

**--This paper has been accepted for publication in the
Journal of Child Language, please do not distribute without
permission--**

Targeted adaptation in infants following live exposure to an accented talker

Melissa PAQUETTE-SMITH^{a*}

Angela COOPER^b

Elizabeth K. JOHNSON^b

^aDepartment of Psychology, University of California Los Angeles, 502 Portola Plaza, Los Angeles, CA, 90095, USA

^bDepartment of Psychology, University of Toronto, 3359 Mississauga Rd., Mississauga, ON, L5L 1C6, CANADA

*Corresponding author: paquettesmith@psych.ucla.edu

We would like to thank Chen Peng, Keren Smith, Shukri Nur, Alyssa Baumgartner, Yazad Bhathena, Gresha Shah and Lisa Hotson, as well as the other members of the Child Language and Speech Studies Lab for their support. This work was supported by grants from the Social Sciences and Humanities Research Council, Natural Sciences and Engineering Research Council, and the Canada Research Chairs program. Portions of this work were presented at the 3rd Workshop for Infant Language Development, the 58th Annual Meeting of the Psychonomic Society and the 21st International Congress on Infant Studies.

Keywords: *accent adaptation; speech perception; perceptual learning*

Abstract

Infants struggle to understand familiar words spoken in unfamiliar accents. Here, we examine whether accent exposure facilitates accent-specific adaptation. Two types of pre-exposure were examined: video-based (i.e., listening to pre-recorded stories; Experiment 1) and live interaction (reading books with an experimenter; Experiments 2 and 3). After video-based exposure, Canadian English-learning 15- to 18-month-olds failed to recognize familiar words spoken in an unfamiliar accent. However, after face-to-face interaction with a Mandarin-accented talker, infants showed enhanced recognition for words produced in Mandarin English compared to Australian English. Infants with live exposure to an Australian talker were not similarly facilitated, perhaps due to the lower vocabulary scores of the infants assigned to the Australian exposure condition. Thus, live exposure can facilitate accent adaptation, but this ability is fragile in young infants and is likely influenced by vocabulary size and the specific mapping between the speaker and the listener's phonological system.

Introduction

Adult listeners excel at flexibly adapting to pronunciation variations produced by speakers of different accents (e.g., Clarke & Garrett, 2004; Cutler & Broersma, 2005; see Cristia et al., 2012 for a review). Infants and toddlers, on the other hand, appear to struggle with unfamiliar accents (e.g., Best, Tyler, Gooding, Orlando, & Quann, 2009; Schmale, Cristia, Seidl, & Johnson, 2010). How then do young children improve their adaptation to unfamiliar accents, and under what circumstances does this take place? Prior work on the development of adaptive processes in young children has shown mixed results as to when children can successfully adapt to speech variation. Moreover, this research has indicated that the type of exposure can play a significant role in predicting how successful young toddlers will be at adapting to an unfamiliar accent (e.g., van Heugten & Johnson, 2014). In the present work, we investigate these issues by providing toddlers with either live or video-taped exposure to an unfamiliar accented talker. Then, in the test phase, we examine the impact of exposure type on children's ability to recognize words produced by that same talker versus another talker speaking in a different unfamiliar accent.

Acoustic-phonetic variation can arise from a multitude of sources, including differences in talkers' vocal tract size, affect, speaking rate, as well as their linguistic background. In order to understand non-native or regional-accented talkers¹, both children and adults need to learn to accommodate pronunciations that differ from their native-accented norms. However, relative to adults, children have smaller vocabularies, less robust phonemic categories, and less linguistic experience overall that would help them to determine what variation is phonologically relevant

¹ Regional accents refer to variation that arises from learning a first language in a particular location/region. In contrast, non-native accents refer to variation that arises when a speaker learns a second language later in life. In this case, the speaker might import some of the features of their first language (L1) into their second language (L2) (see Bent, Atagi, Akbik, & Bonifield, 2016).

and what variation can be ignored. Moreover, findings from several studies suggest that phonological constancy – where a word is spontaneously recognized in the face of acoustic-phonetic variation (Best et al., 2009) – does not develop until approximately 19 months of age (Best et al., 2009; Mulak, Best, Tyler, Kitamura, & Irwin, 2013). Evidence for the lack of phonological constancy prior to 19 months of age comes from studies examining toddlers' recognition of words spoken in familiar and unfamiliar accents (Best et al., 2009; van Heugten, Paquette-Smith, Krieger, & Johnson, 2018). For example, 15-month-old American children listen longer to known words (e.g., *ball*, *baby*) over unknown words when they are spoken in a familiar American accent; however, they fail to recognize these same words when they are produced in an unfamiliar Jamaican accent. Comparable results have been reported with Australian learning toddlers presented with words in a familiar Australian versus an unfamiliar Jamaican accent (Mulak et al., 2013), and Canadian English-learning 15-month-olds presented with words spoken in a familiar Canadian versus an unfamiliar Australian or French accent (van Heugten & Johnson, 2014; van Heugten et al., 2018). Although younger infants struggle to recognize familiar words when they were spoken in an unfamiliar accent, infants over 19 months of age seem to be successful in all of these studies (Best et al., 2009; Mulak et al., 2013; van Heugten & Johnson, 2014; van Heugten et al., 2018). 19-month-olds' ability to recognize words in unfamiliar accents has been attributed to the increase in vocabulary size that occurs between 15 and 19 months, which has been suggested to enable children to develop phonological constancy (i.e., to begin to discern phonological invariants across novel word forms; Mulak et al., 2013).

However, the literature also contains some inconsistencies regarding the age at which children develop the ability to handle accent variation, as other studies have found that children

over the age of 19 months still continue to struggle with variant pronunciations (e.g., Floccia, Delle Luche, Durrant, Butler, & Goslin, 2012; Schmale, Hollich, & Seidl, 2011; van Heugten, Krieger, & Johnson, 2015). For instance, in an eye-tracking task Canadian English-learning 25-month-olds, but not 20-month-olds, were able to recognize words spoken in Scottish-accented English (van Heugten et al., 2015). Additionally, Schmale et al. (2011) reported that 24-month-old children trained on a set of novel words produced by a talker of a familiar American English accent were unable to generalize recognition of these newly-learned words when they were produced by a Spanish-accented English talker. Other studies have argued that the development of the ability to comprehend words produced by accented talkers follows a protracted developmental trajectory (Creel, Rojo, & Paullada, 2016). In line with this view, studies with young children have shown that even 5- to 7-year-olds still do not exhibit adult-like performance with unfamiliar accents (e.g., Bent & Atagi, 2017; Bent & Holt, 2018).

Although there seems to be considerable variation in infants' ability to spontaneously recognize words in unfamiliar accents, there is some evidence that pre-exposure to an accent may help young toddlers overcome these difficulties (Schmale, Cristia, & Seidl, 2012; van Heugten & Johnson, 2014). For example, White and Aslin (2011) found (using an artificially constructed accent) that 19-month-olds were able to accommodate to an unfamiliar pronunciation variant following training with that specific accent. And a study with Dutch-learning 24-month-olds demonstrated that adaptation to an unfamiliar regional variety of the native language can take place within two minutes (van der Feest & Johnson, 2016). Other studies have suggested that pre-exposure to multiple different accents can also, in some circumstances, facilitate adaptation. For example, in a recent listening preference study, Potter and Saffran (2017) demonstrated that 18-month-olds but not 15-month-olds were able to

recognize known words in a novel, unfamiliar accent (British English) after exposure to a mixture of different unfamiliar accents (but note that pre-exposure to British English on its own was not sufficient to elicit recognition).

While the abovementioned studies have shown that pre-exposure can facilitate adaptation in infants over 18 or 19 months of age, only one study has shown that infants under 18-months can benefit from exposure. In this study, 15-month-old Canadian English-learning infants were exposed to a video-recorded story read by an Australian English talker. In the test phase, they were tested on their ability to recognize words produced by that same talker (van Heugten & Johnson, 2014). Critically, the ability to capitalize on this pre-exposure phase was contingent on their being highly familiar with the words in the story (“The Very Hungry Caterpillar”). Only the children whose Canadian-accented caregivers had read the story to them at least once a day for the 14 days prior to their lab visit recognized the Australian-accented words, presumably because only those children who were highly familiar with the story were able to benefit from the Australian pre-exposure provided in the lab. These findings highlight the importance of familiarity with the linguistic context (e.g., lexical content), suggesting it can enable infants to more efficiently extract information about an accent. This work may also indicate that either phonological constancy exists in some form earlier than other studies have posited or that phonological constancy may not be necessary for adaptive processes to take place. Young infants may be capable of employing an alternative adaptive strategy that does not require them to make targeted adjustments to specific phonemic categories but rather involves loosening their criteria for what is considered a permissible match between input and stored representations. Schmale, Seidl, and Cristià (2015) posited that in contexts lacking strong top-down information (or in infants with an impoverished lexicon), infants may utilize a “general

expansion” strategy, loosening their categories to be more accepting of pronunciations that deviate from their own native-accented norms.

Although van Heugten & Johnson (2014) suggests that video exposure can be effective in facilitating adaptation in some circumstances (i.e., when children are highly familiar with the story), other studies have shown that brief video exposure may not be enough to help even older 20-month-olds adapt (van Heugten et al., 2015). It is possible that under more ecologically valid conditions exposure might be more effective in facilitating adaptation than previous work has suggested. The type of video exposure that children are given in the lab is a far cry from the rich face-to-face interactions that children have with speakers in real life. There is considerable evidence that face-to-face social interaction may be key to facilitating other types of language learning (i.e., Barr, 2013; DeLoache et al., 2010; Krcmar, Grela, & Lin, 2007; Kuhl, Tsao, & Liu, 2003; Lytle, Garcia-sierra, & Kuhl, 2018; Roseberry, Hirsh-Pasek, & Golinkoff, 2014). For example, children learned new vocabulary items better when taught by a caregiver compared to when they watched baby media designed to teach those words (DeLoache et al., 2010). Similarly, live interaction with a Mandarin speaker but not TV exposure seemed to help American infants maintain a non-native Mandarin phonemic contrast (Kuhl et al., 2003). Despite evidence suggesting that live interaction may be particularly effective in helping infants to adapt to unfamiliar accents, no research to date has examined this possibility.

Thus, upon examining the diverse findings reported in the existing body of work on accent adaptation in toddlers, a more nuanced picture of accented speech comprehension in young children appears to be necessary; clearly, considerations such as task difficulty (e.g., familiarity preference, familiar word recognition, word learning), the type of exposure (e.g., live vs video) and the characteristics of the particular accents being presented (and their relationship

to the listener's native accent) play a key role in children's success in coping with phonetic variation. Moreover, most prior studies have presented children with the same accent at exposure and test. This can make it difficult to ascertain what kind of adaptive mechanisms children are employing when confronted with an unfamiliar accent, and whether they are making targeted adjustments for a specific accent or whether they are increasing their willingness to accept deviant pronunciations more generally. Only one study to date has investigated whether young children's accent adaptation generalizes to a novel accent, but this study did not show evidence of accent-specific adaptation, as children were only able to adapt to the novel accent after exposure to a variety of different accents (Potter & Saffran, 2017). In addition, much of the previous work has focused on young infants' ability to recognize familiar words produced by a regionally-accented talker (instead of a non-native accented talker; van Heugten et al., 2018) – and there is evidence to suggest that the processing of some non-native accented speech may differ from the processing of regionally-accented speech. For example, previous studies have indicated that non-native speakers' realization of segments may be more variable than regional or L1 speakers' productions (Baese-Berk & Morrill, 2017; Wade, Jongman, & Sereno, 2007), which could make it more difficult for infants (and adults) to make the appropriate phonetic adjustments during adaptation.

The goal of the present study was to investigate whether video and/or live exposure can facilitate infants' adaptation to unfamiliar accents. To test this, in Experiment 1, we provided 15- to 24-month-old Canadian English-learning infants with video exposure to a Mandarin-accented speaker. Then, using a preferential looking task, we tested the impact of this exposure on their ability to recognize words in a Mandarin and an Australian accent. In Experiments 2 and 3, we examine whether a more ecologically valid type of exposure – namely, live interaction – might

be more effective in facilitating adaptation. Taken together, our results suggest that under certain circumstances live exposure to an accented speaker can facilitate comprehension of words produced by that particular accented speaker.

Experiment 1 – Recorded exposure

Previous work has suggested that it is not until 19 months of age that infants are able to spontaneously understand known words when they are spoken in an unfamiliar accent (Best et al., 2009; van Heugten et al., 2018). However, in some circumstances pre-exposure to the unfamiliar accent seems to facilitate adaptation in children under 19 months of age. For example, one study showed that 15-month-olds are able to recognize words in an unfamiliar Australian accent after brief video exposure to the Australian speaker reading a familiar storybook (van Heugten & Johnson, 2014). Although there is some evidence that pre-exposure can be effective, it not clear what infants are doing in response to this exposure. Is pre-exposure helping infants to adapt to the specific characteristics of the speaker’s accent or is exposure simply prompting infants to become more accepting of pronunciations that differ from their native variant of English? In the current study, we begin to disambiguate between these two possibilities by exposing infants to a Mandarin-accented speaker and testing their ability to recognize words in the same Mandarin speaker’s voice and an unfamiliar Australian speaker’s voice. Unlike previous work which has looked at the role of pre-exposure on adaptation to a native (i.e., British or Australian) accent, this is the first study to examine the role of exposure in facilitating adaptation to a non-native accent (i.e., Mandarin-accented English).

In the pre-exposure phase we presented 15- to 18-month-old Canadian English-learning infants with a 10-minute pre-recorded video of a Mandarin-accented talker reading a series of children’s books. Importantly, all children were highly familiar with the storybooks that were

read in the video, as their parents were sent copies of the books and asked to read at least two of those books to their child every day for at least two weeks prior to visiting the lab. After the exposure phase, children were tested on their recognition of words in the familiar Mandarin talker's voice and an unfamiliar Australian-accented talker's voice.

If exposure facilitates accent adaptation, then infants under 19 months of age should be able to recognize items produced by the Mandarin-accented English talker at test. As in previous adaptation studies using an eye tracking paradigm, this would manifest as a greater proportion of looking time to the labeled images for the pre-exposed talker. Given prior work showing that 15- to 18-month-olds have difficulty spontaneously recognizing items produced in an unfamiliar accent (e.g., Best et al., 2009; van Heugten & Johnson, 2014), we predict they would not be above chance in recognizing the target words labeled by the unexposed Australian-accented talker.

Methods

Participants

Twenty-four 15- to 18-month-old infants were tested ($M_{age} = 509.41$ days, range = 464-555, 9 males). Parents reported no hearing impairments or recent ear infections at the time of testing. Children were raised in households where English was spoken at least 90% of the time and had no previous exposure to Australian- or Mandarin-accented English according to a language questionnaire administered prior to testing. The participants we tested were ethnically diverse and came from a variety of cultural backgrounds. Caregivers also completed the MacArthur-Bates Communicative Development Inventories (CDI; Fenson et al., 2007) Words and Gestures form to provide an approximate index of their child's expressive vocabulary size.

On average children in this sample produced 45.45 words ($SD = 66.97$)². An additional 8 infants were tested but were excluded due to fussiness ($n = 6$), experimenter error ($n = 1$) or prior exposure to one of the accents in the study ($n = 1$).

Stimuli

The exposure stimuli consisted of a 10-minute pre-recorded video of a Mandarin-accented talker reading three books (i.e., *Good Night Moon*, *Each Peach Pear Plum* and *The Very Hungry Caterpillar*) in an infant-directed manner. The Mandarin-accented talker began learning English at the age of 7 in school in China and moved to Canada as an adult. She was a highly fluent speaker but had a readily detectable accent. These three books were selected because they are best-selling books in this age group, and they do not contain many of the target words included in the test phase. When the speaker encountered a target word she substituted the word with a synonym (e.g., substituting the word ‘kitten’ for ‘cat’).

The auditory stimuli in the test phase consisted of 24 nouns considered to be familiar to 16-month-old infants. Based on the vocabulary norms reported in the Wordbank database, the average frequency of the target words in 16-month-olds’ receptive vocabulary was approximately 71% (range: 37-98%; Frank et al., 2016). According to parental report, the children in this study understood an average of 69.79% ($SD = 21.29$) of the target words, in line with the norms reported above. Note that it is possible that parents may have, as other studies have suggested, actually underestimated their child’s receptive vocabulary size (Furey, 2011; Houston-Price, Mather, & Sakkalou, 2007). That is, the children in our study may have comprehended more target words than their caregivers indicated. The target words were

² $N = 22$; The caregivers of two children did not submit their vocabulary forms.

produced in sentence frames (e.g., *Look at the [target]!* and *Where's the [target]?*) by the female Mandarin-accented talker (from the video exposure) and a female Australian-accented talker (see Appendix A for phonetic transcriptions of the target words). Additionally, auditory attention-getters, which included *hey*, *look*, and *wow*, were recorded by the talkers to precede the target carrier sentence. To maintain the infants' interest, statements and questions about the stimuli (*Can you see it? Do you like it? How cute! Isn't it pretty?*) were recorded and presented following the target phrase. Talkers were instructed to produce the items in a child-directed manner.

Procedure

In the exposure phase, 15- to 18-month-olds watched the 10-minute pre-recorded video of the Mandarin-accented speaker reading the three storybooks. To ensure children were relatively familiar with the stories, caregivers indicated which of the three books (if any) they read to their child. If they did not have at least two of the three books used in the recording, they were mailed one or two book(s) prior to their appointment. All parents were asked to read at least two of the three books to their child every day for two weeks (caregivers kept a log of their reading pattern for those two weeks). Parents read two books per day on average ($SD = .07$)³. The mean number of story readings that occurred over the two-week period was 28.38 ($SD = 4.04$). The video exposure mirrored a typical television watching experience where the child sat on a couch with their caregiver facing the TV screen. Immediately after the video, children were brought into an IAC sound-attenuated booth for the test phase.

During the preferential looking test phase, children sat on their caregiver's lap. To prevent caregivers from biasing their child's responses, caregivers were asked to wear

³ N = 21; three families did not submit their reading log.

headphones playing masking music (i.e., a combination of music and speech stimuli from the study). Children were presented with pairs of images against a white background on a screen. One image was a named target, while the other was an unnamed distracter. Each image was presented twice throughout the experiment, serving once as a target and once as a distracter. Half of the items were produced by the Australian-accented talker, and the other half by the Mandarin-accented talker. Four blocks of 6 trials each were presented (24 trials in total), and the trials were blocked by talker. The blocks alternated between the Australian- and Mandarin-accented talkers. Which items were produced by which talker, whether Australian- or Mandarin-accented blocks were presented first, and on which side of the screen the labeled target image was presented was counterbalanced across participants.

Each trial lasted 6000 ms, with the pair of pictures looming for the duration of the trial. An auditory attention-getter to capture the child's interest (e.g., *Hey!*) was presented 300 ms after the pictures were displayed. A target phrase was then played (e.g., *Where's the [target]?*), with the target word occurring 3000 ms after trial onset, followed by one of the positive comments (e.g., *How cute!*). Each session was videotaped at 30 frames per second with a remote-controlled camera to allow for frame-by-frame offline coding using SuperCoder (Hollich, 2005). Each frame was coded for whether the toddler looked at the left image, the right image or neither image by coders who were unaware of the auditory or visual content of the trials. Inter-coder agreement was consistently high between the two coders (mean correlation = 0.99 for 4 subjects). Following prior work (e.g., Delle Luche, Durrant, Poltrock, & Floccia, 2015), the proportion of looking time to the target picture in the post-naming window was calculated (looking time to target / total looking time to target + distracter).

Results and discussion

The proportion of looking time to the target was analyzed for the entire post-naming period starting from 250 ms post-word onset (Figure 1). We began the window of analysis 250 ms after word onset to give infants time to program an eye movement. Fixations that occur before this time point are unlikely to be driven by hearing the target word (see Buckler, Oczak-Arsic, Siddiqui, & Johnson, 2017; Fernald, Perfors, & Marchman, 2006; Swingley & Aslin, 2000; van Heugten et al., 2015 for a similar window onset). Trials where the participant did not fixate on either the target or distracter during the post-naming period were excluded. A linear mixed-effects regression model was constructed (Baayen, Davidson, & Bates, 2008) using the lme4 (Bates, Mächler, Bolker, & Walker, 2015) and LmerTest (Kuznetsova, Brockhoff, & Christensen, 2017) packages in R. The model included the contrast-coded fixed effect of Accent Familiarity (Unfamiliar vs. Familiar) and Participant Age in months (a continuous variable). The model also included the interaction between Accent Familiarity and Participant Age. The maximal random effects structure that would converge was implemented (Barr, Levy, Scheepers, & Tily, 2013), which included random intercepts for subjects and items, and a by-subject random slope for Accent Familiarity. T-test and p-values were estimated using Satterthwaite's approximation for degrees of freedom. Cohen's *d* was estimated using the EMAtools package in R (Kleiman, 2017).

There was no main effect of Accent Familiarity, $b = -0.40$, $SE = 0.59$, $t(20.27) = -0.68$, $p = .507$, $d = -0.30$ indicating that infants' did not perform significantly better when the words were produced by the Familiar compared to the Unfamiliar talker. Furthermore, their performance was not above chance (.5) for either talker ($t < 1.55$, $p > .136$; Mandarin-accented proportion = .53; Australian-accented proportion = .53; See Figure 1). There was no main effect

of Age in months, $b = 0.02$, $SE = 0.02$, $t(20.99) = 1.06$, $p = .303$, $d = 0.46$ and no interaction between Accent Familiarity and Age, $b = 0.02$, $SE = 0.03$, $t(20.25) = 0.68$, $p = .505$, $d = 0.30$.

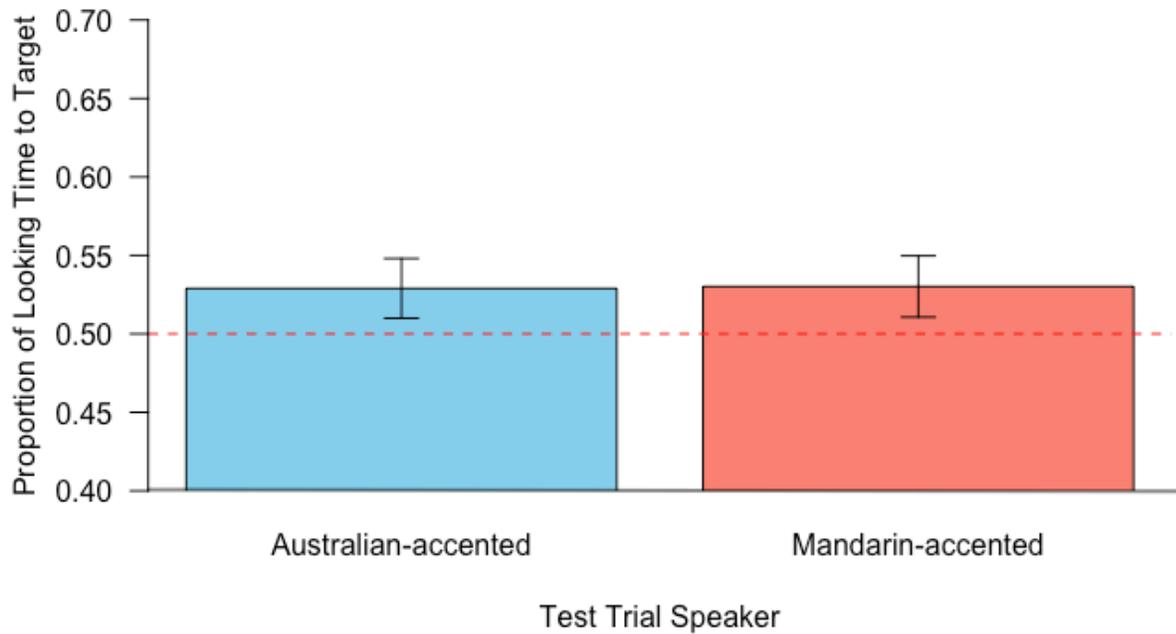


Figure 1. Proportion of looking time to target in the post-naming window starting 250 ms after word onset for the Australian- and Mandarin-accented test talkers following video pre-recording. Error bars denote +/- 1 standard error. Dashed horizontal line indicates chance performance.

DRAFT

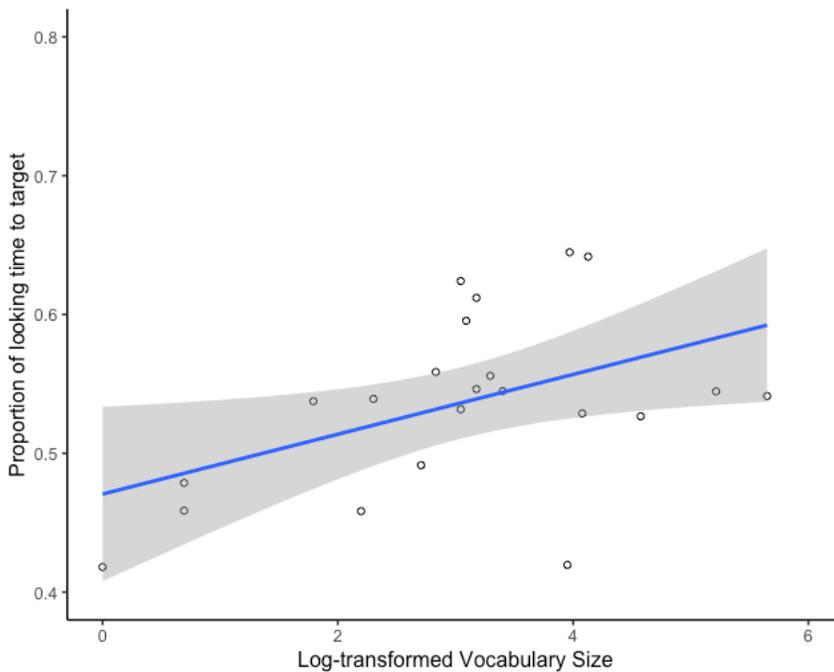


Figure 2. Mean proportion of looking time to the target versus log-transformed vocabulary size (as reported on the MacArthur-Bates Words and Gestures form). The area in grey represents the 95% CI.

Previous work has suggested that vocabulary size may be a better predictor of adaptation than age (Mulak et al., 2013; van Heugten et al., 2015). In order to test this, we conducted a post-hoc regression analysis to assess the effect of vocabulary size on adaptation ability (while controlling for age). A Shapiro-Wilk test indicated that vocabulary scores were not normally distributed, $W(22) = 0.64$, $p < .001$. Thus, similar to previous work, vocabulary size was log-transformed⁴ (Mulak et al., 2013; van Heugten et al., 2015). Overall, the model was not a good fit for the data, $F(2,19) = 2.81$, $p = .085$. That being said, vocabulary size, $b = 0.02$, $p = .040$,

⁴ Log-transformed scores were computed using the formula: $\text{Score} = \text{Log}(1+x)$. 1 was added to each value prior to taking the natural log as some participants had a reported vocabulary score of 0.

seemed to be a better predictor of infants' portion of looks to the target than age, $b = 0.01$, $p = .708$.

Taken together we find no evidence that video exposure facilitated 16.5-month-olds' adaptation to an unfamiliar Mandarin-accented talker. Indeed, only one study has shown that brief video exposure to an unfamiliar-accented speaker can facilitate adaptation in children under 18 months of age (van Heugten & Johnson, 2014). That study tested Canadian children's ability to understand words in an unfamiliar Australian accent by measuring their preferences for known words over unknown words. In contrast, the other study in which children were pre-exposed to a video of an accented speaker showed no evidence of adaptation after exposure, even in children as old as 20 months of age (van Heugten et al., 2015). This could suggest that the impact of video exposure may be fragile. It is also plausible that differences in the methods used to assess word recognition may have contributed to these discrepancies. The preferential looking paradigm used in the present experiment (and in van Heugten et al., 2015) may be more difficult for young children as it requires them to access the meaning of the word (and look at the correct image) rather than simply identifying whether a string of words sounds familiar or unfamiliar. In addition, in the present study, children were given exposure to a non-native accent (i.e., Mandarin-accented English), which may be more difficult than the regional accents used in previous work (e.g., Australian English in van Heugten & Johnson, 2014). Compared to regional-accented speakers, non-native speakers have been argued to show greater variability in their production of phonemes (Floccia, Goslin, Girard, & Konopczynski, 2006; Goslin, Duffy, & Floccia, 2012; see however Vaughn, Baese-Berk, & Idemarub, 2019) which may have made the exposure phase less effective in helping children to make the appropriate phonetic adjustments during adaptation.

Given that video exposure did not seem to facilitate adaptation in Experiment 1, in Experiment 2 we attempt to strengthen the impact of exposure by having children interact (face-to-face) with the unfamiliar accented talker prior to test. It has been argued that language learning, especially in infancy, may be socially gated (Kuhl, 2007). There have been numerous studies demonstrating that infants learn more in situations where there is social interaction compared to television or video exposure alone (Barr, 2013; Deloache et al., 2010; Krcmar et al., 2007; Kuhl et al., 2003; Lytle et al., 2018; Roseberry et al., 2014). Given the benefits of live exposure in learning, in Experiment 2, we test whether 10-minutes of live exposure to a non-native (Mandarin accented) speaker or a regional (Australian accented) speaker may help children to recognize known words produced by those speakers.

Experiment 2

In Experiment 1, we found no evidence that video exposure to a Mandarin-accented speaker facilitated 15- to 18-month-olds' recognition of words in that speaker's voice. In Experiment 2, we examine whether infants might benefit from a more ecologically valid face-to-face pre-exposure phase. Numerous studies have highlighted that language learning under the age of 3, from foreign phonemic contrasts (Kuhl et al., 2003) to novel vocabulary items (Krcmar et al., 2007; Roseberry et al., 2014), is less successful via non-socially contingent video when compared to conditions with live interaction. Considering that prior work on accent adaptation has almost exclusively relied on pre-recorded audio-visual exposure phases, they may have underestimated the potential value of exposure to accented speech and the extent to which young children can adapt to an unfamiliar accent.

In Experiment 2, children interacted with either the Mandarin-accented talker or the Australian-accented talker who recorded the test phase stimuli. In order to maximize the likelihood that we would observe adaptation, we tested a wider range of ages (children from 15 to 24 months of age) than we did in Experiment 1. These ages were selected because they enabled us to examine children's performance under and over the age of 19 months, which has been considered a critical turning point in the development of phonological constancy (Best et al., 2009; van Heugten et al., 2018).

Instead of having parents read a specific storybook to their child at home, parents were asked to bring in their child's favorite storybooks to be read with the unfamiliar accented speaker in the lab. This was done to increase the ecological validity of the task, and also to ensure that the speaker read books that the child knew well and would be interested in. After 10 minutes of live interaction, children were tested using the preferential looking task from Experiment 1. To examine the types of adaptation strategies that children employ as a result of exposure (i.e., general expansion or accent specific adaptation), children were tested on both the Australian and the Mandarin speaker's accent at test. If live exposure is effective in facilitating accent specific adaptation, then children will look more to the target image when the target word is produced in the accent they were exposed to compared to the accent they were not exposed to. In contrast, if live exposure leads children to loosen their category boundaries (i.e., employ a general expansion strategy), then exposure may similarly facilitate children's recognition of words in both the exposed and the unexposed accents. As in previous work, we also predict that there will be age-related changes in children's ability to understand accented speech. Previous studies have shown that children above the age of 19 months can spontaneously recognize unfamiliar accent variants (e.g., Best et al., 2009); thus, children over 19 months may demonstrate above chance

recognition for both the Mandarin-accented and Australian-accented English talkers (regardless of who they interacted with). Regardless of older children's spontaneous ability to cope with accented speech, interacting with a talker may still result in higher recognition accuracy for the items produced by that talker compared to the items produced by the unfamiliar talker.

Methods

Participants

Sixty-four Canadian English-learning infants were tested ($M_{age} = 571.45$ days, range = 458-732 days, 30 males) who satisfied the same criteria as in Experiment 1. For children between 15 and 18 months, caregivers completed the Words and Gestures version of the CDI (Mean words produced = 34.60, $SD = 36.55$)⁵. For children between 18 and 24 months, caregivers completed the Words and Sentences version (Mean words produced = 208.13, $SD = 152.01$). An additional 20 infants were excluded from the study due to fussiness ($n = 16$), caregivers not bringing books for the storybook reading ($n = 1$), experimenter/equipment error ($n = 2$), and prior exposure to one of the accents in the study ($n = 1$).

Stimuli

For the storybook reading, parents were asked to bring a few of the child's favorite books, to ensure children would be interested in the books and would be sufficiently familiar with the words in the story. The books were read by either the Mandarin or the Australian-accented talker that were heard in the test phase. Prior to the reading, the talkers noted any target words from the test phase that occurred in the books so as to avoid producing them during the

⁵ The caregivers of two children did not submit their vocabulary forms.

reading. The test phase was identical to Experiment 1. According to parental report, the 15- to 18-month-old children in this study understood 69.92% ($SD = 19.88$) of the target words presented in the test phase and the children between 19 and 24 months understood 90.63% ($SD = 16.29$).

Procedure

For the storybook reading, children were seated with their caregiver and interacted with one of the two talkers from the preferential looking task. The talker read the books in a lively, engaging manner for approximately 10 minutes. Immediately after the storybook reading, children were seated on their caregiver's lap in a sound-attenuated booth and the test phase began.

Results and Discussion

Similar to Experiment 1, the proportion of looking time to the target was calculated for the entire post-naming period beginning 250 ms after word onset (Figure 2) and the same trial exclusion measures were taken. A linear mixed-effects regression model was constructed with the exposure phase Story Reader (Australian-accented talker vs. Mandarin-accented talker; contrast coded) and Accent Familiarity (Unfamiliar accented talker vs Familiar accented talker; contrast coded) and Age in months (continuous variable) entered as fixed effects. We also included the interaction between Story Reader and Accent Familiarity and the interaction between Age and Accent Familiarity. The maximal random effects structure was implemented which included random intercepts for subjects and items, and by-item slopes for Age and by-subject slopes for Accent Familiarity. LMER follow-up tests were conducted using the emmeans package in R (Lenth, 2019).

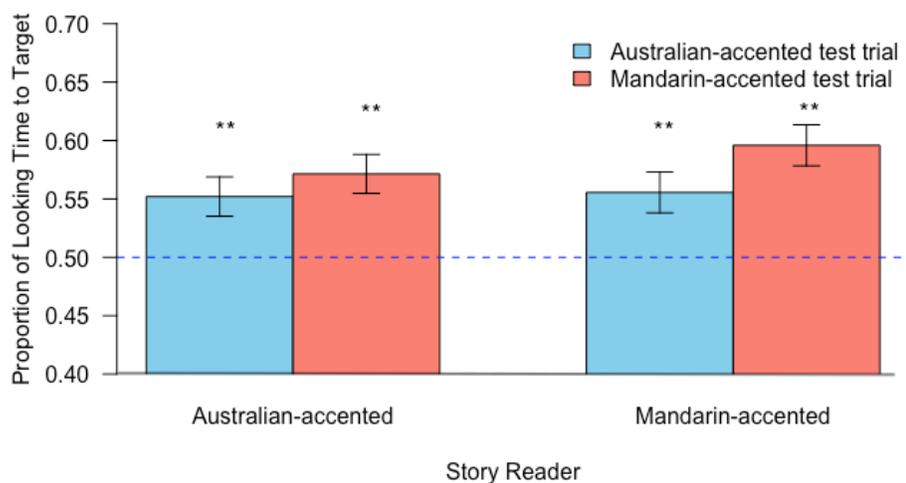


Figure 3. Infants' recognition of words produced by the Australian and Mandarin-accented speaker after 10 minutes of live exposure to either the Australian or the Mandarin-accented speaker. Proportion of looking time to target in the post-naming window starting 250 ms after word onset. Error bars denote +/- 1 standard error. Dashed horizontal line indicates chance performance. Asterisks denote performance significantly above chance (i.e., * $p < .05$, ** $p < .01$).

DRAFT ONLY

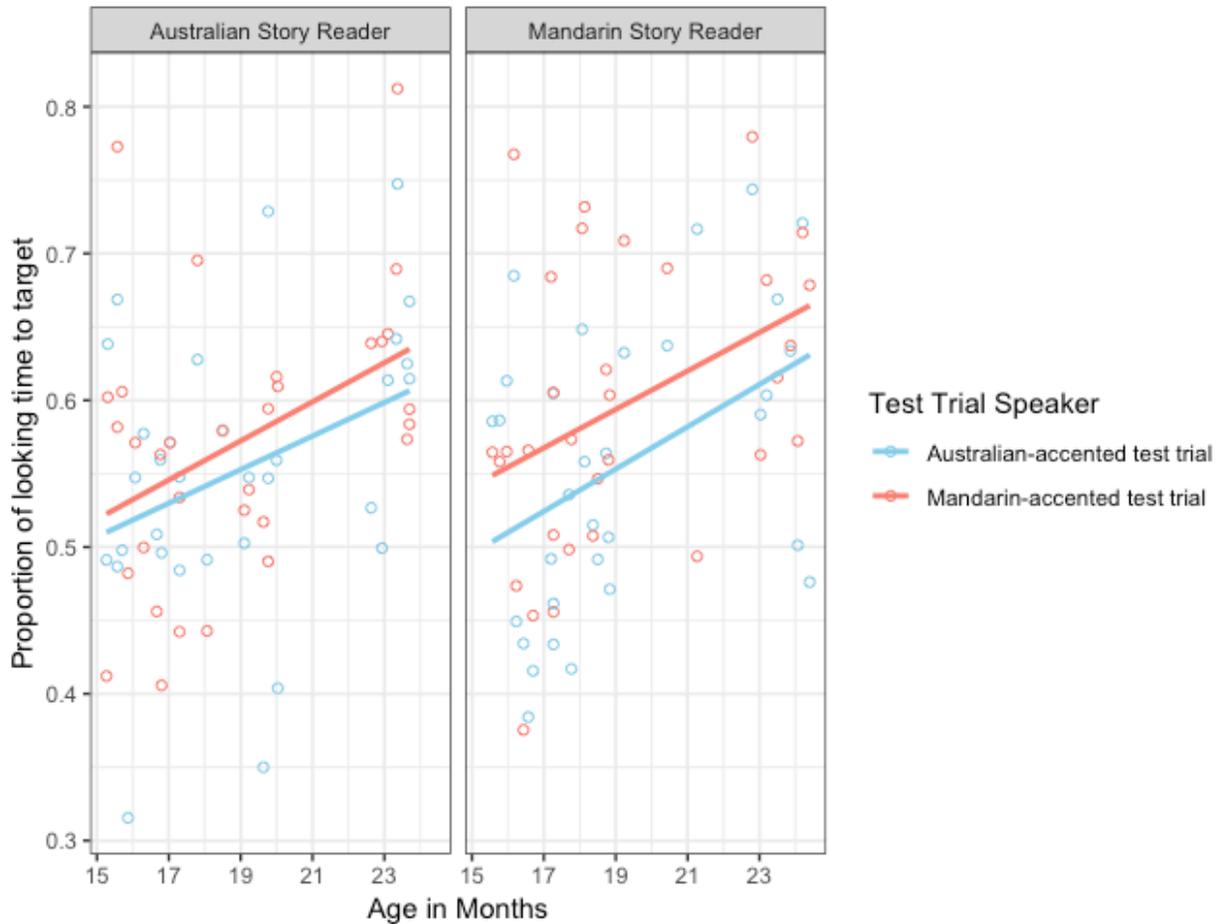


Figure 4. Proportion of looking time to target in the post-naming window starting 250 ms after word onset, broken down by age.

There was no main effect of Story Reader meaning that overall children who interacted with the Australian speaker did not perform significantly better in the test phase compared to children who interacted with the Mandarin speaker, $b = 0.01$, $SE = 0.02$, $t(61.58) = 0.52$, $p = .602$, $d = 0.13$. We also did not see a main effect of Accent Familiarity, meaning that children did not perform significantly better on the trials produced by the talker they interacted with (the Familiar talker) compared to trials produced by the Unfamiliar talker, $b = 0.06$, $SE = 0.10$, $t(1337.11) = 0.59$, $p = .553$, $d = 0.03$.

As predicted, we observed a main effect of Age, $b = 0.01$, $SE = 0.003$, $t(53.56) = 3.66$, $p < .001$, $d = 1.00$ such that older children looked more to the target image when cued compared to

younger children (see Figure 3). There was no interaction between Accent Familiarity and Age, $b = -0.002$, $SE = 0.01$, $t(1336.12) = -0.46$, $p = .645$, $d = -0.03$ meaning that the impact of exposure did not change as a function of our continuous variable age. There was a weak but significant interaction between Story Reader and Accent Familiarity, $b = 0.06$, $SE = 0.03$, $t(1335.37) = 2.04$, $p = .042$, $d = 0.11$. Post-hoc pairwise comparisons suggest that children showed greater recognition of words in a Mandarin accent after exposure to the Mandarin speaker, $b = 0.04$, $t(1333) = 2.12$, $p = .035$, but they were no better at recognizing words in the Australian accent after exposure to the Australian speaker, $b = -0.02$, $t(1338) = -0.78$, $p = .434$. This is surprising, as we would predict that exposure to the Mandarin and Australian speakers should be equally effective in facilitating adaptation to that specific speaker's accents. It is possible that given the non-random assignment of children into groups (i.e., they were assigned based on the story reader's availability, with the Mandarin story reader available primarily on weekdays and the Australian available primarily on weekends), there could have been some unintended systematic differences in the children assigned to the Mandarin vs. the Australian Story Reader conditions. Indeed, there may also be differences in the vocabulary size of the children exposed to the Australian-accented speaker compared to the Mandarin-accented speaker. In the younger half of the sample (under 19 months of age), that completed the Words and Gestures form, the group that interacted with the Australian speaker had a smaller vocabulary ($M = 20.53$, $SD = 20.88$) than the group that interacted with the Mandarin speaker ($M = 48.67$, $SD = 43.68$), $t(20.08) = -2.25$, $p = .036^6$. In the older half of the sample (that completed the CDI: Words and Sentences form), the children that interacted with the Australian speaker ($M = 175.56$ words, $SD = 116.36$) were not statistically different (in terms of their vocabulary size) compared to those who

⁶ Note: $N=30$; Welch's t-test, due to in-equality of error variances.

interacted with the Mandarin speaker ($M = 240.69$, $SD = 178.74$), $t(30) = -1.22$, $p = .231$.

Previous work has suggested that vocabulary size can play a substantive role in infants' ability to adapt to an unfamiliar accent (see Mulak et al., 2013; van Heugten et al., 2015). Thus, the vocabulary size difference in the younger half of the sample may have contributed to differences in the effectiveness of the pre-exposure phase in the Australian compared to Mandarin exposed infants.

To begin to assess the contribution of vocabulary size in predicting adaptation in this study, we conducted an exploratory post-hoc regression analysis examining the impact of children's age and log-transformed vocabulary size⁷ on their mean proportion of looking time to the target. Overall, the model accounted for 20.53% of the variance in task performance, $F(2,59) = 8.88$, $p < .001$. In contrast to Experiment 1, age seemed to be a better predictor of performance, $b = 0.01$, $p = .077$, than vocabulary size, $b = 0.01$, $p = .164$. This may be due to that fact that we tested a wider age range in this experiment (including infants over 19 months of age). It is difficult to tease apart the unique effects of age and log-transformed vocabulary size given that these two variables are correlated, $r(60) = .71$, $p < .001$.

⁷ Note. Because two different versions of the CDI were used in this sample (i.e., the CDI: Words and Gestures and the CDI: Words and Sentences). Vocabulary scores were converted into estimates of total vocabulary size using a Vocabulary size calculator: (<https://www.bcbl.eu/cdi/>; Mayor & Plunkett, 2011). Comparisons using raw vocabulary scores follow the same pattern of results. Two children were excluded from this analysis, as their caregivers did not submit their vocabulary forms. The total Vocabulary size variable was log transformed, as it violates assumptions of normality (Shapiro-Wilk test $W(62) = 0.63$, $p < .01$)

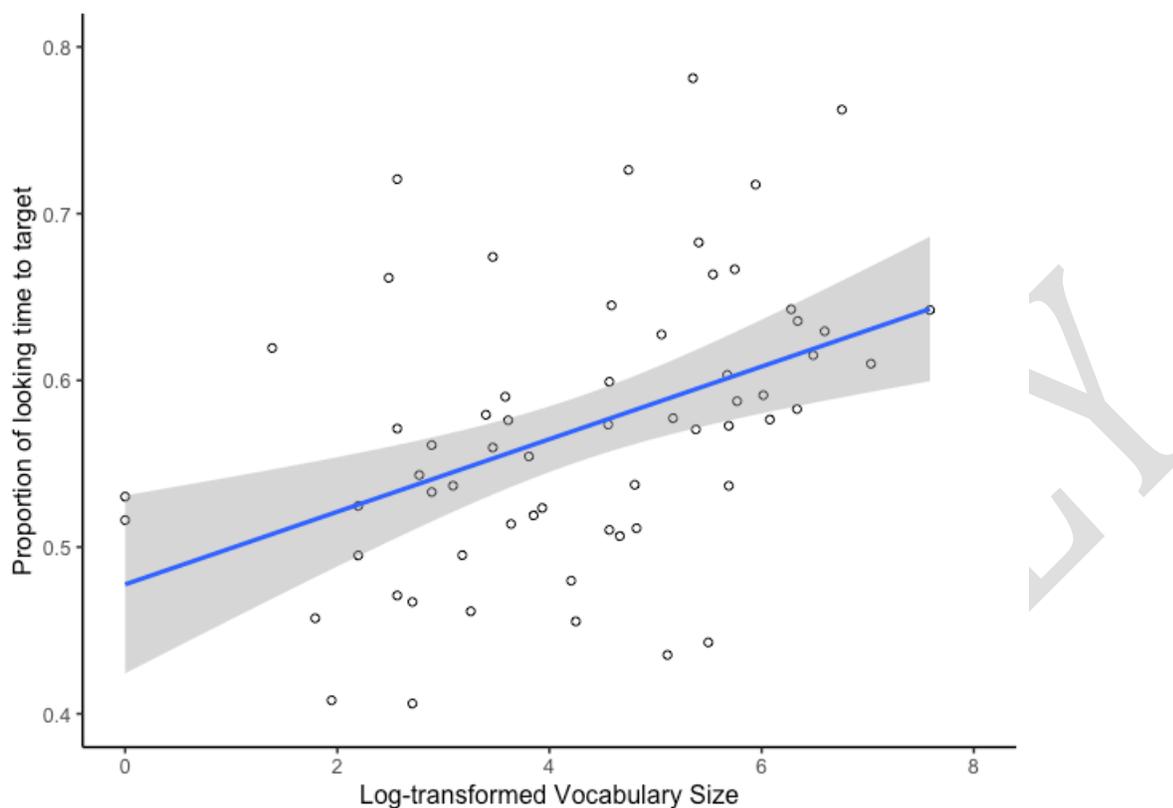


Figure 5. Mean proportion of looking time to the target versus log-transformed total vocabulary size. The area in grey represents the 95% CI.

As a group, the 15- to 24-month-olds tested in this experiment were above chance (.5) in recognizing words produced by both speakers regardless of who they interacted with in the exposure phase (all t s > 3.09, all p s < .004; see Figure 2). This makes sense given that half the sample ($n = 32$) is over 19 months of age and may (due to the development of phonological constancy; Best et al., 2009; van Heugten et al., 2018) already be able to spontaneously recognize pronunciation variants that deviate from their own native-accented norms. In order to better compare our findings with Experiment 1 (which tested 15- to 18-month-olds) and examine children who cannot spontaneously comprehend accented speech, we conducted a planned follow-up analysis looking at the effect of pre-exposure on the 32 children who were between 15 and 18 months. Similar to the LMER results reported above, in this younger subset of the sample

(Mean Age = 16.7 months) infants exposed to the Mandarin speaker trended towards performing better on the Mandarin ($M = .57$, $SD = 0.11$) compared to the Australian trials ($M = .52$, $SD = 0.09$), $t(15) = 1.95$, $p = .071$, but children exposed to the Australian speaker did not perform significantly better on the Australian ($M = .53$, $SD = 0.08$) compared to the Mandarin test trials ($M = .54$, $SD = 0.10$), $t(15) = -0.39$, $p = .705$.⁸ The group who was exposed to the Mandarin speaker was above chance in recognizing words in the Mandarin accent ($M = .57$), $t(15) = 2.48$, $p = .026$, but not the unfamiliar Australian accent ($M = .52$), $t(15) = 0.81$, $p = .433$. Although this subset of the sample is small ($N = 16$), these findings are consistent with the notion that in some circumstances live exposure might be able to help young children adapt to an unfamiliar accent.

Taken together, this is the first study to examine the effect of live exposure on infants' accent adaptation. The results of the LMER model above suggest that after interacting with the Mandarin-accented talker for 10-minutes, infants were better able to recognize words produced by that talker relative to the Australian-accented talker. Importantly, this experience with the Mandarin-accented talker appears to have enabled 15 to 24-month-olds to more accurately recognize words they had never heard her utter before. During the storybook reading, care was taken to ensure that the story reader would avoid or substitute words included in the test phase. This suggests that their success in the test phase was not due to recognizing the same words heard during exposure but arose from participants generalizing their exposure to novel items. However, this pre-exposure to a specific accent did not seem to generalize to a novel, unfamiliar (Australian) accent.

⁸ We conducted the same analysis on the group of children over 19-months. As expected (given their age) children in this group were above chance in recognizing words in both the Mandarin $t(31) = 8.19$, $p < .001$, and Australian accent $t(31) = 4.75$, $p < .001$. In this group the infants exposed to the Mandarin speaker did not perform better on the Mandarin ($M = .62$, $SD = 0.08$) compared to the Australian trials ($M = .59$, $SD = 0.09$), $t(15) = 1.36$, $p = .194$, and the children exposed to the Australian speaker did not perform better on the Australian test trials ($M = .57$, $SD = 0.11$) compared to the Mandarin test trials ($M = .60$, $SD = 0.08$), $t(15) = -1.17$, $p = .262$. This is in line with work suggesting that 25-month-olds (who were already able to spontaneously recognize words in accented speech) did not show greater recognition after pre-exposure (van Heugten et al., 2015).

Although we see some evidence that children performed better on the Mandarin accent test trials after exposure to the Mandarin speaker, we do not see the same benefit for Australian-accented trials after exposure to the Australian speaker. There are a number of possible explanations for this. As discussed above, differences in vocabulary size between the children assigned to interact with the Mandarin versus the Australian speaker may have played a role. It is also possible, given evidence that children are more likely to accept accented pronunciations from speakers of different races (Weatherhead & White, 2018), that the race of the Mandarin accented story reader (i.e., Asian vs. Caucasian) might have also contributed to these differences (a possibility that we discuss in greater depth in the general discussion). Finally, it is conceivable that the Australian-accented talker was simply less intelligible than the Mandarin-accented talker, making the task of adapting to her accent more challenging. It is this last possibility that we explore in Experiment 3.

Prior research with adult listeners adapting to non-native accented speech has suggested that adaptation was slower or required more exposure to low- versus high-intelligibility talkers (Bradlow & Bent, 2008). Indeed, we see that children in this sample seem to be better at understanding words in the Mandarin accent ($M = .58$, $SD = .10$) compared to the Australian accent ($M = .55$, $SD = .10$), $t(63) = 2.52$, $p = .014$ (regardless of exposure). Is it the case that the Australian speaker was less intelligible than the Mandarin speaker? To investigate this possibility, in Experiment 3 we compared the performance of young infants with no pre-exposure to the Mandarin- and Australian-accented talkers used in Experiments 1 and 2.

Experiment 3

In Experiment 3, we examined whether there were inherent intelligibility differences between the Australian- and Mandarin-accented talkers. That is, if infants received no unfamiliar

accent exposure and only heard Canadian-accented English during the storybook reading, how well would they comprehend the two talkers during the test phase? Given previous work suggesting that children over 19-months are able to spontaneously recognize words in an unfamiliar accent without exposure (e.g., Best et al., 2009; Mulak et al., 2013; van Heugten et al., 2015, 2018) and the fact that the older children in our experiment succeeded without exposure, only younger 16-month-olds were recruited for Experiment 3.

Method

Participants

Sixteen 16-month-old infants from the same population as Experiments 1 and 2 were included ($M_{age} = 500$ days, range = 470-531, 7 males). According to parental report on the CDI- Words and Gestures from, participants in this experiment produced an average of 29.13 words ($SD = 30.30$). This vocabulary size is comparable to the vocabulary size of the 16-month-olds who were given television exposure in Experiment 1, $t(36) = -0.91, p = .370$, and the 16-month-olds exposed to the Mandarin speaker in Experiment 2, $t(29) = -1.46, p = .156$. An additional 12 infants were tested in this experiment but were excluded due to fussiness ($n = 6$), prior accent exposure ($n = 2$), caregivers not bringing books for the storybook reading ($n = 2$), or being more than three weeks premature ($n = 2$).

Stimuli & Procedure

For the storybook reading, the children were assigned to be read to by one of two female Canadian-accented talkers. The same stimuli and procedure were used in the test phase as in Experiments 1 and 2. According to parental report, the children in this study understood 66.67% ($SD = 17.28$) of the target words presented in the test phase.

Results and Discussion

The proportion of looking time to target was analyzed for the same time window as Experiments 1 and 2 (Figure 5). A linear mixed-effects regression model was constructed with contrast-coded fixed effects of Test Trial Speaker (Australian-accented vs. Mandarin-accented), Age in months and their interaction. We included random intercepts for subjects and items, as well as a by-subject random slope for Test Trial Speaker. There was no main effect of Test Trial Speaker, $b = -0.51$, $SE = 0.76$, $t(37.48) = -0.67$, $p = .505$, $d = -0.22$. Thus, our data does not support the notion that it was easier for infants to recognize words in the Mandarin-accented ($M = .53$, $SD = 0.09$) compared to the Australian-accented ($M = .56$, $SD = 0.07$) talkers. There was also no significant effect of Age, $b = 0.02$, $SE = 0.02$, $t(25.31) = 0.75$, $p = .460$, $d = 0.30$ or its interaction with Test Speaker, $b = 0.03$, $SE = 0.05$, $t(37.24) = 0.63$, $p = .532$, $d = 0.21$). Although performance on the Mandarin-accented vs. Australian accented test trials was not statistically different, the proportion of looking time to target was above chance for the Australian-accented voice, $t(15) = 3.37$, $p = .004$. However, listeners were not significantly above chance for the Mandarin-accented talker, $t(15) = 1.06$, $p = .305$. As in the previous two experiments, we conducted a post-hoc regression analysis to examine the relationship between age, log-transformed vocabulary size and performance. Given the small sample size, this model does not seem to be a good fit for the data $F(2,13) = 1.35$, $p = .292$. In this experiment, neither age, $b = 0.02$, $p = .288$ nor vocabulary size, $b = 0.01$, $p = .205$ predicted the proportion of looking time to target.

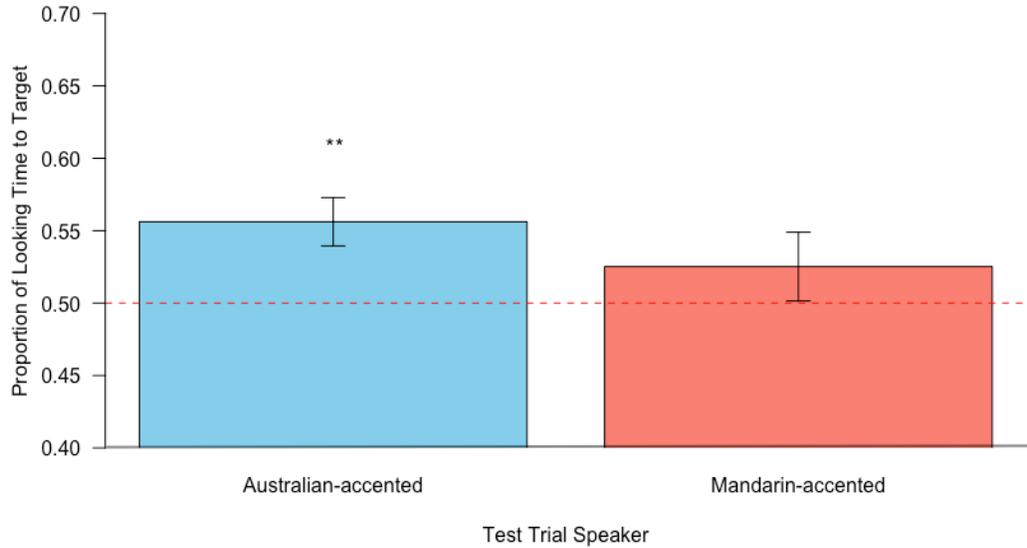


Figure 6. Proportion of looking time to target for post-naming window starting 250 ms after word onset for the Australian- and Mandarin-accented test talkers following a storybook reading with a Canadian-accented talker. Error bars denote +/- 1 standard error. Dashed horizontal line indicates chance performance. Asterisks denote performance significantly above chance (i.e., * $p < .05$, ** $p < .01$).

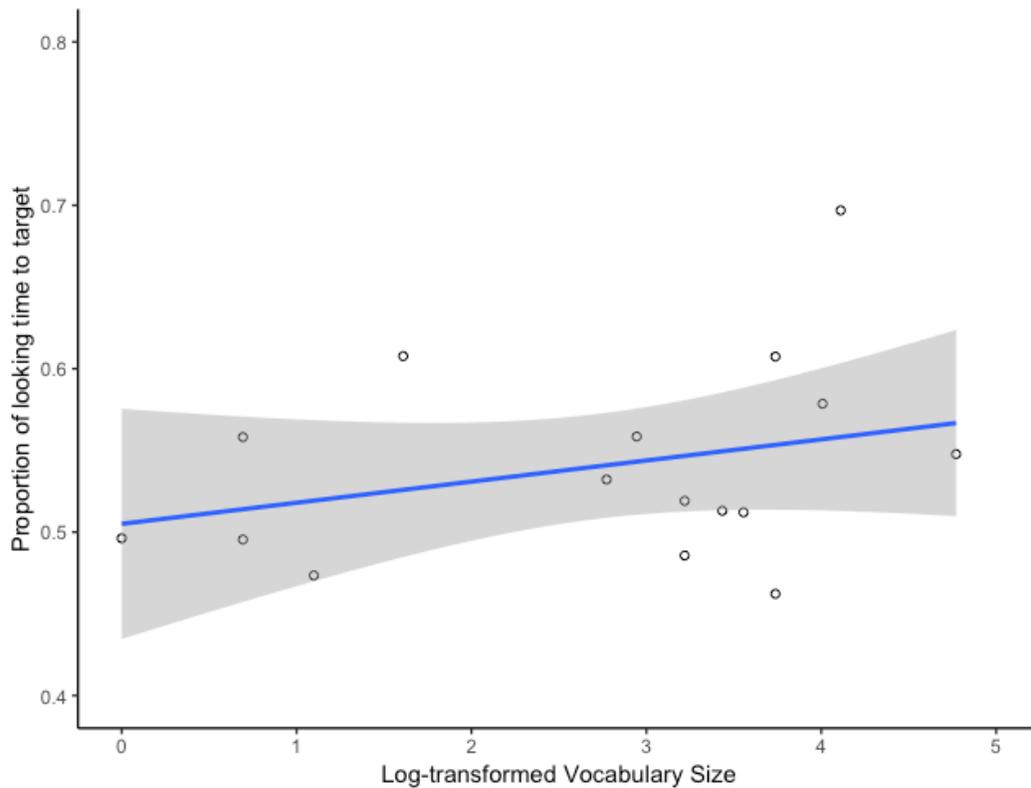


Figure 7. Mean proportion of looking time to the target versus log-transformed vocabulary size. The shaded grey area represents the 95% CI.

Overall, the findings of Experiment 3 suggest that infants' inability to adapt to Australian-accented English in Experiment 2 was not as a result of the Australian talker being inherently less intelligible than the Mandarin-accented talker. Indeed, we find no evidence that there are differences in infant's ability to recognize words produced by the Australian versus the Mandarin talker. Interestingly, the 16-month-olds in this study were able to recognize the words produced by the unfamiliar Australian-accented speaker. This the first study to show comprehension of an unfamiliar accented speaker in 16-month-olds, who have not been given any pre-exposure to the accent.

General Discussion

The need to accommodate variant pronunciations of familiar words is a common challenge for infants and adults alike. Listeners often encounter talkers whose phonetic realization of words, as a result of a non-native or regional accent, differs from their own. While adults have been shown to accommodate very quickly to foreign-accented speech (e.g., Clarke & Garrett, 2004), infants often struggle to understand familiar words spoken in unfamiliar accents (Best et al., 2009; Floccia et al., 2012; Mulak et al., 2013). In this study, we examined the impact of pre-exposure on toddlers' adaptation to two unfamiliar accents (Mandarin- and Australian-accented English). In Experiment 1, toddlers were presented with 10 minutes of video exposure to a Mandarin-accented speaker. In Experiment 2, the pre-exposure phase consisted of 10 minutes of live face-to-face interaction with either the Mandarin or the Australian-accented talker. Finally, in Experiment 3 toddlers interacted with a Canadian-accented talker prior to completing the test phase. Our results suggest that in some circumstances, live exposure can facilitate adaptation. However, the impact of exposure seems to vary depending on the specific

characteristics of the speakers (e.g., accent characteristics) and the type of exposure phase (video vs. live). Only the children who had face-to-face interaction with the Mandarin speaker (in Experiment 2) showed improvements in their ability to understand words spoken in that speaker's accent.

Consistent with prior studies, our results demonstrate that children's ability to rapidly accommodate pronunciation variation increases with age. In Experiment 1, 16-month-olds who were exposed to a video of an unfamiliar Mandarin-accented speaker struggled to recognize words in that accent. In Experiment 2, we tested a wider age range of children after live interaction with either the Mandarin or the Australian speakers from the test phase. Here, children achieved above chance word recognition for both an unfamiliar regional- and non-native-accented talker, regardless of who they interacted with during the storybook reading phase. These results are likely driven by the older children in the sample, as children's age in months was a strong predictor of their proportion of looking time to the target. This provides further evidence in support of the claim that infants establish phonological constancy before their second birthday (Best et al., 2009). That is, they possess lexical representations that are sufficiently abstract to be able to efficiently handle phonetic variation that arises from accent differences.

The results of Experiment 2 are compatible with the possibility that engaging, live interaction with the Mandarin-accented talker may have helped to facilitate the subsequent recognition of familiar words that they had never heard that talker produce. This is evidenced by the fact that children who interacted with the Mandarin speaker were better at recognizing the Mandarin-accented test items compared to the Australian-accented items. Furthermore, in an analysis focusing on just the younger children in Experiment 2, we find evidence that the

younger children (15- to 18-month-olds) who interacted with the Mandarin speaker were able to recognize words in that speaker's accent. At the same time, infants who were either not exposed to that accent (receiving Australian- or Canadian-accent exposure), or who received exposure via a pre-recorded video were not statistically above chance in their performance on the Mandarin-accented words. Although this subset of our sample is relatively small, these findings could suggest that providing a more ecologically valid context for adaptation (e.g., interacting with a real person reading their favorite books) might be beneficial for infants under 19 months. It is also possible that methodological differences between Experiment 1 (in which parents were told which books to read at home) and Experiment 2 (in which parents brought in their children's favorite books to be read by the speaker), might have bolstered the effectiveness of live exposure. Nonetheless, these findings provide the first evidence of successful adaptation in children as young as 16 months of age to a non-native accented talker.

The current study is also unique amongst accent adaptation studies with children under 19 months because we attempt to investigate whether adaptation (at this age) is accent-specific. As discussed in the Introduction, the majority of prior work has tested infants' adaptation using the same accent to which they were initially exposed (e.g., Schmale et al., 2012; van Heugten & Johnson, 2014; van Heugten et al., 2015; White & Aslin, 2011). As a result, little is known about the nature of these adaptive processes, whether it involves making targeted phonemic adjustments for a specific accent or talker or whether infants are simply increasing their tolerance for deviant pronunciations more generally. The latter might predict that infants in the present work would be able to generalize their accent exposure to a novel, unfamiliar accent (in this case, Australian-accented English), despite only being exposed to the Mandarin talker. In our study, evidence of facilitation was only found for the Mandarin-accented talker heard during the

storybook reading and not for the unfamiliar Australian talker. Given that exposure was manipulated between subjects, it is difficult to determine to what degree the Mandarin exposure may have facilitated (or even hindered) recognition for the Australian accent. However, we are fairly certain given the results of Experiment 3 (in which 15- to 18-month-olds did not show recognition of words in the Mandarin accent after exposure to a Canadian speaker), that the younger children in Experiment 2 would not have recognized words in the Mandarin accent without live exposure to the Mandarin speaker. These results may suggest that adaptation, at least in this context, could involve making targeted, accent-specific shifts as a result of exposure to shifts perceived in the input (see Kleinschmidt & Jaeger, 2015 for a discussion). This is in line with previous adult literature (e.g., Bradlow & Bent, 2008) that did not find cross-accent generalization following exposure to a single talker of an accent. This is not to say that an adaptation strategy involving a more relaxed tolerance for pronunciation deviations would never be employed (see for example Potter & Saffran, 2017); it may be dependent on the amount of talker and accent variation presented during exposure. Indeed, in this experiment, infants heard a single speaker producing one non-native accent; however, perhaps such a strategy would be utilized in contexts with greater talker or accent variation (Baese-Berk, Bradlow, & Wright, 2013; Schmale et al., 2015). Future work should examine in greater detail the circumstances in which these different adaptation mechanisms may be used.

Interestingly, the benefit of live interaction found for the Mandarin speaker was not similarly found for the Australian speaker. That is, infants who participated in the storybook reading with the Australian-accented talker were not significantly better at recognizing words produced by that talker during the test phase. One possible explanation could be that the Australian-accented talker was less intelligible than the Mandarin talker, which could have made

adaptation more challenging, even with live exposure. Indeed, the phonetic transcriptions (reported in Appendix A), indicate that the Australian accent deviated from Canadian English on more of the target words compared to Chinese-accented English speaker. That being said, phonetic transcriptions do not fully capture pronunciation differences between accents, and we know young children are sensitive to sub-phonemic differences in the pronunciation of speech sounds (Paquette-Smith, Fecher, & Johnson, 2016). Moreover, the results of Experiment 3 do not support the notion that the Mandarin speaker was easier to comprehend. When infants received Canadian-accented exposure in the exposure phase, they did not perform better on the Mandarin- vs. Australian-accented words, which suggests that the findings of Experiment 2 are unlikely to be driven by intelligibility differences between the two talkers or the differences in the mapping between Canadian children's native accent and Australian versus Mandarin accents. An alternative possibility, that we favor, is that given the non-random assignment of children to story reader (i.e., children exposed to the Australian speaker were tested during week because of experimenter's availability), the group of infants that were exposed to the Mandarin vs. the Australian speakers may have differed systematically in a way that influenced vocabulary size. Indeed, the 15- to 18-month-old Australian-exposed children tested in Experiment 2 had significantly lower vocabulary sizes than the Mandarin exposed children. However, there were no differences in vocabulary size between the Mandarin exposed children in Experiment 2 and the other groups of 15- to 18-month-olds tested in Experiments 1 and 3.

An alternative possibility is that race may have influenced the extent to which infants adapted to the talkers. In the present work, our Australian-accented talker was a Caucasian woman, while our Mandarin-accented talker was an Asian woman, who did not match the race of the majority of our infants. Recent findings have indicated that 16-month-old infants recognized

familiar words when produced by an unfamiliar accent if produced by an other-race talker but not a same-race talker (Weatherhead & White, 2018). That is, seeing an other-race talker may have provided a cue that they were interacting with someone not from their language community, promoting them to rapidly make adjustments to accommodate the pronunciation variation. It is also possible that, similar to Schmale, Seidl, Cristia (2015), simply being exposed to some form of visual diversity (whether it be race or gender/age) might have prompted accommodation. This racial discrepancy in the current study could have supported adaptation for infants interacting with the other-race, Mandarin-accented talker and inhibited adaptation for those interacting with the same-race, Australian-accented talker. Taken together, one or a combination of these considerations may have conspired to constrain infants' ability to benefit from live exposure to the Australian speaker. In order to begin to draw firm conclusions about the effect of live interaction on adaptation, a replication with a larger sample size and stricter control over variables such as the race of the speaker and the time of the week children were tested is needed.

In sum, the present work demonstrated adaptation following live interaction to a non-native talker in infants under 2 years of age. Moreover, these adaptive processes appeared to be targeted to the specific accent to which they were exposed, as infants did not generalize to an unfamiliar, regionally accented talker. However, a separate group of infants were not found to be able to adapt to the regionally accented talker, perhaps due to a confluence of factors such as the talker's race and their overall vocabulary size. Taken together, these findings highlight that it is no longer sufficient to ask at what age a child can adapt to an unfamiliar accent but under what conditions adaptation is possible, considering not only age but vocabulary size, task demands, and the specific characteristics of the accent (e.g., the number and size of pronunciation

deviations relative to the native accent). Moreover, further research is needed examining the mechanisms underlying adaptation in young children by examining whether these adaptive processes involve making targeted phoneme-specific shifts in response to shifts perceived in the input or a more general relaxing of criteria for what constitutes an acceptable match between input and representation (e.g., Schmale et al., 2015), and critically, the circumstances which might induce the use of these different processes. In short, this work highlights the need to study adaptation in children more thoroughly and systematically to better elucidate the key factors enabling children's comprehension of unfamiliar accents. By doing so, researchers can generate better and more comprehensive models of spoken language development.

DRAFT ONLY

References

- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. <https://doi.org/10.1016/j.jml.2007.12.005>
- Baese-Berk, M. M., Bradlow, A. R., & Wright, B. A. (2013). Accent-independent adaptation to foreign accented speech. *JASA Express Letters*, *133*(3), EL174–EL180. <https://doi.org/10.1121/1.4789864>
- Baese-Berk, M., & Morrill, T. H. (2017). Speaking rate consistency in native and non-native speakers of English. *The Journal of the Acoustical Society of America*, *138*, EL223–EL228. <https://doi.org/10.1121/1.4929622>
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Barr, R. (2013). Memory Constraints on Infant Learning From Picture Books, Television, and Touchscreens. *Child Development Perspectives*, *7*(4), 205–210. <https://doi.org/10.1111/cdep.12041>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Bent, T., & Atagi, E. (2017). Perception of Nonnative-accented Sentences by 5- to 8-Year-olds and Adults: The Role of Phonological Processing Skills. *Language and Speech*, *60*(1), 110–122. <https://doi.org/10.1177/0023830916645374>
- Bent, T., Atagi, E., Akbik, A., & Bonifield, E. (2016). Classification of regional dialects,

- international dialects, and nonnative accents. *Journal of Phonetics*, 58, 104–117.
<https://doi.org/10.1016/j.wocn.2016.08.004>
- Bent, T., & Holt, R. (2018). Shhh... I Need Quiet! Children's Understanding of American, British, and Japanese-accented English Speakers. *Language and Speech*, 1–17.
<https://doi.org/10.1177/0023830918754598>
- 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. <https://doi.org/10.1111/j.1467-9280.2009.02327.x>
- Bradlow, A. R., & Bent, T. (2008). Perceptual Adaptation to Non-Native Speech. *Cognition*, 106(2), 707–729. <https://doi.org/10.1016/j.cognition.2007.04.005>
- Buckler, H., Oczak-Arsic, S., Siddiqui, N., & Johnson, E. K. (2017). Input matters: speed of word recognition in two-year-olds exposed to multiple accents. *Journal of Experimental Child Psychology*, 164, 87–100. <https://doi.org/10.1016/j.jecp.2017.06.017>
- Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *The Journal of the Acoustical Society of America*, 116(6), 3647–3658.
<https://doi.org/10.1121/1.1815131>
- Creel, S. C., Rojo, D. P., & Paullada, A. N. (2016). Effects of contextual support on preschoolers' accented speech comprehension. *Journal of Experimental Child Psychology*, 146, 156–180. <https://doi.org/10.1016/j.jecp.2016.01.018>
- Cristia, A., Seidl, A., Vaughn, C., Schmale, R., Bradlow, A., & Floccia, C. (2012). Linguistic processing of accented speech across the lifespan. *Frontiers in Psychology*, 3(NOV), 1–15.
<https://doi.org/10.3389/fpsyg.2012.00479>
- Cutler, A., & Broersma, M. (2005). Phonetic Precision in Listening. In W. J. Hardcastle & J.

- Mackenzie Beck (Eds.), *A Figure of Speech* (pp. 64–91). Mahwah, NJ: Lawrence Erlbaum Associates.
- Delle Luche, C., Durrant, S., Poltrock, S., & Floccia, C. (2015). A methodological investigation of the Intermodal Preferential Looking paradigm: Methods of analyses, picture selection and data rejection criteria. *Infant Behavior and Development, 40*, 151–172.
<https://doi.org/10.1016/j.infbeh.2015.05.005>
- Deloache, J. S., Chiong, C., Sherman, K., Islam, N., Vanderborght, M., Troseth, G. L., ... Doherty, K. O. (2010). Do Babies Learn From Baby Media ?
<https://doi.org/10.1177/0956797610384145>
- Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates Communicative Development Inventories: User's guide and technical manual* (2nd ed.). Baltimore, MD: Paul H. Brookes.
- Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology, 42*(1), 98–116. <https://doi.org/10.1037/0012-1649.42.1.98>
- 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*(1), 95–100. <https://doi.org/10.1016/j.cognition.2012.03.011>
- Floccia, C., Goslin, J., Girard, F., & Konopczynski, G. (2006). Does a regional accent perturb speech processing? *Journal of Experimental Psychology: Human Perception and Performance, 32*(5), 1276–1293. <https://doi.org/10.1037/0096-1523.32.5.1276>
- Furey, J. E. (2011). Production and maternal report of 16- and 18-month-olds' vocabulary in low- and middle-income families. *American Journal of Speech-Language Pathology, 20*(1),

38–46. [https://doi.org/10.1044/1058-0360\(2010/09-0073\)](https://doi.org/10.1044/1058-0360(2010/09-0073))

Goslin, J., Duffy, H., & Floccia, C. (2012). Brain & Language An ERP investigation of regional and foreign accent processing. *Brain and Language*, *122*, 92–102.

<https://doi.org/10.1016/j.bandl.2012.04.017>

Hollich, G. (2005). Supercoder: A program for coding preferential looking (Version 1.5) [Computer Software]. West Lafayette: Purdue University.

Houston-Price, C., Mather, E., & Sakkalou, E. (2007). Discrepancy between parental reports of infants' receptive vocabulary and infants' behaviour in a preferential looking task. *Journal of Child Language*, *34*(4), 701–724. <https://doi.org/10.1017/S0305000907008124>

Kleiman, E. (2017). EMAtools: Data Management Tools for Real-Time Monitoring/Ecological Momentary Assessment Data. R package version 0.1.3. <https://CRAN.R-project.org/package=EMAtools>

Kleinschmidt, D. F., & Jaeger, T. F. (2015). Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review*, *122*, 148–203.

<https://doi.org/10.1037/a0038695>

Krcmar, M., Grela, B., & Lin, K. (2007). Can toddlers learn vocabulary from television? An experimental approach. *Media Psychology*, *10*(1), 41–63.

<https://doi.org/10.1080/15213260701300931>

Kuhl, P. K. (2007). Is speech learning “gated” by the social brain? *Developmental Science*, *10*(1), 110–120. <https://doi.org/10.1111/j.1467-7687.2007.00572.x>

Kuhl, P. K., Tsao, F.-M., & Liu, H.-M. (2003). Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(15), 9096–9101.

<https://doi.org/10.1073/pnas.1532872100>

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82(13).

<https://doi.org/10.18637/jss.v082.i13>

Lenth, R. (2019). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.4.3.01. <https://CRAN.R-project.org/package=emmeans>

Lytle, S. R., Garcia-sierra, A., & Kuhl, P. K. (2018). Two are better than one : Infant language learning from video improves in the presence of peers, *115*(40), 1–8.

<https://doi.org/10.1073/pnas.1611621115>

Mayor, J., & Plunkett, K. (2011). A statistical estimate of infant and toddler vocabulary size from CDI analysis. *Developmental Science*, 14(4), 769–785. <https://doi.org/10.1111/j.1467-7687.2010.01024.x>

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.

<https://doi.org/10.1111/cdev.12087>

Paquette-Smith, M., Fecher, N., & Johnson, E. K. (2016). Two-year-olds' sensitivity to subphonemic mismatch during online spoken word recognition. *Attention, Perception, and Psychophysics*, 78(8). <https://doi.org/10.3758/s13414-016-1186-4>

Potter, C. E., & Saffran, J. R. (2017). Exposure to multiple accents supports infants' understanding of novel accents. *Cognition*, 166, 67–72.

<https://doi.org/10.1016/j.cognition.2017.05.031>

Roseberry, S., Hirsh-Pasek, K., & Golinkoff, R. M. (2014). Skype me! Socially contingent

interactions help toddlers learn language. *Child Development*, 85(3), 956–970.

<https://doi.org/10.1111/cdev.12166>

Schmale, R., Cristia, A., & Seidl, A. (2012). Toddlers recognize words in an unfamiliar accent after brief exposure. *Developmental Science*, 15(6), 732–738.

<https://doi.org/10.1111/j.1467-7687.2012.01175.x>

Schmale, R., Cristia, A., Seidl, A., & Johnson, E. K. (2010). Developmental Changes in Infants' Ability to Cope with Dialect Variation in Word Recognition. *Infancy*, 15(6), 650–662.

<https://doi.org/10.1111/j.1532-7078.2010.00032.x>

Schmale, R., Hollich, G., & Seidl, A. (2011). Contending with foreign accent in early word learning. *Journal of Child Language*, 38(5), 1096–1108.

<https://doi.org/10.1017/S0305000910000619>

Schmale, R., Seidl, A., & Cristia, A. (2015). Mechanisms underlying accent accommodation in early word learning: evidence for general expansion. *Developmental Science*, 18(4), 664–670. <https://doi.org/10.1111/desc.12244>

Swingley, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, 76(2), 147–166. [https://doi.org/10.1016/s0010-0277\(00\)00081-0](https://doi.org/10.1016/s0010-0277(00)00081-0)

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.

<https://doi.org/10.1080/10489223.2015.1047096>

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.

<https://doi.org/10.1037/a0032192>

van Heugten, M., Krieger, D. R., & Johnson, E. K. (2015). The Developmental Trajectory of

- Toddlers' Comprehension of Unfamiliar Regional Accents. *Language Learning and Development*, 11(1), 41–65. <https://doi.org/10.1080/15475441.2013.879636>
- van Heugten, M., Paquette-Smith, M., Krieger, D. R., & Johnson, E. E. K. (2018). Infants' recognition of foreign-accented words: Flexible yet precise signal-to-word mapping strategies. *Journal of Memory and Language*, 100(100), 51–60. <https://doi.org/10.1016/j.jml.2018.01.003>
- Vaughn, C., Baese-Berk, M., & Idemarub, K. (2019). Re-Examining Phonetic Variability in Native and Non-Native Speech. *Phonetica*, 76, 349–357. <https://doi.org/10.1159/000487269>
- Wade, T., Jongman, A., & Sereno, J. A. (2007). Effects of Acoustic Variability in the Perceptual Learning of Non-Native-Accented Speech Sounds. *Phonetica*, 64(2–3), 122–144. <https://doi.org/10.1159/000107913>
- Weatherhead, D., & White, K. S. (2018). And then I saw her race: Race-based expectations affect infants' word processing. *Cognition*, 177, 87–97. <https://doi.org/10.1016/j.cognition.2018.04.004>
- White, K. S., & Aslin, R. N. (2011). Adaptation to novel accents by toddlers. *Developmental Science*, 14(2), 372–384. <https://doi.org/10.1111/j.1467-7687.2010.00986.x>

Appendix A: Table of Phonetic Transcriptions

	Mandarin-accented English	Australian-accented English	Canadian-accented English
<i>baby</i>	'beɪbi	'beɪbi	'beɪbi
<i>ball</i>	bɑ:l	bɔ:l	bɑ:l
<i>bear</i>	bɪ	be:ə	bɛəɪ
<i>bee</i>	bi:	bi:	bi:
<i>bird</i>	bə'dɪ	bɛ:d	bə'd
<i>boat</i>	bɔ ^w t	bɔut	bɔ ^w t
<i>bunny</i>	'bɑ'ni	'bʌni	'bʌni
<i>bus</i>	bʌs	bɛs	bʌs
<i>cake</i>	k ^h eɪk	k ^h eɪk	k ^h eɪk
<i>car</i>	k ^h ɑ	k ^h a:	k ^h ɑ
<i>cat</i>	k ^h æt	k ^h æt	k ^h æt
<i>cow</i>	k ^h aw	k ^h aw	k ^h aw
<i>dog</i>	dɑk	dɔg	dɑ:g
<i>duck</i>	dɑ:k	dak	dʌk
<i>fish</i>	fɪʃ	fɪʃ	fɪʃ
<i>foot</i>	fʊt	fʊt	fʊt
<i>hand</i>	hænd	hænd	hænd
<i>hat</i>	hæt	hɑ:t	hæt
<i>sheep</i>	ʃi:p	ʃi:p	ʃi:p
<i>shoe</i>	ʃu: ^w	ʃu: ^w	ʃu: ^w
<i>teeth</i>	t ^h iθ	t ^h əiθ	t ^h iθ
<i>toes</i>	t ^h o ^w s	t ^h əʊz	t ^h o ^w z
<i>tree</i>	t ^h ri:	t ^h ri:	t ^h ri:
<i>truck</i>	t ^h ɹɛk	t ^h ɹɛk	t ^h ɹʌk