%	0	0%
cap	1	5%	0	0%
hut	0	0%	0	0%
hurt	5	25%	1	5%
birth	0	0%	0	0%

Note. * There were 20 participants in total.

B. Experiment 2 - 24 months

	Children Had Understood		Children Had Used	
Word	Count*	Percentage	Count*	Percentage
1. Words used as familiar stimulus in Mispronunciation conditions				
cup	20	100%	17	85%
bus	18	90%	16	80%
bath	20	100%	16	80%
heart	14	70%	12	60%
hat	19	95%	20	100%
tap	17	85%	11	55%
Mean	*18*	*90%*	*15.3*	*76.7%*
2. Words used as familiar stimulus in Correct/Novel Pronunciation conditions				
sock	20	100%	20	100%
foot	20	100%	18	90%
book	20	100%	20	100%
fork	20	100%	16	80%
Mean	*20*	*100%*	*18.5*	*92.5%*
Grand Mean	***18.8***	***94%***	***16.6***	***83%***
3. Some actual words used as Mispronunciations				
carp	1	5%	0	0%
cap	6	30%	3	15%
hut	1	5%	1	5%
hurt	17	85%	10	50%
birth	0	0%	0	0%

Note. * There were 20 participants in total.

Appendix II: Language Background Questionnaire

Child's Name ____________________ [Participant No.______]

1. Where do you live? (your suburb, your neighborhood)
____

2. How long have you lived in that neighborhood? _____________________________________

3. Where did the child's principle caregiver (Circle one: mother/father/other________) grow up?
_____________________________________ (Specify: state, city/town, suburb, country)

4. Where did the child's second caregiver (Circle one: mother/father/other_________) grow up?
_____________________________________ (Specify: state, city/town, suburb, country)

5. What is the child's principle caregiver's highest level of educational attainment?
(Circle one: primary school, high school year 10, high school year 12, TAFE Certificate or equivalent, university degree, postgraduate degree)

6. What is the child's second caregiver's highest level of educational attainment?
(Circle one: primary school, high school year 10, high school year 12, TAFE Certificate or equivalent, university degree, postgraduate degree)

7. Does your child have siblings? If so, how old are they?
____

8. Is your child exposed to languages other than English? YES/NO
Language: _____________________________________ Hours: ____________
Language: _____________________________________ Hours: ____________

9. How many hours a week is your child in daycare/preschool outside the home? ___________

10. How many hours a week does your child watch TV and in what languages?
Language: _____________________________________ Hours: ____________
Language: _____________________________________ Hours: ____________

11. How many hours a week is your child read to (at home and daycare)?
Home - Language: _____________________________ Hours: ____________
Daycare - Language: _____________________________ Hours: ____________

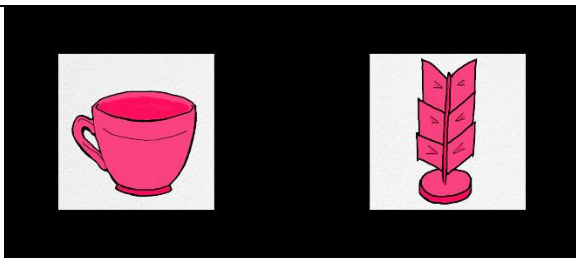
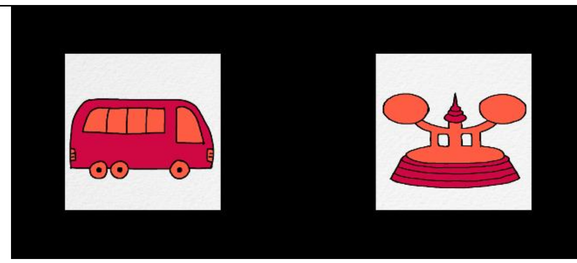
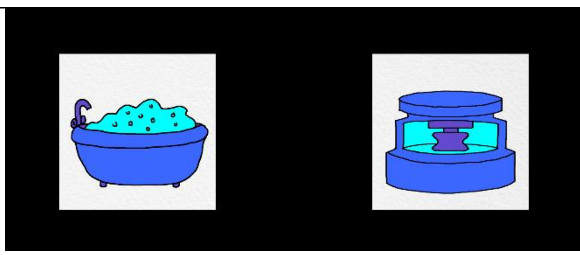
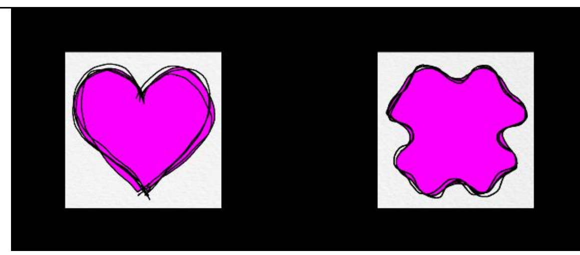
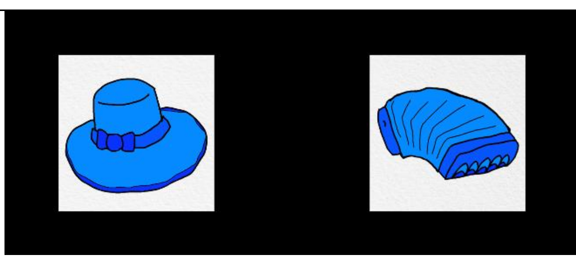
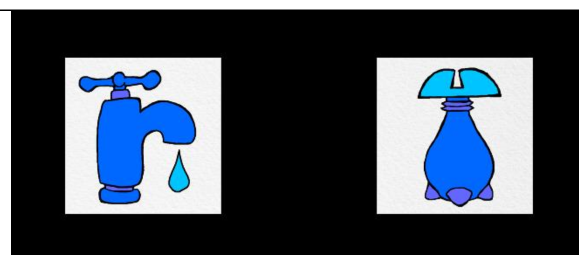
12. Does your child exhibit any language difficulties? ___________________________________

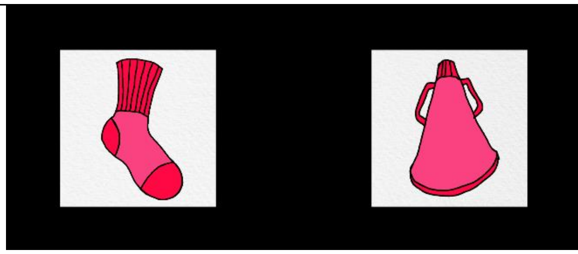
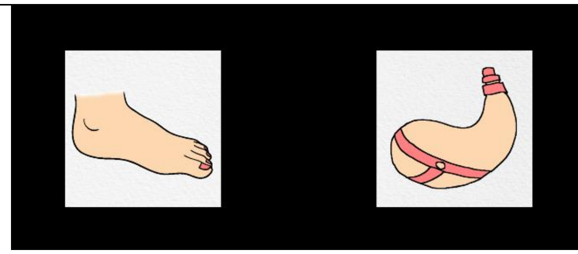
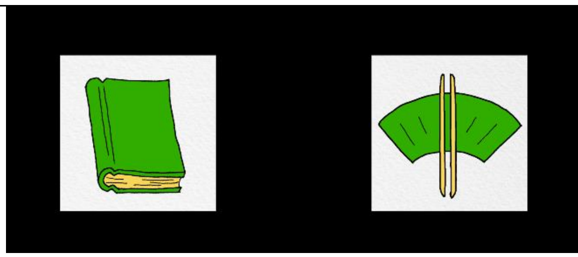
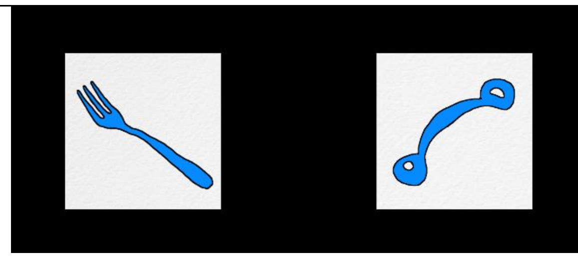
13. Does anyone else in your family exhibit language, hearing or reading difficulties? Yes/NO
If so, please explain ____

Thank you!

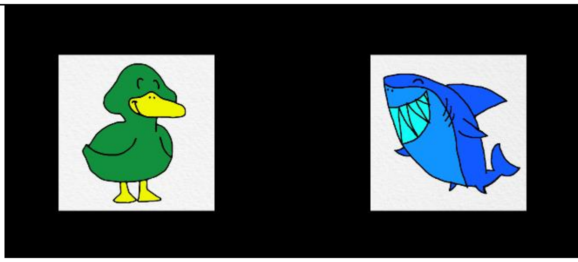
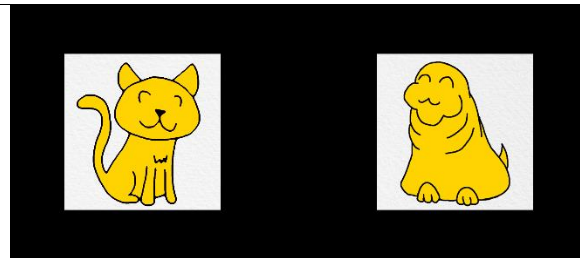
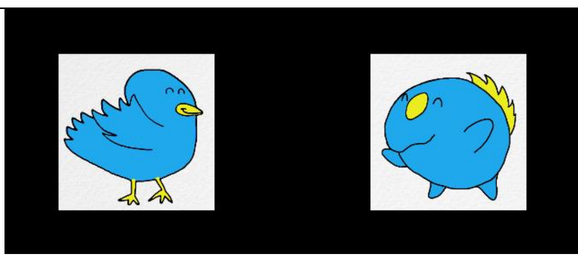
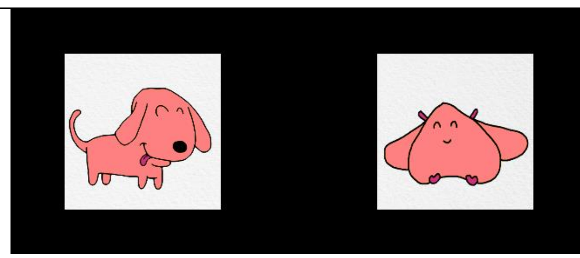
Appendix III: Familiar-Novel Picture Pairs Used as Visual Stimuli in the Experiments

A. Ten picture pairs for test items (demonstrated with all familiar items on the left)

Mispronunciations	
	
cup /kɐp/ - ***carp*** **[kɐːp]** or ***cap*** **[kæp]**	bus /bɐs/ - ***barss*** **[bɐːs]** or ***bass*** **[bæs]**
	
bath /bɐːθ/ – ***buth*** **[bɐθ]** or ***birth*** **[bɜːθ]**	heart /hɐːt/ – ***hut*** **[hɐt]** or ***hirt*** **[hɜːt]**
	
hat /hæt/ – ***het*** **[het]** or ***hut*** **[hɐt]**	tap /tɐp/ – ***tep*** **[tep]** or ***tup*** **[tɐp]**

Correct and Novel Pronunciations	
	
sock /sɔk/ – ***sock*** **[sɔk]** or ***nart*** **[nɐːt]**	foot /fʊt/ – ***foot*** **[fʊt]** or ***gup*** **[gɐp]**
	
book /bʊk/ – ***book*** **[bʊk]** or ***geap*** **[giːp]**	fork /foːk/ – ***fork*** **[foːk]** or ***noss*** **[nɔs]**

B. Four picture pairs for training and filler items

Training	
Correct Pronunciation	Novel Pronunciation
	
duck/shark – ***duck*** or ***shark***	cat – ***nep***
Filler	
Correct Pronunciation	Novel Pronunciation
	
bird – ***bird***	dog – ***dirk***

Chapter Four: Effect of Early Dialectal Exposure on Adult Perception of Phonemic Vowel Length

This chapter is based on the following paper in submission for publication:

Chen, H., Xu Rattanasone, N., Cox, F., & Demuth, K. (in submission). Effect of early dialectal exposure on adult perception of phonemic vowel length. *Journal of the Acoustical Society of America*.

All components of this paper, both experimental and written, have been completed by me, with advice from the co-authors (my supervisors) when needed.

Abstract

Attunement to native phonological categories and the specification of relevant phonological features in the lexicon occur early in development for mono-lingual/dialectal speakers. However, few studies have investigated whether and how early exposure to two dialects of a language continuously might influence the development of phonological categories, especially when a phonemic contrast exists only in one dialect. This study therefore compared perceptual sensitivity to mispronunciations in phonemic vowel length in Australian English adult listeners with and without early exposure to another English dialect that did not have this contrast. The Intermodal Preferential Looking paradigm and an eye-tracker were used to record looking behavior. The results showed that, while both mono- and bi-dialectal groups were sensitive to mispronunciations in vowel length, the bi-dialectal adults were more likely to accept a mispronunciation in vowel length as the target compared to mono-dialectal adults. The bi-dialectal group were also more tolerant to mispronunciations in vowel length than in vowel height and vowel backness. These results reveal that the bi-dialectal Australian English adults have a contrastive vowel length feature that is less specified in the lexicon compared to mono-dialectal adults. The findings suggest a complex influence of early exposure to another dialect on the development of phonological categories.

Key words: bi-dialectal adults, perception, phonemic vowel length, Australian English

Effect of Early Dialectal Exposure on Adult Perception of Phonemic Vowel Length

Introduction

It is widely established that infants have become attuned to native phonetic categories by the end of the first year of life and have attenuated their ability to discriminate a wide-range of nonnative speech sounds (Burnham, 1986; Burnham, Earnshaw, & Quinn, 1987; Kuhl, Williams, Lacerda, Steven, & Linblom, 1992; Polka & Werker, 1994; Werker & Polka, 1993; Werker & Tees, 1983, 1984a, 1984b; among others). When entering the word learning stage, infants begin to learn spoken word forms and display perceptual sensitivity to the phonetic variations that can change word meaning in their native language, i.e., phonemic contrasts. Previous studies have shown that infants have tuned into the contrastive function of native vowels and consonants by 14-15 months (Fennell & Waxman, 2010; Fennell & Werker, 2003; Mani & Plunkett, 2007; Swingley & Aslin, 2002; among others), and their sensitivity becomes reliably robust by 18-20 months (Mani, Coleman, & Plunkett, 2008; Mani & Plunkett, 2007; Nazzi, 2005; Swingley, 2003, 2009; Swingley & Aslin, 2000; White & Morgan, 2008). However, to develop a mature phonological system, language learners not only need to be able to recognize which phonetic variations are contrastive, but also need to understand which phonetic variants do NOT alter word meaning, i.e., *phonological constancy* (Best, Tyler, Gooding, Orlando, & Quann, 2009). Recent studies have revealed that infants develop phonological constancy only by 19 months, which is evidenced by their ability to recognize familiar words in a nonnative regional/artificial accent (Best et al., 2009; Mulak, Best, Tyler, Kitamura, & Irwin, 2013; White & Aslin, 2011; but see also van Heugten & Johnson, 2014 for a suggestion that it might be achieved earlier at 15 months).

If the above is taken as the typical developmental trajectory of phonological categories in monolingual and mono-dialectal children, it is of theoretical interest to understand whether and how continuous exposure to two dialectal accents of one language simultaneously at an early age might influence the development of long-term phonological categories, especially when one dialect has a phonemic contrast that is not contrastive in the other. However, little is known about this issue as most research has focused on early phonological development on infants (Durrant, Delle Luche, Cattani, & Floccia, 2015; Floccia, Delle Luche, Durrant, Butler, & Goslin, 2012; van der Feest & Johnson, 2016). The current study therefore aims to examine the effect of early exposure to single versus multiple dialects on the development of vowel contrasts in adult listeners.

Exposure to native language input leads to the development of language-specific phonological representations. As far as the perception of vowel categories is concerned, infants' perception pattern shifts from language-general to language-specific by as young as six months of age, and around 10 months for native consonant categories (Kuhl et al., 1992; Polka & Werker, 1994). For instance, it has been shown that English-learning 6-month-olds could not discriminate the German vowel contrasts that 4-month-olds could (Polka & Werker, 1994). This indicates that infants have tuned into their native vowel categories and lost the sensitivity to nonnative contrasts after being exposed to six months of native input. On the other hand, some native contrasts are enhanced in infants' perception. For example, Japanese infants become sensitive to vowel duration contrasts between around 7-9.5 months (Sato, Sogabe, & Mazuka, 2010), showing the facilitative effect of native input (see Burnham et al. (1987) for a review on consonants).

When infants begin to learn words, they display sensitivity to phonetic detail first with *familiar* words compared to novel words. For example, previous studies have shown that American and British English-learning 14-15-month-olds can recognize both vowel and consonant changes in familiar words using the Intermodal Preferential Looking (IPL, Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987) paradigm (Mani & Plunkett, 2007; Swingley & Aslin, 2002). At 18 months, British English-learning infants can detect a 1-feature vowel quality change (i.e., vowel height or backness) in familiar words (Mani et al., 2008), and Japanese-learning infants become sensitive to a vowel length change in newly learned words (Chen, Yamane, Xu Rattanasone, Demuth, & Mazuka, 2015). Thus, by 18 months, infants have well specified phonological contrasts for most vowel features in their lexicon.

Although infants' recognition of cross-category variation and mastery of phonological contrasts are reported to be a critical developmental step, the ability to attenuate within-category variation and the understanding of phonological constancy are also regarded to be a milestone in developing abstract phonological categories (Best, 1994; Best et al., 2009). While infants gradually tune into specific phonological contrasts with the aid of exposure to native language input in the early years, they also develop top-down abstraction ability as they learn more words, which helps them to not over-specify those within-category variants.

During their first year of development, infants' sensitivity to the variation between native and nonnative dialectal accents is attenuated at around 6-9 months. At 5 months, English infants can discriminate native and nonnative dialects (e.g., American English vs. British English, Nazzi, Jusczyk, & Johnson, 2000; or South-West English vs. Welsh English, Butler, Floccia, Goslin, & Panneton, 2011), but fail if both dialects are unfamiliar (e.g., Welsh English vs. Scottish English to South-West English infants, Butler et al., 2011). However, it is reported that infants lose the differential attention to spontaneous speech produced in their native dialect and an unfamiliar nonnative dialect (e.g., American English vs. Australian English, or Australian English vs. South African English) at 8-9 months, in either a visual fixation preference or visual habituation task (Kitamura, Panneton, & Best, 2013; Kitamura, Panneton, Diehl, & Notley, 2006). If the nonnative dialect is familiar to the infants (e.g., American English to Australian infants), the loss of differential attention happens even earlier, at 6 months (Kitamura et al., 2006; Kitamura et al., 2013). These findings indicate that, by the end of the first year, infants have the ability to recognize their native language even if it is spoken in an unfamiliar accent with which they have had little experience.

Infants also develop the ability to ignore within-category phonetic variation in the second year of life when they are learning words. Best et al. (2009) examined 15- and 19-month-old American English infants' preference for familiar vs. unfamiliar words in their native dialect and in a nonnative Jamaican English. They found that while both groups preferred listening to familiar words in the native dialect, only the 19-month-olds demonstrated the same preference while listening to the Jamaican dialect, indicating the older infants could identify familiar words in an unfamiliar dialect. Mulak et al. (2013) later tested Australian English infants at the same ages in a word recognition task using the IPL paradigm, and again found that only 19-month-olds accepted the Jamaican English words as target familiar items, while the 15-month-olds did not. These results indicate that, only by 19 months, have infants developed phonological constancy, permitting within-category variation along with increased normalization abilities. This is consistent with the results from White and Aslin (2011) who manipulated vowel quality from /ɑ/ to /æ/ for familiar words, and tested 19-20-month-old American English infants using a word recognition task in the IPL paradigm. The infants who had exposure to the vowel shift before the test accepted the mispronunciations, whereas those who had no such previous exposure or had exposure to other vowel shifts did not. Furthermore, these effects extended from the particular items heard in the exposure phase to words sharing the same vowels in the test phase. Interestingly, a recent study has also found that if the 15-month-olds were first exposed to familiar stories read in an unfamiliar accent prior to test, their inability of recognizing the accented words can readily be overcome (van Heugten & Johnson, 2014). Taken together, these findings indicate that infants have gradually developed stronger ability for top-down generalization during 15-19 months of age, and can adapt to a novel accent.

Taken together, these results provide a developmental trajectory for the acquisition of native phonological categories in mono-lingual/dialectal infants. However, it remains unclear whether infants with prolonged and repeated natural exposure to two dialects simultaneously will develop the same phonological categories of a target language as mono-dialectal populations. This is an especially interesting question when the two dialects employ different phonemic contrasts to distinguish a set of words; it is unclear whether the bi-dialectal population will specify *both* phonological contrasts represented in the two native accents, specify only *one* set of the contrasts, or not specify *any* of the contrasts given the inconsistent category boundaries in the input. Therefore, understanding how bi-dialectal populations specify competing phonological features in the lexicon can help inform the role that native input plays in the development of both phonological and lexical representations.

Nevertheless, there have been very few studies on the phonological development of bi-dialectal populations. Floccia and colleagues (2012) were one of the first to compare phonological specification in word recognition in infants who have consistent exposure to one versus two dialectal variants. They examined mono-dialectal and bi-dialectal 20-month-old English infants who were raised in a rhotic community, and compared their sensitivity to rhoticity using a word recognition task and the IPL paradigm. They found that infants who had been exposed to both a rhotic and a non-rhotic accent could recognize the target familiar words only when hearing them spoken in the rhotic accent, but not the non-rhotic one. This result was therefore similar to that of the mono-dialectal infants. They thus suggested that bi-dialectal infants with exposure to accentual variations of words encoded only a single canonical form in their lexicon, and failed to recognize the non-canonical variant. However, rhoticity is an allophonic variation rather than a contrastive feature in English dialects, i.e., it does not typically change the lexical identity of an English word.

In contrast to Floccia et al. (2012), Durrant et al. (2015) used a *mispronunciation detection* task together with the IPL paradigm to investigate whether infants with mono- versus bi/multi-dialectal exposure would have a similar degree of phonological specificity for familiar words. They tested two groups of English 20-month-olds with familiar words, either correctly pronounced, or mispronounced in the initial consonant or vowel, by the same speaker of a local South West English dialect. The mispronunciations were all phonemically contrastive with the correct pronunciations in the local South West dialects, as well as in the other English dialects that the bi/multi-dialectal infants had been exposed to. While the mono-dialectal group could consistently detect a mispronunciation in the familiar words, the bi/multi-dialectal group did not differ in looking times in the correct versus mispronunciation conditions. This suggests that early representations of familiar words for bi-dialectal 20-month-olds might be phonologically less well specified, or contain phonetically more relaxed boundaries, due to the impact of consistent exposure to dialectal variability. However, it is unclear whether this is due to a delayed mastery of phonological contrasts compared to their mono-dialectal peers, or the result of category broadening as a strategy to accommodate any possible variation.

A recent study by van der Feest and Johnson (2016) has reported different findings which suggest that bi-dialectal infants are able to adapt their speech processing strategies to suit speakers of different dialects. Again using the mispronunciation detection task and the IPL paradigm, van der Feest and Johnson looked at four groups of 24-month-old Dutch infants. The two mono-dialectal groups were only exposed to the community dominant dialect where fricative voicing contrasts have disappeared word initially and become with-category variations, while the two bi-dialectal groups were *also* exposed to the dialects where the contrasts are maintained cross-categorially. The results showed that, for the two bi-dialectal groups, the one tested with a local devoicing dialect ignored the fricative voicing contrast, whereas the one tested with a dialect maintaining the contrasts was able to detect a mispronunciation in fricative voicing. These results suggest that the bi-dialectal infants do not simply treat fricative voicing as allophonic, as did their mono-dialectal peers, but are very adaptive in associating the contrasts with word meaning according to the different dialects.

The findings of van de Feest and Johnson (2016) suggest that by 2 years of age, bi-dialectal infants are able to underspecify a phonemic feature (fricative voicing) in one dialect, but well specify the feature in another dialect. This suggests flexibility in signal-to-word mapping processes more than flexibility in forming categorical boundaries. This seems to be contrary to Durrant and colleagues' findings showing a general phonetic boundary relaxation in bi-dialectal infants at a younger age. However, given that infants at either 20 or 24 months are still young and might not have a fully developed phonological system, it would be interesting to know whether the observed effects persist, and how a mature bi-dialectal population might encode the phonological features in the two native dialects. Note that fricative voicing in Dutch dialects is either contrastive or neutralized, so there are no competing contrastive features involved in the two dialectal variants. Hence, it remains unclear how a bi-dialectal population would specify competing phonological contrasts in the lexicon. Furthermore, the perception of vowels is less categorical than that of consonants, and might show more gradient effects.

Studies from L2 adult listeners have shown that vowel duration might be a default cue for adults (if available) to differentiate vowels with small spectral differences that are not relevant in their native language (Bohn, 1995; Bohn & Polka, 2001). However, as L2 adults do not have early and continuous exposure to the other language, their phonological perception might be different from those who have native input from birth. To better explore this issue, the present study therefore compared the responses of native Australian English adult listeners with and without early exposure to other dialects of English, focusing on their perception of phonemic vowel length in Australian English.

English dialects differ greatly in the number and type of vowels in the phonemic inventory (Wells, 1982). For instance, Standard Southern British (SSB) English has 11 monophthongs (excluding schwa), and has several overlaps between vowels (see Figure 1), whereas General American English has 10 (Ladefoged, 2001; Peterson & Barney, 1952) or 12 monophthongs (Hillenbrand, Getty, Clark, & Wheeler, 1995; see Figure 2). In the SSB accent of English, the low vowels /ɑː/ vs. /ʌ/ have a durational difference, but with an additional prominent spectral quality distinction (Williams & Escudero, 2014). Likewise, in General American English, /ɑ/ is phonetically longer than /ʌ/, but again with a primary difference in spectral quality (Clopper, Pisoni, & de Jong, 2005). Australian English, however, differs from many other English dialects; it has a phonemic vowel length contrast with minimal spectral difference for a subset of vowels – /ɐ/ vs. /ɐː/ and /e/ vs. /eː/[1] (Bernard, 1967; Cox, 2006; Cox & Palethorpe, 2007; Watson & Harrington, 1999; see also Figure 3). For example, the word *hut* /ɐ/ can be differentiated from the word *heart* /ɐː/ by duration only. Similarly, the contrast between *shed* /e/ and *shared* /eː/ is also based on vowel duration alone for many speakers.

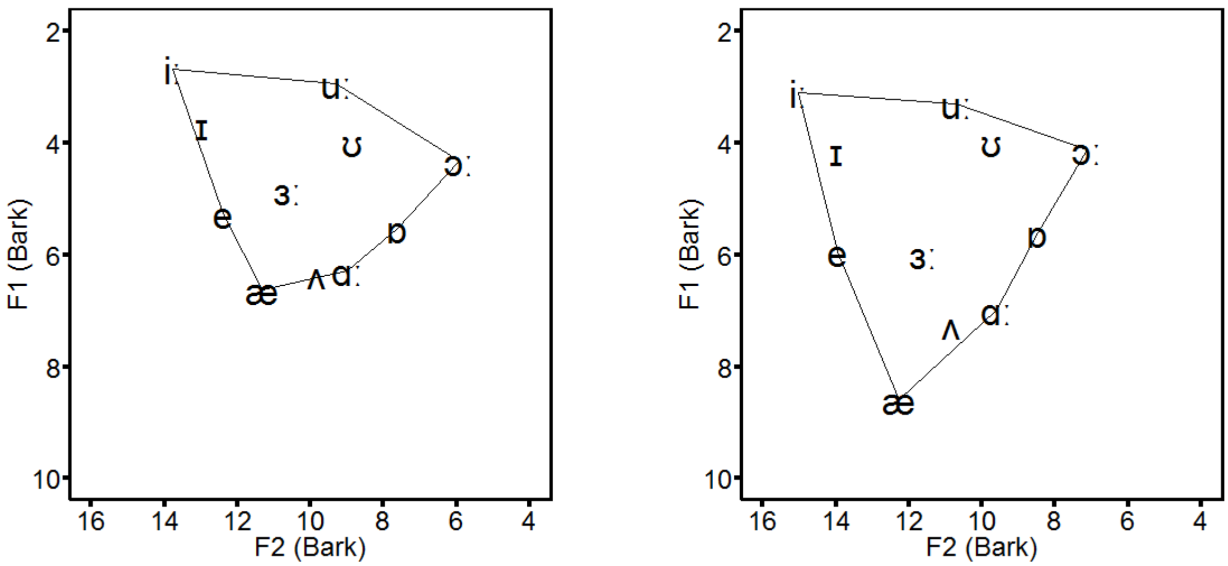


Figure 1. British English: Monophthongs vowel formant plots of Standard Southern British based on citation form /hVd/ words from 8 male speakers (left) and 8 female speakers (right) (from values of Deterding (1990) presented in Deterding (1997)).

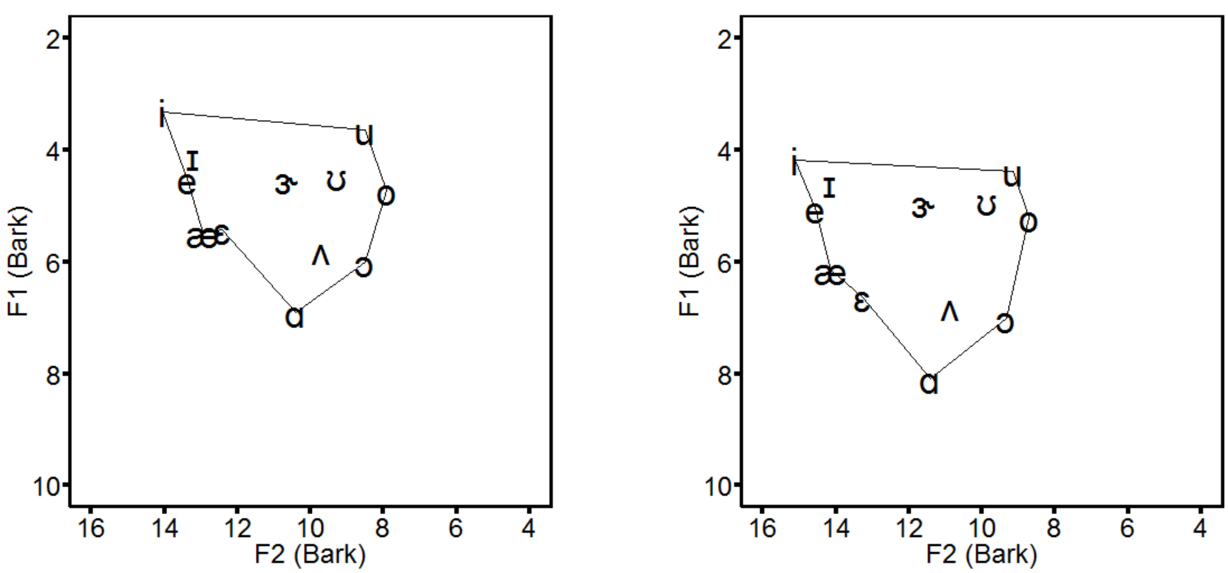


Figure 2. American English: Monophthong vowel formant plots of General American based on citation form /hVd/ words from 45 male speakers (left) and 48 female speakers (right) (from values presented in Hillenbrand et al. (1995)).

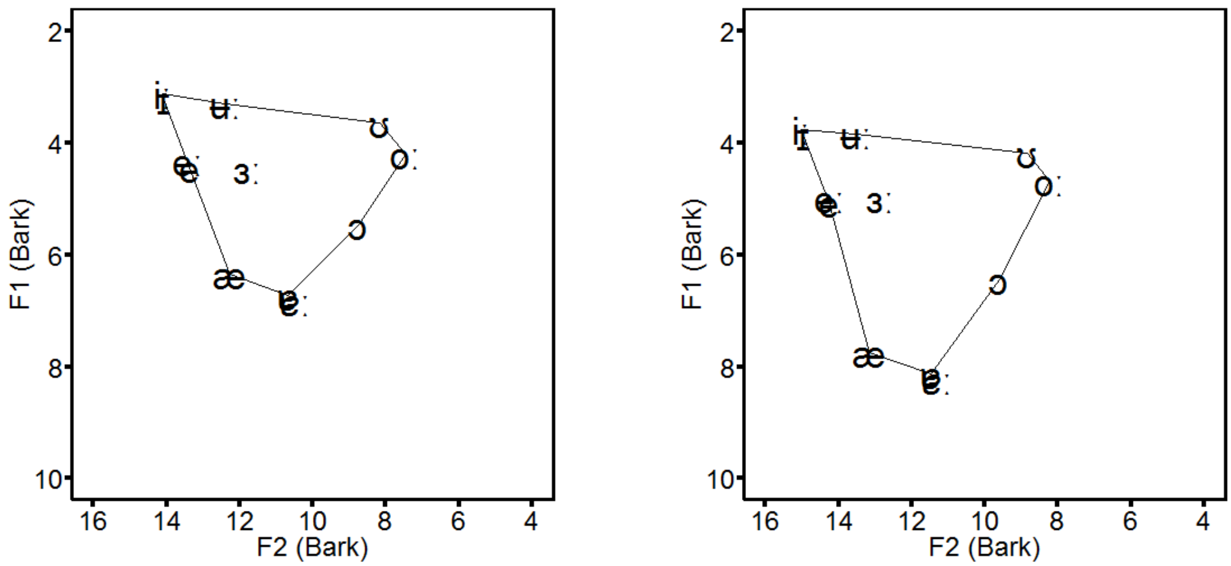


Figure 3. Australian English: Monophthong vowel formant plots based on citation form /hVd/ words from 60 male speakers (left) and 60 female speakers (right) (from values presented in Cox (2006)).

A recent study has shown that mono-dialectal Australian English-learning infants can identify mispronunciations in vowel length in familiar words by 24 months in an IPL task (Chen, Xu Rattanasone, Cox, & Demuth, 2015). They were also sensitive to a set of well-controlled contrasts in vowel height (/e/ - /æ/ and /ɜː/ - /ɐː/) and vowel backness (/æ/ - /ɐ/), both with a small (1-degree) quality difference. The current study employed a similar method to test Australian English adult listeners' sensitivity to mispronunciations in these three vowel features. The IPL paradigm has successfully used to test adults' phonological sensitivity (e.g., Shatzman & McQueen, 2006), so it is appropriate for our study on adults as well. As mature language users, Australian English adult listeners have a well-developed phonological system, regardless of whether they had early exposure to another English dialect or not. By comparing their perceptual responses, we wanted to answer the following questions: a) whether the bi-dialectal group were sensitive to phonemic vowel length contrasts; b) whether the bi-dialectal group were more tolerant to mispronunciations in phonemic vowel length compared to the mono-dialectal group; and c) whether the bi-dialectal group were more tolerant to mispronunciations in phonemic vowel length than vowel height and backness compared to the mono-dialectal group. The answers will help us better understand how early exposure to a dialect without contrastive vowel length might influence sensitivity to vowel length. This in turn will provide further insight into the role of native input in the development of phonological categories.

Method

Participants

Two groups of Australia-born Australian English-speaking adults were recruited and participated in Sydney, Australia. One group (n = 15, 4 males, 11 females; M_{age} = 21 years, range = 18-31 years), identified as 'AusE', were monolingual Australian English speakers whose parents were both born in the Greater Sydney area and spoke only Australian English in daily life. The other group (n = 10, 5 males, 5 females; M_{age} = 22 years, range = 18-26 years), identified as 'AusE+', were also monolingual Australian English speakers but with at least one parent who was a native speaker of another English dialect (either rhotic, or having prominent spectral differences in short/long vowel pairs). These included American English (n = 3), British English (n = 3), Scottish English (n = 1), Irish English (n = 1), Maltese English (n = 1), and Singapore English (n = 1). Three additional participants were tested but excluded from the final analysis due to low eye-tracking sample rate (=1%, n = 1; from AusE+), and inability to provide data on at least one trial per condition tested (n = 2; 1 for each group). All participants reported that they spoke Australian English only and had very limited knowledge of a second language. Participants did not have any reported language or hearing problems, and had minimal exposure to a second language other than English (< 1 hr per day).

Stimuli and Design

Participants were tested with 20 monosyllabic CVC items with a voiceless coda consonant (see Table 1). Four were correct pronunciations (CP) of familiar English words. Another twelve items were mispronunciations of familiar English words with only a single vowel feature change – four in vowel height (VH_{MP}), four in vowel backness (VB_{MP}), and four in vowel length (VL_{MP}). The remaining four items were novel pronunciations (NP) of familiar words, with multiple segments different from that of the familiar word.

All test items were presented as the labels of objects with a definite article *the* in the experiment, such that they were interpreted as nouns. All the mispronunciations and novel items resulted in either nonce words, low frequency words, or typically used as verbs.

In addition, two familiar stimuli (*bird, bed*) were used in the training, with a novel pronunciation *nep* paired with *bed*. Another four familiar nouns (*sheep, rat, cheese, ball*) were used as fillers and always pronounced correctly during the test, so participants heard eight correct pronunciations in total during the entire experiment. This was done so that throughout the test each participant was equally likely to hear mispronunciations and correct pronunciations. Performance on these filler items was not analyzed.

Table 1. *Pronunciation conditions, familiar words and test (mis)pronunciations*

Pronunciation Condition	Familiar	Test	Familiar	Test
Correct Pronunciation (CP)	(mat)	mat [mæt]	(foot)	foot [fʊt]
	(fork)	fork [foːk]	(nurse)	nurse [nɜːs]
Vowel Height Misp. (VH_{MP})				
/æ/ → /e/	(cat)	ket [ket]	(hat)	het [het]
/ɐː/ → /ɜː/	(heart)	hurt [hɜːt]	(park)	pirk [pɜːk]
Vowel Backness Misp. (VB_{MP})				
/ɐ/ → /æ/	(bus)	bass [bæs]	(duck)	dack [dæk]
/æ/ → /ɐ/	(bat)	butt [bɐt]	(tap)	tup [tɐp]
Vowel Length Misp. (VL_{MP})				
/ɐ/ → /ɐː/	(pup)	parp [pɐːp]	(nut)	nart [nɐːt]
/ɐː/ → /ɐ/	(bath)	buth [bɐθ]	(shark)	shuck [ʃɐk]
Novel Pronunciation (NP)	(book)	geap [giːp]	(sock)	mirt [mɜːt]
	(horse)	dirk [dɜːk]	(shirt)	gup [gɐp]

Note. VH_{MP} = Vowel Height Mispronunciation; VB_{MP} = Vowel Backness Mispronunciation; VL_{MP} = Vowel Length Mispronunciation. Australian English is non-rhotic and therefore none of these words is produced with an /ɹ/ or rhoticised vowel.

Auditory stimuli. The auditory stimuli were recorded by a native Australian English female speaker in a child-friendly manner[2]. Recordings were made in an acoustically shielded recording booth, sampled at 48 kHz. Both the familiar words and the test items were recorded three times in a carrier sentence: "*Look at the [Target]!*" The best token for each target item was chosen based on the judgments of all four authors (the third is a native speaker of Australian English), and used as the final auditory stimulus in the experiment.

The acoustics of the vowels used in the mispronunciation conditions (and correct pronunciations) are shown in Table 2. The mean duration of the short vowels was 237 ms (*SD* = 18 ms, range = 210-260 ms), while the mean duration of the long vowels was 483 ms (*SD* = 34 ms, range = 423-543 ms). The short and long vowels maintained a constancy of 1:2 durational contrast with no overlap. These durations were also consistent with those reported for child-directed speech in a previous study on Australian English (Yuen, Cox, & Demuth, 2014). The F1 and F2 values in the target vowel quality contrasts were also consistent with previous studies (e.g., Cox, 2006). Acoustic measurements were completed using Praat (Boersma & Weenink, 2012). Due to the very high pitch used in the child-friendly manner (max. 330-375 Hz), formant tracking in Praat was largely influenced by the first harmonic (F0) and unreliable. Therefore, the formant measurements for each token were based on the spectral slice obtained from the point where the formants were resolved once the pitch returned to a value below 300 Hz. The spectral slice was taken no later than 70% of the vowel.

Table 2. *Mean durations (ms) and F1 and F2 values (Hz) for the vowel stimuli used in the mispronunciation conditions (with range)*

Short			Long		
Vowel	Duration	Formants	Vowel	Duration	Formants
/e/	225 (210-239)	*F1*: 780 (771-789) *F2*: 2456 (2446-2456)	/ɜː/	460 (458-461)	*F1*: 769 (744-749) *F2*: 1573 (1461-1685)
/æ/	254 (248-260)	*F1*: 1125 (1102-1147) *F2*: 1954 (1864-2043)			
/ɐ/	234 (218-253)	*F1*: 1058 (1013-1108) *F2*: 1622 (1593-1702)	/ɐː/	465 (423-506)	*F1*: 1036 (1013-1058) *F2*: 1506 (1505-1506)
Grand Mean	237 (210-260)		Grand Mean	483 (423-543)	

Note. The vowel /ɐ/ had four tokens; all others had two.

Visual stimuli. Cartoon style clip art pictures were created for each familiar stimulus and judged by the authors and their colleagues as typical exemplifiers of the corresponding objects. A novel object was created for each familiar stimulus and matched in style, colour, shading, size and visual complexity. These were yoked as familiar-novel pairs. All the novel objects were also assessed for their novelty. During the experiment, the yoked familiar-novel cartoon pictures were depicted against two off-white background frames, and displayed side-by-side horizontally on a black screen (see example in Figure 4). A full set of the 20 picture pairs for the test items and the six pairs for training and fillers is shown in Appendix I.

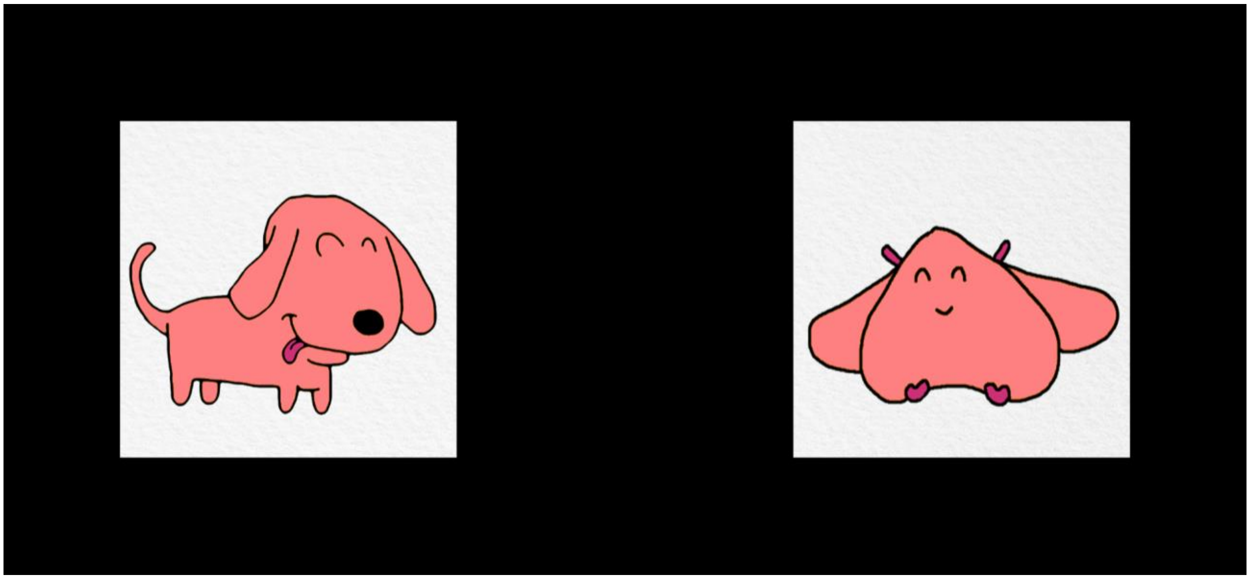


Figure 4. An example of the yoked cartoon pictures: *pup – parp [ɐː]*

Apparatus

The entire experiment took place in a sound attenuated test room. A Tobii TX300 Eye Tracker was used to record participants' looking behaviours. The eye-tracker was down-sampled to 120 Hz[3], tilted at 30°, collecting gaze data from both eyes with a 10-165 ms recovery time for lost tracking. Visual stimuli were shown in the original 23" screen unit containing the built-in eye-tracker, and displayed at 1920 x 1080 pixels. In the IPL task, the two off-white background frames containing the yoked stimulus pictures were displayed at a size of 13.4 x 13.4 cm each on the screen, at a distance of 13.4 cm from each other, providing a minimal 11° and a maximal 32° horizontal gaze angle from the participant who sat approximately 70 cm in front of the screen. The auditory stimuli were delivered at a conversation level (≈ 65 dBA) from two computer speakers located on both sides of the screen. A Panasonic digital video camera with a zoom lens was positioned beside the eye-tracker to record participant's' looking behaviour as a secondary recording of the sessions.

The IPL task was conducted using the Tobii Studio software which controlled the presentation of stimuli and test trials. Each trial was delivered in a video (.AVI) file encoded in JPEG codec 3.2.4 at 24 frames-per-second. The video files were created using the FinalCut Pro software on a Macintosh computer by integrating the corresponding auditory and visual stimuli.

After the test, participants completed a language background questionnaire (Appendix II).

Procedure

The study comprised of a familiarization session and an IPL test session. Before starting the sessions, all participants were first instructed that they were going to participate in a study of Australian English.

Familiarization. To ensure participants could easily and correctly identify the familiar pictures with the target familiar words (e.g., an image of a dog often referred to as 'doggie' in child-directed speech was referred to as '*pup'* in this study), they were first familiarized with real photos and labels of all the tested familiar stimuli. Participants' were sitting approximately 65 cm in front of the eye-tracker and presented with photos of the familiar stimuli one-by-one on the screen in 5-second intervals. Real photos were used for this familiarization phase to maintain the visual novelty of the cartoon pictures in the later test. Participants also listened to the recording of each familiar word while looking at the corresponding photo, which was excised to contain only "*the [Target]*". This familiarized participants with each picture and the associated familiar word label. Each photo and associated recording was played to the participants only once.

Test. The test session contained 26 trials, including two training trials, four test trials for each pronunciation condition (i.e., 4 VH_{MP}, 4 VB_{MP}, 4 VL_{MP}, 4 CP, 4 NP), and four filler trials (4 CP). In the two training trials, participants saw the yoked pair of a familiar object paired with a novel object on the screen. They heard the correct pronunciation of the familiar object in the first trial and a novel pronunciation in the second. At the end of the training trials, the correct picture danced to cheerful music while the incorrect one stayed still. This feedback helped to cue participants to doing the task, by orienting their looks to a familiar object when hearing a familiar word, and to a novel object when hearing something other than the label of a familiar object.

In the test trials, participants saw pictures of a familiar and a novel object presented side-by-side on the screen. After 4 seconds, the two pictures were replaced with a looming red ball in the middle of the black screen. The gaze-centering period lasted for 1 second before the looming ball disappeared and the same set of pictures were displayed on the screen again. The auditory stimulus then started: "*Look at the [Target]!*", with the vowel onset of the target test item aligned with the offset of 5.875 seconds into the trial. The trial ended 4 seconds after the vowel onset of the target test item, with the entire trial lasting for 9.875 seconds and the entire test session lasted about 5 minutes for each participant (see Figure 5 for an example of a test trial).

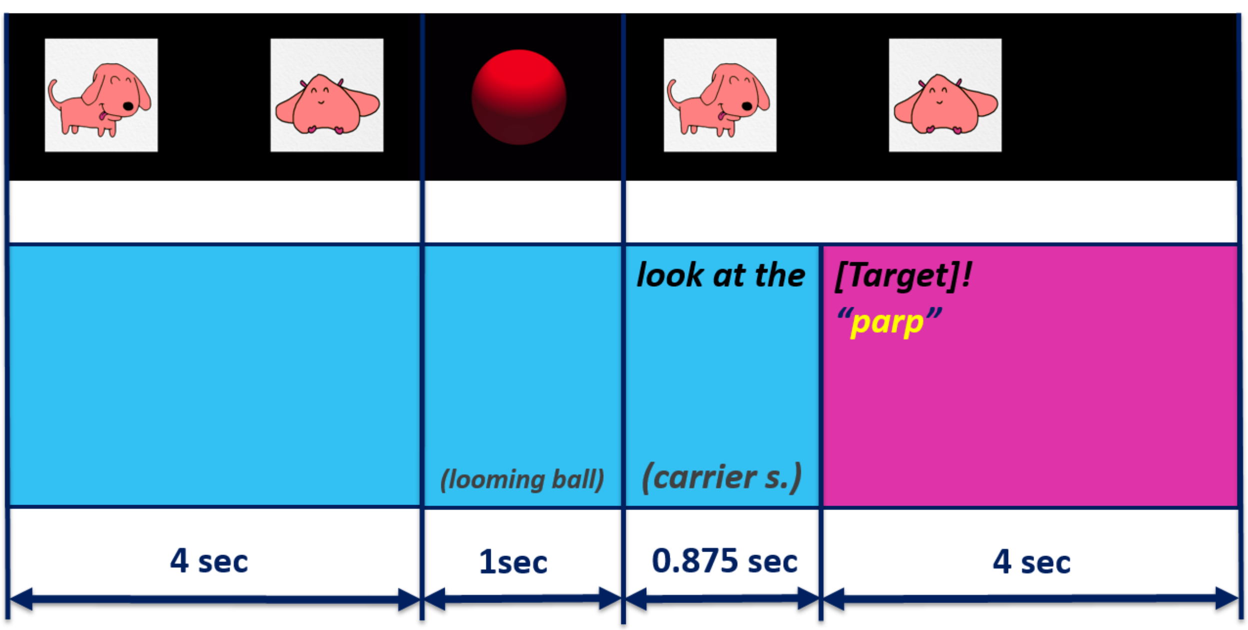


Figure 5. An example of a test trial: "*Look at the parp [ɐː]!*"

Participants were randomly assigned to four test versions. The order of the test trials in each version was pseudo-randomized so that no two consecutive trials had the same condition. The trial orders in version 3 and 4 were the revised orders of versions 1 and 2, respectively. Familiar and novel pictures appeared equally often on the left and right side of the screen in each version, and were counterbalanced across all four versions. The familiar pictures did not appear more than twice on the same side in consecutive trials.

Data Analysis

Gaze position on each trial was recorded by the Tobii TX300 Eye Tracker approximately every 8.33 ms. Recordings of the looking data were converted into fixations using the I-VT fixation filter in Tobii Studio (version 3.2.3). Missing data points were interpolated for sections having a duration below 60 ms, and fixations less than 75 ms were discarded. Areas of Interest (AOIs) were defined as two 17 x 17 cm squares covering the off-white background frames in each trial, given the typical 1-0.5° accuracy for remote eye-trackers. The AOIs were marked as *Familiar* or *Novel* in each trial according to the corresponding cartoon pictures.

In order to determine whether participants' looking behaviours differed as a function of the mispronunciation conditions, Proportion of Familiar Fixation (PFF) was used as the dependent measure in the current analysis. This has been used in previous IPL studies to investigate adults' and infants' looking preference after hearing the auditory stimulus (Swingley & Aslin, 2000, 2002, 2009; among others). For each phase, we calculated the proportion of fixation duration on the familiar object (i.e., PFF) by dividing the fixation duration on the *Familiar* AOI by the total fixation duration on both the *Familiar* and the *Novel* AOIs. Participants' total fixation duration on each AOI was calculated for the time-window from 233 ms after the vowel onset of the target test item to 2233 ms. We used 233 ms (8.33 ms x 28 f) because 200 ms reaction time is typically reported in lexical decision tasks.

Only trials in which the sum duration of the AOI fixations equaled more than 65% of the 2-second window were included. This strict criterion was set to ensure that we only analyzed those trials where participants were largely attending and their looking behaviours were reasonably relevant to an expected response following the auditory stimulus. Across all the participants, 50 trials (10%) of the overall 500 trials were excluded.

Results

Average PFFs for each of the five pronunciation conditions in both the AusE and AusE+ groups are shown in Figure 6. We expected to see a larger PFF in a trial if participants regarded the label they heard as the correct label for the familiar object. A larger difference in PFFs between a mispronunciation condition and CP would indicate a bigger mispronunciation effect.

With alpha set at .05, a mixed-design ANOVA with *pronunciation condition* (CP, VH_{MP}, VB_{MP}, VL_{MP}, NP) as a within-subjects factor and *group* (AusE, AusE+) as a between-subjects factor was conducted to evaluate the PFF responses. This ANOVA revealed a significant main effect of *pronunciation condition*, $F(4, 92) = 83.570$, $p < .001$, $\eta_p^2 = .784$, suggesting that the responses differed for the five pronunciation conditions irrespective of group. Bonferroni adjusted *post-hoc* pairwise comparisons showed that the responses in CP was significantly different from all other conditions (all *p*s < .001), with larger PFF in CP ($M = .937$) than in the mispronunciations (all *M*s < .597) and NP ($M = .155$). This suggests that the participants were performing the task – they consistently looked more towards the familiar objects when hearing correct labels, looked less towards familiar objects after listening to the three types of vowel mispronunciations, and looked more towards novel objects after hearing the novel pronunciations. Further, the responses in VL_{MP} was significantly different from VH_{MP} and NP (both *p*s < .001), but did not significantly differ from VB_{MP} ($p = .254$) when collapsing across groups, with larger PFF in VL_{MP} ($M = .597$) than VH_{MP}

(*M* = .289) and NP (*M* = .155). This suggests that the participants in general fixated less on familiar objects when labels were mispronounced in vowel height and novel words than when mispronounced in vowel length. However, both groups performed similarly to mispronunciations in vowel length and backness. The responses in VB_{MP} and VH_{MP} ($p < .001$), VB_{MP} and NP ($p < .001$) were significantly different, with larger PFF in VB_{MP} ($M = .507$) than in VH_{MP} ($M = .289$) and NP ($M = .155$). This suggests that the participants looked less towards the familiar objects when labels were mispronounced in vowel height or as novel words than when they were mispronunciations in vowel backness. The difference between VH_{MP} and NP was not significant, $p = .136$, suggesting that participants had similar responses for vowel height mispronunciations and novel words in general.

In addition, while the main effect of *group* was not significant, $F(1, 23) = 2.117$, $p = .159$, $\eta_p^2 = .084$, there was a significant interaction between *pronunciation condition* and *group*, $F(4, 92) = 2.972$, $p = .023$, $\eta_p^2 = .114$, suggesting that the effect of different conditions differed in the AusE vs. AusE+ groups. To investigate the interaction and understand whether the two groups showed different sensitivities to a mispronunciation condition, planned comparisons were conducted to examine the *condition* x *group* effect between CP and each of the mispronunciation conditions, with a Bonferroni adjusted alpha (.0125). The analyses revealed a significant interaction between CP and VL_{MP} in the two groups, $F(1, 23) = 9.573$, $p = .005$, $\eta_p^2 = .294$, with larger PFF difference between CP and VL_{MP} in the AusE group than in the AusE+ group. No other contrasts were significant (CP vs. VH_{MP}, $F(1, 23) = 3.009$, $p = .096$, $\eta_p^2 = .116$; CP vs. VB_{MP}, $F(1, 23) = .328$, $p =.572$, $\eta_p^2 = .014$; CP vs. NP, $F(1, 23) = 5.346$, $p = .030$, $\eta_p^2 = .189$). This indicated the AusE group were more sensitive to mispronunciations in VL than the AusE+ group, but did not differ in performance on other mispronunciation conditions.

To address the question of whether the AusE+ group was more tolerant to mispronunciations in vowel length than other vowel features compared to the AusE group, planned comparisons were conducted to evaluate the *condition* x *group* interactions between VL_{MP} and each of the two mispronunciations, with a Bonferroni corrected alpha (.025). A significant interaction was found in the two groups between VL_{MP} and VB_{MP}, $F(1, 23) = 6.118$, $p = .021$, $\eta_p^2 = .210$, with larger PFF differences between VL_{MP} and VB_{MP} in the AusE+ group than in the AusE group. The interaction between VL_{MP} and VH_{MP} was not significant in the two groups, $F(1, 23) = 1.304$, $p = .265$, $\eta_p^2 = .054$. This revealed that the AusE+ group was at least more tolerant to mispronunciations in vowel length than in vowel backness compared to the AusE group.

To further examine whether the AusE+ group was less sensitive to VL_{MP} than VH_{MP} and VB_{MP} compared to the AusE group, the PFF responses were compared to chance (.5) for each mispronunciation condition in both groups. Two-tailed *t*-tests indicated that for the AusE group, the responses were significantly below chance in VH_{MP}, $t(14) = -4.011$, $p = .001$, but not significant in either VB_{MP}, $t(14) = .211$, $p = .836$, or VL_{MP}, $t(14) = .081$, $p = .937$. This analysis revealed that the AusE group more consistently looked to the novel objects when the labels were mispronounced in vowel height, but looked randomly when the labels were mispronounced in vowel backness and length. On the other hand, for the AusE+ group, their PFF responses were significantly above chance in VL_{MP}, $t(9) = 4.643$, $p = .001$, but not significant in either VH_{MP}, $t(9) = -1.652$, $p = .133$, or VB_{MP}, $t(9) = .026$, $p = .979$. The results showed that the AusE+ group more consistently looked to the familiar objects when the labels were mispronounced in vowel length, but looked randomly between the two objects when the labels were mispronounced in vowel height and backness. Together these results indicated that the AusE+ group was more tolerant of mispronunciations in vowel length than to mispronunciations in vowel height/backness compared to the AusE group.

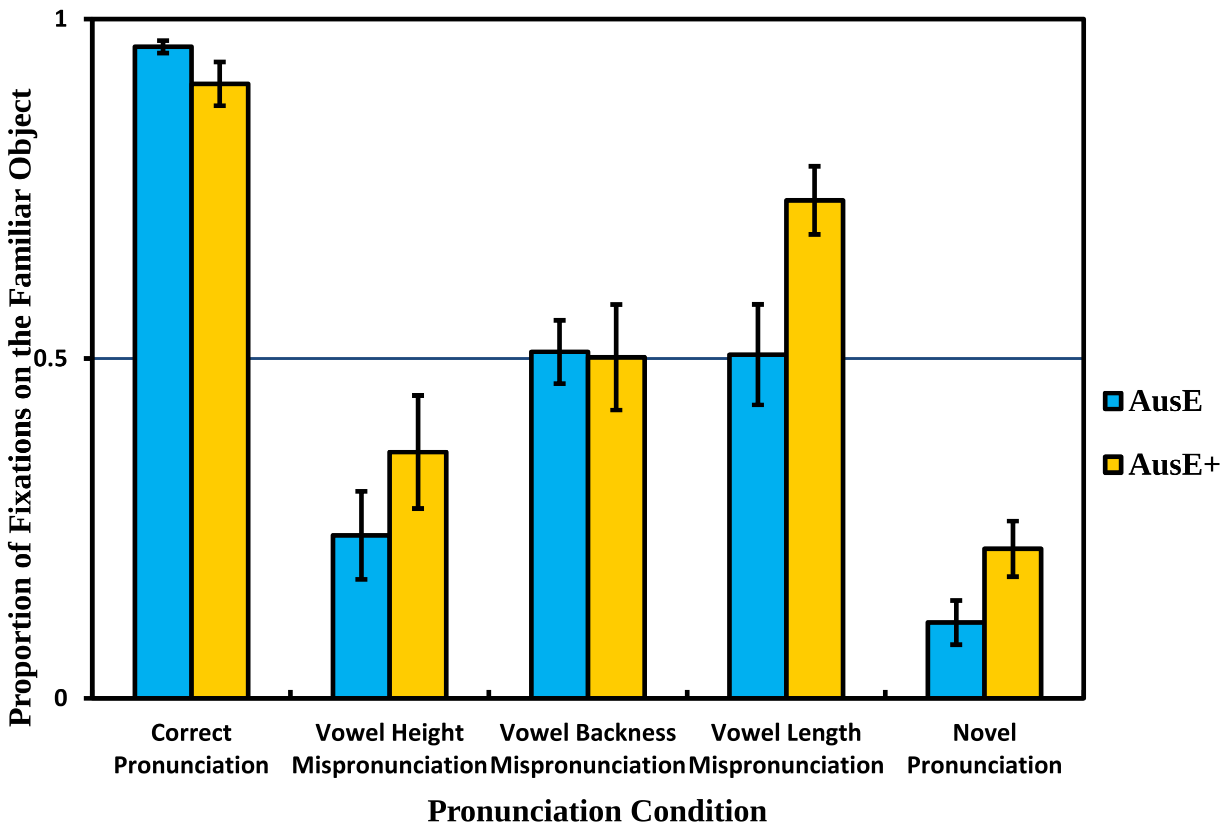


Figure 6. Average proportions of fixation duration on the *Familiar* AOI in the five pronunciation conditions. AusE = Australian English-speaking adults without early exposure to other English dialects; AusE+ = Australian English-speaking adults with early exposure to another English dialect; Responses observed in 233-2233 ms time-window from the vowel onset of the target test word; Error bars indicate +/- 1.00 standard error of mean.

To summarize, these results indicated that all the participants consistently looked more towards the familiar object in the CP condition whereas they applied the novel label to the novel object in the NP condition, showing that they performed the task and followed the mutual exclusivity principle. Responses in the mispronunciation conditions were all different from that of the CP, indicating that the participants were sensitive to all mispronunciations. Furthermore, the difference in responses between CP and VL_{MP} was bigger in the AusE group than in the AusE+

group, indicating that the AusE group was less tolerant to VL_{MP} errors than the AusE+ group. Moreover, the difference in responses between VL_{MP} and VB_{MP} was larger in the AusE+ group than in the AusE group, revealing that the AusE+ group was less sensitive to VL_{MP} errors than VB_{MP} errors compared to the AusE group. In addition, the responses of the AusE+ group were not significantly different from chance in VH_{MP} and VB_{MP} conditions, but were significantly above chance in VL_{MP}. This suggest that the AusE+ group was more tolerant of VL_{MP} errors than to the other mispronunciations, compared to the AusE group where responses were not significantly different from chance in both VL_{MP} and VB_{MP} conditions.

Discussion

In this study we investigated the perceptual sensitivity to phonemic vowel length, vowel height and vowel backness in Australian English adult listeners with and without early exposure to English dialects that do not have vowel length contrasts. We found that all participants looked significantly less at the familiar objects when the label was mispronounced in vowel height, backness and length than when the label was correctly pronounced, irrespective to whether they had early exposure to other dialectal accent or not. This indicates that mono-dialectal and bi-dialectal Australian English adults were both sensitive to mispronunciations in these three types of single vowel feature change. However, the difference in looking patterns between the correct pronunciation and the vowel length mispronunciation condition was significantly smaller in the bi-dialectal group than in the mono-dialectal group, showing that bi-dialectal Australian English adults are more tolerant to vowel length changes in words than the mono-dialectal adults. Furthermore, when comparing looking patterns in the three mispronunciation conditions between the two groups, we found that the difference between vowel length and vowel backness mispronunciation conditions was significantly bigger in the bi-dialectal group than the mono-dialectal group, suggesting the bi-dialectal Australian English adults were less sensitive to a change in vowel length than to a change in vowel backness for familiar words, compared to their mono-dialectal peers. Also, we found that the bi-dialectal group more consistently looked towards the familiar objects when the label was mispronounced in vowel length, but looked randomly between the familiar and novel objects when the label was mispronounced in vowel height and backness. In contrast, the mono-dialectal group looked randomly between the two objects in both vowel length and vowel backness mispronunciation conditions, and looked more towards the novel objects when the label was mispronounced in vowel height. This pattern of results indicates that bi-dialectal Australian English adults are more tolerant to situations when words are changed in vowel length than when words are changed in vowel height or backness, compared to mono-dialectal Australian English adults.

Our results also suggest that the bi-dialectal group might still have sensitivity to the phonemic vowel length contrast, but that it may be less well specified in their lexicon compared to the mono-dialect peers. Their looking pattern following a vowel length mispronunciation was different from the correct pronunciation, showing they *do* have some kind of sensitivity to changes in phonemic vowel length. However, they more consistently looked at the familiar object, which indicates bi-dialectal adults maybe more willing to accept vowel length mispronunciations as the target. This suggests the length feature is not well specified in their lexicon. A similar example can be found in Dutch adults, who produce contrastive /ɑ/ and /aː/ vowels that involve both vowel quality and length differences (Adank, van Hout, & Smits, 2004; Nooteboom & Doodeman, 1980). Both behavioural and ERP studies have revealed that monolingual Dutch adult listeners are sensitive only to the shortening of long vowel in words, but not to the lengthening of a short vowel (Chládková, Escudero, & Lipski, 2013, 2015; Escudero et al., 2009; van der Feest & Swingley, 2011). This suggests that Dutch adults specify the vowel length feature only in words with an /aː/

vowel, while the vowel length feature in words with /ɑ/ is under-specified. Likewise, bi-dialectal Australian English adults might still be sensitive to the alternation of vowel length categories, given that Australian English is the dominant accent in the community, but they probably under-specify the vowel length feature in real words, as they have experience with another dialectal accent where the vowel length feature is unreliable for lexical identity.

Our findings therefore reflect the complex influence of native language input on a bi-dialectal population's phonological representations. In contrast to the infant findings in Floccia et al. (2012) and in van der Feest and Johnson (2016), our bi-dialectal Australian English adults' performance was not the same as the mono-dialectal group. The fact that bi-dialectal Australian English adults are more tolerant to mispronunciations in vowel length than the mono-dialectal adults indicates they have may have developed more flexible phonological categories with respect to vowel length. Bi-dialectal Australian English adults receive speech input from two native dialects, one with phonemic vowel length contrasts and the other without (but contrastive in other features), so vowel length does not consistently contrast in what they hear from the input in general. As a result, it could be more difficult for bi-dialectal individuals to develop robust sensitivity to the vowel length feature and form clear phonemic categories, compared to the mono-dialectal population who receive more consistent and reliable contrastive input. Such flexible phonological boundaries might be necessary and advantageous for the bi-dialectal population. Vowel length categories that are too rigidly specified would filter out potential lexical contrasts, leading to potential misunderstandings. Given the variability in the overall input, it is necessary for bi-dialectal Australian English adults to allow flexibility in vowel length boundaries, and become more tolerant to alternations of vowel length in the lexicon.

Words which involve phonemic vowel length in Australian English employ other features to distinguish meanings in the other dialects, i.e., vowel quality and/or rhoticity. Our results suggest that bi-dialectal adults may be unlikely to specify both phonological contrasts in the lexicon as this would violate the economy principle (Best, 1994). Rather, they might specify only one of the competing contrastive features, or not specify any, resulting in vague categories. Although our study was not designed to test whether bi-dialectal adults specify vowel length or vowel quality/rhoticity feature in their lexicon, we did show that they may not specify the vowel length feature, even though Australian English is the dominant accent in the ambient environment (cf. Floccia et al., 2012). For example, bi-dialectal populations might more reliably specify spectral quality cues, given that both vowel quality and rhoticity are acoustically salient (with changes in F1, F2 and F3), and vowel length often interacts with higher level prosody (e.g., phrase-final lengthening) that is easily influenced by discourse/pragmatic factors (cf. Yuen et al., 2014). Vowel length might therefore not be as reliable a cue as spectral quality cues, especially when intrinsic vowel length is variable in the input. This may be especially the case in a language like Australian English, where only a small subset of vowels exhibit phonemic vowel length contrasts. Thus, for those growing up in a bi-dialectal English environment in Australia, vowel length may not be as robust a cue as vowel quality in establishing phonemic contrasts within vowel system. Our results therefore suggest the bi-dialectal populations who have been exposed to dialectal variants of the same word may not always specify the contrastive feature cues of those variants in their lexicon, even though they have some sensitivity to the feature. Exploring the nature of the bi-dialectal lexicon would therefore be an interesting area for further research.

The fact that the bi-dialectal group in this study were more tolerant of situations where a label was mispronounced in vowel length compared to vowel height or backness, provides important information about how they organize their within-category and cross-category contrasts, and how these might be integrated as phonological categories. The Australian English vowel height and backness contrasts tested here might have slightly different F1/F2 values in other dialects, but the alternations remain cross-category. Instead of broadening all phonological categories (cf. Durrant et al., 2015), the bi-dialectal adults maintained a robust sensitivity to these contrasts and were less tolerant to the mispronunciations. However, the vowel length contrast is a cross-category phonemic contrast in Australian English but a within-category phonetic variation in other dialects of English. Bi-dialectal adults might therefore establish this as a more abstract within-category phonetic variation so as to allow for successfully processing of both dialectal accents. Therefore, the acquisition task for bi-dialectal adults is complex, necessitating integration of phonetic information from two types of native input to establish their vowel categories.

For future study, it would be important to include more participants to increase the statistical power of the analysis and to evaluate whether the bi-dialectal Australian English adults might specify the spectral quality cues implemented in the other dialectal accent. Moreover, it would be interesting to test whether and when bi-dialectal infants would specify the contrastive features in the two dialects, how the developmental trajectory proceeds, and whether and how they would recover the information in situations with limited knowledge of context and discourse/pragmatics. Another area of interest would be to investigate whether bilingual and various L2 populations would exhibit similar sensitivities to Australian English phonemic vowel length as found here with the bi-dialectal adults. This could have important implications for understanding how other learners of Australian English specify vowel length features in their lexicon.

In summary, our results suggest that bi-dialectal Australian English adults with early exposure to another English dialect without phonemic vowel length may not well specify vowel length as a feature in the lexicon, though they might still have sensitivity to the contrast. They are more tolerant to vowel length mispronunciations than the mono-dialectal group and appear to accept labels with altered vowel length as the targets. This suggests that bi-dialectal adults have developed more flexible phonological categories of vowel length than the mono-dialectal population. Our findings highlight the complex effect of native input on the development of phonological categories and the generalization ability required in phonological acquisition. In contrast to some infant studies exploring dialectal variation (e.g., Floccia et al., 2012; van der Feest & Johnson, 2016), our findings suggest that the canonical feature is not always specified in the lexicon, especially when another phonological contrast is employed in the non-canonical accent. The specification of phonological features is crucially determined by the specific characteristics of the features in the systems. Our study tested bi-dialectal adults with the IPL paradigm and successfully demonstrated the flexibility in phonological categories in bi-dialectal adults. The findings therefore deepen our understanding of the role of native input in the development of phonological categories and the mechanism of abstract phonological organization, with implications for better understanding phonological development in bilingual populations.

Footnotes

1 The vowel /eː/ presents quite variably in Australian English, from monophthongal - when existing in some bi-morphemic words, to fully diphthongal, and often disyllabic in open syllables (Cox, 2006; Harrington, Cox, & Evans, 1997).

2 The stimuli were created also for another child study, so child-friendly design was employed.

3 The eye-tracker was down-sampled so that the study could be comparable with another study using Tobii X120 Eye Tracker.

Acknowledgments

We thank all the Macquarie University students involved in the study for their participation. We also thank Benjamin Davies, Kelly Miles, Ivan Yuen, Fabia Andronos, and the Child Language Lab at Macquarie University for their helpful assistance and feedback. This research was partially supported by funding from the first author's iMQRES at Macquarie University and the following grants: ARC Centre for Cognition and its Disorders CE110001021 and ARC FL130100014. Portions of this study were presented at the Australian Linguistic Society Annual Conference 2015 (ALS2015) - Western Sydney University, Sydney, 9-11 December 2015.

References

Adank, P., van Hout, R., & Smits, R. (2004). An acoustic description of the vowels of Northern and Southern standard Dutch. *Journal of the Acoustical Society of America*, *116*, 1729-1738. doi:10.1121/1.1779271

Bernard, J. (1967). Length and identification of Australian English vowels. *Journal of the Australasian Universities Modern Language Association*, *27*, 37-58.

Best, C. T. (1994). Learning to perceive the sound pattern of English. In C. Rovee-Collier & L. Lipsitt (Eds.), *Advances in Infancy Research* (Vol. 9, pp. 217-304). Hillsdale, NJ: Ablex Publishers.

Best, C. T., Tyler, M. D., Gooding, T. N., Orlando, C. B., & Quann, C. A. (2009). Development of phonological constancy: Toddlers' perception of native- and Jamaican-accented words. *Psychological Science*, *20*(5), 539-542. doi:10.1111/j.1467-9280.2009.02327.x

Boersma, P., & Weenink, D. (2012). *Praat: doing phonetics by computer* [Computer program]. Version 5.3.29, retrieved 4 October 2012 from http://www.praat.org/

Bohn, O.-S. (1995). Cross-language speech perception in adults: L1 transfer doesn't tell all. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (pp. 275-300). M.D.: York, Timonium.

Bohn, O.-S., & Polka, L. (2001). Target spectral, dynamic spectral and duration cues in infant perception of German vowels. *Journal of the Acoustical Society of America*, *110*(1), 504-515.

Buder, E. H., & Stoel-Gammon, C. (2002). American and Swedish children's acquisition of vowel duration: Effects of vowel identity and final stop voicing. *Journal of the Acoustical Society of America*, *111*(4), 1854-1864.

Burnham, D. K. (1986). Developmental loss of speech perception: Exposure to an experience with a first language. *Applied Psycholinguistics*, *7*(3), 207-240. doi:10.1017/S0142716400007542

Burnham, D. K., Earnshaw, L. J., & Quinn, M. C. (1987). The development of categorical identification of speech. In B. E. McKenzie & R. H. Day (Eds.), *Perception in infancy: Problems and issues* (pp. 237-275). Hillsdale, NJ: Lawrence Erlbaum.

Butler, J., Floccia, C., Goslin, J., & Panneton, R. (2011). Infants' discrimination of familiar and unfamiliar accents in speech. *Infancy*, *16*(4), 392-417. doi:10.1111/j.1532-7078.2010.00050.x

Chládková, K., Escudero, P., & Lipski, S. C. (2013). Pre-attentive sensitivity to vowel duration reveals native phonology and predicts learning of second-language sounds. *Brain & Language*, *126*, 243-252.

Chládková, K., Escudero, P., & Lipski, S. C. (2015). When "AA" is long but "A" is not short: Speakers who distinguish short and long vowels in production do not necessarily encode a short-long contrast in their phonological lexicon. *Frontiers in Psychology*, *6*, 438.

doi:10.3389/fpsyg.2015.00438

Chen, H., Xu Rattanasone, N., Cox, F., & Demuth, K. (under review). Understanding the acquisition of phonemic vowel length contrasts in Australian English-learning 18- and 24-month-olds. *Language Learning and Development.*

Chen, H., Yamane, N., Xu Rattanasone, N., Demuth, K., & Mazuka, R. (2015). *Japanese infants are aware of phonemic vowel length in novel words at 18 months*. Paper presented at the 40th Boston University Conference of Language Development (BUCLD40), Boston University, Boston.

Clopper, C. G., Pisoni, D. B., & de Jong, K. (2005). Acoustic characteristics of the vowel systems of six regional varieties of American English. *Journal of the Acoustical Society of America*, *118*(3), 1661-1676. doi:10.1121/1.2000774

Cox, F. (2006). The acoustic characteristics of /hVd/ vowels in the speech of some Australian Teenagers. *Australian Journal of Linguistics*, *26*(2), 147-179.

Cox, F., & Palethorpe, S. (2007). Australian English. *Journal of the International Phonetic Association*, *37*, 341-350. doi:10.1017/S0025100307003192

Deterding, D. (1990). *Speaker normalization for automatic speech recognition* (Unpublished doctoral dissertation). Cambridge University, United Kingdom.

Deterding, D. (1997). The formants of monophthong vowels in Standard Southern British English pronunciation. *Journal of the International Phonetic Association*, *27*, 47-55. doi:10.1017/S0025100300005417

Durrant, S., Delle Luche, C., Cattani, A., & Floccia, C. (2015). Monodialectal and multidialectal infants' representation of familiar words. *Journal of Child Language*, *42*(02), 447-465. doi: doi:10.1017/S0305000914000063

Escudero, P., Benders, T., & Lipski, S. (2009). Native, non-native and L2 perceptual cue weighting for Dutch vowels: The case of Dutch, German, and Spanish listeners. *Journal of Phonetics*, *37*, 452–465. doi:10.1016/j.wocn.2009.07.006

Fennell, C. T., & Waxman, S. R. (2010). What paradox? Referential cues allow for infant use of phonetic detail in word learning. *Child Development*, *81*(5), 1376-1383. doi:10.1111/j.1467-8624.2010.01479.x

Fennell, C. T., & Werker, J. F. (2003). Early word learners' ability to access phonetic detail in well-known words. *Language and Speech*, *46*(2-3), 245-264. doi:10.1177/00238309030460020901

Floccia, C., Delle Luche, C., Durrant, S., Butler, J., & Goslin, J. (2012). Parent or community: Where do 20-month-olds exposed to two accents acquire their representation of words? *Cognition*, 124, 95-100. doi:10.1016/j.cognition.2012.03.011

Golinkoff, R., Hirsh-Pasek, K., Cauley, K., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, *14*, 23- 45.

Harrington, J., Cox, F., & Evans, Z. (1997). An acoustic phonetic study of broad, general, and cultivated Australian English vowels. *Australian Journal of Linguistics*, *17*, 155-184. doi:10.1080/07268609708599550

Hillenbrand, J. M., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*, *97*, 3099-3111.

Kitamura, C., Panneton, R., & Best, C. T. (2013). The development of language constancy: attention to native versus nonnative accents. *Child Development*, *84*(5), 1686-1700. doi:10.1111/cdev.12068

Kitamura, C., Panneton, R., Diehl, M., & Notley, A. (2006). Attuning to the native dialect: When more means less. In P. Warren & C. I. Watson (Eds.), *Proceedings of the 11th Australian International Conference on Speech Science and Technology* (pp. 124-129). Canberra:

Australian Speech Science & Technology Association Inc.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, 255, 606-608.

Ladefoged, P. (2001) *Vowels and consonants: An introduction to the sounds of languages*. Oxford: Blackwell Publishers.

Mani, N., Coleman, J., & Plunkett, K. (2008). Phonological specificity of vowel contrasts at 18-months. *Language and Speech*, *51*(1&2), 3-21.

Mani, N., & Plunkett, K. (2007). Phonological specificity of vowels and consonants in early lexical representations. *Journal of Memory and Language*, *57*(2), 252-72.

Mani, N., & Plunkett, K. (2008). 14-month-olds pay attention to vowels in novel words. *Developmental Science*, *11*(1), 53–59.

Mani, N., & Plunkett, K. (2011). Does size matter? Subsegmental cues to vowel mispronunciation detection. *Journal of Child Language*, *38*, 606-627.

Mulak, K. E., Best, C. T., Tyler, M. D., Kitamura, C., & Irwin, J. R. (2013). Development of phonological constancy: 19-month-olds, but not 15-month-olds, identify words in a non-native regional accent. *Child Development*, *84*(6), 2064-2078. doi:10.1111/cdev.12087

Nazzi, T. (2005). Use of phonetic specificity during the acquisition of new words: Differences between consonants and vowels. *Cognition*, *98*, 13-30.

Nazzi, T., Jusczyk, P. W., & Johnson, E. K. (2000). Language Discrimination by English-Learning 5-Month-Olds: Effects of Rhythm and Familiarity. *Journal of Memory and Language*, *43*(1), 1-19. doi:10.1006/jmla.2000.2698

Nooteboom, S. G., & Doodeman, G. J. N. (1980). Production and perception of vowel length in spoken sentences. *Journal of the Acoustical Society of America*, *67*, 276–287.

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*, *24*(2), 175-184.

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.

Roach, P., Knowles, G., Varadi, T., & Arnfield, S. (1993). MARSEC: A machine-readable spoken English corpus. *Journal of the International Phonetic Association*, *23*, 47-54.

Sato, Y., Sogabe, Y., & Mazuka, R. (2010). Discrimination of phonemic vowel length by Japanese infants. *Developmental Psychology*, *46*(1), 106-199. doi:10.1037/a0016718

Shatzman, K. B., & McQueen, J. M. (2006). Segment duration as a cue to word boundaries in spoken-word recognition. *Perception & Psychophysics*, *68*(1), 1-16. doi:10.3758/BF03193651

Swingley, D. (2003). Phonetic detail in the developing lexicon. *Language and Speech*, *46*(2-3), 265-294. doi:10.1177/00238309030460021001

Swingley, D. (2009). Onsets and codas in 1.5-year-olds' word recognition. *Journal of Memory and Language*, *60*(2), 252-269. doi:10.1016/j.jml.2008.11.003

Swingley, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, *76*, 147-166.

Swingley, D., & Aslin, R. N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, *13*(5), 480-484. doi: 10.1111/1467-9280.00485

van der Feest, S. V. H., & Johnson, E. K. (2016). Input-driven differences in toddlers' perception of a disappearing phonological contrast. *Language Acquisition*, *23*(2), 89-111. doi: 10.1080/10489223.2015.1047096

van der Feest, S. V. H., & Swingley, D. (2011). Dutch and English listeners' interpretation of vowel duration. *Journal of the Acoustical Society of America*, *129*(3), EL57-EL63. doi:10.1121/1.3532050.

van Heugten, M., & Johnson, E. K. (2014). Learning to contend with accents in infancy: Benefits of brief speaker exposure. *Journal of Experimental Psychology: General*, *143*(1), 340-350. doi:10.1037/a0032192

Watson, C. I., & Harrington, J. (1999). Acoustic evidence for dynamic formant trajectories in Australian English vowels. *Journal of the Acoustical Society of America*, *106*(1), 458-468. doi:10.1121/1.427069

Wells, J. C. (1982). *Accents of English*. Cambridge: Cambridge University Press.

Werker, J. F., & Polka, L. (1993). Developmental changes in speech perception: New challenges and new directions. *Journal of Phonetics*, *21*(1-2), 83-101.

Werker, J. F., & Tees, R. C. (1983). Developmental changes across childhood in the perception of non-native speech sounds. *Canadian Journal of Psychology*, *37*(2), 178-286.

Werker, J. F., & Tees, R. C. (1984a). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *7*, 49-63.

Werker, J. F., & Tees, R. C. (1984b). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America*, *75*, 1866-1878.

White, K. S., & Aslin, R. N. (2011). Adaptation to novel accents by toddlers. *Developmental Science*, *14*(2), 372-384. doi:10.1111/j.1467-7687.2010.00986.x

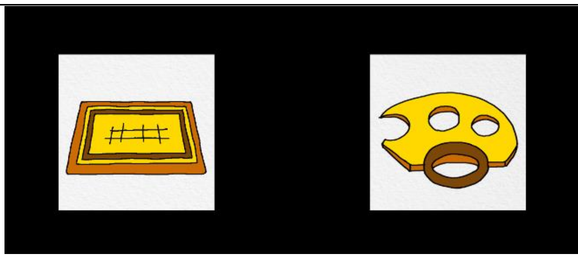
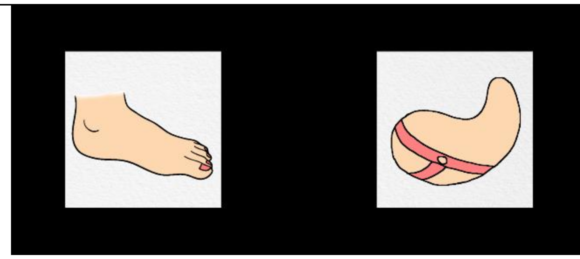
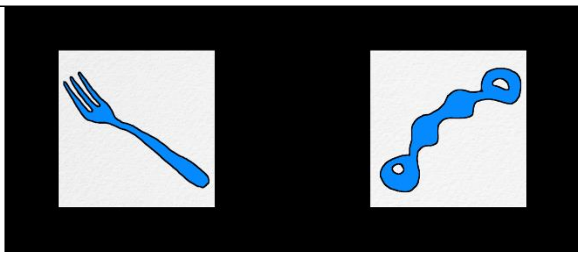
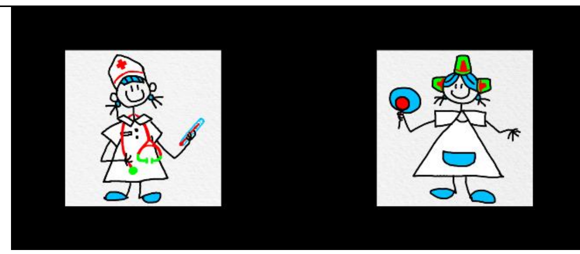
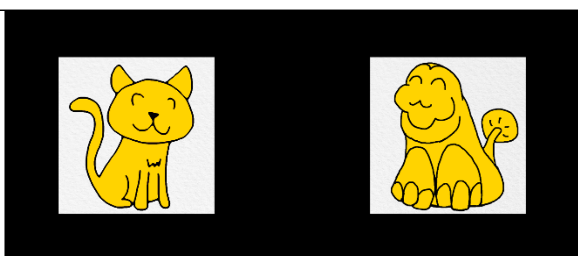
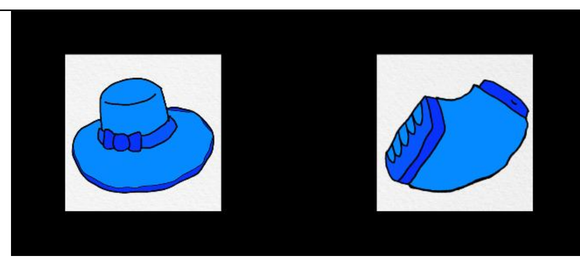
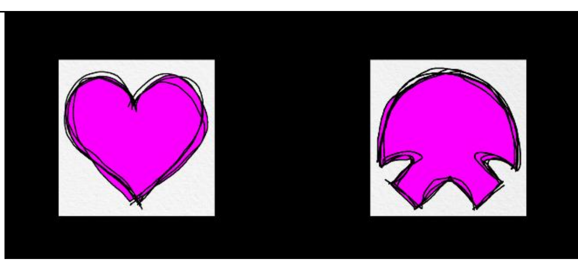
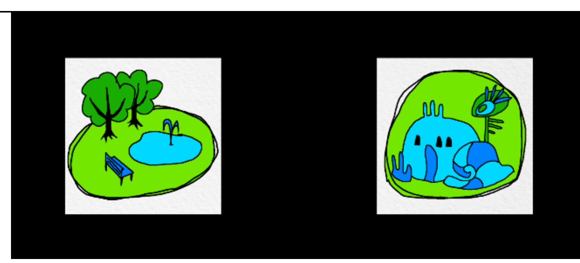
White, K. S., & Morgan, J. L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language*, *59*, 114-132.

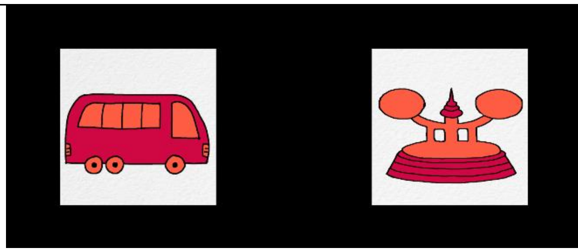
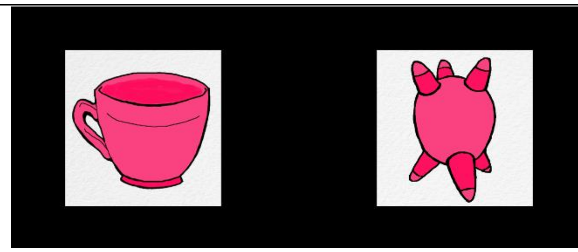
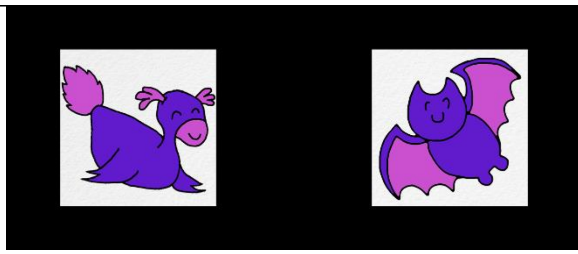
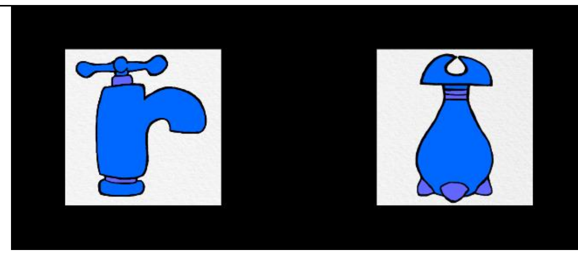
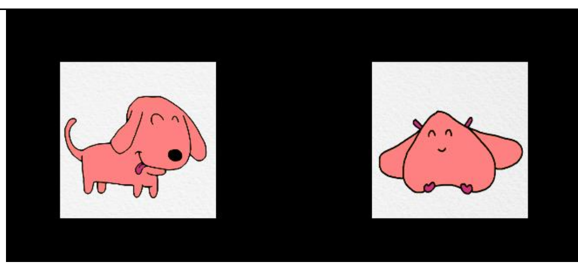
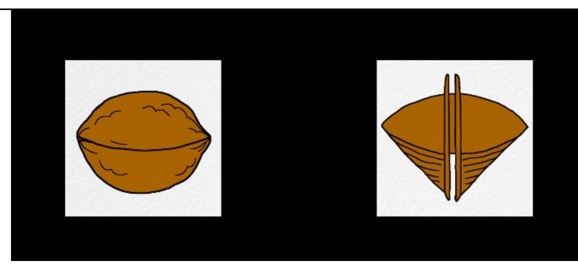
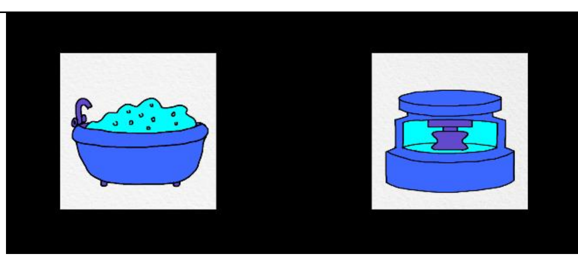
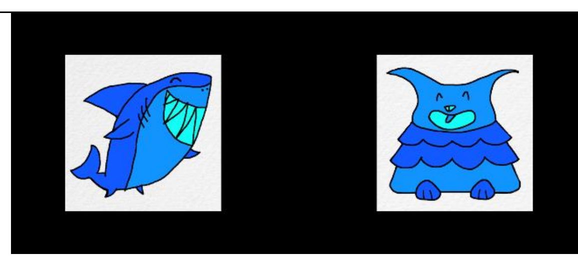
Williams, D., & Escudero, P. (2014). A cross-dialectal acoustic comparison of vowels in Northern and Southern British English. *Journal of the Acoustical Society of America*, *136*(5), 2751-2761. doi:10.1121/1.4896471

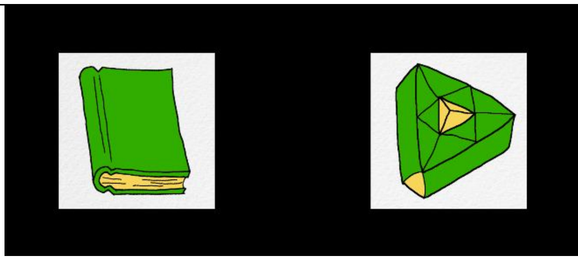
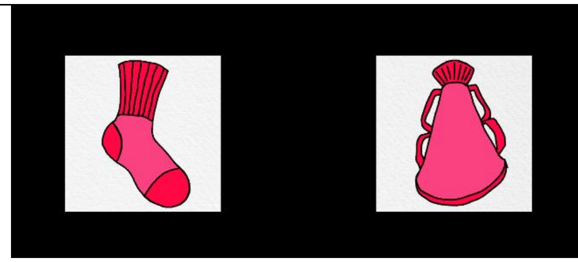
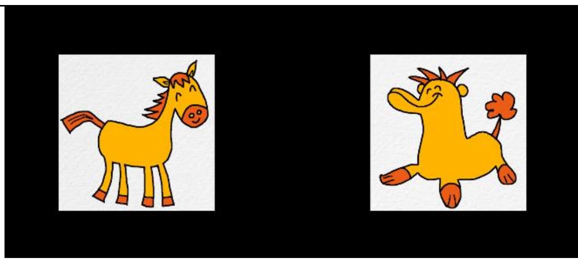
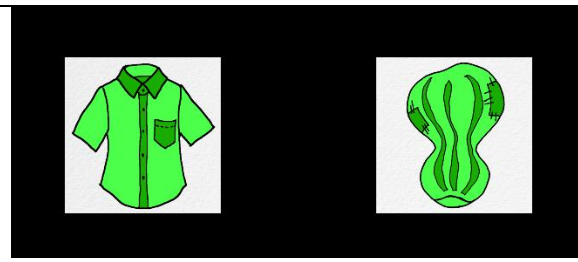
Yuen, I., Cox, F., & Demuth, K. (2014). Three-year-olds' production of Australian English phonemic vowel length as a function of prosodic context. *Journal of the Acoustical Society of America*, *135*(3), 1469-1479.

Appendix I: Familiar-Novel Picture Pairs Used as Visual Stimuli in the Experiment

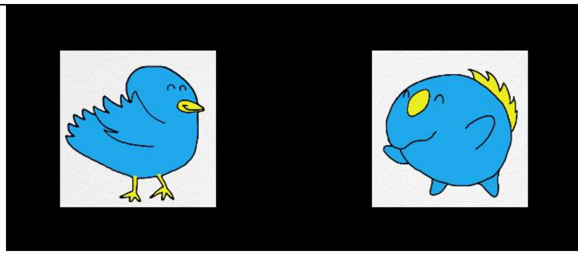
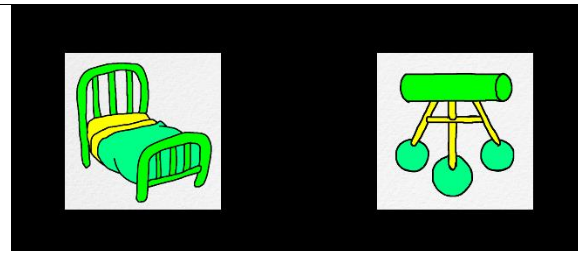
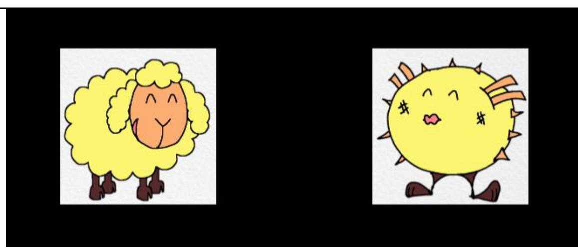
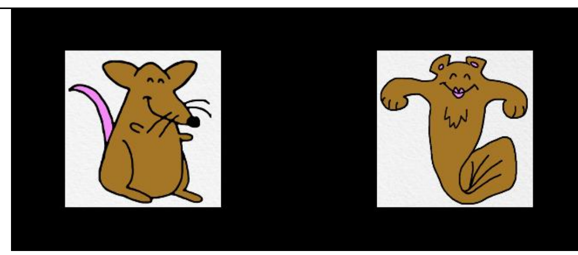
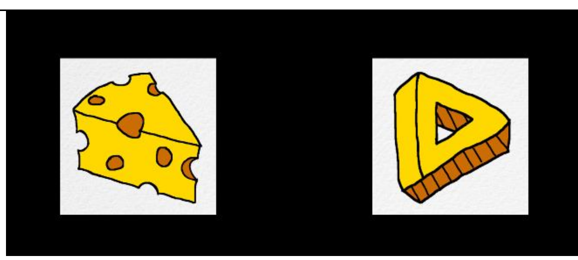
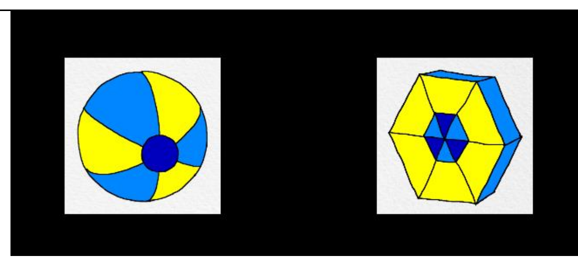
A. Twenty picture pairs for test items (demonstrated with all familiar items on the left)

Correct Pronunciations	
	
mat /mæt/ – ***mat*** **[mæt]**	foot /fʊt/ – ***foot*** **[fʊt]**
	
fork /foːk/ – ***fork*** **[foːk]**	nurse /nɜːs/ – ***nurse*** **[nɜːs]**
Vowel Height Mispronunciations	
	
cat /kæt/ – ***ket*** **[ket]**	hat /hæt/ – ***het*** **[het]**
	
heart /hɐːt/ – ***hurt*** **[hɜːt]**	park /pɐːk/ – ***pirk*** **[pɜːk]**

Vowel Backness Misrpronunciations	
	
bus /bɐs/ – ***bass*** **[bæs]**	duck /dɐk/ – ***dack*** **[dæk]**
	
bat /bæt/ – ***but*** **[bɐt]**	tap /tæp/ – ***tup*** **[tɐp]**
Vowel Length Mispronunciations	
	
pup /pɐːp/ – ***parp*** **[pɐːp]**	nut /nɐːt/ – ***nart*** **[nɐːt]**
	
bath /bɐːθ/ – ***buth*** **[bɐθ]**	shark /ʃɐːk/ – ***shuck*** **[ʃɐk]**

Novel Mispronunciations	
	
book /bʊk/ – ***geap*** **[giːp]**	sock /sɔk/ – ***mirt*** **[mɜːt]**
	
horse /hoːs/ – ***dirk*** **[dɜːk]**	shirt /ʃɜːt/ – ***gup*** **[gɐp]**

B. Six picture pairs for training and filler items

Training	
Correct Pronunciation	Novel Pronunciation
	
bird – ***bird***	bed – ***nep***
Filler	
Correct Pronunciation	
	
sheep – ***sheep***	rat – ***rat***
	
cheese – ***cheese***	ball – ***ball***

Appendix II: Language Background Questionnaire

Participant's Name ________________ [No.________] Gender: M / F Age ____________

1. Where do you live? (your suburb, your neighborhood)
____

2. How long have you lived in that neighborhood? _____________________________________

3. Residential history. Please state where you have lived and the approximate ages in each place (Please specify: state, city/town, suburb & country):
____
____

4. Previous High schools and grades attended: ____
____

5. What is your first language? ____

6. How often do you use this language, in terms of hours per day? _________________________

7. What other languages do you speak/understand? _____________________________________

8. At what age(s) did you learn these languages? ______________________________________

9. How often do you use these languages, in terms of hours per day?
   Language: ______________________________________ Hours: ____________
   Language: ______________________________________ Hours: ____________

10. Do you have any speech or hearing problems? _____________________________________

11. Mother's place of birth and first language: ____

12. Father's place of birth and first language: ____

13. Describe the occupation of your parents/guardians. (If you are mature age, give your own occupation as well) Female: ______________________________

Male: ______________________ Your Own: ______________________

Which category best describes these occupations? (Please tick)

Female	Male	Your own
Sales/Personal Services	Sales/Personal Services	Sales/Personal Services
Home duties	Home duties	Home duties
Professional	Professional	Professional
Unemployed	Unemployed	Unemployed
Trades	Trades	Trades
Labourer	Labourer	Labourer
Clerical	Clerical	Clerical
Management	Management	Management
Plant/Machine Operator	Plant/Machine Operator	Plant/Machine Operator

14. Which category best describes your parent/guardian's education, as well as your own education? (Please tick; you can write a note such as "attended but did not finish" if that's the case)

Female	Male	Your own
Primary school	Primary school	Primary school
High school before your 10	High school before year 10	High school before year 10
High school – year 10	High school – year 10	High school – year 10
High school – year 12	High school – year 12	High school – year 12
Trade certificate or TAFE	Trade certificate of TAFE	Trade certificate of TAFE
College diploma	College diploma	College diploma
University degree	University degree	University degree
Master's degree	Master's degree	Master's degree
Doctor's degree	Doctor's degree	Doctor's degree

15. Please provide your contact information below.

Email: ______________________

Phone: ______________________

Thank you for your participation!

Chapter Five: General Discussion

Overall Conclusions and Implications

This thesis aimed to investigate how phonemic vowel length contrasts are represented in infants and adults. There have been many studies on the learning of vowel quality contrasts (e.g., Mani, Coleman, & Plunkett, 2008; Mani & Plunkett, 2007, 2008, 2011; Nazzi, 2005), but studies on phonemic vowel length have been limited to production data (e.g., Buder & Stoel-Gammon, 2002; Fikkert, 1994; Kehoe & Lleó, 2003; Ota, 1999, 2001; Yuen, Cox, & Demuth, 2014) and discrimination data (Mugitani, Pons, Fais, Dietrich, Werker, & Amano, 2009; Sato, Sogabe, & Mazuka, 2010). Phonemic vowel length is not only important for understanding word meaning in some languages, but also crucial for mastering prosodic structure (e.g., syllable structure, stress system, etc.) in many languages. We therefore explored how and when sensitivity of phonemic vowel length contrasts develop in infants, and how they are represented in adults. To achieve this, we looked at infants' development of perceptual sensitivity to phonemic vowel length contrasts in two languages, one which represents this contrast systematically (Japanese, Chapter Two), and the other that does not (Australian English, Chapter Three). We also examined this issue in Australian English adults with and without exposure to another English dialect that does not have phonemic vowel length contrasts (Chapter Four). The findings of these studies are discussed below.

In Study 1 (Chapter Two), we examined whether monolingual Japanese infants are sensitive to phonemic vowel length contrasts by 18 months. Infants who were taught with novel words containing a long vowel succeeded in learning the novel words and detecting shortening mispronunciations, whereas infants who were taught with novel words containing short vowels failed to learn the novel words. We thought these results might be due to the acoustic salience of long vowels, the greater variability of short vowel durations, and the preference infants for listening to the high frequency Heavy-Light word forms in Japanese IDS. The findings provide evidence for a well-specified phonemic length category in long vowels in infants who are leaning a language where vowel length categories are relevant to word learning. Even with limited input for long vowels (e.g., 6% of all words in IDS; Bion, Miyazawa, Kikuchi, & Mazuka, 2013) acoustic salience may facilitate the early development of phonological categories for long vowels. The study therefore provides a significant contribution to the literature on phonological development, revealing an important developmental achievement in Japanese infants' representations of phonemic vowel length by 18 months. This then raised questions about how and when phonemic vowel length contrasts may be represented in other languages. In Japanese, phonemic vowel length contrasts systematically for all five vowels. However, languages like Australian English employ phonemic vowel length contrasts in only a subset of vowels. We therefore anticipated it might be more challenging for Australian English infants to represent phonemic vowel length at the same age, and this contrast might be later acquired.

To explore this issue, we conducted Study 2 (Chapter Three) in two IPL experiments. The results showed that, while the 18-month-olds did not seem to perform the task, the 24-month-olds successfully displayed sensitivity to mispronunciations of vowel height, vowel backness, and more importantly, vowel length. The findings revealed that Australian English infants are aware of phonemic vowel length contrasts at least by 24 months. The study provides an important complement to Study 1 (Chapter Two) for a more complete picture of early development of phonemic vowel length across different vowel systems with systematic vs. non-systematic phonemic vowel length contrasts. The study also demonstrates a well-controlled comparison of infants' sensitivity to phonemic vowel length and other 1-feature vowel quality contrasts for the first time, and therefore contributes to our understanding of the relationship between vowel length and vowel quality acquisition. Although it is not clear why the 18-month-olds did not do the task, this could be for methodological reason. It would be therefore interesting to employ the same novel word learning and mispronunciation detection task as that used in Study 1 with the Japanese infants, to explore this issue with Australian English-learning infants as well.

The findings of the infants' studies raised many questions regarding adults' representations of phonemic vowel length contrasts, especially for those who have received variable phonological categories from two types of native input. The Study 3 (Chapter Four) therefore compared the perceptual sensitivity to mispronunciations of phonemic vowel length in Australian English adults with and without early exposure to another English dialect that does not have this contrast, using an IPL task similar as that used in Study 2. The study has revealed three general findings. First, both the mono-dialectal and bi-dialectal adults showed sensitivity to mispronunciations of phonemic vowel length. Second, the bi-dialectal adults were more tolerant to mispronunciations of phonemic vowel length compared to the mono-dialectal adults. Third, the bi-dialectal adults were more tolerant to mispronunciations of phonemic vowel length than vowel height/backness compared to the mono-dialectal group. This suggests that bi-dialectal adults have developed more flexible phonological categories of vowel length and may not necessarily specify a phonemic vowel length feature in their lexicon. The results also suggest that a canonical phonemic feature is not always specified in the lexicon when another phonological feature is exploited in the non-canonical dialect. Specification of phonological features is closely related to acoustic saliency and specific characteristics of the features in the systems. This study therefore deepens our understanding of the complex effects of native input on the development of phonological representations, with implications for phonological development in bilingual populations.

Overall, this thesis demonstrated that, 1) infants acquiring a language with systematic phonemic vowel length contrasts have developed sensitivity to phonemic vowel length contrasts by 18 months, similar to the developmental time course of other 1-feature vowel quality contrasts (i.e., vowel height and backness; Mani et al., 2008; Mani & Plunkett, 2007); 2) infants acquiring a language with non-systematic phonemic vowel length appear to have a later mastery of phonemic vowel length contrasts by 24 months; and 3) bi-dialectal adults exposed to dialects with and without phonemic vowel length contrasts establish more flexible phonological categories for vowel length. Together these findings highlight the important and complex role of native input for the development of phonological representations, where the systematicity of the phonological features in the native system and variability in exposure to different types of native phonological systems both affect the specification of phonological features early in development.

Study Limitations and Future Directions

The major limitations of this thesis included several aspects. For Study 1, infants did not learn the novel words with a short vowel in the first syllable, therefore, longer training sessions should be employed to ensure better association between novel labels and objects. Also, it would be useful to have better control of the variability of different prosodic contexts in which the short vowels are presented. As discussed in the study, the variability in the phonetic realization of the novel words with short vowels in different prosodic contexts is not related to the vowel category *per se*, but it might make learning lexical items more difficult for very young infants. For Study 2, the 18-month-olds appeared to be unable to perform the task, probably due to the limited vocabulary at this age. The small vocabulary in very young infants may have made the mispronunciation detection a less sensitive task in examining complex issues such as the comparison of infants' specifications amongst different vowel features in their lexicon. For Study 3, a larger sample will be needed to provide more robust results.

For future research, it would be interesting to test Japanese infants at 24 months to examine whether older infants have acquired the phonemic representation of short vowels. Also, the same task used in Study 1, i.e., the novel word learning and mispronunciation detection paradigm, could be also used to directly compare the performance of infants learning systematic and non-systematic phonemic vowel length contrasts. In addition, infants learning other languages with systematic vowel length contrasts, such as Arabic, Finnish, Estonian, etc., could be tested to provide further insight into the development of phonological representations. Further, it would be important to directly test the perception of phonological categories in bi-dialectal adults with exposure to two different phonological systems with and without phonemic vowel length contrasts. It would also be very interesting to explore both the perception and production of phonemic vowel length in the same speakers; one might anticipate that those with better perceptional representations might also exhibit this in production. Finally, it would be beneficial to extend these studies into bilingual infants and adults in exploring bilingual phonological systems with different phonological features on vowels.

To conclude, this thesis brings together three studies exploring the perceptual sensitivity of phonemic vowel length contrasts in infants and adults. The overall results emphasize the important role of the systematicity and stability of a phonological feature in the development of phonological representations across languages and populations. This in turn may inspire further studies on phonemic vowel length contrasts, and provide implications for the development of vowel representations more generally.

References

Bion, R. A. H., Miyazawa, K., Kikuchi, H., & Mazuka, R. (2013). Learning phonemic vowel length from naturalistic recordings of Japanese infant-directed speech. *PLoS ONE*, *8*(2), e51594. doi:10.1371/journal.pone.0051594

Buder, E. H., & Stoel-Gammon, C. (2002). American and Swedish children’s acquisition of vowel duration: Effects of vowel identity and final stop voicing. *Journal of the Acoustical Society of America*, *111*(4), 1854-1864.

Fikkert, P. (1994). *On the acquisition of prosodic structure* (Doctoral dissertation, HIL dissertations 6, Leiden University). The Hague: Holland Academic Graphics.

Kehoe, M. M., & Lleó, C. (2003). The acquisition of nuclei: A longitudinal analysis of phonological vowel length in three German-speaking children. *Journal of Child Language*, *30*, 527-556.

Mani, N., Coleman, J., & Plunkett, K. (2008). Phonological specificity of vowel contrasts at 18-months. *Language and Speech*, *51*(1&2), 3-21.

Mani, N., & Plunkett, K. (2007). Phonological specificity of vowels and consonants in early lexical representations. *Journal of Memory and Language*, *57*(2), 252-72.

Mani, N., & Plunkett, K. (2008). 14-month-olds pay attention to vowels in novel words. *Developmental Science*, *11*(1), 53–59.

Mani, N., & Plunkett, K. (2011). Does size matter? Subsegmental cues to vowel mispronunciation detection. *Journal of Child Language*, *38*, 606-627.

Mugitani, R., Pons, F., Fais, L., Dietrich, C., Werker, J. F., & Amano, S. (2009). Perception of vowel length by Japanese- and English-learning infants. *Developmental Psychology*, *45*(1), 236-247.

Nazzi, T. (2005). Use of phonetic specificity during the acquisition of new words: Differences between consonants and vowels. *Cognition*, *98*, 13-30.

Ota, M. (1999). *Phonological theory and the acquisition of prosodic structure: Evidence from child Japanese* (Unpublished doctoral dissertation). Georgetown University, Washington D. C.

Ota, M. (2001). Phonological theory and the development of prosodic structure: Evidence from child Japanese. *Annual Review of Language Acquisition*, *1*, 65-118.

Sato, Y., Sogabe, Y., & Mazuka, R. (2010). Discrimination of phonemic vowel length by Japanese infants. *Developmental Psychology*, *46*, 106-199.

Yuen, I., Cox, F., & Demuth, K. (2014). Three-year-olds' production of Australian English phonemic vowel length as a function of prosodic context. *Journal of the Acoustical Society of America*, *135*(3), 1469-1479.

Ethics Approvals

MACQUARIE University

HUI CHEN <hui.chen15@students.mq.edu.au>

RE: Ethics Application – Final Approval (Subject to Condition/s) (Ref.5201200814)

Fhs Ethics <fhs.ethics@mq.edu.au> Mon, Jan 7, 2013 at 11:48 AM
To: Prof Katherine Demuth <katherine.demuth@mq.edu.au>
Cc: A/Prof Felicity Cox <felicity.cox@mq.edu.au>, Dr Nan Xu <nan.xu@mq.edu.au>, Miss Hui Chen <hui.chen15@students.mq.edu.au>

Dear Prof Demuth,

RE: 'The representations of phonemic vowel length in child language development ' (Ref: 5201200814)

Thank you for your recent correspondence. Your response has addressed the issues raised by the Faculty of Human Sciences Human Research Ethics Sub-Committee. Approval of this application has been granted and you may now proceed with your research.

This approval is subject to the following condition/s:

1. Please forward correspondence from the specific sources/centres indicating approval to the Ethics Sub-Committee when they become available.

This research meets the requirements of the National Statement on Ethical Conduct in Human Research (2007). The National Statement is available at the following web site:

http://www.nhmrc.gov.au/_files_nhmrc/publications/attachments/e72.pdf.

The following personnel are authorised to conduct this research:

A/Prof Felicity Cox
Dr Nan Xu
Miss Hui Chen
Prof Katherine Demuth

NB. STUDENTS: IT IS YOUR RESPONSIBILITY TO KEEP A COPY OF THIS APPROVAL EMAIL TO SUBMIT WITH YOUR THESIS.

Please note the following standard requirements of approval:

1. The approval of this project is conditional upon your continuing compliance with the National Statement on Ethical Conduct in Human Research (2007).

2. Approval will be for a period of five (5) years subject to the provision of annual reports.

Progress Report 1 Due: 7th January 2014
Progress Report 2 Due: 7th January 2015
Progress Report 3 Due: 7th January 2016
Progress Report 4 Due: 7th January 2017
Final Report Due: 7th January 2018

NB. If you complete the work earlier than you had planned you must submit a Final Report as soon as the work is completed. If the project has been discontinued or not commenced for any reason, you are also required to submit a Final Report for the project.

Progress reports and Final Reports are available at the following website:

http://www.research.mq.edu.au/for/researchers/how_to_obtain_ethics_approval/human_research_ethics/forms

3. If the project has run for more than five (5) years you cannot renew approval for the project. You will need to complete and submit a Final Report and submit a new application for the project. (The five year limit on renewal of approvals allows the Committee to fully re-review research in an environment where legislation, guidelines and requirements are continually changing, for example, new child protection and privacy laws).

4. All amendments to the project must be reviewed and approved by the Committee before implementation. Please complete and submit a Request for Amendment Form available at the following website:

http://www.research.mq.edu.au/for/researchers/how_to_obtain_ethics_approval/human_research_ethics/forms

5. Please notify the Committee immediately in the event of any adverse effects on participants or of any unforeseen events that affect the continued ethical acceptability of the project.

6. At all times you are responsible for the ethical conduct of your research in accordance with the guidelines established by the University. This information is available at the following websites:

http://www.mq.edu.au/policy/
http://www.research.mq.edu.au/for/researchers/how_to_obtain_ethics_approval/human_research_ethics/policy

If you will be applying for or have applied for internal or external funding for the above project it is your responsibility to provide the Macquarie University's Research Grants Management Assistant with a copy of this email as soon as possible. Internal and External funding agencies will not be informed that you have final approval for your project and funds will not be released until the Research Grants Management Assistant has received a copy of this email.

If you need to provide a hard copy letter of Final Approval to an external organisation as evidence that you have Final Approval, please do not hesitate to contact the FHS Ethics at the address below.

Please retain a copy of this email as this is your official notification of final ethics approval.

Yours sincerely,

Dr Peter Roger
Chair
Faculty of Human Sciences
Human Research Ethics Sub-Committee

____

Faculty of Human Sciences - Ethics
Research Office
Level 3, Research HUB, Building C5C
Macquarie University
NSW 2109

Ph: +61 2 9850 4197
Fax: +61 2 9850 4465

MACQUARIE University

HUI CHEN <hui.chen15@students.mq.edu.au>

RE: Ethics Amendment 1 - Approved (Ref No. 5201200814)

Fhs Ethics <fhs.ethics@mq.edu.au> Tue, Apr 30, 2013 at 3:25 PM
To: Professor Katherine Demuth <katherine.demuth@mq.edu.au>
Cc: A/Prof Felicity Cox <felicity.cox@mq.edu.au>, Dr Nan Xu <nan.xu@mq.edu.au>, Miss Hui Chen <hui.chen15@students.mq.edu.au>

Dear Prof Demuth,

RE: 'The representations of phonemic vowel length in child language development ' (Ref: 5201200814)

Thank you for your recent correspondence regarding the amendment request. We apologies for the delay in responding.

The amendments have been reviewed and we are pleased to advise you that the amendments have been approved.

This approval applies to the following amendments:

1. To include Tobii X120 Eye Tracker in the Perception Study to record participants' looking behaviours;
2. Revised Information and Consent forms.

Please note that the researcher/s may wish to check the Information and Consent form for missing spaces in Paragraph 3 (e.g. "alooking game", "lookingbehaviours") before distribution.

Please accept this email as formal notification that the amendments have been approved. Please do not hesitate to contact us in case of any further queries.

All the best with your research.

Kind regards,

FHS Ethics
****
Faculty of Human Sciences - Ethics
Research Office
Level 3, Research HUB, Building C5C
Macquarie University
NSW 2109

Ph: +61 2 9850 4197
Fax: +61 2 9850 4465

Email: fhs.ethics@mq.edu.au

http://www.research.mq.edu.au/

MACQUARIE University

HUI CHEN <hui.chen15@students.mq.edu.au>

RE: Ethics Amendment 2 - Approved (Ref No. 5201200814)

Fhs Ethics <fhs.ethics@mq.edu.au> Thu, Jul 11, 2013 at 4:36 PM
To: Professor Katherine Demuth <katherine.demuth@mq.edu.au>
Cc: A/Prof Felicity Cox <felicity.cox@mq.edu.au>, Dr Nan Xu <nan.xu@mq.edu.au>, Hui Chen <hui.chen15@students.mq.edu.au>

Dear Prof Demuth,

RE: 'The representations of phonemic vowel length in child language development ' (Ref: 5201200814)

Thank you for your recent correspondence regarding the amendment request. The amendment has been reviewed and we are pleased to advise you that the amendment has been approved.

This approval applies to the following amendment:

1. To add a new advertisement flyer to recruit infant participants.

Please accept this email as formal notification that the amendment has been approved. Please do not hesitate to contact us in case of any further queries.

All the best with your research.

Kind regards,

FHS Ethics
****

Faculty of Human Sciences - Ethics
Research Office
Level 3, Research HUB, Building C5C
Macquarie University
NSW 2109

Ph: +61 2 9850 4197
Fax: +61 2 9850 4465

Email: fhs.ethics@mq.edu.au

http://www.research.mq.edu.au/

MACQUARIE University

HUI CHEN <hui.chen15@students.mq.edu.au>

RE: Ethics Amendment 3 – Approved with Condition/s (5201200814)

Fhs Ethics <fhs.ethics@mq.edu.au> Tue, Dec 10, 2013 at 3:17 PM
To: Professor Katherine Demuth <katherine.demuth@mq.edu.au>
Cc: A/Prof Felicity Cox <felicity.cox@mq.edu.au>, Dr Nan Xu <nan.xu@mq.edu.au>, Hui Chen <hui.chen15@students.mq.edu.au>

Dear Prof Demuth,

RE: 'The representations of phonemic vowel length in child language development ' (Ref: 5201200814)

Thank you for your recent correspondence regarding the amendment request. The amendment request has been reviewed and I am pleased to advise you that the amendments have been approved.

This approval applies to the following amendments:

1. Additional data collection – a Receptive Vocabulary added to the original checklist and a vocabulary questionnaire added to the project;
2. Change in testing time – now one hour (instead of 30 minutes);
3. Change in payment – $30 cash instead of $20 gift card;
4. Additional recruitment – Posting flyers on related Facebook pages/groups, setting up a stand at Macquarie Centre and distributing flyers via baby/toddler production shops in shopping centres.

Please note that this approval is subject to the following condition:

1. Please forward all relevant permission letters/emails (e.g. from Macquarie Centre) for our records;
2. Please forward the revised Information and Consent form/s for review (e.g. indicating the change in payment and testing time).

Please accept this email as formal notification that the amendments have been approved.

Please do not hesitate to contact us in case of any further queries.

All the best with your research.

Kind regards,

FHS Ethics

****
Faculty of Human Sciences – Ethics
Research Office
Level 3, Research HUB, Building C5C
Macquarie University
NSW 2109

Ph: +61 2 9850 4197
Fax: +61 2 9850 4465

Email: fhs.ethics@mq.edu.au

http://www.research.mq.edu.au/

MACQUARIE University

HUI CHEN <hui.chen15@students.mq.edu.au>

RE: HS Ethics Amendment 4 ‑ Approved (Ref No. 5201200814)

Fhs Ethics <fhs.ethics@mq.edu.au> Thu, Jul 24, 2014 at 12:24 PM
To: Professor Katherine Demuth <katherine.demuth@mq.edu.au>
Cc: A/Prof Felicity Cox <felicity.cox@mq.edu.au>, Dr Nan Xu <nan.xu@mq.edu.au>, Hui Chen <hui.chen15@students.mq.edu.au>

Dear Prof Demuth,

RE: 'The representations of phonemic vowel length in child language development ' (Ref: 5201200814)

Thank you for your amendment request on the 16th June 2014. The amendments were reviewed and approved on the 24th June 2014. We apologise for the delay in sending this formal notification as the administrator was on leave.

This approval applies to the following amendments:

1. Monolingual Australian English study
a) Changes in the Information Statement and Child Consent form, as described in section 6;
b) Changes in Information Statement and Adult Consent form, as stated in section 6;

2. Additional recruitment ‑ To include local churches, libraries, parents in local playgrounds and parks by using flyers and advertisement;

3. Additional groups of bilingual speakers (adults and children) ‑ Mandarin‑English and Japanese‑English bilinguals, as proposed in section 6;

4. Documents noted and approved
1) Updated Information and Consent Form (ICF) for Monolingual Children;
2) Updated ICF for Monolingual Adults;
3) ICF for Bilingual Children;
4) ICF for Bilingual Adults;
5) Language Questionnaire for Bilingual Adults;
6) Vocabulary Questionnaire for Bilingual Adults;
7) Recruitment Flyer for Bilingual Children;
8) Recruitment Flyers for Bilingual Adults.

Please accept this email as formal notification that the amendments have been approved. Please do not hesitate to contact us in case of any further queries.

All the best with your research.

Kind regards,

FHS Ethics
****

Faculty of Human Sciences ‑ Ethics
Research Office
Level 3, Research HUB, Building C5C
Macquarie University
NSW 2109

Ph: +61 2 9850 4197
Fax: +61 2 9850 4465

MACQUARIE University

HUI CHEN <hui.chen15@students.mq.edu.au>

RE: HS Ethics Amendment 5 - Approved (Ref No. 5201200814)

Fhs Ethics <fhs.ethics@mq.edu.au> Thu, Nov 13, 2014 at 11:14 AM
To: Professor Katherine Demuth <katherine.demuth@mq.edu.au>
Cc: A/Prof Felicity Cox <felicity.cox@mq.edu.au>, Dr Nan Xu <nan.xu@mq.edu.au>, Hui Chen <hui.chen15@students.mq.edu.au>

Dear Prof Demuth,

RE: 'The representations of phonemic vowel length in child language development ' (Ref: 5201200814)

Thank you for your recent correspondence regarding the amendment request.

The amendments have been reviewed and we are pleased to advise you that the amendments have been approved.

This approval applies to the following amendments:

1. Additional recruitment - To recruit one group of monolingual Japanese-acquiring children (aged 17 - 19 months, as stated in Section 6;

2. Documents in relations to this amendment

a) Appendix B: Research to be Undertaken outside Australia;
b) Attachment 1: Memorandum of Understanding between Macquarie and RIKEN;
c) Attachment 2: Preliminary Notice of Acceptance issued by RIKEN;
d) Attachment 3: Approval Letter from Laboratory for Language Development at RIKEN;
e) Attachment 4: Information Statement and Parent/Child Consent Form from RIKEN (English);
f) Attachment 5: Information Statement and Parent/Child Consent Form from RIKEN (Japanese).

Please accept this email as formal notification that the amendments have been approved. Please do not hesitate to contact us in case of any further queries.

All the best with your research.

Kind regards,

FHS Ethics
****
Faculty of Human Sciences - Ethics
Research Office
Level 3, Research HUB, Building C5C
Macquarie University
NSW 2109

Ph: +61 2 9850 4197
Fax: +61 2 9850 4465

Email: fhs.ethics@mq.edu.au

http://www.research.mq.edu.au/

Memorandum of Understanding
between
Department of Linguistics, Macquarie University (Australia)
and
Laboratory for Language Development, RIKEN Brain Science Institute (Japan)

Department of Linguistics, Macquarie University and Laboratory for Language Development, RIKEN Brain Science Institute hereby conclude this Memorandum of Understanding (MOU) to promote international collaboration in education and research, particular in the field of Language Acquisition.

1. Both parties aim to encourage the following activities, and thereby promote international academic collaboration:

   (a) Exchange of research and academic materials, and publications;
   (b) Exchange of faculty and research scholar including student;
   (c) Holding meetings for education and research including seminar and symposia;
   (d) Establishment of joint research projects

2. This MOU shall become effective as of the date of signatures of both parties and remain valid for a period of five (5) years thereafter. However, this MOU may be terminated by either party with a minimum of ninety (90) days written notice to the other. Activities in progress at the time of termination of this MOU shall be permitted to complete as planned under the separate specific agreements unless otherwise agreed.
   This MOU may be further extended for an additional period of five (5) years upon the written agreement of both parties.

3. Any modification or amendment of this MOU shall be made on the basis of mutual written understanding.

Each Party shall sign two identical copies of this MOU, and retain one copy.

Signed for Macquarie University

[signature]

Prof. PRETORIUS Sakkie
Deputy Vice Chancellor, Research
Macquarie University

Date: 24·06·2014

Signed for RIKEN

[signature]

Prof. MAZUKA Reiko
Laboratory Head
Brain Science Institute RIKEN

[signature]

IWABUCHI Haruyuki
Director,
Brain Science Planning Office RIKEN

Date: June 18, 2014