Learning to Contend with Accents in Infancy: Benefits of Brief Speaker Exposure

Marieke van Heugten and Elizabeth K. Johnson

Learning to Contend with Accents in Infancy: Benefits of Brief Speaker Exposure

Marieke van Heugten and Elizabeth K. Johnson

Department of Psychology, University of Toronto

3359 Mississauga Road North, Mississauga, ON, L5L 1C6, Canada

Author Note

Marieke van Heugten, Department of Psychology, University of Toronto; Elizabeth K. Johnson, Department of Psychology, University of Toronto

MvH is now at the Laboratoire de Sciences Cognitives et Psycholinguistique, CNRS/EHESS/DEC-ENS, Paris, France.

We thank Julie Kow and Victoria Ng for producing our speech materials as well as Craig Chambers for constructive comments on an earlier draft of this paper. This work was supported by grants from CFI and NSERC awarded to EKJ and an OGS grant and SRCD Dissertation Funding Award to MvH. Part of this work was reported in the first author's Ph.D. dissertation, and has been presented at the 12th International Congress for the Study of Child Language, Montreal, Canada; the 162nd Meeting of the Acoustical Society of America, San Diego, CA, and the 18th International Conference on Infant Studies, Minneapolis, MN.

Correspondence should be addressed to Elizabeth K. Johnson, Department of Psychology, University of Toronto Mississauga, 3359 Mississauga Road North, Mississauga, Ontario, Canada, L5L 1C6. E-mail: elizabeth.johnson@utoronto.ca. Phone: 905-569-4785. Fax: 905-569-4850.

Word count: 9158

Abstract

While adults rapidly adjust to accented speakers' pronunciation of words, young children appear to struggle when confronted with unfamiliar variants of their native language (e.g., American English-learning 15-month-olds cannot recognize familiar words spoken in Jamaican English; Best et al., 2009). It is currently unclear, however, why this is the case, or how infants overcome this apparent inability. Here, we begin to address these crucial questions. Experiments 1 and 2 confirm with a new population that infants are initially unable to recognize familiar words produced in unfamiliar accents. That is, Canadian-English learning infants cannot recognize familiar words spoken in Australian English until they near their second birthday. However, Experiments 3 and 4 show that this early inability to recognize accented words can readily be overcome when infants are exposed to a story read in the unfamiliar accent prior to test. Importantly, this adaptation only occurs when the story is highly familiar, consistent with the idea that top-down lexical feedback may enable the adaptation process. We conclude that infants, like adults, have the cognitive capacity to rapidly deduce the mapping between their own and an unfamiliar variant of their native language. Thus, the essential machinery underlying spoken language communication is in place much earlier than previous studies have suggested.

Keywords: language development, word recognition, infant speech perception, accent adaptation

Abstract Word Count: 207

Learning to Contend with Accents in Infancy: Benefits of Brief Speaker Exposure

At the crux of human communication lies the ability to map the physical speech signal onto stored representations in our mental lexicon. This task is far from trivial, in large part due to the prevalence of substantial variation in the pronunciation of words by speakers of different accents. For example, the word *ball* produced by an Australian-English speaker might sound more like *bowl* than *ball* to an American-English speaker's ears, and the Canadian pronunciation of the word *about* may sound more like *a boot* to an American listener. Thus, achieving parity with a conversational partner requires more than just speaking the same language, it also requires that listeners know enough about their interlocutors' speaking style to be able to accurately map distinct pronunciations onto the same underlying lexical representation. How do we, as listeners, solve this many-to-one mapping problem? And even more perplexingly, how do young children, who are just acquiring the phoneme inventory of their language and building up a rudimentary vocabulary, first begin to cope with speaker-related differences in the pronunciations of words? To date, this latter foundational issue in speech perception remains unsolved.

Listeners of all ages exhibit processing costs when confronted with accent-induced deviations in the pronunciation of words. For example, adults' word recognition abilities tend to be more accurate and efficient when target words are produced by a native speaker as opposed to a foreign-accented speaker (Adank, Evans, Stuart-Smith, & Scott, 2007; Munro & Derwing, 1995). The additional demands imposed by accent deviation, however, do not prevent fully competent adult language users from recognizing distinctly realized word tokens as referring to the same lexical item (Floccia, Goslin, Girard, & Konopczynski, 2006).

In contrast to adults, infants appear to exhibit far greater difficulty in coping with accent and other speaker-related variations in the realization of words (see Cristià et al., 2012, for an overview). For example, children under a year of age struggle to map tokens of the same word produced in different accents onto the same underlying word form representation (Schmale & Seidl, 2009; Schmale, Cristià, Seidl, & Johnson, 2010). And while North-American 15-month-olds readily recognize highly familiar words[1] such as *mommy* and *ball* presented in isolation in their own variant of English, they fail to do so when these words are produced in a Jamaican accent, indicating that unfamiliar accents pose serious difficulties for word recognition at this age (Best, Tyler, Gooding, Orlando, & Quann, 2009). It is not until several months later that American-English learning infants are able to recognize isolated familiar word forms produced in the unfamiliar Jamaican accent.

The picture that has emerged from these studies is that young infants are, at least initially, less adept than adults at dealing with phonetic divergence in the realization of words. In line with these data, models such as WRAPSA (Jusczyk, 1997) and PRIMIR (Werker & Curtin, 2005) have proposed that infants' word representations are episodic in nature. Although these models differ in exactly how speech is initially analyzed, they both suggest that word recognition is achieved by comparing the incoming speech signal against existing traces stored in the mental lexicon, in a fashion similar to exemplar-models of adult speech perception (e.g., Goldinger, 1996; 1998). These traces are activated in proportion to their acoustic similarity to the signal, such that an unstored *echo* (the combined output of all traces) arises during retrieval. According to these models, infants struggle with divergence in the realization of words because indexical details such as the speaker's voice and accent are preserved in the traces. As a result, phonetic deviations from

previously heard word tokens cannot easily be linked to these stored instances and are hence not recognized as referring to the same underlying representation. Only once listeners experience a substantial amount of variation in the pronunciation of a given word, and therefore have had a chance to store multiple phonetically distinct exemplars in their mental lexicon, will they be able to recognize variable word tokens as referring to the same word (see Rost & McMurray, 2009 for an example in infant word learning). In short, with more exemplars stored in memory (and activated during retrieval), the common properties among those traces are highlighted whereas the less consistent properties are de-emphasized. This way, episodic models can account for patterns of generalization without requiring prototypes or abstraction.

Although substantial evidence for exemplar-based storage of word forms has amassed in the past two decades (e.g., Bradlow, Nygaard, & Pisoni, 1999; Goldinger, 1998; Palmeri, Goldinger, & Pisoni, 1993), more recent approaches to adult language perception appear to converge on the inclusion of abstract components in listeners' mental word representations (Cutler, 2008; Goldinger, 2007; Pisoni & Levy, 2007), acknowledging the importance of both types of information. This shift is, at least in part, due to findings demonstrating that adults require only brief speaker exposure to adapt to the speaker's accent (Bradlow & Bent, 2008; Clarke & Garrett, 2004; Dahan, Drucker, & Scarborough, 2008; Eisner & McQueen, 2005; Kraljic & Samuel, 2005; Maye, Aslin, & Tanenhaus, 2008; Mitterer & McQueen, 2009, Trude & Brown-Schmidt, 2012), using top-down lexical knowledge to work out the mapping between acoustically variable spoken words and their corresponding underlying lexical representations (e.g., Norris, McQueen, & Cutler, 2003). Listeners presented with an accented speaker, for example, become increasingly efficient at

understanding that person's speech over time (e.g., Bradlow & Bent, 2008; Clarke & Garrett, 2004; Mitterer & McQueen, 2009). This experience with the speaker does not need to be extensive; sometimes as little as two to four sentences can be sufficient for listeners to accommodate that speaker's characteristics (Clarke & Garrett, 2004). Listeners thus rapidly become familiar with the specific characteristics of a speaker (or a group of speakers) and immediately exploit this information to tune in to the speaker they are listening to, such that speaker comprehension is enhanced. For this reason, understanding familiar speakers tends to be easier than understanding unfamiliar speakers (Bradlow & Pisoni, 1999; Nygaard & Pisoni, 1998; Nygaard, Sommers, & Pisoni, 1994).

What are the mechanisms inducing such speaker tuning? Studies showing that listeners accommodate changes in the realization of segmental information have claimed that the remapping takes place at the prelexical level (e.g., Norris et al., 2003; also see Cutler, 2008). According to this view, adjustments in the links between the acoustic signal and a phonemic category ensure that the intended phoneme category is accessed, even if the realization of that phoneme is different from that of typical speakers. The phonemic information then feeds forward to the lexical level, where the phonological form of the word can be activated. Such processes cannot easily be handled by purely exemplar models. Models such as WRAPSA (Jusczyk, 1997) and Goldinger's (1998) implementation of MINERVA 2 (Hintzman, 1986) have postulated that unstored echoes activated at retrieval could essentially simulate much of human speech perception without the need for prototypes or prelexical units. In the absence of any abstraction in the mental lexicon, prelexical phonemic remapping effects such as those observed during speaker accommodation are difficult to explain. For this reason, a mental lexicon consisting

exclusively of individual episodes has been argued to be insufficient to account for speaker adaptation (e.g., Cutler, 2008; Cutler, Eisner, McQueen, & Norris, 2010).

The finding that adults readily accommodate accented speakers and hence quickly overcome accent-imposed difficulties during speech perception stands in apparent contrast to the reported literature on infants' difficulty contending with unfamiliar accents. Previous studies showing that phonetic dissimilarity impedes infants' word recognition (Best et al., 2009), however, have not entertained the possibility that that infants, too, could potentially benefit from brief speaker exposure. By presenting infants with lists of isolated word tokens rather than fluent speech, past infant studies may have failed to provide young language users with the rich information normally contained within the speech signal encountered in everyday listening conditions. This lack of exposure to the speaker producing words in the unfamiliar accent may have prevented infants from adapting to the unfamiliar accent, much as exposure to only isolated words hinders adult word recognition (see Van Heugten & Johnson, 2012 for a related argument concerning infant speech perception).

Evidence for the view that various types of context might help drive adaptation by young listeners is provided by a word recognition study with North-American 19-month-olds (White & Aslin, 2011). In this study, toddlers were presented with a speaker speaking an accent differing minimally from the children's own accent. That is, the accent simply involved a single vowel shift from [a] to [æ], such that words such as *bottle* and *block* were produced as *battle* and *black*. By shifting only one vowel, the researchers were able to carry out a carefully controlled examination of whether accent exposure would help children adapt to unfamiliar pronunciations of words. Prior to the test phase, toddlers were shown

pictures of familiar objects on a TV screen, each of which was separately and unambiguously labeled. Crucially, one group of toddlers was presented with labels containing the [a] → [æ] vowel shift used in the test phase whereas another group of toddlers was presented with the same words produced with the original vowel. Given evidence that children implicitly access the names of visually presented objects (Mani & Plunkett, 2010), the exposure phase of this paradigm likely provided the 19-month-olds presented with the vowel shift with a visual key to work out the mapping between their own variant of English and the vowel-shifted variant they were being presented with in the lab. At test, infants who heard the unshifted vowels in the familiarization phase failed to show any increase in their looks toward the referent after the target had been named (much like mispronunciations have previously been shown to negatively affect word recognition; e.g., Swingley & Aslin, 2000; White & Morgan, 2008). In contrast, children who had been presented with the accented pronunciations in the training phase did recognize the test words with the shifted vowels, suggesting they had adapted to the speaker. This adaptation generalized to items not heard during training and was not the result of infants simply becoming more tolerant of all vowel deviations: test items in which the vowel [a] had been replaced by the vowel [I] were not recognized, suggesting that infants had shifted the [a] category to include [æ], but not expanded the category to include all deviant vowels. This study was the first to demonstrate that children as young as 19 months of age can adapt to unfamiliar pronunciations of words and is consistent with the notion that by 19 months of age, children have stored abstract phonological representations in their developing mental lexicon, enabling them to work out the mapping between their native and an unfamiliar variant of English (Best et al., 2009; also see Floccia, Delle Luche, Durrant, Butler, &

Goslin, 2012 and Johnson & Oczak, 2012, for related work examining word recognition in 24-month-old children with routine exposure to multiple regional accents). It is unclear, however, when and how the ability to cope with accent variation develops.

While the finding that toddlers adapt to a speaker's single vowel shifts provides us with the crucial insight that young children possess the fundamental ability to accommodate speakers, it does not answer the question of whether children would similarly adapt to accent variation in more naturalistic settings. Differences between regional accents, for example, often concern more than just a single phoneme shift. In addition, the majority of the speech infants and toddlers hear consists of multiword utterances (Aslin, 1993). Exposure to a speaker producing only isolated words that all contain a single vowel shift is thus not very likely to occur. Moreover, in the visual and auditory complexity of the real world, children are not typically presented with clear one-to-one mappings between words and their referents (Medina, Snedeker, Trueswell, & Gleitman, 2011). The question hence arises whether exposure to a speaker with a distinct unfamiliar accent similarly enables adaptation. In other words, can children deal with more extreme accent-induced phonetic mismatches in the realization of words when given appropriate naturalistic exposure to the speaker? And finally, when does this ability arise? Can infants only adapt to novel accents at around 19 months, once they have reportedly first developed phonologically abstract representations (as has been suggested by Best et al., 2009)? Or does the ability to cope with accent variation begin to emerge earlier, in line with the idea that perhaps younger infants might also be capable of at least some degree of abstraction?

Here, we examine the possibility that exposure to fluent Australian-accented speech enables Canadian English-learning 15-month-olds to work out the mapping between their

native accent and the unfamiliar Australian accent. If adaptation abilities such as those described above also surface for more distinct accents, this would indicate that infants can overcome previously observed difficulties with recognizing words in unfamiliar accents. That is, although 15-month-olds may have trouble creating signal-to-word maps for a distinct regional accent when words are presented out of context (Best et al., 2009), they may at the same time possess the ability to contend with this type of speaker variability when the listening conditions are more ecologically valid. Evidence for speaker adaptation in infancy might call into question the view that infants' early word representations are purely exemplar-based and would be consistent with the idea that that speaker information may be encoded prelexically from very early on.

Experiment 1

In Experiment 1, we test Canadian English learning 15-month-olds' ability to recognize word forms in Australian-accented English. Australian English is a variant of English that is phonetically and prosodically distinct from North-American English (Wells, 1982). Compared to North-American English, for example, Australian vowels tend to be more raised and fronted and [I] and [ə] are often merged in unstressed syllables. In addition, intonation patterns can also be vastly different between the two variants. Would this dissimilarity prevent Canadian 15-month-olds to recognize word forms in the unfamiliar accent, just like Jamaican-accented word forms are not recognized by American 15-month-olds (Best et al., 2009)?

Best and colleagues tested infants' ability to recognize accented words using an auditory-only Visual Fixation Procedure. To ensure consistency between their and our work, and to focus on the potential phonetic-to-phonemic remapping between accents in the

absence of visual information, the current study similarly utilized an auditory-only paradigm in which infants were presented with lists of multiple known (e.g, *bottle*) and lists of multiple nonsense words (e.g., *shammy*). In Experiment 1a, infants listened to the word lists presented in their own variant of English (i.e. Canadian English) whereas infants in Experiment 1b heard these words produced in an unfamiliar Australian English accent. Earlier studies using similar designs have shown that infants prefer listening to known over nonsense words when these words are produced in their native accent (Best et al., 2009, Hallé & De Boysson-Bardies, 1994; Swingley, 2005; Vihman, Nakai, DePaolis, & Hallé, 2004). However, at 15 months of age, no such preference is observed when words are produced in an unfamiliar accent (Best et al., 2009). We thus predict that in this study only infants in Experiment 1a (and not 1b) will listen longer to known than to nonsense words.

Method

Participants.

A total of 32 normally developing English-learning 14.5- to 15.5-month old infants from the Greater Toronto Area were tested, 16 in each of Experiments 1a and 1b (age range: 447-477 days; 17 boys). None of the parents reported any hearing issues or recent ear infections and none of the infants in Experiment 1b had had any substantial exposure to Australian-accented English, as established by a language questionnaire at the end of the lab visit. An additional 4 infants (2 in each of Experiment 1a and 1b) were tested, but excluded from the analyses due to extreme fussiness. All participating infants in this and subsequent experiments received a certificate and a small gift.

Stimuli.

Infants were presented with eight word lists. Half of these lists (i.e. the known word lists) consisted of words generally known by infants at 15 months of age (*daddy, bottle, diaper, mommy, grandma, kitty, ball, dog, bath, kiss, cup, shoe*), as evidenced by an average word understanding of 90.1% (range 68.7%-100%) by 15-month-olds in the Lexical Development Norms for English (Dale & Fenson, 1996). The other half of the lists, referred to as the nonsense word lists (reflecting their status in the infants' lexicons), contained low-frequency words or phonotactically legal nonsense words expected to be unknown to the 15-month-olds (*koddy, dimma, dapper, mitty, guttle, shammy, bog, bap, deuce, kie, koth, brall*). In order to exclude potential biases due to preferences for specific speech sounds, the two types of word lists were matched in phonemes. Each list contained all twelve known or nonsense words, repeated twice for a total of 24 words. Monosyllabic and bisyllabic words were alternated and the position of each word varied across lists. For Experiment 1a, words were produced by a native female English speaker from the Greater Toronto Area. The word lists in Experiment 1b were produced by a native female English speaker born and raised in Sydney, Australia. Speakers were instructed to speak in an infant-directed fashion. For each speaker, known and nonsense words were matched in terms of word length and average pitch level of the stressed vowel. The Canadian speaker's known words were on average 559 ms long and the nonsense words 579 ms. The average pitch of the vowel was 358 Hz for the known words and 378 Hz for the nonsense words. For the Australian speaker, the average known word length was 516 ms and the average nonsense word length 518 ms. The average vowel pitch was 301 Hz for known words and 292 Hz for nonsense words. Words were equated for loudness and interspersed with silences of approximately 800 ms. All lists lasted 34.5 s.

Procedure.

Infants were tested individually using a variant of the Headturn Preference Procedure similar to the one previously used by Swingley (2005). Participants were seated on their caregivers' lap in the center of a dimly-lit double walled sound-attenuated IAC test booth. A red light was mounted at eye level on the panel in front of the infant. Each of the side panels held a blue light, with loudspeakers positioned directly underneath. First, the center light flashed. Once the infants oriented towards this light, the experimenter, who monitored the infants' behavior via a muted TV screen outside the booth, pressed a button to terminate the center light. This automatically initiated the flashing of one of the sidelights. The presentation of the word lists was contingent upon the infants' looking behavior; lists started playing once the infant oriented towards the sidelight and were presented until the infant looked away for two consecutive seconds or until the maximum trial length of 34.5 s was reached. Infants' orientation time to known and nonsense words was measured. The eight lists were presented in random order, each subject order created separately by our custom-developed software, with the restriction that there could be no more than two known or two nonsense word lists in a row. To avoid incurring spurious biases towards known words in our results, parents were naïve to the experimental predictions and listened to masking music mixed with experimental stimuli occurring at irregular intervals over closed headphones throughout the whole experiment. The experiment lasted approximately 2-3 minutes.

Results and Discussion

In line with past studies (Best et al., 2009), infants only listened longer to words than to nonsense words in Experiment 1a, when presented with a speaker of their own

accent, but not in Experiment 1b, where the speaker spoke in an unfamiliar accent (Figure 1, left panel). Specifically, all infants in Experiment 1a listened longer to lists of known words (average: 11.55 s) than to nonsense word lists (average: 6.29 s). Infants tested in Experiment 1b listened to known words for 9.39 s on average and to nonsense words for 10.06 s, with 6 out of 16 infants listening longer to the known words. A 2 x 2 ANOVA including the data of both Experiments 1a and 1b with word status (known vs. nonsense word) as a within-participant factor and accent familiarity (native vs. unfamiliar) as a between-participant factor revealed a main effect of word status ($F(1,30) = 7.039$; $p = 0.013$; $\eta_p^2 = .190$), indicating that infants preferred to listen to known over nonsense words. Crucially, this main effect was modulated by the interaction between word status and accent familiarity ($F(1,30) = 11.698$; $p = 0.002$; $\eta_p^2 = .281$). That is, as revealed by two-tailed planned paired samples t-tests, only infants presented with a Canadian speaker in Experiment 1a preferred to listen to known over nonsense words ($t(15) = 4.224$; $p = .001$; $r = .737$; mean difference = 5.26 s, 95% confidence interval (CI) [2.61, 7.92]). No such effect was obtained for infants tested with an Australian speaker in Experiment 1b ($t(15) = -.552$; $p = .589$; mean difference = -0.66 s, 95% CI [-3.23, 1.90]). Thus, while infants at 15 months of age recognize word forms in their own accent, the phonetic dissimilarity of an unfamiliar accent prevents them from recognizing isolated words in unfamiliar accents. These findings conceptually replicate previous work examining American infants' understanding of Jamaican-accented English (Best et al., 2009) with a different population and a different unfamiliar accent.

Experiment 2

Without any prior exposure to the speaker, the 15-month-olds in Experiment 1 failed to recognize word forms in a variant of English distinct from their own native accent. While accent-induced deviations in phonetic-to-phonemic mapping impede speech processing in adult listeners as well, adults can nonetheless recognize words in unfamiliar accents, albeit with slight delays compared to native-accented speech (e.g., Floccia et al., 2006). This raises the question of how and when infants develop more proficient word recognition abilities. Previous studies using a different population and a different accent (Jamaican English) have suggested that such competences develop between 15 and 19 months of age (Best et al., 2009). Of course, different accents vary in how distinct they are from the infants' target accent (see Clopper & Pisoni, 2004 for an applied example) and it is possible that Jamaican-accented English is more or less distinct from Connecticut English than Australian-accented English is from Canadian English. In order to verify that the accent and stimuli used in the current study gives rise to a similar developmental pattern and to more precisely assess when infants are able to overcome these phonetic differences, Experiment 2 examines the developmental trajectory of infants' word recognition in Australian-accented English using the same materials as in Experiment 1b. If the ability to recognize word forms in Australian English in the absence of speaker information develops as a function of linguistic maturity, the older the infants are, the greater their preference may be to listen to known over nonsense words.

Method

Participants.

An additional 32 normally developing English-learning infants from the Greater Toronto Area were tested. Sixteen of these infants were between 17 and 18 months of age

(age range: 523-548 days; 12 boys). The remaining 16 infants were between 21.5 and 22.5 months of age (age range: 662-686 days; 6 boys). As was the case in Experiment 1b, no hearing issues or recent ear infections were reported and none of the infants had had any substantial exposure to Australian-accented English. One additional 17.5-month-old and three additional 22-month-olds were tested, but were excluded from the analyses due to extreme fussiness.

Stimuli and Procedure.

The stimuli and procedure are identical to those in Experiment 1b.

Results and Discussion

To examine how Canadian English-learning infants' ability to recognize word forms in an Australian accent develops over time, infants' performance on this task was assessed using infants from all three age groups (including the 15-month-olds from Experiment 1b). The orientation times to known and nonsense words were submitted to a repeated measures ANCOVA with word status (known vs. nonsense word) as a within-participant factor and age (in days) as the covariate. Although there were no main effects of word status and age, the interaction between word status and age was significant ($F(1,46) = 4.605$; $p = .037$), suggesting that over time, infants start preferring to listen to known over nonsense words (see Figure 2). To facilitate comparison between these results and previous research, and to examine infants' abilities to recognize Australian-accented words at particular ages, follow-up analyses were conducted separately for each of the age groups. Two-tailed planned paired samples t-tests indicated that the effect of word status only reached significance at 22 months of age ($t(15) = 2.663$; $p = .018$; $r = .567$; mean difference = 3.43 s, 95% CI [0.68, 6.18] for 22-month-olds; $t(15) = .281$; $p = .783$; mean difference = 0.46 s, 95% CI

[-3.01, 3.92] for 17.5-month-olds; see Figure 3). Similar to previous studies (Best et al., 2009) children gain the ability to cope with unfamiliar accents during word recognition sometime between 15 months of age and their second birthday. The finding that the words in the Australian accent can be recognized by just slightly older Canadian English-learning infants than those tested in Experiment 1 shows that the Australian English words used in the present study can be understood by young children (just like the Jamaican English words were understood by 19-month-olds). This raises the question of whether this shift in infants' ability to recognize word forms in an unfamiliar accent is necessarily caused by a qualitative shift in the nature of infants' word representations, as argued by Best and colleagues (2009). Previous work has shown that 19-month-olds can adapt to an accent consisting of a single vowel shift when known words are explicitly labeled in the new accent (White & Aslin, 2011). If the reason for success at this task is that children first develop abstract phonological representations only around 19 months of age, then younger children should be unable to deduce the phonetic-to-phonological mapping of an unfamiliar accent. Experiment 3 addresses this possibility by providing infants with exposure to the characteristics of the accent prior to test. If 15-month-olds are indeed unable to accommodate unfamiliar accents, they should fail to recognize accented words regardless of any exposure to the characteristics of the Australian accent. In contrast, if 15-month-olds, like adults, are capable of learning the mapping between an unfamiliar accent and their own native accent, speaker exposure may enable young infants to recognize the accented word forms in the test phase.

Experiment 3

Experiments 1 and 2 reveal that infants experience difficulty recognizing words in an unfamiliar Australian variant of English until at least 18 months of age. Infants in those experiments, however, only heard a total of twelve isolated words with no context. This presents them with little information regarding characteristics of the accent, potentially making it impossible for them to adapt to the accent. Given that only brief experience with a speaker can help adult listeners accommodate that speaker's accent (Bradlow & Bent, 2008; Clarke & Garrett, 2004; Dahan et al., 2008; Eisner & McQueen, 2005; Kraljic & Samuel, 2005; Maye et al., 2008; Mitterer & McQueen, 2009, Norris et al., 2003; Trude & Brown-Schmidt, 2012), it is possible that infants, too, might be better able to cope with unfamiliar accents after having benefited from speaker exposure. That is, prior access to the characteristics of an accent may help infants work out the inter-accent signal-to-word maps, which they can later use to recognize the test items. Thus, in Experiment 3, we investigate whether Canadian 15-month-olds, who experience difficulty recognizing words in Australian English, can adapt to the speaker's accent, thereby overcoming their inability to recognize word forms in an unfamiliar accent. In order to examine this, Canadian-English-learning 15-month-olds were tested on the same Australian-accented test stimuli used in Experiments 1b and 2. This time, however, the test phase was preceded by a two-minute recorded reading of 'The Very Hungry Caterpillar' storybook (Carle, 1969), read by the same Australian speaker who had recorded the test items. Crucially, the story did not contain any of the words used in the test phase. In order for infants to succeed at recognizing the known words in the test phase, simply memorizing the Australian pronunciation of the story words would thus not suffice. Instead, infants would need to

create a more generalized remapping strategy that could also be applied to the test words, none of which they had never heard spoken by the Australian speaker. If the exposure video indeed induces such form of accent adaptation, infants should prefer to listen to the known word list over the nonsense word list in the test phase.

Method

Participants.

Sixteen normally developing English-learning 14.5- to 15.5-month old infants from the Greater Toronto Area were tested (age range: 449-473 days; 10 boys). As in Experiments 1b and 2, none of the parents reported any hearing issues or recent ear infections and none of the infants had had any substantial exposure to Australian-accented English. An additional 4 infants were tested, but were excluded from the analyses due to extreme fussiness.

Stimuli and Procedure.

Test phase stimuli were identical to those in Experiment 1b and 2. Prior to test, however, infants were presented with a video recording of 'The Very Hungry Caterpillar' story read in an infant-directed fashion by the same Australian speaker who also recorded the isolated words. The two-minute exposure video displayed the speaker from the shoulders up against a green background. Infants sat on their caregivers' lap and watched the video on a large TV screen. The movie continuously played until the end, after which the test phase started.

Results and Discussion

The results of this experiment indicate that 15-month-olds failed to adapt to the speaker's accent after hearing a 2-minute story produced by the same Australian-accented

speaker (see Figure 1, middle panel). That is, infants listened to lists of known words for an average of 8.93 s and to lists of novel words for an average of 8.09 s, with 8 out of 16 infants listening longer to the known words. These orientation times are not significantly different from one another ($t(15) = .627$; $p = .540$; mean difference = 0.84 s, 95% CI [-2.03, 3.72]).

Experiment 3 shows that the two minutes of exposure to 'The Very Hungry Caterpillar' story did not enable 15-month-old infants to tune into the inter-accent mapping between their own native and an unfamiliar Australian accent, at least not to an extent that it would help them recognize previously unheard words. There are at least two possible explanations for why this may be the case. One possibility is that speaker exposure may not be beneficial at this age. Perhaps, in line with previous reports (Best et al., 2009; Jusczyk, 1997; Werker & Curtin, 2005), infants' early word representations may lack the abstract components needed to accommodate unfamiliar accents and accent exposure may only be used to work out the signal-to-word maps once infants have developed more abstract representations. If this were the case, this would imply a discontinuity between the word representations early in life and those later in life and may leave one wondering how early episodic representations would develop into more mature representations that carry abstract components (Cutler, 2008).

Alternatively, exposure to 'The Very Hungry Caterpillar' story may not have induced adaptation because the story contains many key words 15-month-olds are unlikely to know (e.g., *caterpillar, pickle, cocoon*). In fact, only 32 of the 107 story words occur on the Words and Gestures Communicative Development Inventory (CDI) and only 7 of these items occur within the 150 most frequently known words, as indicated by Lexical

Development Norms for English (Dale & Fenson, 1996). Productive vocabulary scores collected from parental reports further indicated that the 15-month-olds in Experiment 3 on average produced less than 2% of these relatively easy 32 story words included in the CDI (also see Van Heugten, 2012, for findings that even 20-month-old Canadian English learners drawn from the same population experience difficulty understanding the easiest of the storybook words such as *strawberry* and *butterfly* in their own accent). Both adult listeners and school-aged children utilize lexical information to infer the intended phonological category of an ambiguous sound and hence to tune in to the speaker's realization of sound patterns (Eisner & McQueen, 2005; McQueen, Tyler, & Cutler, 2012; Norris et al., 2003). If lexical feedback is a prerequisite for accent adaptation even in infancy, infants' inability to accommodate the Australian accent may be due to an inability to access the words in the exposure phase. Not recognizing words in the exposure phase, in other words, would be comparable to adults listening to nonsense words and would prevent tuning into an unfamiliar accent (Norris et al., 2003). Accent adaptation may thus only be only educed once infants recognize a sufficient number of words spoken in the unfamiliar accent. Experiment 4 examines this possibility.

Experiment 4

In Experiment 4, we test the possibility that it was infants' inability to recognize word forms in the exposure phase of Experiment 3 that prevented them from adaptation. If this were the case, then increased familiarity with the words in the story may allow infants to access those words in the exposure phase, even when produced in Australian-accented English. Accessing the words in the exposure phase may in turn allow infants to use the phonological code needed to guide accent accommodation. In order to accomplish this and

yet test the same population under the same listening conditions, infants would essentially have to gain greater knowledge of the words in the story. Given that storybook reading can prompt word learning in children (Ganea, Allen, Butler, Carey, & DeLoache, 2009; Horst, Parsons, & Bryan, 2011), one way for infants to learn words occurring in the story would be to present them with 'The Very Hungry Caterpillar' storybook before being exposed to and tested on the same materials as those used in Experiment 3. To familiarize infants with the story, parents were asked to read the book at home to their child once a day for the two weeks prior to their lab visit. Note that in order for the storybook reading to assist speaker accommodation, infants do not need to learn the meaning of these words. That is, knowledge of just the sound patterns of the words may be sufficient to create a phonological representation of the word. By comparing the Australian word forms to these stored representations, infants may be able to deduce the phonetic-to-phonological mapping of the unfamiliar accent. During their visit, at which point infants had heard the story read to them a minimum of 14 times in their own Canadian English accent, infants were exposed to the exact same Australian-accented exposure video and test trials from Experiment 3. Thus, just like the infants in Experiment 3, infants in the current experiment heard the story in Australian English for the first time in the lab. If speaker exposure enables infants to work out the mapping between their native and the unfamiliar accent, but only when they have access to lexical information, then being familiar with the story should cause infants to listen longer to the word list than the nonsense word list in the test phase.

Method

Participants.

Sixteen normally developing English-learning 14.5- to 15.5-month old infants from the Greater Toronto Area were tested in Experiment 4 (age range: 448-469 days; 7 boys). As in Experiments 1b, 2, and 3, none of the parents reported any hearing issues or recent ear infections and none of the infants had had any substantial exposure to Australian-accented English. An additional 3 infants were tested, but excluded from the analyses due to extreme fussiness. To maximize the likelihood that parents did read the storybook to their infants on a daily basis, they were asked to complete a diary, which inquired about the time of day the story was read and the person who read the story. The data from an additional two infants were replaced due to the parents' failure to read the book at home once a day for two weeks. According to the diaries, all infants in the final sample listened to the story at least 14 times before participating in the story. Reading was, for the most part, carried out by direct family members.[2]

Stimuli and Procedure.

Stimuli and procedure were identical to Experiment 3.

Results and Discussion

As can be seen in Figure 1 (right panel), infants in Experiment 4 preferred to listen to the known words over the nonsense words. Infants listened to known words for an average of 15.23 s and to nonsense words for an average of 8.90 s, with 14 out of 16 infants listening longer to the known words. This difference is statistically significant ($t(15) = 3.616$; $p = .002$; $r = .682$; mean difference = 6.33 s, 95% CI [2.62, 10.04]), indicating that infants do adapt to the Australian speaker's accent after hearing the familiar story.

In order to verify that infants in Experiment 4 did behave differently from those in Experiment 3, a mixed model ANOVA with word status as a within-participant factor and

exposure story familiarity as a between-participant factor was conducted. This yielded a main effect of word status ($F(1,30) = 10.624$; $p = 0.003$; $\eta_p^2 = .262$). Importantly, however, this finding was qualified by familiarity to the storybook ($F(1,30) = 6.212$; $p = 0.018$; $\eta_p^2 = .172$), demonstrating that only infants familiarized with the story at home preferred to listen to known over nonsense word lists.

General Discussion

Accent-related variation in the phonetic realization of words is ubiquitous in everyday human communication. Previous research has shown that such between-speaker variability in the realization of words may greatly hamper infants' early word recognition (Best et al., 2009; Schmale & Seidl, 2009; Schmale, et al., 2010). However, in order for children to develop mature communicative abilities, they must learn to contend with this type of variation. Here, we examined whether Canadian infants, at the early stages of language development, could deduce the signal-to-word mapping between their own Canadian and an unfamiliar Australian variant of their language. The results of this study indicate that access to characteristics of a speaker's accent can indeed assist infants' early word recognition abilities. With only two minutes of exposure to the speaker, 15-month-olds can accommodate a speaker's natural accent, as observed by infants' subsequent preference for known over nonsense words. Just like adults (e.g., Bradlow & Bent, 2008; Clarke & Garrett, 2004) and older children (McQueen et al., 2012; Schmale, Cristià, & Seidl, 2012; White & Aslin, 2011), 15-month-olds are thus able to exploit experience with an accented speaker to tune in to the speaker's realization of words. This suggests that from a very early age, infants are able to contend with accent idiosyncrasies and hence possess the fundamental skills for making them proficient communicators.

Crucially, infants in Experiment 4 recognized words spoken in an unfamiliar accent in the test phase even though none of those words occurred in the exposure story, indicating that the adaptation to the speaker's word realizations in the exposure phase generalized to previously unheard words (in that accent). In line with abstractionist models in the adult literature, then, this may mean that the early lexicon consists of word representations that are suitably abstract to deal with phonetic variability due to between-accent differences. According to these models, processes such as speaker adaptation do play an important role during speech encoding, but these speaker-dependent signal-to-word mapping strategies are established at a prelexical level. Once sounds have been categorized, segmental patterns are sent to the lexical level, and lexical items can be accessed. Even if infants' early word representations are not yet fully mature, the results from this study are consistent with the notion that infants store abstract linguistic representations in their mental lexicon and that these representations are sufficiently robust to allow infants to accommodate unfamiliar accents.

Episodic models of early speech perception also have the potential to explain the formation of links between the story words produced in the infants' own accent and the ones produced in an unfamiliar Australian accent. For example, by retrieving the read-at-home versions of the story from memory during the exposure phase, infants can – at least theoretically – compare the two variants, co-activate the Canadian and Australian story word exemplars, and map them onto one another. However, purely exemplar-based models that do not include a pre-lexical layer would have difficulty explaining how infants establish phonetic-to-phonemic mappings between accents that are sufficiently generalizable to test items not occurring during exposure (as observed in Experiment 4). In

order for such prelexical adjustments to take place, speech sounds would necessarily have to be analyzed at a segmental level, which, in the absence of recoding speech into abstract units prior to word access, may be challenging to explain. While our results thus call into question the viability of extreme exemplar models (Goldinger, 1996; 1998), they do, of course, not preclude the possibility that early word representations contain episodic information. A model proposing a combination of both exemplar and abstract components would, however, align nicely with the recently proposed hybrid nature of adult representations (Goldinger, 2007).

The effect of speaker accommodation in the current study is consistent with the idea that infants deduce the mapping between linguistic properties of the Canadian and Australian accents. An alternative interpretation of our results, however, could be that infants in Experiment 4 simply relaxed their criteria for word access (see Schmale et al., 2012 for a related discussion). Although previous work with toddlers and adults has argued that listeners are fairly specific in their adaptation skills, and the ability to recognize accented words is not simply a product of vowel expansion (Maye et al., 2008; White & Aslin, 2011; Trude & Brown-Schmidt, 2012), it is possible that the younger infants in our study noticed the dramatic divergence of the Australian English speaker from the Canadian standard, but were unable to determine the exact manifestation of these differences. To nonetheless accommodate the speaker, infants could have become more tolerant of phonetic deviation by broadening the mapping between the speaker's phoneme realizations and the underlying linguistic representations, by expanding their lexical prototypes, or by simply lowering the activation threshold needed to access words. Pure episodic models by definition neither feature phoneme representations nor prototypes, but could potentially

adjust the echo intensity threshold for word recognition and might hence be able to account for word recognition in Experiment 4. Note, however, that although we cannot fully rule out this interpretation, any decrease in activation threshold should result not only in accepting linguistically irrelevant deviations from the norm, but also in accepting deviations crossing phonological boundaries. This should, in turn, lead to a reduced detection of mispronunciations. Because a good part of our nonwords differ in just a single phoneme from highly frequent words infants at 15 months of age may know (i.e. *mitty-kitty, bog-dog, bap-bath, deuce-do, kie-I/my, brall-ball*), one may expect that tolerance for deviation diminishes the difference in orientation time between known and nonsense words in Experiment 4 compared to Experiment 1a. Our results show that this is not the case. Even so, future work using paradigms that measure label-object mapping such as the Preferential Looking Procedure should examine whether infants are indeed more vulnerable to accepting mispronunciations in distinct unfamiliar accents than in their native accent (see White & Aslin, 2011, for evidence regarding toddlers' sensitivity to mispronunciations in novel accents). In addition, by testing whether exposure to Australian English may allow infants to generalize to speakers of other – distinct – accents, further experiments could clarify whether the accommodation of accented speakers involves a directional shift in phonemic mapping between the signal and the linguistic representation or whether infants simply tolerate more deviation. If infants accommodate the specific deviations from their own native accent, no cross-accent generalization should be observed (as is the case for adults; Bradlow & Bent, 2008). If, in contrast, infants relax their criteria for word recognition, experience with one accent may help word recognition in another accent.

Despite the vast amount of work on perceptual learning in speech in adulthood (e.g., Eisner & McQueen, 2005; Kraljic & Samuel, 2005; Norris et al., 2003) and the theoretical implications infants' ability to accommodate accents may have on the nature of early word recognition, only few studies to date have examined the possibility that speaker exposure allows infants or young children to develop an experience-induced strategy for decoding linguistic information conveyed by the speaker (McQueen et al., 2012; Schmale et al., 2012; White & Aslin, 2011). None of these studies have directly examined the mechanism responsible for such adaptation processes. The finding in the current study that speaker exposure only assists word recognition with increased familiarity to the story, however, allows us to speculate what these mechanisms may be. At least two explanations emerge. First, as discussed before, it is possible that without recent routine exposure to the story, the number of words infants knew (and accessed) in the exposure phase was insufficient to allow for lexically-guided accent adaptation. This is plausible given that even children older than 15 months of age, drawn from the same Canadian English learning population, have been found to experience difficulty recognizing the easiest of the story words (Van Heugten, 2012). The unfamiliar words in the story would essentially be treated the way adult listeners treat nonsense words and would hence fail to prompt adaptation. Infants who were read the storybook for a minimum of 14 times, in contrast, presumably became familiar with many of the previously unfamiliar words in the story. Even if they did not learn the exact meaning of these words, the knowledge of the sound patterns of these words may have allowed more word forms to be accessed during the exposure phase, which in turn may have induced adaptation to the unfamiliar accent. Alternatively, encoding a familiar story may be less demanding than encoding a less familiar story, such that more

processing resources may have been available to recognize deviant pronunciations of words (see e.g., Fennell, 2012; Fennell & Werker, 2003 for evidence showing that decreased processing load can dramatically alter young children's ability to process linguistic materials). For example, it is possible that recognition of a story activates word forms associated with that story. These increased activation levels (in Experiment 4) due to decreased processing load (compared to Experiment 3) would arguably facilitate word recognition, such that words can be recognized despite the phonetic divergence in surface forms. Note that although it is in theory also possible that decreased processing load enables infants to better compute the properties of the speaker's phoneme distributions without lexical information being involved, the purely bottom-up extraction of phonological regularities has been argued to be insufficient to establish phonological categories in the complexity of speech in naturalistic settings (e.g., Feldman, Griffiths, & Morgan, 2009; Martin, Peperkamp, & Dupoux, 2013). It is thus likely that lexical information is essential for infants to adapt to accented speech. The notion that speaker accommodation may be lexically mediated is consistent with adult work showing similar effects (Davis, Johnsrude, Hervais-Adelman, Taylor, & McGettigan, 2005; Eisner & McQueen, 2005; Mitterer & McQueen, 2009; Norris et al., 2003) and although not explicitly tested, aligns with the finding that 19-month-olds adapt to a speaker's vowel shift after exposure to labeled pictures including that shift (White & Aslin, 2011). If 15-month-olds, whose vocabularies are far from mature, indeed require lexical access to accommodate unfamiliar accents, this would highlight the importance of lexically-guided adaptation strategies across the life span. In ongoing work, we are further examining the role of lexical information for adapting to accented speech. In addition, we also aim to

explore how much and what type of exposure is necessary for infants to induce accent accommodation.

The current set of findings not only raises the question of *how* infants accommodate accented speech but also exactly *what* infants are adapting to. In Experiment 3 and 4, the exposure story and the test items were produced by the same Australian-accented speaker. This means that infants' ability to recognize the Australian-accented words in Experiment 4 could have been the result of either speaker adaptation or accent adaptation. Specifically, appropriate exposure to one Australian speaker induced a better understanding of that same accented speaker, but it is unclear whether it would also generalize to a different Australian speaker. The adult literature suggests that adaptation to one speaker does not always transfer to another speaker of the same accent and that only exposure to multiple different speakers of the same accent induces full speaker-independent accent adaptation (Bradlow & Bent, 2008). Whether infants would follow a similar generalization pattern is an empirical question and should be addressed in future work. Either way, infants' ability to overcome difficulty understanding speakers in unfamiliar accents after only two minutes of exposure suggests that in everyday life, they might be able to quickly accommodate accented speakers of their own language, even in the absence of long-term exposure to their accent. This makes speech perception in infancy remarkably efficient.

The finding that infants accommodate unfamiliar accents after brief experience with a speaker's accent also has implications for children growing up in environments where multiple regional variants of their native language are spoken. Recent work suggests that at the early stages of word recognition, children with routine exposure to more than one accent are somewhat slower to recognize familiar words in the regionally dominant version

of their language than children with exposure to only the regionally dominant accent (Johnson & Oczak, 2012; see Floccia et al., 2012, for related work). However, if children exposed to multiple accents are able to rapidly adapt to the regionally dominant version of the language after limited speaker experience, practical consequences of mixed accent input may be minimal. In fact, the early challenge to recognize words in the regional accent outside their homes and the continuous need for linking word pronunciations across different accents may result in more beneficial speaker adaptation abilities over time. This may hold even for unfamiliar accents, such that children receiving multi-accent input early in life may be better able to contend with accent variation at a later age. Studies in our lab are currently addressing this possibility.

Experiment 2 clearly displays a developmental trajectory in infants' ability to recognize Australian-accented words in the absence of prior speaker experience. Together with the finding that North-American infants learn to recognize Jamaican-accented words over time (Best et al., 2009), this provides convergent evidence that infants start recognizing isolated words in unfamiliar accents in the second half of their second year of life, even if they are not familiar with the accent at hand. What causes this developmental pattern to occur? In other words, why do 15-month-olds require prior speaker exposure to recognize word forms in unfamiliar accents, when only slightly older children are able to recognize the same accented word forms without having heard the speaker before? A previous proposal has put forward the idea that the development of the ability to recognize word forms in unfamiliar accents is triggered by a general qualitative shift in infants' early word representations when word forms become phonemic in nature (Best et al., 2009). However, given our findings that infants *can* recognize accented words after a mere two

minutes of experience with the speaker, it is possible that there is no drastic change in the nature of linguistic representations during the latter half of the second year of life. Rather, what might distinguish these older children from the younger ones is their enhanced (more precise) storage of the items presented to them at test (though see Mani & Plunkett, 2007; Swingley & Aslin, 2000; White & Morgan, 2008 for findings suggesting that early word representations are phonologically specific from early on). Simply by virtue of being older, infants' prior experience with words (including the test items) increases and this may further refine the underlying phonological forms of the words (cf. Mattock, Amitay, & Moore, 2010). Note that although this would involve a change in the word form representations, the shift should be viewed as no more than a fine-tuning process of the approximate linguistic categories established in the preceding months (e.g., Kuhl, 1991; Polka & Werker, 1994; Werker & Tees, 1984). Alternatively, older infants may have had more exposure to unfamiliar accents and their previous experience accommodating different accents could potentially help them recognize the Australian words 'on the fly', without exposure to speaker information being necessary (cf. Schmale et al., 2010). While prior access to the characteristics of the accent likely remains helpful (as supported by the findings that even adults' word recognition is enhanced after familiarity with an accented speaker; Bradlow & Bent, 2008; Clarke & Garrett, 2004), it is no longer needed to exceed the activation threshold required to access the words in this study. Either explanation would allow for greater continuity between infants' early word form representations and their mature counterparts later in life.

Mapping phonetically variable signals onto the same underlying linguistic representations is arguably one of the most impressive cognitive feats accomplished by

humans. While adults can recognize words despite ample variability in the pronunciation of words across accents, infants had been found to experience serious difficulty contending with accent variation. This study has shown that online perceptional learning can help infants confront this challenge. Specifically, speaker exposure allows even infants, who are only at the initial stages of learning to speak, to adapt to unfamiliar accents and hence work out sophisticated speaker-dependent signal-to-word maps. This could be taken as evidence for the view that word representations are sufficiently abstract to deal with phonetic variability from very early on and gives rise to the possibility that the transformation of word recognition from infancy into adulthood may be a gradual development that is solely quantitative in nature.

References

Adank, P., Evans, B. G., Stuart-Smith, J., Scott, S. K. (2009). Comprehension of familiar and unfamiliar native accents under adverse listening conditions. *Journal of Experimental Psychology Human Perception and Performance, 35*, 520-529. doi: 10.1177/0956797610389192

Aslin, R. N. (1993). Segmentation of fluent speech into words: Learning models and the role of maternal input. In B. Boysson-Bardies, S. de Schonen, P. Jusczyk, P. MacNeilage and J. Morton (Eds.), *Developmental Neurocognition: Speech and Face Processing in the First Year of Life* (pp. 305-315). Dordrecht: Kluwer Academic 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*, 539–542. doi: 10.1111/j.1467-9280.2009.02327.x

Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition, 106*, 707–729. doi: 10.1016/j.cognition.2007.04.005

Bradlow, A. R., Nygaard, L. C., & Pisoni, D. B. (1999). Effects of talker, rate, and amplitude variation on recognition memory for spoken words. *Perception & Psychophysics, 61*, 206–219. doi: 10.3758/BF03206883

Bradlow, A. R. & Pisoni, D. B. (1999). Recognition of spoken words by native and non-native listeners: Talker-, listener-, and item-related factors. *Journal of the Acoustical Society of America, 106*, 2074–2085.

Carle, E. (1969). *The very hungry caterpillar.* New York: Philomel Books.

Clarke, C. M., & Garrett, M. (2004). Rapid adaptation to foreign accented English. *Journal of the Acoustical Society of America, 116*, 3647–3658. doi: 10.1121/1.1815131

Clopper. C. G. & Pisoni, D. B. (2004). Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics, 32*, 111-140. doi: 10.1016/S0095-4470(03)00009-3

Cristià, A., Seidl, A., Vaughn, C., Schmale, R., Bradlow, A., & Floccia, C. (2012). Linguistic processing of accented speech across the lifespan. *Frontiers in Psychology, 3*, 479. doi: 10.3389/fpsyg.2012.00479

Cutler, A. (2008). The abstract representations in speech processing. *Quarterly Journal of Experimental Psychology, 61*, 1601–1619. doi: 0.1080/13803390802218542

Cutler, A., Eisner, F., McQueen, J. M., & Norris, D. (2010). How abstract phonemic categories are necessary for coping with speaker-related variation. In C. Fougeron, B. Kühnert, M. d'Imperio & N. Vallée (Eds.), *Laboratory Phonology 10: Variation, Phonetic Detail and Phonological Representation* (pp. 91–111). Berlin: Mouton de Gruyter.

Dahan, D., Drucker, S. J., & Scarborough, R.A. (2008). Talker adaptation in speech perception: adjusting the signal or the representations? *Cognition, 108*, 710–718. doi: 10.1016/j.cognition.2008.06.003

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers, 28*, 125–127. doi: 10.3758/BF03203646

Davis, M. H., Johnsrude, I. S., Hervais-Adelman, A. G., Taylor, K. J., & McGettigan, C. (2005). Lexical information drives perceptual learning of distorted speech: Evidence

From the comprehension of noise-vocoded sentences. *Journal of Experimental Psychology: General, 134*, 222–241. doi: 10.1037/0096-3445.134.2.222

Eisner, F., & McQueen, J. M. (2005). The specificity of perceptual learning in speech processing. *Perception & Psychophysics, 67*, 224–238. doi: 10.3758/BF03206487

Feldman, N. H., Griffiths, T. L., & Morgan, J. L. (2009). Learning phonetic categories by learning a lexicon. In N. A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 2208–2213). Austin, TX: Cognitive Science Society.

Fennell, C. T. (2012). Object familiarity enhances infants' use of phonetic detail in novel words. *Infancy, 17,* 339–353. doi: 10.1111/j.1532-7078.2011.00080.x

Fennell, C. T., & Werker, J. F. (2003). Early word learners' ability to access phonetic detail in well-known words. *Language and Speeech 46*, 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

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,* 1276–1293. doi: 10.1037/0096-1523.32.5.1276

Ganea, P. A., Allen, M. L., Butler, L., Carey, S., & DeLoache, J. S. (2009). Toddlers' referential understanding of pictures. *Journal of Experimental Child Psychology. 104*, 283–295. doi: 10.1016/j.jecp.2009.05.008

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1166–1183. doi: 10.1037/0278-7393.22.5.1166

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review, 105*, 251–279. doi: 10.1037/0033-295X.105.2.251

Goldinger, S. D. (2007). A complementary-systems approach to abstract and episodic speech perception. In J. Trouvain & W. J. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences* (pp. 49–54). Dudweiler, Germany: Pirrot.

Hallé, P. A., & De Boysson-Bardies, B. (1994). Emergence of an early receptive lexicon: Infants' recognition of words. *Infant Behavior and Development, 17*, 119–129. doi: 10.1016/0163-6383(94)90047-7

Hintzman, D.L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review, 93*, 411–428. doi: 10.1037//0033-295X.93.4.411

Horst, J. S., Parsons, K. L., & Bryan, N. M. (2011). Get the story straight: Contextual repetition promotes word learning from storybooks. *Frontiers in Psychology, 2*, 17. doi: 10.3389/fpsyg.2011.00017

Johnson, E. K. & Oczak, S. (2012). *Input matters: Speed of word recognition in toddlers receiving regular exposure to accented speech.* Talk presented at the 18th International Conference on Infant Studies, Minneapolis, MN.

Jusczyk, P. W. (1997). *The discovery of spoken language.* Cambridge, MA: MIT Press.

Kraljic, T. & Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology, 51*, 141–178. doi: 10.1016/j.cogpsych.2005.05.001

Kuhl, P. K. (1991). Human adults and human infants show a 'perceptual magnet effect' for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics, 50*, 93–107. doi: 10.3758/BF03212211

Mani, N. & Plunkett, K. (2007). Phonological specificity of consonants and vowels in early lexical representations. *Journal of Memory and Language, 57*, 252–272.

Mani, N. & Plunkett, K. (2010). In the infant's mind's ear: Evidence for implicit naming in 18-month-olds. *Psychological Science, 21,* 908-913. doi: 10.1177/0956797610373371

Martin, A., Peperkamp, S., & Dupoux, E. (2013). Learning phonemes with a proto-lexicon. *Cognitive Science, 37*, 103–124. doi: 10.1111/j.1551-6709.2012.01267.x

Mattock, K., Amitay, S., & Moore, D. R. (2010). Auditory development and learning. In C. Plack (Ed.), Oxford Handbook of Auditory Science (pp. 297-324). Oxford: Oxford University Press.

Maye, J., Aslin, R. N., & Tanenhaus, M. (2008). The weckud wetch of the wast: Rapid adaptation to a novel accent. *Cognitive Science, 32*, 543–562. doi: 10.1080/03640210802035357

McQueen, J. M., Tyler, M., & Cutler, A. (2012). Lexical retuning of children's speech perception: Evidence for knowledge about words' component sounds. *Language Learning and Development, 8*, 317–339. doi: 10.1080/15475441.2011.641887

Medina, T. N., Snedeker, J., Trueswell, J. C., & Gleitman, L. R. (2011). How words can and cannot be learned by observation. *Proceedings of the National Academy of Sciences, 108*, 9014–9019. doi: 10.1073/pnas.1105040108

Mitterer, H., & McQueen, J. M. (2009). Foreign subtitles help but native-language subtitles harm foreign speech perception. *PLoS One, 4*, A146–A150. doi: 10.1371/journal.pone.0007785

Munro, M. J., & Derwing, T. M. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech, 38*, 289–306. doi: 10.1177/002383099503800305

Norris, D., McQueen, J.M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology, 47*, 204–238. doi: 10.1016/S0010-0285(03)00006-9

Nygaard, L. C., & Pisoni, D. B. (1998). Talker-specific learning in speech perception. *Perception & Psychophysics, 60*, 355–376. doi: 10.3758/BF03206860

Nygaard, L. C., Sommers, M. S., & Pisoni, D. B. (1994). Speech perception as a talker-contingent process. *Psychological Science, 5*, 42–46. doi: 10.1111/j.1467-9280.1994.tb00612.x

Palmeri, T. J., Goldinger, S. D., & Pisoni, D. B. (1993). Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 309–328. doi: 10.1037/0278-7393.19.2.309

Pisoni, D. B., & Levi, S. V. (2007). Some observations on representations and representational specificity in speech perception and spoken word recognition. In M. G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 3–18). Oxford, UK: Oxford University Press.

Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 421–435. doi: 10.1037/0096-1523.20.2.421

Rost, G. C., & McMurray, B. (2009) Speaker variability augments phonological processing in early word learning. *Developmental Science, 12*, 339–7349. doi: 10.1111/j.1467-7687.2008.00786.x

Schmale, R., Cristià, A., & Seidl, A. (2012). Toddlers recognize words in an unfamiliar accent after brief exposure. *Developmental Science, 15*, 732–738. doi: 10.1111/j.1467-7687.2012.01175.x

Schmale, R., Cristià, A., Seidl, A., & Johnson, E. K. (2010). Developmental changes in infants' ability to cope with dialect variation in word recognition. *Infancy, 15*, 650–662. doi: 10.1111/j.1532-7078.2010.00032.x

Schmale, R., & Seidl, A. (2009). Accommodating variability in voice and foreign accent: flexibility of early word representations. *Developmental Science, 12*, 583–601. doi: 10.1111/j.1467-7687.2009.00809.x

Swingley, D. (2005). 11-month-olds' knowledge of how familiar words sound. *Developmental Science, 8*, 432–443. doi: 10.1111/j.1467-7687.2005.00432.x

Swingley, D., & Aslin, R. N. (2000) Spoken word recognition and lexical representation in very young children. *Cognition, 76*, 147–166. doi: 10.1016/S0010-0277(00)00081-0

Trude, A. M., & Brown-Schmidt, S. (2012). Talker-specific perceptual adaptation during online speech perception. *Language and Cognitive Processes, 27*, 979-1001. doi: 10.1080/01690965.2011.597153

Van Heugten, M. (2012). *Speech perception in early childhood: Contending with speaker and accent variation* (Unpublished doctoral dissertation). University of Toronto, Toronto, Canada.

Van Heugten, M., & Johnson, E. K. (2012). Infants exposed to fluent natural speech succeed at cross-gender word recognition. *Journal of Speech, Language, and Hearing Research, 55*, 554–560. doi: 10.1044/1092-4388(2011/10-0347)

Vihman, M. M., Nakai, S., DePaolis, R. A., & Hallé, P. A. (2004). The role of accentual pattern in early lexical representation. *Journal of Memory and Language, 50*, 336–353. doi: 10.1016/j.jml.2003.11.004

Wells, J. C. (1982). *Accents of English, volume 3: Beyond the British Isles*. Cambridge: Cambridge University Press.

Werker, J. F., & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language Learning and Development, 1*, 197–234. doi: 10.1080/15475441.2005.9684216

Werker, J., & Tees, R. (1984). Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant behavior and development, 7*, 49–63. doi: 10.1016/S0163-6383(84)80022-3

White, K.S. & Aslin, R.N. (2011). Adaptation to novel accents by toddlers. *Developmental Science, 14,* 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. doi: 10.1016/j.jml.2008.03.001

Footnotes

[1]Throughout this paper, we use the phrase ‘word recognition’ and ‘word form recognition’ to be synonymous, referring only to the recognition of the phonological form of a word. Of course, there are more – equally crucial – word components that play a role during word recognition (e.g., semantic or grammatical levels). While we certainly do not wish to diminish their importance, this paper solely focuses on the phonological layer and hence refrains from addressing word access on any other level.

[2]A total of 15 out of our 16 infants heard (at least) 14 repetitions produced by a native Canadian English speaker (though 2 infants heard additional versions in a nonnative accent). The one remaining infant consistently heard the story in a slight Trinidadian accent. Note however, that in order to recognize the Australian-accented speaker, this still required the infant to develop inter-accent signal-to-word maps between their family member’s English and our Australian-accented test speaker.

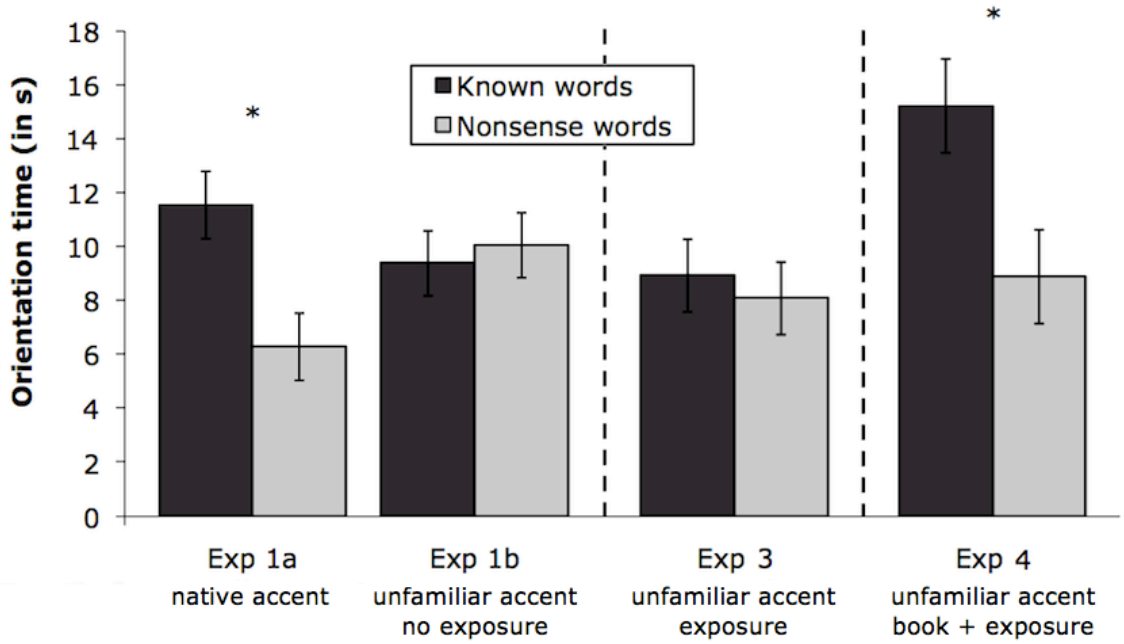

Figure 1. Orientation time in seconds (error bars indicate standard errors of the mean difference scores) for 15-month-olds to known and nonsense words in Experiments 1a (native accent), 1b (unfamiliar accent), 3 (unfamiliar accent with exposure), and 4 (unfamiliar accent with book familiarization and exposure).

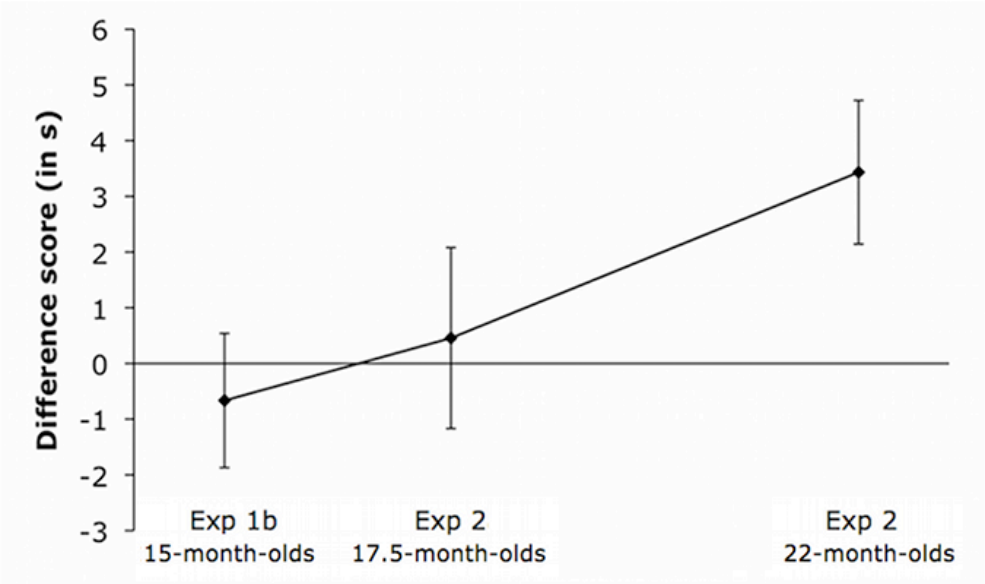

Figure 2. Orientation time difference between known and nonsense words (produced in Australian-accented English) in seconds (error bars indicate standard errors of the mean difference scores) in Experiments 1b (15-month-olds) and 2 (17.5-month-olds and 22-month-olds, respectively).

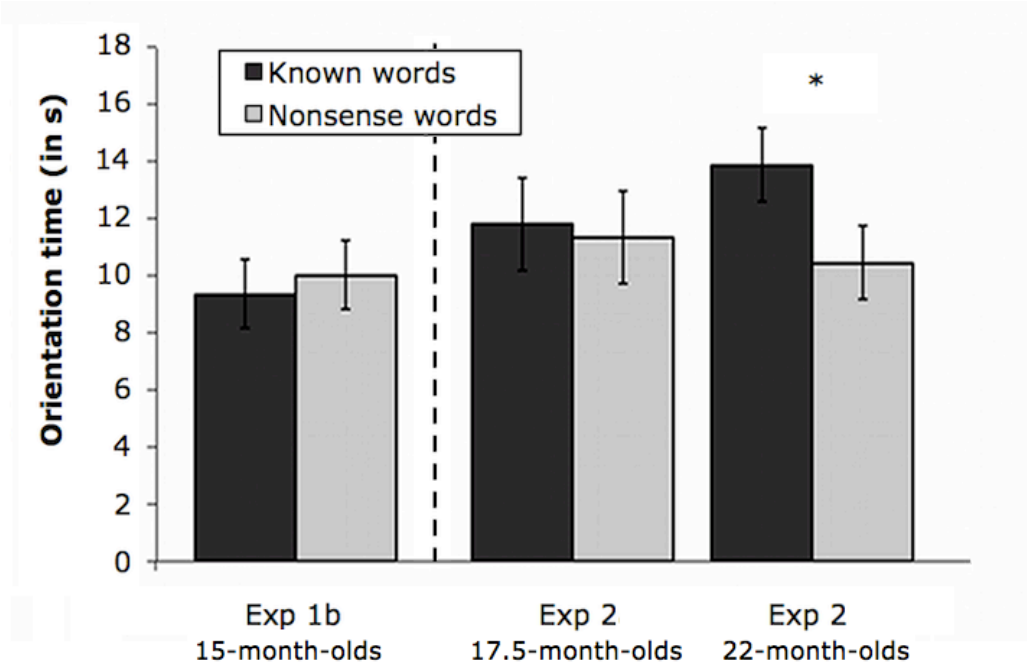

Figure 3. Orientation time in seconds (error bars indicate standard errors of the mean difference scores) to known and nonsense words in Australian-accented English in Experiments 1b (15-month-olds), and 2 (17.5- and 22-month-olds, respectively).